frontiers | Research Topics

From clinical trials to real-world data sciences: Evidence-based medicine for value in health

Edited by
Kevin Lu and Bin Jiang

Published in
Frontiers in Pharmacology
Frontiers in Public Health
Frontiers in Medicine

FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-83251-768-0
DOI 10.3389/978-2-83251-768-0

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

From clinical trials to real-world data sciences: Evidence-based medicine for value in health

Topic editors
Kevin Lu — University of South Carolina, United States
Bin Jiang — Peking University, China

Citation
Lu, K., Jiang, B., eds. (2023). *From clinical trials to real-world data sciences: Evidence-based medicine for value in health*.
Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83251-768-0

Table of contents

08 Editorial: From randomized clinical trials to real-world data and big data sciences: Generating evidence-based medicine for value in western and herbal medicines
Xiaomo Xiong, Gang Lv, Bin Jiang and Kevin Lu

11 Association Between the Health Locus of Control and Medication Adherence: An Observational, Cross-Sectional Study in Primary Care
Hanna-Maria E. Gerland and Tino Prell

20 Impact of Multimorbidity Subgroups on the Health Care Use and Clinical Outcomes of Patients With Tuberculosis: A Population-Based Cohort Analysis
Qin Chen, Yang Che, Yue Xiao, Feng Jiang, Yanfei Chen, Jifang Zhou and Tianchi Yang

30 Trends and Patterns of Antibiotics Use in China's Urban Tertiary Hospitals, 2016–19
Yulei Zhu, Yang Qiao, Rouli Dai, Xin Hu and Xin Li

47 Short Practical Regimen of Acupuncture for Melasma: A Prospective Cohort Study in a Tertiary Hospital in Thailand
Thanan Supasiri, Nuntida Salakshna and Krit Pongpirul

54 Methodological Characteristics of Clinical Trials Supporting the Marketing Authorisation of Advanced Therapies in the European Union
Carolina Iglesias-Lopez, Antònia Agustí, Antonio Vallano and Merce Obach

68 The Transition From Ambrisentan to Macitentan in Patients With Pulmonary Arterial Hypertension: A Real-word Prospective Study
Yusi Chen, Jun Luo, Jingyuan Chen, Eugene Kotlyar, Zilu Li, Wenjie Chen and Jiang Li

78 Corrigendum: The Transition From Ambrisentan to Macitentan in Patients With Pulmonary Arterial Hypertension: A Real-Word Prospective Study
Yusi Chen, Jun Luo, Jingyuan Chen, Eugene Kotlyar, Zilu Li, Wenjie Chen and Jiang Li

80 Comparison of Systemic Treatments for Metastatic Castration-Resistant Prostate Cancer After Docetaxel Failure: A Systematic Review and Network Meta-analysis
Junru Chen, Yaowen Zhang, Xingming Zhang, Jinge Zhao, Yuchao Ni, Sha Zhu, Ben He, Jindong Dai, Zhipeng Wang, Zilin Wang, Jiayu Liang, Xudong Zhu, Pengfei Shen, Hao Zeng and Guangxi Sun

89 Methylxanthine Treatment in Patients Hospitalized for Acute Exacerbation of Chronic Obstructive Pulmonary Disease in China: A Real-World Study Using Propensity Score Matching Analysis
Zijie Zhan, Yiming Ma, Ke Huang, Chen Liang, Xihua Mao, Yaowen Zhang, Xiaoxia Ren, Jieping Lei, Yan Chen, Ting Yang and Chen Wang

100 Effect of 0.01% Atropine on Accommodation in Myopic Teenagers
Huixia Li, Liying Zhang, Hong Tian, Song Zhang, Xueyan Zhang, Han Zhang, Yujing Chen, Wenping Qi, Xiaoying Wu, Hongmei Jiang, Hailong Yang, Yajun Yang, Lei Liu and Guisen Zhang

106 Efficacy and Safety of Collagenase Clostridium Histolyticum in the Treatment of Peyronie's Disease: An Evidence-Based Analysis
Dehong Cao, Jinze Li, You Lu, Yin Huang, Bo Chen, Zeyu Chen, Yinzhi Shen, Liangren Liu and Qiang Wei

114 The Impact of a Multifaceted Pharmacist-Led Antimicrobial Stewardship Program on Antibiotic Use: Evidence From a Quasi-Experimental Study in the Department of Vascular and Interventional Radiology in a Chinese Tertiary Hospital
Jinhui Xu, Jian Huang, YanXia Yu, Dayong Zhou, Ying Wang, Sudong Xue, Erning Shang, Jiantong Sun, Xinyuan Ding, Lu Shi, Lufen Duan, Lian Tang, Qin Zhou and Xin Li

126 Neoplasm Risk in Patients With Rheumatoid Arthritis Treated With Fostamatinib: A Systematic Review and Meta-analysis
Yuehong Chen, Huan Liu, Yunru Tian, Zhongling Luo, Geng Yin and Qibing Xie

137 Prediction of Atrial Fibrillation in Hospitalized Elderly Patients With Coronary Heart Disease and Type 2 Diabetes Mellitus Using Machine Learning: A Multicenter Retrospective Study
Qian Xu, Yan Peng, Juntao Tan, Wenlong Zhao, Meijie Yang and Jie Tian

146 Associations of Fentanyl, Sufentanil, and Remifentanil With Length of Stay and Mortality Among Mechanically Ventilated Patients: A Registry-Based Cohort Study
Wen Wang, Qiao He, Mingqi Wang, Yan Kang, Peng Ji, Shichao Zhu, Rui Zhang, Kang Zou and Xin Sun

154 One-Bed-One-Team—Does an Integrated General Hospital Inpatient Model Improve Care Outcomes and Productivity: An Observational Study
Yi Feng Lai, Shi Qi Lee, Yi-Roe Tan, Zheng Yi Lau, Jason Phua, See Meng Khoo, Satya Pavan Kumar Gollamudi, Cher Wee Lim and Yee Wei Lim

161 Statin Therapy for Hyperlipidemic Patients With Chronic Kidney Disease and End-Stage Renal Disease: A Retrospective Cohort Study Based on 925,418 Adults in Taiwan
Fung-Chang Sung, Ying-Chin Jong, Chih-Hsin Muo, Chih-Cheng Hsu, Wen-Chen Tsai and Yueh-Han Hsu

171 The Relationship Between Short-Term Surrogate Endpoint Indicators and mPFS and mOS in Clinical Trials of Malignant Tumors: A Case Study of Approved Molecular Targeted Drugs for Non-Small-Cell Lung Cancer in China
Mingjun Rui, Zijing Wang, Zhengyang Fei, Yao Wu, Yingcheng Wang, Lei Sun, Ye Shang and Hongchao Li

185 Effects of Motherwort Injection Versus Intramuscular Oxytocin for Preventing Postpartum Hemorrhage Among Women Who Underwent Cesarean Section
Ming-xi Li, Chun-rong Liu, Meng Chen, Hong-cai Shang, Wen Wang, Xiao-chao Luo, Ling Li, Ya-na Qi, Yi-quan Xiong, Shi-yao Huang, Jing Wang, Kang Zou, Xing-hui Liu, Jing Tan and Xin Sun

193 The Incidence and Characteristics of Perinatal Stroke in Beijing: A Multicenter Study
Qianqian Xia, Zhao Yang, Yao Xie, Ying Zhu, Zixin Yang, Mingyan Hei, Yingxue Ding, Weijing Kong, Limin Kang, Shengnan Yang, Yabo Mei, Zhichun Feng, Li Zhang, Yanzhe Lei, Ping Wang, Jingyu Dong, Li Yang, Jun Ju, Hesheng Chang, Shina Zhan, Jinqian Yu, Peng Zhang, Ran Wang, Hui Guo, Xinli Liu, Huaqing Tan, Yakun Liu, Zhenzong Zhang, Lixia Zhen, Jinting Yan, Zhan Liu, Chunxia Yang, Qingqing Wang, Jinfeng Wang, Lifang Sun, Huihui Zeng, Jing Li, Wenwen Qu, Xuemei Wang, Guiying Liu, Xi Yang, Xinxin Liu, Xuefeng Zhang, Xiaojing Xu, Yanan Gu, Hui Long, Li Zhang, Lili Liu, Zezhong Tang and Xinlin Hou

205 The Use of Oral Anticoagulation Is Not Associated With a Reduced Risk of Mortality in Patients With COVID-19: A Systematic Review and Meta-Analysis of Cohort Studies
Meng-Fei Dai, Si-Tong Guo, Yi-Jun Ke, Bao-Yan Wang, Feng Yu, Hang Xu, Zhi-Chun Gu and Wei-Hong Ge

216 Reliability of the Evidence to Guide Decision-Making in Acupuncture for Functional Dyspepsia
Jinke Huang, Jiali Liu, Zhihong Liu, Jing Ma, Jinxin Ma, Mi Lv, Fengyun Wang and Xudong Tang

225 Retrospective Analysis of the Accuracy of High-Frequency Ultrasound for Pancreaticobiliary Maljunction in Pediatrics at a Single Center
Qiuchen Xu, Min Liu, Qiumei Wu, Wen Ling and Shan Guo

234 Economic Burden and Quality of Life of Hepatocellular Carcinoma in Greater China: A Systematic Review
Huimin Zou, Meng Li, Qing Lei, Zejun Luo, Yan Xue, Dongning Yao, Yunfeng Lai, Carolina Oi Lam Ung and Hao Hu

250 The Dose-Related Efficacy of Acupuncture on Endometrial Receptivity in Infertile Women: A Systematic Review and Meta-Analysis
Xiaoyan Zheng, Siyi Yu, Liying Liu, Han Yang, Fangge Wang, Hongmei Yang, Xingyu Lv and Jie Yang

272 Vancomycin, Daptomycin, Antistaphylococcal β-Lactam, and Trimethoprim-Sulfamethoxazole Monotherapy and Combination Therapy in the Management of Methicillin-Resistant *Staphylococcus aureus*: A Network Meta-Analysis
Xiaonan Xu, Ni Lu, Pan Song, Mingzhen Zhou, Yuanxiao Li, Zirui Wang and Xin Gao

281 Efficacy of the Integrative Acupuncture and Moxibustion Treatment in Patients With Major Depressive Disorder: The Study Protocol for a Multicenter, Single-Blinded, Randomized Trial in China
Yuan Zhang, Yamin Liu, Baile Ning, Luda Yan, Lihua Wu, Delong Zhang, Changhong Li, Wenwei Ouyang, Shengyong Su, Shuo Jiang, Guangcai Zhang, Junfeng Xu, Zhen Wang, Zhong Zheng, Dong Zheng, Shan Chen, Lu Sun and Wenbin Fu

292 Family Socioeconomic Position and Lung Cancer Risk: A Meta-Analysis and a Mendelian Randomization Study
Xusen Zou, Runchen Wang, Zhao Yang, Qixia Wang, Wenhai Fu, Zhenyu Huo, Fan Ge, Ran Zhong, Yu Jiang, Jiangfu Li, Shan Xiong, Wen Hong and Wenhua Liang

302 Clinical Effectiveness of Conservative Treatments on Lumbar Spinal Stenosis: A Network Meta-Analysis
Xuanwei Chen, Zhizhan Zheng and Jianhua Lin

311 Influencing Factors of Health Technology Assessment to Orphan Drugs: Empirical Evidence in England, Scotland, Canada, and Australia
Na Zhou, Hong Ji, Zheng Li, Jun Hu, Jia-Hua Xie, Yu-Heng Feng and Ni Yuan

320 Chronic Noise Exposure and Risk of Dementia: A Systematic Review and Dose-Response Meta-Analysis
Linghao Meng, Yang Zhang, Shushan Zhang, Fugui Jiang, Leihao Sha, Yajia Lan and Lei Huang

331 The ABC-Stroke Score Refines Stroke Risk Stratification in Patients With Atrial Fibrillation at the Emergency Department
Jan Niederdöckl, Julia Oppenauer, Sebastian Schnaubelt, Filippo Cacioppo, Nina Buchtele, Alexandra-Maria Warenits, Roberta Laggner, Nikola Schütz, Magdalena S. Bögl, Gerhard Ruzicka, Sophie Gupta, Martin Lutnik, Safoura Sheikh Rezaei, Michael Wolzt, Harald Herkner, Hans Domanovits, Anton N. Laggner, Michael Schwameis and Ziad Hijazi

338 Superior gait performance and balance ability in Latin dancers
Yen-Ting Liu, Ang-Chieh Lin, Szu-Fu Chen, Chih-Jen Shih, Tien-Yun Kuo, Fu-Cheng Wang, Pei-Hsin Lee and Adeline Peiling Lee

354 Efficacy and safety of pharmacological interventions for pruritus in primary biliary cholangitis: A systematic review and meta-analysis
Chenyi Xu, Rensong Yue, Xuelian Lv, Shengnan Wang and Mengmeng Du

370 Cost-accuracy and patient experience assessment of blood pressure monitoring methods to diagnose hypertension: A comparative effectiveness study
Luis González-de Paz, Belchin Kostov, Xavier Freixa, Carmen Herranz, Elena Lagarda, María Ortega, Elisa Pérez, Silvia Porcar, Eva Sánchez, Montserrat Serrato, Ingrid Vidiella and Antoni Sisó-Almirall

frontiers | Frontiers in Public Health

TYPE Editorial
PUBLISHED 13 February 2023
DOI 10.3389/fpubh.2023.1129399

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Hai Fang,
Peking University, China

*CORRESPONDENCE
Kevin Lu
✉ lu32@email.sc.edu
Bin Jiang
✉ binjiang@hsc.pku.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

RECEIVED 21 December 2022
ACCEPTED 17 January 2023
PUBLISHED 13 February 2023

CITATION
Xiong X, Lv G, Jiang B and Lu K (2023) Editorial: From randomized clinical trials to real-world data and big data sciences: Generating evidence-based medicine for value in western and herbal medicines.
Front. Public Health 11:1129399.
doi: 10.3389/fpubh.2023.1129399

COPYRIGHT
© 2023 Xiong, Lv, Jiang and Lu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Editorial: From randomized clinical trials to real-world data and big data sciences: Generating evidence-based medicine for value in western and herbal medicines

Xiaomo Xiong[1†], Gang Lv[2†], Bin Jiang[3*] and Kevin Lu[1*]

[1]Department of Clinical Pharmacy and Outcomes Sciences, University of South Carolina, Columbia, SC, United States, [2]Department of General Surgery, The First Medical Center of Chinese PLA General Hospital, Beijing, China, [3]Department of Administrative and Clinical Pharmacy, School of Pharmaceutical Sciences, Health Science Center, Peking University, Beijing, China

KEYWORDS

real-world data, randomized controlled trial (RCT), evidence based medicine, value in health, herbal medicine, Traditional Chinese Medicine

Editorial on the Research Topic
From randomized clinical trials to real-world data and big data sciences: Generating evidence-based medicine for value in western and herbal medicines

Evidence-based medicine (EBM) differs from traditional medical practices in that it places emphasis on the use of published evidence and rigorous evaluation (1). The main idea of EMB is that medical decisions should be based on objective research findings as much as possible (1, 2), and EBM is increasingly accepted in clinical practice and has a wide range of implementation (3). The goal of this Research Topic is to provide evidence-based medicine by assessing the effectiveness, safety, and economics of medicines that matter to patients, where we have collected a series of studies related to the topic. The collected studies under this topic cover a wide range of evidence-based studies related to clinical effectiveness, including comparative effectiveness research using real-world data, research related to the coronavirus disease of 2019 (COVID-19) pandemic, and systematic reviews and meta-analysis to summarize the best evidence-based medicine available in the literature. We expect that the research on this topic or the evidence they provide will act an informative role in clinical practice, to guide clinical treatment or formulate intervention measures.

Most clinical interventions that are marketed or approved have undergone a rigorous process of the gold standard of EBM, i.e., the randomized controlled trial (RCT), to measure clinical efficacy (4–7). However, in RCTs, the setting is rigorous, and the strict inclusion and exclusion criteria for patients may affect the external study of their findings (6, 8). Therefore, evidence derived from real-world studies or clinical effectiveness is warranted to guide day-to-day clinical practice (9). Meanwhile, many alternative treatments, such as herbal medicine and traditional medicine, are used in clinical practices, however, with a lack of evidence on efficacy from RCTs (10). As such, evidence of their clinical effectiveness is crucial. Therefore, being able to obtain more evidence about these studies is of great significance for informing practice.

Although EBM informing clinical practice and public health policy decision-making is well-documented in the literature, the role of EBM in some aspects of clinical effectiveness, such as clinical trial designs, medication use, quality of life, and the economic burden has received less attention. Several articles under this Research Topic may complement this point. First, for clinical trial designs, Iglesias-Lopez et al. conducted a systematic review, which found that most authorized advanced therapy medicinal products (ATMPs) are based on small, open-label,

uncontrolled, and single-arm trials. The findings of this research suggest that in future studies, more methodologically sound randomized-controlled trials for ATMPs are needed. Meanwhile, a case study by Rui et al. found a relationship between short-term surrogate endpoint indicators and median progression-free survival (mPFS) and median overall survival (mOS) in clinical trials of malignant tumors. Second, for medication use, based on the hospital database with 3,422,710 outpatient and emergency visits and 26, 118, 436 inpatient hospitalizations from 2016 to 2019, Zhu et al. studied the trends and patterns of use of antibiotics in China, which might help inform clinical guidance in the future. Finally, in terms of quality of life and economic burden, Zou H. et al. systematically reviewed the economic burden and quality of life of patients with hepatocellular carcinoma (HCC) in Greater China, and found that HCC had a negative impact on the quality of life among the patients, mainly related to the aspects of physical, cognitive, and social functions. At the same time, this study revealed the impact of HCC on the economic burden of patients, which showed that HCC significantly increased patients' health-related costs (Zou H. et al.).

As the highest-level evidence in EBM that is able to integrate the results of multiple studies, meta-analysis can implement the generalization of RCTs and enhance the guidance of medical practice (11, 12). This Research Topic includes several systematic reviews and studies of meta-analysis that can help inform clinical practice. Specifically, these studies provide clinical evidence for several chronic conditions. Chen X. et al. systematically reviewed the clinical efficacy of conservative treatments on lumbar spinal stenosis, which found that limaprost might have better efficacy. In another review about the safety of fostamatinib among patients with rheumatoid arthritis, Chen Y. et al. found that compared to the placebo, fostamatinib was associated with a higher risk of malignant neoplasms at 52 weeks, with an odds ratio of 4.49 (95% CI 1.03–19.60). In a study with more than 1,000 patients with Peyronie's disease by Cao et al., the authors found that collagenase clostridium histolyticum has a significant effect on treating Peyronie's disease. In addition to chronic diseases, this topic has collected several studies on evidence generated in systematic reviews for patients with cancer. Based on five trials with more than 3,000 patients, Chen J. et al. found that cabazitaxel is effective in post-docetaxel settings but associated with a high risk of serious adverse events (SAEs). Using meta-analysis with 13 studies, Zou X. et al. found that a poorer socioeconomic status was associated with an increased risk of lung cancer in adulthood, with an OR of 1.25 (95% CI 1.10–1.43).

In the context of the pandemic, this Research Topic also collects a study about the evidence on the effectiveness of using oral anticoagulation for coronavirus disease 2019 (COVID-19). Specifically, Dai et al. conducted a systematic review and meta-analysis and found that chronic oral anticoagulation could not reduce mortality or intensive care unit (ICU) admission rate for patients with COVID-19. This Research Topic also contains a research protocol for a further clinical trial to assess the efficacy of integrative acupuncture and moxibustion treatment in patients with major depressive disorder (MDD; Zhang et al.).

We believe that the studies included in this Research Topic can serve as a clinical guide to a certain extent, especially as a supplement to randomized clinical trials. At the same time, we also believe that these studies can serve as a methodological inspiration for future research. Ultimately, we hope this Research Topic can support research on the value of health in evidence-based medicine.

Author contributions

KL, XX, GL, and BJ conceived the idea for the editorial and wrote the initial draft. All authors approved the final version of the editorial.

Acknowledgments

Thanks to the authors that submitted their articles to the Research Topic.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Guyatt G. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. New York, NY: McGrawHill Education (2015).

2. Evidence-Based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA*. (1992) 268:2420–5. doi: 10.1001/jama.1992.03490170092032

3. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. *Lancet*. (2017) 390:415–23. doi: 10.1016/S0140-6736(16)31592-6

4. Van Norman GA. Drugs, devices, and the FDA: part 1: an overview of approval processes for drugs. *JACC Basic Transl Sci*. (2016) 1:170–9. doi: 10.1016/j.jacbts.2016.03.002

5. Van Norman GA. Drugs, devices, and the FDA: part 2: an overview of approval processes: FDA approval of medical devices. *JACC Basic Transl Sci*. (2016) 1:277–87. doi: 10.1016/j.jacbts.2016.03.009

6. Bothwell LE, Greene JA, Podolsky SH, Jones DS. Assessing the Gold Standard–Lessons from the History of RCTs. *N Engl J Med*. (2016) 374:2175–81. doi: 10.1056/NEJMms1604593

7. Meldrum ML. A brief history of the randomized controlled trial. From oranges and lemons to the gold standard. *Hematol Oncol Clin North Am*. (2000) 14:745–60, vii. doi: 10.1016/s0889-8588(05)70309-9

8. Grossman J, Mackenzie FJ. The randomized controlled trial: gold standard, or merely standard?. *Perspect Biol Med*. (2005) 48:516–34. doi: 10.1353/pbm.2005.0092

9. Garrison LP Jr, Neumann PJ, Erickson P, Marshall D, Mullins CD. Using real-world data for coverage and payment decisions: the ISPOR Real-World Data Task Force report. *Value Health*. (2007) 10:326–35. doi: 10.1111/j.1524-4733.2007.00186.x

10. Fung FY, Linn YC. Developing traditional Chinese medicine in the era of evidence-based medicine: current evidences and challenges. *Evid Based Complement Alternat Med*. (2015) 2015:425037. doi: 10.1155/2015/425037

11. Leucht S, Chaimani A, Cipriani AS, Davis JM, Furukawa TA, Salanti G. Network meta-analyses should be the highest level of evidence in treatment guidelines. *Eur Arch Psychiatry Clin Neurosci*. (2016) 266:477–80. doi: 10.1007/s00406-016-0715-4

12. Wormald R, Evans J. What makes systematic reviews systematic and why are they the highest level of evidence?. *Ophthalm Epidemiol*. (2018) 25:27–30. doi: 10.1080/09286586.2017.1337913

frontiers in Medicine

ORIGINAL RESEARCH
published: 16 August 2021
doi: 10.3389/fmed.2021.705202

Association Between the Health Locus of Control and Medication Adherence: An Observational, Cross-Sectional Study in Primary Care

Hanna-Maria E. Gerland[1] *and Tino Prell*[1,2]*

[1] Department of Neurology, Jena University Hospital, Jena, Germany, [2] Center for Healthy Ageing, Jena University Hospital, Jena, Germany

Background: Medication non-adherence is an important healthcare issue and a common problem. Many predictors of non-adherence have been found in different settings and cohorts.

Objective: Evaluate the impact of the health locus of control (HLC) on unintentional/intentional non-adherence in primary care.

Methods: In this observational, cross-sectional study, 188 patients (mean age 63.3 ± 14.9 years) were recruited from three primary care practices in Jena, Germany, over 4 months. The study assessed demographic data, self-reported adherence (German Stendal adherence to medication score, SAMS), HLC, and depression.

Results: According to the SAMS total score, 44 (27.5%) were fully adherent, 93 (58.1%) were moderately non-adherent, and 23 (14.4%) were clinically significantly non-adherent. The most common reasons for non-adherence were forgetting to take the medication or lacking knowledge about the prescribed medication. Multiple linear regression revealed that adherence was good in people with external HLC and poor in internal HLC. In particular, intentional non-adherence was positively associated with internal HLC and negatively with fatalistic external HLC. Depression had a negative influence on both intentional and unintentional non-adherence.

Conclusion: HLC is an independent predictor of medication non-adherence and is a promising target for interventions that enhance adherence.

Keywords: adherence, depression, community-dwelling population, locus of control, drugs, compliance

OPEN ACCESS

Edited by:
Yu Heng Kwan,
Duke-NUS Medical School, Singapore

Reviewed by:
Jiekie Phang,
Singapore General Hospital, Singapore
Dan Blalock,
Duke University Health System, United States

Correspondence:
Tino Prell
Tino.prell@med.uni-jena.de

Specialty section:
This article was submitted to Family Medicine and Primary Care, a section of the journal Frontiers in Medicine

Received: *04 May 2021*
Accepted: *26 July 2021*
Published: *16 August 2021*

Citation:
Gerland H-ME and Prell T (2021) Association Between the Health Locus of Control and Medication Adherence: An Observational, Cross-Sectional Study in Primary Care. Front. Med. 8:705202. doi: 10.3389/fmed.2021.705202

INTRODUCTION

Health locus of control (HLC) refers to whether individuals think they can take responsibility for their health or not. It reflects the degree to which someone attributes the consequences of their behavior to internal or external factors. The original concept was developed in 1954 by Julian B. Rotter in his Social Learning Theory (1) and expanded by Wallston et al. (2). HLC is considered internal (a belief that one can control one's health) or external (a belief that external factors control

health). External factors are external social control (control by other persons, physicians) and fatalistic external control (control by fate, god, chance). Therefore, people with a high internal HLC should be more likely to engage in health-promoting and -sustaining activities. In contrast, people who believe that their health is beyond their control are probably less likely to implement recommended health behavior (3).

HLC also plays an important role in medication adherence (3). Adherence describes the extent to which patients take their medication as agreed with their health care providers (4). Patient-related factors, condition-related factors, and therapy-related factors influence adherence (5, 6). Non-adherence is common in patients with chronic disorders and contributes to higher morbidity, higher mortality, and increased healthcare costs. Non-adherence can be intentional (e.g., if a patient deliberately decides not to take medication) or unintentional (e.g., forgetting to take medication) (7).

However, HLC's contribution to non-adherence in general, and intentional or unintentional non-adherence, needs clarification. One study of women on hormone therapy reported that patients who deliberately missed their medication viewed themselves as having less influence over their health than patients who reported forgetting their medication. Moreover, women who reported forgetting their medication also had a stronger belief that powerful others (such as health professionals, family, friends) are relevant to improving their health than patients who intentionally chose not to take their hormones (8). However, this study's results cannot be generalized; therefore, we aimed to explore the relationship between HLC and unintentional/intentional non-adherence in a community-dwelling population. We hypothesized that internal HLC is associated with intentional non-adherence and that external HLC is associated with less intentional non-adherence.

METHODS

Study Design and Participants

The ethics committee of Jena University Hospital (5290-10/17) and the Landesärztekammer Thüringen approved the study. In this cross-sectional study, several questionnaires were distributed to three randomly selected family doctors' practices in Jena, Germany, between September 2019 and December 2019. Patients completed the questionnaires anonymously in the waiting room and placed them in a box. Of the 200 questionnaires distributed, 188 were completed. The inclusion criteria were (1) age over 18, (2) at least one permanent medication, and (3) the absence of dementia according to medical records, and (4) the ability to fill out the questionnaires independently.

Questionnaires

The self-reported demographic data included age, gender, marital status (single, divorced/widowed or married), and level of education (high: German *Abitur* or university; medium: German *Realschule* or General Certificate of Secondary Education; low: German *Hauptschule* or no school). Additionally, patients were asked about their diseases and the total daily number of medications administered. Adherence was assessed using the German Stendal adherence to medication score (SAMS) (9), which is based on 18 items forming a cumulative scale (0–72), where 0 represents complete adherence and 72 complete non-adherence (10). The SAMS allows assessment of three common patterns or reasons for non-adherence: modification of medications, forgetting to take the medications, and lacking knowledge about the medications (7). The entire SAMS and the manual are available online (CC BY NC 3.0 license; https://data.mendeley.com/datasets/ny2krr3vgg/1) (9, 11).

To examine the patients' perceptions of their autonomy support through the doctor and assess the patient-physician relationship, the German version of the Health Care Climate Questionnaire (HCCQ) was used. Autonomy support consists of providing patients with effective treatment options, supporting patient initiatives, eliciting and considering patients' views, and minimizing control and judgment. The HCCQ consists of 15 items on a seven-point Likert scale ranging from 1 = strongly disagree to 7 = strongly agree. Only item 13 is reverse-coded. The individual score is calculated with the mean score after reversing item 13. Higher mean scores indicate a higher level of perceived autonomy support (12).

To determine the patients' HLC, we used the German questionnaire "Kontrollüberzeugung zu Krankheit und Gesundheit" (KKG), which assessed the overall control belief for disease and health. It is based on the Multidimensional HLC Scale (MHLC). Its 21 items are statements with a six-point Likert scale from "Does not apply at all" (value 1) to "Very much applies" (value 6). From the item response, a cumulative score is formed for three subscales (from min. 7 to max. 42): internal, social external, and fatalistic external HLC. High sum values indicate high HLC and low sum values low HLC. Internality represents the belief in their self-control of health-related events. Social externality means that a person thinks that their physical condition depends largely on the actions of others (e.g., physicians). A fatalistic externality is a conviction that fate and chance influence their state of health.

The Patient Health Questionnaire (PHQ)-9 questionnaire was used to detect depression. Its 9 items cover DSM-IV criteria for the diagnosis of depression. The cumulative score can range from 0 to 27 and assigns patients to the categories "no depression" (<10), "mild depression" (10–19), and "manifest depression" (20–27) (13, 14).

Due to severely incomplete or uncompleted questionnaires ($n = 14$) and contradictory response behavior ($n = 14$), 28 patients were excluded from the analysis. Excluded patients tended to be younger (median = 59, IQR = 25.75), most had a high educational level ($N = 13$, 46.5%), and took an average of 3 medications per day (IQR = 2). The number of drugs per day in the excluded patients did not significantly differ from the included patients (median = 4, IQR = 4; $p = 0.23$). The central tendency of the KKG and HCCQ questionnaire responses did not differ from the included patients. The median sum score of the PHQ-9 was 6.0 (IQR = 5.25) in the excluded patients, slightly higher than in the included patients (median = 4.0, IQR = 6.0).

Statistics

SPSS version 23.0 (IBM Corporation, Armonk, NY, USA) and R statistics version 4.0.2 (R Foundation for Statistical Computing, Vienna, Austria were used for the statistical analyses. Descriptive analyses were used to describe the cohort. The cut-off for clinically significant non-adherence was set at a total SAMS of 9 points (80th percentile). Generally, non-adherence becomes clinically significant when <80% of a prescribed medication is taken (15, 16). The cohort was categorized study and sample-specific into adherent (SAMS = 0), moderately non-adherent (SAMS = 1–8), and non-adherent (SAMS ≥ 9) patients. Patients were categorized according to the main reason/factor of non-adherence: medication modifications, missing knowledge about medication, and forgetfulness (7, 9). For this purpose, the items that belonged to a factor were summed and divided by the number of items resulting in three SAMS factors. In the factor' modifications', medications were changed by the patients without consulting their doctors about either experiencing side effects or health improvement. The factor' missing knowledge' included patients who were unaware of the purpose and/or dosage of their medication. The factor' forgetfulness' included patients who forgot their medication frequently.

In addition, patients were categorized into three HLC groups: internal, social external, and fatalistic external HLC. The groups' clinical and sociodemographic parameters were compared by analysis of variance, the Kruskal–Wallis test, and the chi-square test. Four linear regressions were performed with the (1) SAMS total, (2) the SAMS modification factor, (3) the SAMS missing knowledge factor, and (4) the SAMS forgetting factor as the dependent variables and age, gender, education level, number of medications per day, internal HLC subscore, social external HLC subscore, fatalistic external HLC subscore, the HCCQ average score, and the PHQ9 sum score as independent variables (backward selection). The significance levels for variables entering the linear regression model and for removing from the model were set at 0.05 and 0.1, respectively. Autocorrelation (Durbin–Watson) and multicollinearity (variance inflation factor and tolerance) were excluded. However, residuals were not normally distributed for the models with the SAMS factors as dependent variables. Although the multiple linear regression is mostly robust to violations of the normal distribution assumption, we computed additional elastic net models to test the results (17). Elastic net regularization leads to parsimonious models, which are easier to interpret. Variable selection is performed by shrinking the parameters toward zero and attenuating overfitting. Ten-fold cross validation was performed to choose the best model with the lowest mean cross-validated error. Within the elastic net algorithm, variables remain in the model if the prediction error averaged over the ten-fold cross validation samples is reduced. The elastic net algorithm performs well in highly correlated variables, either including all variables with similar regression coefficients or excluding all variable from the best model. Elastic net regularization was performed using the package glmnet in R version 3.6.2 (R Foundation for Statistical Computing, Vienna, Austria). Again using elastic net, four models were calculated with the (1) SAMS total, (2) the SAMS modification factor, (3) the SAMS missing knowledge factor, and (4) the SAMS forgetting factor as the dependent variables and age, gender, educational level, number of medications per day, internal HLC subscore, social external HLC subscore, fatalistic external HLC subscore, the HCCQ average score, and the PHQ9 sum score as independent variables.

Normally distributed values are expressed as means and standard deviation (SD); other values are expressed as medians and interquartile range (IQR). All categorical variables are presented as numbers and percentages. Statistical significance was set at a p-value < 0.05.

TABLE 1 | Characteristics of the cohort.

Characteristic	*n*	%	Missing
Gender			
Female	98	61.3	0
Male	62	38.8	
Education			
Low	32	20.0	0
Middle	65	40.6	
High	63	39.4	
Marital status			
Not married	54	34.2	2
Married	104	65.8	

	Mean	SD	Median	IQR	Missing
Age	63.3	14.9	67.0	23	0
SAMS sum score	4.3	4.9	3	4.9	0
HCCQ mean value	6.3	1.0	6.6	0.73	27
PHQ-9 sum score	5.7	4.8	4	6	4
Subtotal score internal HLC	24.94	6.148	26	8	12
Subtotal score social external HLC	23.89	5.209	24	7.25	18
Subtotal score fatalistic external HLC	18.59	6.684	18	10	20

SAMS, Stendal adherence to medication score; HCCQ, Health Care Climate Questionnaire; PHQ-9, Patient Health Questionnaire-9; HLC, Health locus of control.

RESULTS

Descriptive Statistics

Data from 160 patients were used for the following analyses. Most patients were 60 years or older (65%), female, married, pensioned, and had a middle or higher education level (**Table 1**). The self-reported diseases comprised a wide spectrum; hypertension, diabetes, and thyroid diseases were the most common. The number of medications taken daily was 4.4 (SD = 3.0) (**Table 1**). According to the PHQ-9, 123 (78.8%) reported no, 24 (15.4%) reported mild, 6 (3.8%) moderate, and 3 (1.9%) severe depression (4 missing data). The HCCQ indicated that the patient's perceived support for autonomy through the doctor was very high.

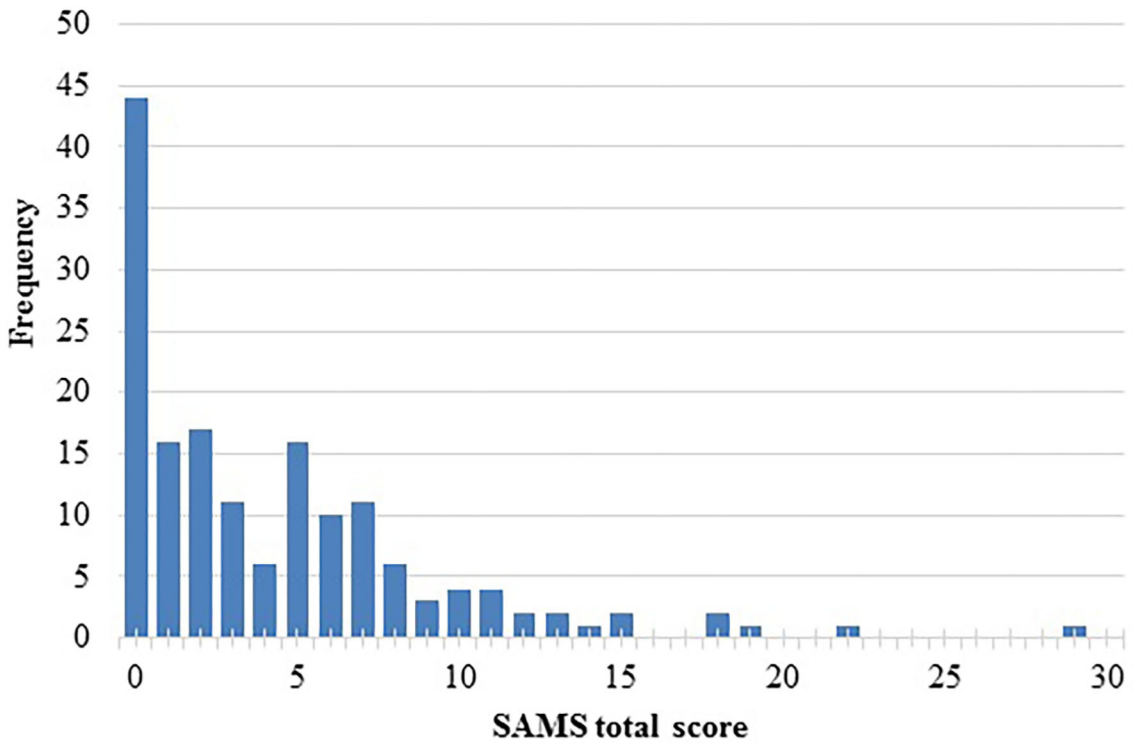

FIGURE 1 | Distribution of Stendal Adherence to Medication Score (SAMS) total score.

Adherence

According to the SAMS, 44 (27.5%) reported being fully adherent (SAMS = 0), 93 (58.1%) reported moderate non-adherence (SAMS = 1–8), and 23 (14.4%) clinically significant non-adherence (SAMS ≥ 9). **Figure 1** displays the SAMS distribution. The main reasons for non-adherence were forgetting ($n = 55$, 47.4%), missing knowledge about prescribed medication ($n = 39$, 33.6%), and modifying prescribed medication ($n = 22$, 19.0%) (refers to the 116 participants with a SAMS > 1). The partial correlation between the SAMS total score and the PHQ-9 score, adjusted for the influence of the variable age, was positive with an $r = 0.201$ ($p = 0.012$).

HLC Beliefs

Table 1 lists the mean sum scores of the three HLC subscales. The highest subscale was determined for every patient; consequently, 30 (18.8%) participants had predominantly fatalistic external HLC, 70 (43.8%) social external HLC, and 11 (6.9%) internal HLC. Furthermore, 28 (17.5%) overlapped with equal values for 2 subscales; 21 subjects could not be categorized due to missing data (a total of 139 analyzed KKG questionnaires).

The three HLC groups did not differ in PHQ-9 ($p = 0.20$), HCCQ ($p = 0.79$), sex ($p = 0.13$), and marital status ($p = 0.09$) (**Table 2**). Patients with internal HLC were younger than patients with external social HLC ($p = 0.024$). Age showed a weak negative correlation with internal HLC ($r = -0.265$, $p = 0.01$). Age did not correlate with social external or fatalistic external HLC.

Adherence and HLC Beliefs

The overall adherence (SAMS total score) did not differ between the three HLC groups ($p = 0.52$) (**Table 2**). The SAMS did not correlate with the sum scores of the three HLC domains. Vice versa, the sum scores of the HLC domains did not differ between adherent (SAMS = 0), moderately non-adherent (SAMS 1 to 8), and clinically significant non-adherent patients (SAMS ≥ 9) (**Table 3**).

Multiple Linear Regression

Regression analyses examined the influence of clinical variables on (1) SAMS total, (2) factor SAMS modification, (3) factor SAMS knowledge, and (4) factor SAMS forgetting. Age, gender, education level, number of medications per day, internal HLC subscore, social external HLC subscore, fatalistic external HLC subscore, the HCCQ average score, and the PHQ9 sum score were entered as independent variables (backward selection).

The PHQ-9, internal HLC, and fatalistic external HLC predicted the total SAMS total score (the final models after backward selection are given in **Table 4**, the detailed models are given as **Supplementary Tables 1–4**). Depression and internal HLC were associated with non-adherence. Fatalistic external HLC was associated with better adherence (**Supplementary Table 1**). However, the explained variance was low. The influence of depression and HLC on the SAMS total score was also demonstrated in the elastic net model (**Supplementary Table 5**). Overall, the strongest influence had depression on the SAMS total.

TABLE 2 | Characteristics of people with predominantly internal, social external, and fatalistic external health locus of control.

Characteristic	Internal health locus of control			Social-external health locus of control			Fatalistic-external health locus of control		
	n	%		*n*	%		*n*	%	
Gender									
Female	9	81.8		36	51.4		19	63.3	
Male	2	18.2		34	48.6		11	36.7	
Education									
Low	1	9.1		9	12.9		8	26.7	
Middle	6	54.5		23	32.9		16	53.3	
High	4	36.4		38	54.3		6	20.0	
Marital status									
Not married	2	20.0		18	26.1		14	46.7	
Married	8	80.0		51	73.9		16	53.3	
	Mean	**95% CI**		**Mean**	**95% CI**		**Mean**	**95% CI**	
Age	52.45	44.32	60.59	65.19	61.79	68.58	63.53	57.65	69.42
SAMS sum score	6.36	0.52	12.21	4.47	3.33	5.62	4.27	2.30	6.24
HCCQ mean value	6.43	6.02	6.83	6.21	5.94	6.48	6.17	5.66	6.67
PHQ-9 sum score	5.82	1.93	9.71	5.14	4.14	6.15	7.03	4.91	9.15
Subscore internal HLC	32.09	30.38	33.80	24.51	23.04	25.98	23.40	21.50	25.30
Subscore social external HLC	19.70	18.01	21.39	26.76	25.68	27.83	20.45	18.64	22.25
Subscore fatalistic external HLC	13.00	9.78	16.22	16.70	15.23	18.16	25.40	23.76	27.04

SAMS, Stendal adherence to medication score; HCCQ, Health Care Climate Questionnaire; PHQ-9, Patient Health Questionnaire-9; HLC, Health locus of control.

SAMS modification was positively associated with internal HLC and depression and negatively associated with fatalistic external HLC. This finding indicates external HLC is related to less intentional non-adherence (**Table 4**, **Supplementary Table 2**). The elastic net model confirmed the influence of depression on the *SAMS modification* (**Supplementary Table 6**). In addition lower age was associated with more modification of medication. However, HLC was not a relevant predictor in the elastic net model. Considering the two models with low coefficients and low explained variance, one can conclude that HLC plays only a minor role for self-reported intentional non-adherence/modification.

The SAMS *missing knowledge* was associated with depression, lower educational level, a higher number of drugs per day, and fatalistic external HLC (**Table 4**, **Supplementary Table 3**). The relevance of fatalistic external HLC and the number of drugs per day on the self-reported knowledge about medication could also be confirmed in the elastic net model (**Supplementary Table 7**). To a lesser extent, gender, depression, and education level were also relevant factors for SAMS *missing knowledge* here.

Forgetting to take medication was more likely reported in young and male persons (**Table 4**, **Supplementary Table 4**). This was in line with the findings in the elastic net model (**Supplementary Table 8**). However, the explained variance was low.

DISCUSSION

This cross-sectional study aimed to analyze the frequency and reasons for non-adherence in patients from general practices and explore the relationship between HLC and unintentional/intentional non-adherence. We hypothesized that internal HLC is associated with intentional non-adherence. Furthermore, we assumed that external HLC is associated with less intentional non-adherence.

In our cohort, about three-quarters (72.5%) reported different degrees of non-adherence. Most were moderately non-adherent, and only a minority were clinically significant non-adherent. The reasons for non-adherence were mostly unintentional (e.g., forgetting to take medication, 47.4%). The prevalence of non-adherence strongly depends on the method of measurement (objective or subjective). Various reviews reported non-adherence ranging from 20 to 80% of cases for certain chronic diseases, especially hypertension and diabetes (5, 18, 19). Our findings are consistent with previous studies showing that *forgetting* is the primary self-reported behavior identified among non-adherent participants (20). In this survey, *forgetting to take medication* was the most frequent reason for non-adherence, followed by a *lack of knowledge* about the prescribed medication and *modification of medication*.

In our study, we also assessed the patients perceived autonomy support through the doctor using the HCCQ. The mean score was very high, implying that patients feel strongly supported by their family doctor. Another reason for the high mean value of HCCQ might be the traditionally close relationship between general practitioners and patients in Germany, which might discourage patients from criticizing their family doctor (with whom they have a long-standing relationship). But how does the doctor-patient relationship affect adherence? With correction for demographic factors, depression, and HLC, we did not observe

TABLE 3 | Characteristics of people with different degrees of non-adherence.

Characteristic	Adherent			Moderately non-adherent			Non-adherent		
	n	%		*n*	%		*n*	%	
Gender									
Female	30	68.2		53	57.0		15	65.2	
Male	14	31.8		40	43.0		8	34.8	
Education									
Low	11	25.0		19	20.4		2	8.7	
Middle	21	47.7		32	34.4		12	52.2	
High	12	27.3		42	45.2		9	39.1	
Marital status									
Not married	19	45.2		29	31.2		6	26.1	
Married	23	54.8		64	68.8		17	73.9	
	Mean	**95% CI**		**Mean**	**95% CI**		**Mean**	**95% CI**	
Age	66.16	62.25	70.07	63.61	60.54	66.68	56.83	49.48	64.17
SAMS sum score	0.00	0.00	0.00	4.00	3.53	4.47	13.52	11.40	15.64
HCCQ mean value	6.39	6.07	6.70	6.19	5.95	6.42	6.38	6.17	6.59
PHQ-9 sum score	5.21	3.60	6.82	5.53	4.58	6.49	7.26	5.34	9.18
Subscore internal HLC	23.18	20.70	25.66	25.43	24.22	26.64	26.09	24.13	28.04
Subscore social-external HLC	22.61	20.77	24.44	24.37	23.23	25.51	24.30	22.32	26.29
Subscore fatalistic-external HLC	18.87	16.53	21.21	18.29	16.84	19.73	19.23	16.18	22.28

SAMS, Stendal adherence to medication score; HCCQ, Health Care Climate Questionnaire; PHQ-9, Patient Health Questionnaire-9; HLC, Health locus of control.

a relevant association between overall adherence (SAMS total) and HCCQ. On the one hand, a trusting relationship could ensure that patients do not modify the prescribed medication regimen and are less intentionally non-adherent. On the other hand, patients might feel uncomfortable reporting medication modification because of this special relationship between the patient and the general practitioner. This is why we used the anonymous design of the study.

In the DEGS-1 study in Germany, between 2008 and 2011, almost 8,000 people were surveyed with the PHQ-9 questionnaire about current depressive symptoms using the same cut-off values as we did. According to self-disclosure, approximately 8% of the German population between 18 and 79 suffer from depressive symptoms (21). In our cohort, 21.2% of the respondents reported depressive symptoms, which is higher than the general prevalence in Germany. Depressive symptoms impair adherence (22), as indicated in our study with a weak effect strength. However, patients with depressive symptoms might be more self-critical when reporting their adherence (22).

HLC was assessed by the German KKG questionnaire. The KKG is a modification of the "multidimensional HLC" (MHLC) (2) questionnaire. It is based on the same theoretical and methodological concept. Still, the KKG has a unique aspect in the formulation of its items: Patients are asked about their perceived ability to influence their physical *conditions* (physical complaints, physical well-being). In contrast, the MHLC uses the terms *illness* and *health*. Yet, chronically ill patients might find it difficult to use the term *health* referring to their situation. Therefore, the rather neutral terms of the KKG questionnaire are more suited for heterogeneous patient populations, as they are found in general practices. Most patients in our survey had an external social control belief. The external social HLC includes all persons who influence his or her physical condition according to the respondent's perception. Therefore, all that can be said is that a high proportion of patients in our study believe that other people strongly influence their physical condition.

There are conflicting results in terms of the relationship between HLC and adherence. Nafradi et al. (3) published a review of 154 studies on the relationship between self-efficacy, HLC, and adherence. Of these, 71 studies refer to the relationship between HLC and adherence. The studies applying subjective adherence measures reported more association either in a positive or negative direction (58.5%) between the HLC dimensions and adherence compared to the studies applying mixed (14.3%) or objective (40%) adherence measures. More than half of the studies report no link between HLC and medication adherence. Doctors' HLC was most likely to show a positive association with adherence (3).

Regarding the link between internal HLC and adherence, which is considered the most predictive HLC dimension in the literature, only 10 studies out of 26 (38%) showed a positive association; 15 studies (58%) showed no association. We found that adherence was better in people with external HLC and worse in people with internal HLC. In particular, *modification of prescribed drugs without physician consultation* was positively associated with internal HLC and negatively with fatalistic external HLC. However, one has to keep in mind that the explained variance in the model was low. This indicates that

TABLE 4 | Summary of multiple linear regressions (final models after backward selection).

	Std. coefficient B	Std. Err	*p*	Overall model
SAMS total (dependent variable)				
Subtotal score internal HLC	0.146	0.080	0.072	*p* 0.003 corrected R^2 0.096
Subtotal score fatalistic external HLC	−0.132	0.078	0.093	
PHQ-9 sum score	0.339	0.101	0.001	
SAMS modification (dependent variable)				
Subtotal score internal HLC	0.016	0.008	0.053	*p* 0.012 corrected R^2 0.076
Subtotal score fatalistic external HLC	−0.016	0.008	0.037	
PHQ-9 sum score	0.025	0.010	0.012	
SAMS missing knowledge (dependent variable)				
Education level (1 = low, 2 = middle, 3 = high)	−0.109	0.052	0.038	*p* 0.007 corrected R^2 0.094
Number of medications per day	0.028	0.013	0.031	
Subtotal score fatalistic external HLC	−0.013	0.006	0.027	
PHQ-9 sum score	0.016	0.008	0.049	
SAMS forgetting (dependent variable)				
Gender (0= female, 1= male)	0.216	0.109	0.052	*p* 0.025 corrected R^2 0.080
Age	−0.008	0.003	0.017	

SAMS, Stendal adherence to medication score; PHQ-9, Patient Health Questionnaire-9; HLC, Health locus of control.

HLC plays a role for adherence, but its influence should not be overestimated.

Some methodological aspects clarify these findings. The German KKG questionnaire that we used to measure HLC is a modification of the MHCL questionnaire by Wallston et al. (2). However, Wallston himself questioned HLC as a predictor of health behavior (23). In a summary of several studies using his MHCL questionnaire, he described that the loci of control could explain on average < 10% of the variance of specific health behavior (23). *Locus of control* is a generalized expectancy, a belief that one can control one's health. It is a kind of personality trait. In contrast, *self-efficacy* is the belief that one can manage a certain situation or behavior. It is state-related instead of trait-related. It is possible to have an internal locus of control regarding one's health. Still, if the conviction in one's capabilities to carry out a specific behavior is missing (as in adherence to a certain medication), one is not likely to engage in that behavior. In this example, there could be a high level of internal HLC and simultaneously low self-efficacy regarding medication adherence (23). Medication adherence means the reliable and correct intake of prescribed medication as agreed by doctor and patient. It is a situational health behavior. Therefore, we suggest that in the context of adherence research, the situation-specific self-efficacy belief should be studied in addition to HLC. Future studies incorporating the locus of control concept should also ask how important adherence and the resulting better outcome are to the respondents.

Moreover, for further studies, it would be interesting to study the relationship between adherence and HLC after subdividing external social HLC into 1. physicians and 2. other influencing persons (24). This relationship could also help to understand the influence of this scale on adherence. The review mentioned above on HLC and adherence by Nafradi et al. also reported that physicians tend to influence adherence (3) positively. Other people, such as family and friends, have an ambivalent influence, depending on whether these others support or discourage the patients to adhere (3). Future research should divide external social HLC into two subgroups so that a total of four dimensions of locus of control can be assessed; i.e., internality, formal help (=physicians), informal help (=other people), and chance to investigate these beliefs appropriately. Furthermore, the formal help dimension would need a broader assessment since it is covered by only two items (24).

Our study has certain limitations. The cross-sectional design does not allow causal statements. The study design creates the possibility of self-report bias in terms of socially desirable answers. We aimed to reduce this bias by using an anonymous design of the study. In addition, the HCCQ questionnaire showed limited informative value due to the very high mean values and their reduced variability. Only patient-related predictors of non-adherence were investigated in our study. In addition, however, there are various other reasons for non-adherence that we did not record with the anonymous design. Thus,

there is a risk of presenting the complex construct of non-adherence in a simplified way. Although it is widely accepted and common to use a threshold of 80% to determine non-adherence, this approach remains arbitrary when there is no objective measure of adherence or other measurable clinical outcome. Therefore, we incorporated group comparisons and correlation analyses in our study to overcome the problem of arbitrary adherence cut-offs. Finally, we cannot make precise statements about the cognition of the patients who participated in our study. The practice staff was asked to distribute the questionnaires according to our inclusion criteria, but this cannot be verified afterward in an anonymous study. Also an analysis taking underlying main disorders and disease durations as well as social networks into account would be a promising approach to study the interaction between adherence and HLC in detail.

CONCLUSIONS

The complex reasons for non-adherence will not be measured or represented in their entirety in any study and cannot be assessed with single questionnaires. It is helpful to obtain and evaluate empirical data on adherence. However, equally important is the willingness in a trusting doctor-patient relationship to work together on an individualized concept of illness and recovery. In this context, patient decision-making also requires adequate health literacy, and not every patient wants to be transferred this competence. Further research is needed to find the right amount of control adjusted to different physical conditions and stages of the disease. For this, qualitative research could also play an important role, as it manages to go into depth and capture complexity by interviewing individual cases.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by ethics committee of Jena University Hospital (5290-10/17) and the Landesärztekammer Thüringen.

AUTHOR CONTRIBUTIONS

H-MG: data acquisition, statistical analyses, and writing first draft. TP: study concept, review and editing. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by a Bundesministerium für Bildung und Forschung grant to TP (01GY1804).

ACKNOWLEDGMENTS

We thank Ulrike Teschner for assistance in study coordination.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2021.705202/full#supplementary-material

REFERENCES

1. Rotter JB. Generalized expectancies for internal versus external control of reinforcement. *Psychol Monogr.* (1966) 80:1–28. doi: 10.1037/h0092976
2. Wallston KA, Wallston BS, DeVellis R. Development of the multidimensional health locus of control (MHLC) scales. *Health Educ Monogr.* (1978) 6:160–70. doi: 10.1177/109019817800600107
3. Nafradi L, Nakamoto K, Schulz PJ. Is patient empowerment the key to promote adherence? A systematic review of the relationship between self-efficacy, health locus of control and medication adherence. *PLoS ONE.* (2017) 12:e0186458. doi: 10.1371/journal.pone.0186458
4. Sabaté E. *Adherence to Long-Term Therapies: Evidence for Action*. Geneva: World Health Organization (2003).
5. Brown MT, Bussell JK. Medication adherence: WHO cares? *Mayo Clin Proc.* (2011) 86:304–314. doi: 10.4065/mcp.2010.0575
6. Yap AF, Thirumoorthy T, Kwan YH. Systematic review of the barriers affecting medication adherence in older adults. *Geriatr Gerontol Int.* (2016) 16:1093–101. doi: 10.1111/ggi.12616
7. Prell T, Grosskreutz J, Mendorf S, Franke GH, Witte OW, Kunze A. Clusters of non-adherence to medication in neurological patients. *Res Social Adm Pharm.* (2019) 15:1419–24. doi: 10.1016/j.sapharm.2019.01.001
8. Atkins L, Fallowfield L. Intentional and non-intentional non-adherence to medication amongst breast cancer patients. *Eur J Cancer.* (2006) 42:2271–6. doi: 10.1016/j.ejca.2006.03.004
9. Franke GH, Nentzl J, Jagla-Franke M. *SAMS. Stendal Adherence to Medication Score. Testmanual.* (2020). Available online at: www.psychometrikon.de (accessed June 28, 2021).
10. Prell T, Schaller D, Perner C, Franke GH, Witte OW, Kunze A, et al. Comparison of anonymous versus nonanonymous responses to a medication adherence questionnaire in patients with Parkinson's disease. *Patient Prefer Adherence.* (2019) 13:151–5. doi: 10.2147/PPA.S186732
11. Prell T, Grosskreutz J, Mendorf S, Witte OW, Kunze A. Data on adherence to medication in neurological patients using the German Stendal Adherence to Medication Score (SAMS). *Data Brief.* (2019) 23:103855. doi: 10.1016/j.dib.2019.103855
12. Schmidt K, Gensichen J, Petersen JJ, Szecsenyi J, Walther M, Williams G, et al. Autonomy support in primary care–validation of the German version of the Health Care Climate Questionnaire. *J Clin Epidemiol.* (2012) 65:206–211. doi: 10.1016/j.jclinepi.2011.06.003
13. Gräfe K, Zipfel S, Herzog W, Löwe B. Screening psychischer Störungen mit dem "Gesundheitsfragebogen für Patienten (PHQ-D)". *Diagnostica.* (2004) 50:171–81. doi: 10.1026/0012-1924.50.4.171
14. Kroenke K. Enhancing the clinical utility of depression screening. *CMAJ.* (2012) 184:281–2. doi: 10.1503/cmaj.112004
15. DiMatteo MR. Variations in patients' adherence to medical recommendations: a quantitative review of 50 years of research. *Med Care.* (2004) 42:200–9. doi: 10.1097/01.mlr.0000114908.90348.f9
16. Offord S, Lin J, Wong B, Mirski D, Baker RA. Impact of oral antipsychotic medication adherence on healthcare resource utilization among

schizophrenia patients with Medicare coverage. *Community Ment Health J.* (2013) 49:625–9. doi: 10.1007/s10597-013-9638-y
17. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J R Stat Soc.* (2005) 67:19. doi: 10.1111/j.1467-9868.2005.00527.x
18. Durand H, Hayes P, Morrissey EC, Newell J, Casey M, Murphy AW, et al. Medication adherence among patients with apparent treatment-resistant hypertension: systematic review and meta-analysis. *J Hypertens.* (2017) 35:2346–57. doi: 10.1097/HJH.0000000000001502
19. Abegaz TM, Shehab A, Gebreyohannes EA, Bhagavathula AS, Elnour AA. Nonadherence to antihypertensive drugs: a systematic review and meta-analysis. *Medicine.* (2017) 96:e5641. doi: 10.1097/MD.0000000000005641
20. Fernandez-Lazaro CI, Adams DP, Fernandez-Lazaro D, Garcia-Gonzalez JM, Caballero-Garcia A, Miron-Canelo JA. Medication adherence and barriers among low-income, uninsured patients with multiple chronic conditions. *Res Social Adm Pharm.* (2019) 15:744–53. doi: 10.1016/j.sapharm.2018.09.006
21. Busch MA, Maske UE, Ryl L, Schlack R, Hapke U. Prevalence of depressive symptoms and diagnosed depression among adults in Germany: results of the German Health Interview and Examination Survey for Adults (DEGS1). *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz.* (2013) 56:733–9. doi: 10.1007/s00103-013-1688-3
22. Grenard JL, Munjas BA, Adams JL, Suttorp M, Maglione M, McGlynn EA, et al. Depression and medication adherence in the treatment of chronic diseases in the United States: a meta-analysis. *J Gen Intern Med.* (2011) 26:1175–82. doi: 10.1007/s11606-011-1704-y
23. Wallston KA. Hocus-pocus, the focus isn't strictly on locus: Rotter's social learning theory modified for health. *Cogn Ther Res.* (1992) 6:183–99. doi: 10.1007/BF01173488
24. Otto C, Bischof G, Rumpf HJ, Meyer C, Hapke U, John U. Multiple dimensions of health locus of control in a representative population sample: ordinal factor analysis and cross-validation of an existing three and a new four factor model. *BMC Med Res Methodol.* (2011) 11:114. doi: 10.1186/1471-2288-11-114

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Gerland and Prell. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Public Health

ORIGINAL RESEARCH
published: 08 October 2021
doi: 10.3389/fpubh.2021.756717

Impact of Multimorbidity Subgroups on the Health Care Use and Clinical Outcomes of Patients With Tuberculosis: A Population-Based Cohort Analysis

Qin Chen[1†], Yang Che[2†], Yue Xiao[3], Feng Jiang[3], Yanfei Chen[3], Jifang Zhou[3] and Tianchi Yang[2*]*

[1] Hwa Mei Hospital, University of Chinese Academy of Sciences, Ningbo, China, [2] Ningbo Municipal Center for Disease Control and Prevention, Institute of Tuberculosis Prevention and Control, Ningbo, China, [3] School of International Pharmaceutical Business, China Pharmaceutical University, Nanjing, China

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Bin Chen,
Zhejiang Center for Disease Control and Prevention (Zhejiang CDC), China
Zhuo Chen,
University of Georgia, United States

***Correspondence:**
Jifang Zhou
1020202613@cpu.edu.cn
Tianchi Yang
57237243@qq.com

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

Received: 11 August 2021
Accepted: 08 September 2021
Published: 08 October 2021

Citation:
Chen Q, Che Y, Xiao Y, Jiang F, Chen Y, Zhou J and Yang T (2021) Impact of Multimorbidity Subgroups on the Health Care Use and Clinical Outcomes of Patients With Tuberculosis: A Population-Based Cohort Analysis. Front. Public Health 9:756717. doi: 10.3389/fpubh.2021.756717

Background: Multimorbidity is defined as the existence of two or more chronic health conditions in the same individual. While patients with tuberculosis commonly have multiple conditions at diagnosis, such as HIV, diabetes, and depression, to the authors' knowledge, there is limited information on the patterns of multimorbidity, and how the types and combinations of conditions could impact the healthcare utilization, expenditure, and TB outcomes.

Methods: An observational cohort study of adult patients diagnosed with tuberculosis was conducted using the Chinese Center for Disease Control and Prevention (CDC)'s National TB Information System (NTBIS) linked to the Ningbo Regional Health Care Database (NRHCD) (2015–2020). Latent class analysis was used to identify comorbidity groups among the subset with ≥2 conditions including TB. Group-level health care use, expenditure, and treatment outcomes were compared with patients without chronic conditions using multivariate regression models.

Results: A total of 9,651 patients with TB were identified, of whom approximately 61.4% had no chronic conditions, 17.4% had 1 chronic condition, and 21.3% had ≥2 chronic conditions. Among those with ≥1 chronic condition other than TB, 4 groups emerged: (1) general morbidity (54.4%); (2) cardiovascular morbidity without complications (34.7%); (3) cardiovascular morbidity with complications (5.0%); (4) respiratory morbidity (5.9%). The respiratory morbidity group experienced the highest expenditures, at 16,360 CNY more overall (95% CI, CNY 12,615–21,215) after adjustment compared with TB patients without chronic conditions. The respiratory morbidity and cardiovascular morbidity with complications group also had the lowest odds of favorable TB outcomes [adjusted odds ratio (aOR), 0.68; 95% CI, 0.49–0.93] and (aOR 0.59, 95% CI 0.42–0.83), respectively. The cardiovascular morbidity without complications group had the highest odds of successful TB treatment (aOR, 1.40; 95% CI, 1.15–1.71).

Conclusions: Multimorbidity is common among patients with TB. The current study identified four distinct comorbidity subgroups, all of which experienced high, yet differential, rates of health care use. These findings highlight the need for urgent reforms to transform current fragmented TB care delivery and improve access to other specialists and financial assistance.

Keywords: tuberculosis, delivery of health care, health expenditures, multimorbidity—coordination of care—general practitioner, latent class analysis

INTRODUCTION

Tuberculosis (TB) is a serious threat to global public health. The World Health Organization (WHO) has estimated that over 10.4 million people are infected by *Mycobacterium tuberculosis*, and that 1.67 million deaths are attributable to TB (1). Multimorbidity is defined as the co-existence of two or more chronic conditions within an individual (2). Due to the nature of chronic infection with TB, patients who are diagnosed with TB may already have chronic conditions such as HIV infection. The presence of multimorbidity has important implications on patient outcomes and healthcare costs.

Chronic communicable diseases such as HIV and hepatitis virus infection are common in TB patients. The risk of developing TB is significantly higher among patients living with HIV infection. Hepatitis B co-infection with TB is common in China, and hepatitis B virus surface antigen (HBsAg)-positive people account for over 7.8% of the population (3). Additionally, the ongoing COVID-19 pandemic is likely to disrupt TB management and lower TB treatment success rates. A recent meta-analysis on the first wave of COVID-19 patients has found that the prevalence of TB ranged between 0.47 and 4.47%, and TB infection was associated with more severe COVID-19 cases (4).

In addition to chronic communicable conditions, global TB eradication initiatives have been hampered by the increasing burden of non-communicable diseases (NCD), especially in low- and middle-income countries. Population aging, urbanization, and unhealthy lifestyles all contribute in a high comorbid disease pattern among patients with TB. A cross-sectional study based on a medical chart review found that ∼1.14% of TB subjects had multimorbidity. Patients with multimorbidity tend to be older, female sex, and to have higher mortality risks (5). Nevertheless, the prevalence of multimorbidity could be underestimated due to poor record-keeping, lack of standardized multimorbidity classification methods, and limited sample size (6).

Traditionally, studies have focused on single conditions such as HIV co-infection, diabetes, and alcohol use disorders that affect the clinical outcomes of TB treatment. There is a need for additional information on chronic condition profiles, as well as care patterns and associated expenditures. Coordination of specialty care and primary services is essential for managing TB patients with multimorbidity. A better understanding of the comorbidity profile could facilitate early detection, improve care coordination, and inform the decision-making relating to patients at greater risk of unfavorable TB treatment outcomes. Until now, no study has compared clinical outcomes and healthcare expenditure in TB patients with varying degrees of multimorbidity.

The objective of this study was to identify and describe chronic morbidity groups among patients with tuberculosis in real-world settings, and to explore differences in demographic and clinical characteristics across subgroups. Finally, we aim to assess the impact of multimorbidity on healthcare utilization, expenditure, and clinical outcomes in real-world settings.

METHODS

Study Design

We performed a population-based retrospective cohort study among patients with TB aged 18 years old or above in Ningbo city, in the People's Republic of China. Tuberculosis cases were confirmed based on the NTBIS. Patients whose diagnosis changed or with insufficient clinical information of healthcare utilization during the baseline period (1 year prior to TB diagnosis) or during the follow-up period were excluded. During the baseline period (1 year preceding their index date, which is the TB diagnosis date), important confounders, namely lifestyle, comorbidities, and TB-related characteristics were defined. The current study was approved by the Institutional Review Board (IRB) of Ningbo Municipal Center for Disease Control and Prevention and the requirement for informed consent was waived. The data were fully de-identified and sensitive information such as patients' names and national ID numbers were anonymized. We thus made re-identification impossible to safeguard patients' privacy. The quality of reporting in the included studies was assessed using the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist (7), which are commonly used for reporting the quality of observational studies.

Data Sources and Data Linkage

Ningbo is a coastal city in the eastern region of the People's Republic of China, with a population of nearly 8.2 million residents in 2018. Since 2005, Ningbo Health and Family Planning Commission initiated a Health Information System covering all healthcare facilities within its jurisdiction. Electronic health records from hospitals and community clinics were collected and integrated to form a comprehensive administrative healthcare database. All records are linked deterministically using the personal identity code. This code is unambiguously assigned to all Chinese citizens as well as registered foreign citizens. The linked records are deidentified before they are further processed

and analyzed. By 2020, the system covered all 65 public hospitals and 154 community health centers or township health centers, including all 12 TB designated hospitals and TB centers of excellence (Ningbo Hwa Mei Hospital).

Information regarding subject lifestyle, vital signs, laboratory results, inpatient and outpatient records, and pharmacy use were available. The outpatient, inpatient, and pharmacy records comprised patient identity code, date of diagnosis, diagnosis name, and diagnosis code (International Classification of Diseases, Tenth Revision, ICD-10), as well as procedure codes, costs, and prescription dispensing details. The data content is updated on a daily basis.

The TB registry provides information on TB diagnosis, treatment, and outcomes. The TB registry records the TB patients' personal identity code, the date of diagnosis, TB clinical classification, and genetic determinants of drug resistance, sputum, and culture results at diagnosis and follow-up were recorded using a standard electronic form by trained staff. TB outcomes were ascertained using standard WHO outcomes definitions for tuberculosis: cure, treatment completion, treatment failure, death from any cause, default, and transfer out. We defined successful outcomes as cure or completion of treatment and unfavorable outcomes such as treatment failure or death. All death records including date of death and causes of death were further confirmed by linkage to the mortality database of the Ningbo Municipal Public Security Bureau.

Cohort Description

The cohort selection is shown in **Figure 1**. Patients were required to be aged $\geq$ 18 years at TB diagnosis. In addition, eligible subjects were required to have at least two healthcare encounters during the baseline period to ensure capturing of important demographic and clinical information prior to TB diagnosis. The reason for requiring eligible subjects to have at least two encounters is to exclude patients who are not continuously followed within the Ningbo health system, which could lead to underestimation of morbidity burdens. Subject characteristics, including chronic condition profiles, were measured over the baseline period. Subjects were followed up from date of TB diagnosis to date of treatment completion, death, end of study (30 June 2020), or loss to follow-up, whichever came first. Healthcare use and expenditure were measured in both the baseline period and follow-up period.

Identification of Chronic Conditions and Generation of Morbidity Groups

ICD-10 diagnosis codes from all inpatient discharge summaries and outpatient encounters were collected and categorized using the Chronic Condition Indicator (CCI) algorithms (v2021.1) developed by Healthcare Cost and Utilization Project (HCUP) (8). The algorithms systematically scan all diagnosis fields and generated classification of distinct chronic conditions (9). Identified chronic conditions that appeared in two outpatient

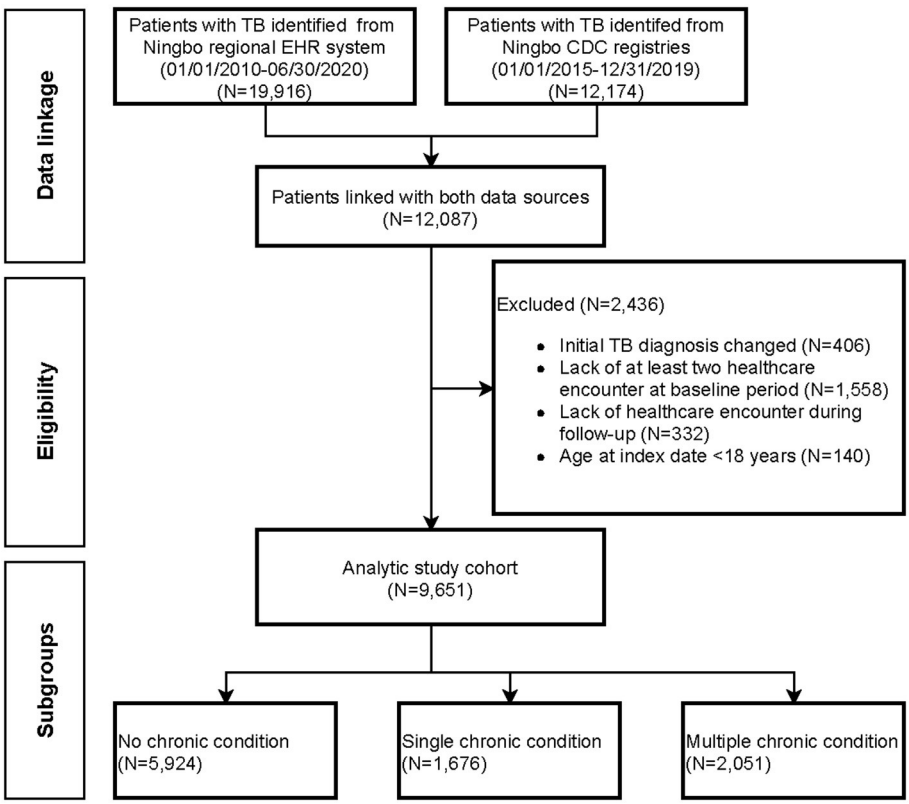

FIGURE 1 | Patient flow diagram documenting the application of selection criteria to arrive at the final analytical cohort.

encounters on different dates or in 1 inpatient hospitalization during the baseline period were retained in the analysis. Patients were categorized as having 0, 1, or multiple (2 and above) chronic conditions, excluding TB.

Latent class analysis (LCA) was performed on the subset of patients with ≥2 chronic conditions. LCA is a data-driven, model-based clustering method that is capable of identifying underlying subgroups of patients with similar multimorbidity patterns. We included all chronic conditions with a prevalence of 1% and above. A series of models were developed and the final model was selected based on statistical convergence, the goodness of fit, interpretability, and parsimony (10–12). A subject-level latent class multimorbidity group was assigned based on the highest predicted probability of membership, calculated from individual chronic condition profiles (13).

Variables

We measured four outcomes during the follow-up period: (1) hospital admission; (2) outpatient visit; (3) monthly total healthcare expenditures; and (4) monthly patient out-of-pocket payment. Total expenditures were calculated based on overall inpatient, outpatient and pharmacy expenditures from a healthcare system's perspective. Estimates were adjusted for inflation using the consumer price index (CPI) and presented in 2014 Chinese Yuans (¥). A hundred CNY was equivalent to 16.3 USD in 2014.

Additional demographic, lifestyle habits, insurance type, TB-related variables, prior resource use were measured during the baseline period. Demographic variables included age, sex, ethnicity, occupation, education level, and marital status.

Statistical Analysis

Patients in each of the morbidity groups were compared with patients without pre-existing chronic conditions at baseline. Categorical variables were compared using chi-square tests and continuous variables were compared using non-parametric Kruskal-Wallis tests accounting for non-normal distribution of variables. Missing data were addressed using missing indicators.

We compared the utilization of inpatient services across latent chronic condition groups using a multivariate logistic lasso regression model, with excess risk for admission expressed in adjusted odds ratio (OR) and 95% Confidence intervals (CIs). We classified healthcare expenditure into 3 types: outpatient expenditure, inpatient expenditure, and outpatient pharmacy expenditure. To account for excess zeros and the skewed non-zero distribution nature of healthcare expenditure data, we used a 2-part regression model to identify the association between morbidity groups and healthcare expenditure. We first fitted a logistic model to estimate the association between morbidity groups with none vs. any utilization of healthcare. Among patients with non-zero expenditure, a second model of the generalized linear model (GLM) using gamma distribution with a log link function (14, 15). We used GLM in the second stage rather than the ordinary least squares of log-transformed healthcare expenditure because the former relaxes the homoscedastic error assumption, which is essential for retransformation to the raw scale (16).

TABLE 1 | Baseline characteristics of tuberculosis patients cohort (*N* = 9,651).

Characteristics	*N*	%
Age		
Mean age at TB diagnosis (SD), y	48.8	18.7
Median age (IQR), y	49	32–63
Age group, y		
18–40	3,641	37.7
41–50	1,465	15.2
51–60	1,603	16.6
61–70	1,570	16.3
71+	1,372	14.2
Female sex	3,073	31.8
Insurance type		
UEBMI	4,128	42.8
URRMI	771	8.0
Other	4,752	49.2
Retreatment status		
Newly diagnosed	8,998	93.2
Retreatment	653	6.8
TB clinical types		
III	8,998	93.2
II	653	6.8
IV	8,758	90.7
I	54	0.6
Extra pulmonary TB	795	8.2
HIV positive	30	0.31
Chronic conditions		
Mean, SD	1.02	1.99
None	5,924	61.4
1	1,676	17.4
2+	2,051	21.3
Charlson comorbidity index		
0	3,152	32.7
1	1,389	14.4
2	1,314	13.6
3	1,323	13.7
4+	2,473	25.6

TB, tuberculosis; SD, standard deviation; IQR, interquartile range; UEBMI, urban employee basic medical insurance; URRMI, urban and rural resident medical insurance; HIV, human immunodeficiency virus.

Analyses were conducted using SAS 9.4 statistical software (SAS Institute Inc., Cary, NC, USA), and open-source statistical software R 4.0.3 (poLCA package, https://CRAN.R-project.org). Two-side $P < 0.05$ was considered as statistically significant.

RESULTS

Sample Description

A total of 9,651 eligible patients with TB were included in the present cohort study, the patient flow diagram is shown in **Figure 1**. The final analytical cohort consisted of 3,073 (31.8%) females, with a mean age of 48.8 years (SD,

18.7) (**Table 1**). Approximately 30.5% of patients were aged above 60 years and 4,128 (42.8%) patients were covered by urban employee basic medical insurance schemes at the time of TB diagnosis. The prevalence of HIV infection was 0.31% and the majority had newly registered TB. Pre-existing chronic conditions were observed in 38.6% of TB patients. The average number of chronic conditions was 1.02 (SD< 1.99), and 2,051 (21.3%) had ≥2 chronic conditions. Essential primary hypertension was the most common chronic condition (14.3%), followed by hyperlipidemia (4.0%), chronic obstructive pulmonary disease (COPD) (2.2%), and atherosclerotic heart disease (2.2%).

TABLE 2 | Characteristics of model-driven chronic condition groups among eligible patients with tuberculosis.

	Respiratory morbidity		Cardiovascular morbidity with complications		Cardiovascular morbidity without complications		General morbidity		*P*-value
Class name									
Sample size	220	5.9%	185	5.0%	1,292	34.7%	2,030	54.5%	–
Mean age (SD)	73.6	10.5	74.5	10.5	65.2	12.1	51.4	16.3	<0.0001
Age group									
18–40	2	0.9%	0	0.0%	34	2.6%	558	27.5%	<0.0001
41–50	3	1.4%	3	1.6%	109	8.4%	368	18.1%	
51–60	12	5.5%	20	10.8%	269	20.8%	465	22.9%	
61–70	60	27.3%	43	23.2%	444	34.4%	376	18.5%	
71+	143	65.0%	119	64.3%	436	33.7%	263	13.0%	
Female sex	36	16.4%	51	27.6%	409	31.7%	673	33.2%	<0.0001
Insurance type									
UEBMI	114	51.8%	64	34.6%	635	49.1%	875	43.1%	<0.0001
URRMI	8	3.6%	19	10.3%	80	6.2%	149	7.3%	
Other	98	44.5%	102	55.1%	577	44.7%	1,006	49.6%	
Retreatment status	34	15.5%	24	13.0%	91	7.0%	166	8.2%	<0.0001
Smoking status									
Current Smoker	16	7.3%	22	11.9%	139	10.8%	156	7.7%	<0.0001
Former smoker	28	12.7%	19	10.3%	113	8.7%	89	4.4%	
No smoker	89	40.5%	69	37.3%	521	40.3%	715	35.2%	
Unknown	87	39.5%	75	40.5%	519	40.2%	1,070	52.7%	
Alcohol consumption status									
Drinker	18	8.2%	10	5.4%	68	5.3%	81	4.0%	<0.0001
Non-drinker	115	52.3%	99	53.5%	691	53.5%	864	42.6%	
Unknown	87	39.5%	76	41.1%	533	41.3%	1,085	53.4%	
Chronic conditions									
Mean, SD	4.0	4.0	8.6	8.5	3.0	2.6	2.5	2.0	<0.0001
Atherosclerotic heart disease	24	10.9%	98	53.0%	87	6.7%	0	0.0%	<0.0001
Cardiac arrhythmia	28	12.7%	40	21.6%	82	6.3%	13	0.6%	<0.0001
Chronic obstructive pulmonary disease	177	80.5%	36	19.5%	2	0.2%	0	0.0%	<0.0001
Chronic pharyngitis	5	2.3%	4	2.2%	48	3.7%	113	5.6%	0.0098
Chronic viral hepatitis B infection	1	0.5%	7	3.8%	13	1.0%	175	8.6%	<0.0001
Emphysema	44	20.0%	8	4.3%	0	0.0%	101	5.0%	<0.0001
Essential primary hypertension	69	31.4%	159	85.9%	1,106	85.6%	43	2.1%	<0.0001
Gastric ulcer	7	3.2%	22	11.9%	36	2.8%	50	2.5%	<0.0001
Gout	0	0.0%	23	12.4%	50	3.9%	39	1.9%	<0.0001
Heart failure	46	20.9%	78	42.2%	12	0.9%	0	0.0%	<0.0001
Hyper lipidemia	8	3.6%	110	59.5%	262	20.3%	10	0.5%	<0.0001
Non-toxic single thyroid nodule	6	2.7%	26	14.1%	42	3.3%	88	4.3%	<0.0001
Other disorders of glycoprotein metabolism	12	5.5%	42	22.7%	17	1.3%	95	4.7%	<0.0001
Sarcoidosis of lung	4	1.8%	11	5.9%	20	1.5%	88	4.3%	<0.0001
Sequelae of respiratory and unspecified tuberculosis	15	6.8%	21	11.4%	2	0.2%	100	4.9%	<0.0001
Presence of ≥1 mental health condition	2	0.9%	8	4.3%	27	2.1%	51	2.5%	0.1261

SD, standard deviation; UEBMI, urban employee basic medical insurance; URRMI, urban and rural resident medical insurance scheme.

TABLE 3 | Association between chronic condition groups and clinical outcomes.

Outcomes	Overall		No chronic conditions		Respiratory morbidity		Cardiovascular morbidity with complications		Cardiovascular morbidity without complications		General morbidity	
Sample size	9,651		5,924		220		185		1,292		2,030	
Overall successful treatment	8,182	84.8%	5,155	87.0%	158	71.8%	122	65.9%	1,066	82.5%	1,681	82.8%
Cured	2,857	29.6%	1,660	28.0%	105	47.7%	52	28.1%	426	33.0%	614	30.2%
Treatment completed	5,325	55.2%	3,495	59.0%	53	24.1%	70	37.8%	640	49.5%	1,067	52.6%
MDR-TB	121	1.3%	69	1.2%	3	1.4%	1	0.5%	11	0.9%	37	1.8%
ADR	40	0.4%	10	0.2%	3	1.4%	0	0.0%	16	1.2%	11	0.5%
Loss to follow-up	945	9.8%	539	9.1%	27	12.3%	37	20.0%	139	10.8%	203	10.0%
Non-TB death	165	1.7%	47	0.8%	15	6.8%	19	10.3%	37	2.9%	47	2.3%
Treatment failure	150	1.6%	87	1.5%	5	2.3%	3	1.6%	15	1.2%	40	2.0%
TB death	48	0.5%	17	0.3%	9	4.1%	3	1.6%	8	0.6%	11	0.5%

TB, tuberculosis; MDR-TB, multi-drug-resistant tuberculosis; ADR, adverse drug reaction.

Identification of Morbidity Groups

The LCA of patients with at least one chronic condition ($N = 3,727$) identified 4 morbidity subgroups. Demographic and clinical characteristics are presented in **Table 2**. The general morbidity group accounted for the largest percentage of TB patients with chronic conditions (54.5%), with an average of 2.5 chronic conditions. The mean age of patients in this group was 54.5 years (SD, 16.3), which is the youngest among all latent class subgroups. The most prevalent chronic condition in this group was chronic viral hepatitis B infection (8.62%) and chronic pharyngitis (5.57%). The cardiovascular morbidity without complications group (34.7% of the current study cohort) experienced a high prevalence of hypertension (85.6%) and hyperlipidemia (20.3%). The mean age was 65.2 years (SD, 12.1 years). The cardiovascular morbidity with complications group (5.0% of the current study group) was among the oldest groups, with a mean age of 74.5 years (SD, 10.5 years) at the time of TB diagnosis. Patients in this group experienced primarily hypertension (85.9%), hyperlipidemia (59.5%), atherosclerotic heart disease (53.0%), and heart failure (42.2%). In this group, the percentage of patients with mental health conditions was the highest (4.3%). Finally, the respiratory morbidity subgroup accounted for 220 (5.9%) of patients, and was predominantly male (83.6%). The mean age was 73.6 years (SD, 10.5 years) and had the highest percentage of patients with retreatment status (15.5%). The most prevalent condition was chronic obstructive pulmonary disease (80.5%) and essential primary hypertension (31.4%).

Association of Morbidity Groups With TB Treatment Outcome

As shown in **Table 3**, a total of 8,182 (84.8%) had successful TB treatment outcomes, with 29.6% cured and 55.2% completing TB treatment. The overall treatment success rate was higher in patients with no chronic conditions (87%), compared with morbidity subgroups. The cardiovascular morbidity without complications group and general morbidity group had overall TB treatment success rates of 82.5 and 82.8%, respectively. Cardiovascular morbidity with complications had the lowest TB treatment success rate of 65.9%, while 71.8% of the patients in the respiratory morbidity group had favorable outcomes overall.

Among the types of unfavorable outcomes, 9.8% of patients were lost to follow-up and 2.2% died during TB treatment. Among the 213 patients who died, 48 death events were attributable to TB, and 165 (1.7%) died from causes other than TB. Approximately 1.3% of patients developed MDR-TB, the percentage of patients who developed MDR-TB were comparable across all subgroups.

After adjusting for age and sex, patients in the cardiovascular morbidity with complications group had a 41% (95% CI 0.46–0.77) lower odds for achieving successful TB treatment outcomes, when compared with patients without chronic conditions. Conversely, patients in the cardiovascular morbidity without complications group had a 27% (95% CI 1.10–1.47) higher odds for achieving successful TB treatment outcomes. After adjusting for additional insurance type, TB retreatment status, and lifestyle variables, patients in the respiratory morbidity group (OR 0.68, 95% CI 0.49–0.93) and cardiovascular morbidity with complications group (OR 0.59, 95% CI 0.42–0.83) were less likely to have favorable TB outcomes. On the other hand, patients from the cardiovascular morbidity without complications group were more likely to achieve successful outcomes (OR 1.40, 95% CI 1.15–1.71) (**Supplementary Table 1**).

Health Resource Use Across Morbidity Groups

Overall healthcare expenditure was assessed from TB diagnosis date until TB treatment termination, death, end of the study period (June 30, 2020), or loss-to-follow-up, whichever came first. The univariate summary of overall healthcare expenditure is presented in **Supplementary Table 2**. On average, the all-cause total cost during TB treatment was 10,052 CNY (SD, 20,976 CNY, in 2014 terms). Differences were also observed with regard to latent class subgroups. Patients with no pre-existing chronic conditions experienced the lowest overall costs at 4,691

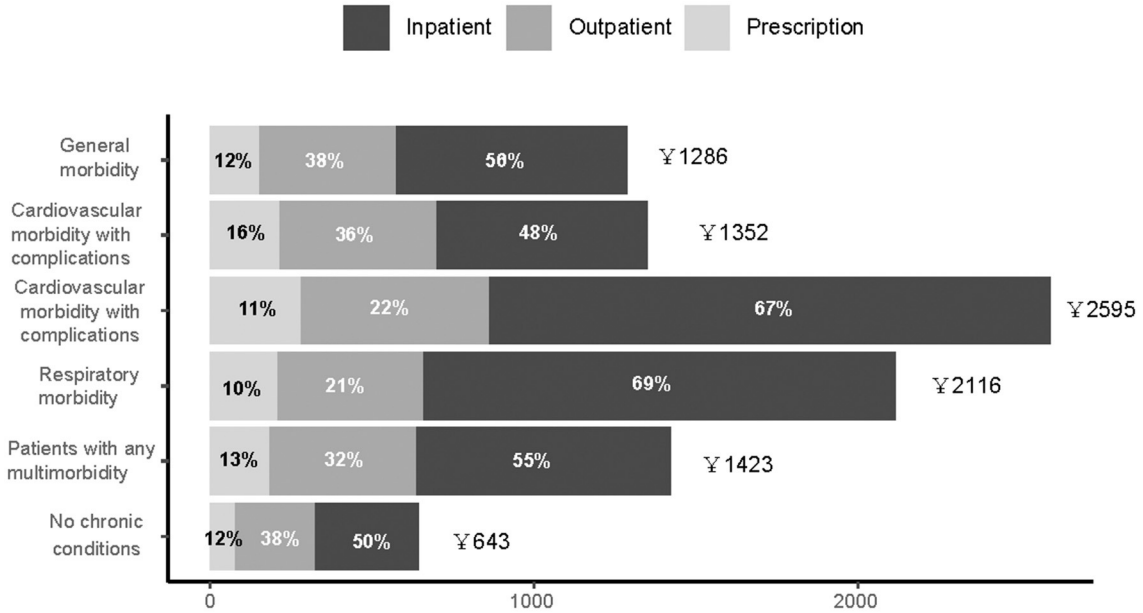

FIGURE 2 | Distribution of average overall individual healthcare expenditures among patients with tuberculosis, stratified by types of services.

CNY, while patients from the cardiovascular morbidity with complications group had the highest expenditure of 19,839 CNY (SD, 40,835 CNY). The excessive healthcare spending in this group was driven by inpatient expenditure.

To account for differential treatment durations, individual monthly healthcare expenditures according to chronic condition groups are presented in **Figure 2**. Monthly average and median healthcare expenditures were higher in the cardiovascular morbidity with complications group, followed by the respiratory morbidity group. However, monthly out-of-pocket (OOP) expenditures ranged between 237 and 531 CNY, with the share of OOP payment ranging between 20.0 and 36.9% (**Supplementary Table 3**).

In multivariable analysis, as shown in **Supplementary Table 4**, after adjusting for relevant demographic and clinical characteristics, patients in the respiratory morbidity group on average incurred an additional of 16,360 CNY (95% CI 12,615–21,215), compared with patients without any chronic conditions. Patients in cardiovascular morbidity with and without complication groups had incremental costs of 10,987 CNY (95% CI 9,542–12,651 CNY) and 10,279 CNY (95% CI 8,999–11,741 CNY), respectively. Patients in the general morbidity group had lower overall expenditure when compared with the reference group (7,832 CNY, 95% CI 6,928–8,853 CNY).

Patients with chronic conditions incurred statistically significantly greater OOP expenditures. Over the course of TB treatment, patients in the respiratory morbidity group had higher expenditure (4,543 CNY, 95% CI 3,163–6,525 CNY), when compared to patients without chronic conditions. Patients in the general morbidity group had a lower adjusted increase in OOP expenditure, costing 3,136 CNY (95% CI 2,632–3,736 CNY) more than the non-chronic condition group (**Supplementary Table 4**).

DISCUSSION

In this retrospective cohort analysis of 9,651 patients with TB, there were three notable findings. First, older patients have higher rates of multimorbidity, compared to previous publications. Second, we identified four distinct latent subgroups based on concurrent diagnoses, and these groups are predictive of TB treatment outcomes. Third, healthcare utilization and expenditure were highly correlated with latent subgroups, thereby such findings are informative in identifying potential targets of continuous care quality improvement.

TB patients with multimorbidity accounted for ∼40% of eligible subjects in this population-based cohort. We successfully linked a majority of patients from the National TB registry with healthcare encounter records. By examining over 10 years' regional electronic health records, we employed a systematic approach in identifying chronic conditions using validated algorithms (9). This method not only enabled us to characterize concurrent conditions at the time of TB diagnosis but also estimate the healthcare use patterns and associated expenditure from both healthcare systems and patients' perspectives. We found significant variability in both acute care utilization, outpatient services, and pharmaceutical products. While patients with cardiovascular conditions and respiratory morbidity experienced worse outcomes, these

patients incurred substantially higher monthly overall and OOP expenditure. The TB treatment success rates were lower among patients with greater comorbidity burdens compared with patients without multimorbidity or within the general morbidity group.

Interestingly, we found that after adjusting for relevant covariates, patients in the cardiovascular morbidity without complications group had higher treatment success rates relative to the non-chronic condition group. A higher proportion of patients in this group achieved cured outcomes (33.0% vs. 28.0%), despite a higher morbidity burden and older age. On the other hand, we found that patients without any comorbid conditions were more likely to be younger patients of the floating population of migrant workers without health insurance coverage. We hypothesized that these patients from the cardiovascular morbidity without complications group were more likely to be covered by a more generous UEBMI scheme and living in Ningbo (17). Thus, patients from this subgroup experienced more continuous health services, allowing for more swift care transitions (18).

Another strength of this study is the use of LCA in providing information regarding clinically meaningful subgroups (19). More simplistic methods employed a fixed list of chronic conditions in defining chronic condition groups, our data-driven approach allowed inclusion of all observed diagnoses to identify distinct subgroups (20–22). This approach takes account for potentially joint comorbidity patterns between more severe chronic conditions such as heart failure and common illnesses like hypertension and gout, our study is capable of elucidating more detailed condition combinations. Although LCA has been widely used to identify clusters of patient-reported symptoms or trajectories of healthcare utilization, to the best of our knowledge, no prior study to date has examined multimorbidity in TB patients using a population-based, real-world dataset. We distinguished subgroups of cardiovascular morbidity with and without complications, has to our knowledge, not been previously reported. The complex comorbidity combinations in a sample subset of patients (respiratory morbidity group 5.9%, and cardiovascular morbidity with complications, 5.0%), with on average 4.0 and 8.6 concurrent chronic conditions warrant further investigation. Patients in this group experienced greater risks for treatment discontinuation due to loss-to-follow-up and death, suggestive of multispecialty care plan in meeting the needs of these patients.

While the overall prevalence of mental health disorders is low, we observed higher psychological burdens in the cardiovascular morbidity with complication subgroup. This is likely an underestimating of actual psychological burdens since patients could develop mental health conditions like depression and anxiety disorders, substance use disorders over the course of TB treatment (23). Our findings highlighted psychological support, active case finding, and early intervention of mental health disorders.

Despite the difference in overall costs, the share of OOP expenditure was much lower in subgroups that incurred more overall expenditure. We found that inpatient services as well as the number of chronic conditions may be the primary driver of expenditures. Since the reimbursement rates tend to be more generous for hospitalization than outpatient services, the OOP expenditures have been offset by insurance reimbursement, which could have a positive impact on care continuation and relieving financial burdens that patients face. However, the one-size-fits-all approach of the current reimbursement schemes oftentimes fails to account for the disparate needs of patients with varying chronic condition profiles. Recent health services research on catastrophic expenditure among patients with TB from the Ningbo region found that 37.1% of TB patients and their households faced catastrophic economic risks (24), therefore, reimbursement rates should be tailored to provide additional support for financially vulnerable patients with greater healthcare needs.

Our study has several limitations. First, we relied on recorded administrative data to identify chronic conditions, which are subject to coding errors and potentially, misclassification bias. Second, our findings may not be generalizable to TB patients in other settings, such as those in HIV endemic countries. Third, our estimates on healthcare expenditure adopted a healthcare system and patients' perspective, which could be different from a payers' perspective. Lastly, there are unmeasured confounding factors that are due to the inherent limitations of our retrospective design, preclude us from establishing a causal relationship between chronic comorbidity combinations and meaningful clinical outcomes, healthcare use, and expenditure.

CONCLUSION

In conclusion, the results of the current study offered much-needed evidence on comorbidity patterns in TB patients from a real-world perspective. The distinct patterns of multimorbidity highlight opportunities to better personalize TB care for patients with pre-existing conditions and a tailored, multidisciplinary approach to managing an increasing older TB population facing multifaceted health requirements.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ningbo Municipal Center for Disease Control and Prevention. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

QC, YC, and JZ designed the study. QC, YC, YX, and FJ were primarily responsible for data analysis and writing the

article. JZ and TY contributed critically to the writing of the article. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Zhejiang Medical Research Project (2018KY733), the Natural Science Foundation of Ningbo (2019A610386 and 2019A610385), and Ningbo Medical and Health Brand Discipline (PPXK2018-10).

ACKNOWLEDGMENTS

The authors would like to thank all participants involved in the present study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2021.756717/full#supplementary-material

REFERENCES

1. WHO. *Global Tuberculosis Report, 2020*. (2020). Available online at: https://apps.who.int/iris/bitstream/handle/10665/336069/9789240013131-eng.pdf (accessed September 16, 2021).
2. The Academy of Medical Sciences, 2018 (2018). Available online at: https://acmedsci.ac.uk/file-download/82222577 (accessed September 16, 2021).
3. Hou J, Wang G, Wang F, Cheng J, Ren H, Zhuang H, et al. Guideline of prevention and treatment for chronic hepatitis B (2015 update). *J Clin Transl Hepatol.* (2017) 5:297–318. doi: 10.14218/JCTH.2016.00019
4. Gao Y, Liu M, Chen Y, Shi S, Geng J, Tian J. Association between tuberculosis and COVID-19 severity and mortality: a rapid systematic review and meta-analysis. *J Med Virol.* (2021) 93:194–6. doi: 10.1002/jmv.26311
5. Reis-Santos B, Gomes T, Macedo LR, Horta BL, Riley LW, Maciel EL. Prevalence and patterns of multimorbidity among tuberculosis patients in Brazil: a cross-sectional study. *Int J Equity Health.* (2013) 12:61. doi: 10.1186/1475-9276-12-61
6. Fortin M, Stewart M, Poitras ME, Almirall J, Maddocks H. A systematic review of prevalence studies on multimorbidity: toward a more uniform methodology. *Ann Fam Med.* (2012) 10:142–51. doi: 10.1370/afm.1337
7. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ.* (2007) 335:806–8. doi: 10.1136/bmj.39335.541782.AD
8. Healthcare Cost and Utilization Project (HCUP). *User Support (HCUP-US)*. (2021). Available online at: https://www.ahrq.gov/cpi/about/otherwebsites/hcupnet.ahrq.gov/index.html (accessed September 16, 2021).
9. *Chronic Condition Indicator (CCI) for ICD-10-CM (Beta Version).* (2021). Available online at: https://www.hcup-us.ahrq.gov/toolssoftware/chronic_icd10/chronic_icd10.jsp (accessed September 16, 2021).
10. Rosellini AJ, Szentkúti P, Horváth-Puhó E, Smith ML, Galatzer-Levy I, Lash TL, et al. Latent classes of posttraumatic psychiatric comorbidity in the general population. *J Psychiatr Res.* (2021) 136:334–42. doi: 10.1016/j.jpsychires.2021.02.013
11. Yang BR, Um HY, Lee MT, Kim MS, Jung SY. Characterizing tramadol users with potentially inappropriate co-medications: a latent class analysis among older adults. *PloS ONE.* (2021) 16:e0246426. doi: 10.1371/journal.pone.0246426
12. Zhou J, Nutescu EA, Han J, Calip GS. Clinical trajectories, healthcare resource use, and costs of long-term hematopoietic stem cell transplantation survivors: a latent class analysis. *J Cancer Survivor Res Pract.* (2020) 14:294–304. doi: 10.1007/s11764-019-00842-1
13. Schreiber JB. Latent class analysis: an example for reporting results. *Res Soc Administr Pharm.* (2017) 13:1196–201. doi: 10.1016/j.sapharm.2016.11.011
14. Manning WG, Mullahy J. Estimating log models: to transform or not to transform? *J Health Econ.* (2001) 20:461–94. doi: 10.1016/S0167-6296(01)00086-8
15. Pregibon D. Goodness of link tests for generalized linear models. *J R Stati Soci Ser C.* (1980) 29:15–24. doi: 10.2307/2346405
16. Belotti F, Deb P, Manning WG, Norton EC. twopm: two-part models. *Stata J.* (2015) 15:3–20. doi: 10.1177/1536867X1501500102
17. Pan Y, Chen S, Chen M, Zhang P, Long Q, Xiang L, et al. Disparity in reimbursement for tuberculosis care among different health insurance schemes: evidence from three counties in central China. *Infect Dis Poverty.* (2016) 5:7. doi: 10.1186/s40249-016-0102-4
18. Yuan B, Li J, Wu L, Wang Z. Multi-level social health insurance system in the age of frequent employment change: the urban unemployment-induced insurance transition and healthcare utilization in China. *Healthcare.* (2019) 7:77. doi: 10.3390/healthcare7020077
19. Nielsen AM, Kent P, Hestbaek L, Vach W, Kongsted A. Identifying subgroups of patients using latent class analysis: should we use a single-stage or a two-stage approach? A methodological study using a cohort of patients with low back pain. *BMC Musculoskelet Disord.* (2017) 18:57. doi: 10.1186/s12891-017-1411-x
20. Omar N, Wong J, Thu K, Alikhan MF, Chaw L. Prevalence and associated factors of diabetes mellitus among tuberculosis patients in Brunei Darussalam: a 6-year retrospective cohort study. *Int J Infect Dis.* (2021) 105:267–73. doi: 10.1016/j.ijid.2021.02.064
21. Araia ZZ, Mesfin AB, Mebrahtu AH, Tewelde AG, Osman R, Tuumzghi HA. Diabetes mellitus and its associated factors in tuberculosis patients in maekel region, eritrea: analytical cross-sectional study. *Diabetes Metab Synd Obes Targets Ther.* (2021) 14:515–23. doi: 10.2147/DMSO.S293557
22. Kang W, Du J, Yang S, Yu J, Chen H, Liu J, et al. The prevalence and risks of major comorbidities among inpatients with pulmonary tuberculosis in China from a gender and age perspective: a large-scale multicenter observational study. *Eur J Clin Microbiol Infect Dis.* (2021) 40:787–800. doi: 10.1007/s10096-020-04077-2
23. Lee G, Scuffell J, Galea JT, Shin SS, Magill E, Jaramillo E, et al. Impact of mental disorders on active TB treatment outcomes: a systematic review and meta-analysis. *Int J Tuberculosis Lung Dis.* (2020) 24:1279–84. doi: 10.5588/ijtld.20.0458
24. Yang T, Chen T, Che Y, Chen Q, Bo D. Factors associated with catastrophic total costs due to tuberculosis under a designated hospital service model: a cross-sectional study in China. *BMC Public Health.* (2020) 20:1009. doi: 10.1186/s12889-020-09136-z

Author Disclaimer: The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Chen, Che, Xiao, Jiang, Chen, Zhou and Yang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 03 November 2021
doi: 10.3389/fphar.2021.757309

Trends and Patterns of Antibiotics Use in China's Urban Tertiary Hospitals, 2016–19

Yulei Zhu[1,2], *Yang Qiao*[3], *Rouli Dai*[4], *Xin Hu*[5]* and *Xin Li*[2,6,7]*

[1]Office of Scientific Research, the Affiliated Stomatological Hospital of Nanjing Medical University, Nanjing, China, [2]School of Health Policy and Management, Nanjing Medical University, Nanjing, China, [3]First School of Clinical Medicine, Nanjing Medical University, Nanjing, China, [4]Medical Office, The Fourth Affiliated Hospital of Nanjing Medical University, Nanjing, China, [5]Department of Pharmacy, Beijing Hospital, National Center of Gerontology, Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing Key Laboratory of Assessment of Clinical Drugs Risk and Individual Application (Beijing Hospital), Beijing, China, [6]Department of Clinical Pharmacy, School of Pharmacy, Nanjing Medical University, Nanjing, China, [7]Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing, China

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Zhuo Wang,
Shanghai Changhai Hospital, China
Xu Ji,
Emory University, United States

***Correspondence:**
Xin Li
xinli@njmu.edu.cn
Xin Hu
huxinbjyy@126.com

Specialty section:
This article was submitted to Drugs Outcomes Research and Policies, a section of the journal Frontiers in Pharmacology

Received: 12 August 2021
Accepted: 17 September 2021
Published: 03 November 2021

Citation:
Zhu Y, Qiao Y, Dai R, Hu X and Li X (2021) Trends and Patterns of Antibiotics Use in China's Urban Tertiary Hospitals, 2016–19. Front. Pharmacol. 12:757309. doi: 10.3389/fphar.2021.757309

Objectives: This study aimed to identify the trends in antibiotics utilization and patients costs, evaluating the effect of the policy and exploring factors associated with the irrational use of antibiotics.

Methods: Based on the Cooperation Project Database of Hospital Prescriptions, data were collected from 89 tertiary hospitals in nine cities in China during 2016–2019. The study sample consisted of prescription records with antibiotics for 3,422,710 outpatient and emergency visits and 26, 118, 436 inpatient hospitalizations.

Results: For outpatients, the proportion of treated with antibiotics declined from 14.72 to 13.92% significantly ($p < 0.01$). The proportion of antibiotic costs for outpatients decreased from 5.79 to 4.45% significantly ($p < 0.01$). For emergency patients, the proportion of treated with antibiotics increased from 39.31 to 43.45% significantly ($p < 0.01$). The proportion of antibiotic costs for emergency patients decreased from 36.44 to 34.69%, with no significant change ($p = 0.87$). For inpatients, the proportion of treated with antibiotics increased from 23.82 to 27.25% significantly ($p < 0.01$). The proportion of antibiotic costs for outpatients decreased from 18.09 to 17.19% with no statistical significance ($p = 0.89$). Other β-lactam antibacterials (1,663.03 ten thousand DDD) far exceeded other antibiotics categories. Stablely ranked first, followed by Macrolides, lincosamide and streptogramins (965.74 ten thousand DDD), Quinolone antibacterials (710.42 ten thousand DDD), and β-lactam antibacterials, penicillins (497.01 ten thousand DDD).

Conclusions: The proportion of treated with antibiotics for outpatients and inpatients meet the WHO standards. The antibiotics use varied by different survey areas, clinical departments, patient gender, patient age and antibiotics categories. More efforts should focus on improving the appropriateness of antibiotics use at the individual level.

Keywords: antimicrobial stewardship, antibiotics use, tertiary hospitals, urban hospitals, China

INTRODUCTION

Antibiotics are medicinal products that kill or inhibit the growth of living microorganisms (European Centre for Disease Prevention and Control, 2018). As one of the most important discoveries in pharmaceuticals, antibiotics have saved millions of lives worldwide. However, when antibiotics are used the emergence of antimicrobial resistance (AMR) is inevitable. Statistically, the consumption of antibiotics increased by 36% from 2000 to 2010 (Van Boeckel et al., 2014). Meanwhile, the number of new antibiotics approved for the market worldwide is decreasing (Liu, 2018). AMR is regarded as one of the greatest threats to global health by the World Health Organization (WHO) (World Health Organization, 2015), which has the potential to affect anyone, of any age, in any country (World Health Organization, 2018). We have now reached a crisis where many antimicrobials are no longer effective against previously curable infectious diseases. As well as the emergence of AMR, prolonged hospitalization, more substantial social and economic burden, increased human suffering, and more premature deaths are the results of inappropriate antibiotics (Gandra et al., 2014; Li et al., 2019). According to a review chaired by the English Economist JIM O'NEILL, if AMR is not curtailed, 10 million people will die of drug-resistant bacterial infection every year by 2050 worldwide, with cumulative economic loss up to 100 trillion USD in total (Jim, 2015). In 2015, the WHO Global Action Plan on AMR was adopted by the member states at the 68th World Health Assembly, which aimed to promote the use of antibiotics (World Health Organization, 2016).

China is one of the largest consumers of antibiotics globally (Van Boeckel et al., 2014), and the abuse of antibiotics is relatively severe in this country (Yin et al., 2013; Wang et al., 2014). In an attempt to optimize the appropriate use of antimicrobial agents, the Chinese government has issued a series of policies and regulations in the last few years. In 2011, China initiated the National Special Rectification Campaign (NSRC) by its ministry of health targeting AMR and rational use of antimicrobial agents at secondary and tertiary hospitals (Chinese Ministry of Health, 2011), which achieved significant effects after several years of implementation (Bao et al., 2015; Xiao, 2018). On August 1, 2012, the strictest regulation yet for antimicrobial stewardship in China was formally taken into effect, which clearly defined all aspects of antibiotic use in health-care institutions (Xiao and Li, 2013). To ensure sustainable progress towards antibiotics' rational use, China also launched its National Action Plan to Curve Antimicrobial Resistance (2016–2020) in 2016. The critical intervention of the action plan for health-care institutions, including 1) to standardize management of the use of antibacterial agents; 2) to optimize the AMR surveillance network in clinics; 3) to enhance the training of professional personnel in the rational use of antibiotics and AMR (National Health and Family Plan Commission of China, 2016; Xiao and Li, 2016).

Although there have been many studies of antibiotic use worldwide, a few studies were multicenter based on continuous data, especially in China. Furthermore, only a few studies on antibiotic prescribing patterns in health-care institutions have been conducted in China, using national or provincial data. Of these, two studies surveys focused on antibiotic prescribing patterns in primary health-care institutions before 2012 (Dong et al., 2008; Wang et al., 2014); one analyzed antibiotic use in specialized hospitals during 2011–12 (Zou et al., 2014), and two focused on antibiotic prescribing and costs in secondary and tertiary health-care institutions before 2016 (Bao et al., 2015; Li et al., 2019). Bao et al. conducted a comparative analysis of antimicrobial use in 65 general public hospitals in China, which found that the National Action Plan for antibiotic stewardship targeting antibiotic misuse had a positive effect in reducing antibiotic consumption and expenditure (Bao et al., 2015). A study based on the Center for Antibacterial Surveillance also demonstrated a significant downward trend in antibiotic use in general hospitals in two provinces of China after implementing restriction. The policy intervention might optimize the use of the antibiotic (Li et al., 2019).

None of these studies conducted a comparative analysis of antibiotic use after implementing the strictest regulation in 2016. Further research is required to identify the effectiveness of policy interventions in China. Former studies demonstrated that most of the antibiotics were procured and consumed by secondary and tertiary hospitals, especially in tertiary hospitals (Lin et al., 2016; Yin et al., 2018). Besides, a few studies focused on comparable antibiotic use data at a population level in tertiary hospitals. Thus, based on the retrospective data from 89 tertiary hospitals in nine cities in China during 2016–2019. We focused on identifying the trends in antibiotics utilization and patients costs, evaluating the effect of the policy and exploring factors associated with the irrational use of antibiotics. We expected that this study's findings could provide a basis for future antibiotic stewardship by policy-makers and health-care institutions.

METHODS

Data Sources

A convenient sampling method was used in this study included 89 tertiary hospitals in China. Information pertaining to the use of antibiotics, including sample city, patient characteristic (gender and age), visit time, type of care (Outpatient/Emergency/Inpatient), clinical department, the number of all Outpatient/Emergency prescriptions or inpatient records with antibiotics, prescribed antibiotics (dosage form, strength, packaging specifications, pharmaceutical manufacturer, and cost) and additional information (the number of all prescriptions/records on each day of sampling, the cost of all pharmaceuticals on each day of sampling) was gathered from the Cooperation Project Database of Hospital Prescriptions (CPDHP). The CPDHP is an extensive database with medicine records procurement covering public tertiary hospitals in nine cities, China (including Beijing, Chengdu, Guangzhou, Hangzhou, Harbin, Shanghai, Shenyang, Tianjin and Zhengzhou). 44.44% of cities total gross domestic product (GDP) was over 15 ten thousand RMB (~21.8 thousand USD) in 2019, and 33.33% was below 10 ten thousand RMB (~14.5 thousand USD). A total of 89 tertiary hospitals were included in

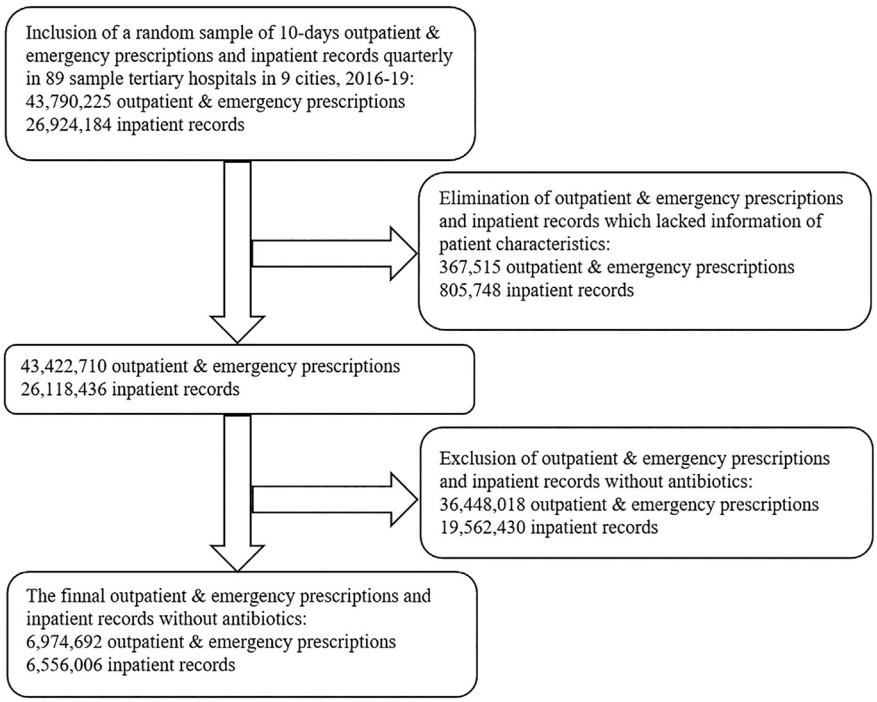

FIGURE 1 | Flow diagram showing the process of selecting prescriptions and records with antibiotics based on the Cooperation Project Database of Hospital Prescriptions.

the current study. Beijing accounted for the largest number of survey hospitals (13 out of 89), and Tianjing was the lowest (6 of 89). The value of Gross Domestic Product varied among the different survey cities, which was highest in Beijing (16.40 ten thousand yuan in 2019) and lowest in Harbin (5.52 ten thousand yuan in 2019). We included the clinical data from the first quarter of 2016 (2016Q1) to the fourth quarter of 2019 (2019Q4). A random sample of 10-days outpatient and emergency prescriptions and inpatient records quarterly in 89 sample tertiary hospitals in nine cities, 2016–19, were aggregated and standardized by engineers for research and decision-making. The events without complete information were excluded, as well as prescription/records without antibiotics (**Figure 1**).

All antibiotics prescribed were coded using Anatomical Therapeutic Chemical (ATC) classification system, which was divided into subgroups: Tetracyclines (J01A); β-lactam antibacterials, penicillins (J01C); Other β-lactam antibacterials (J01D); Macrolides, lincosamide and streptogramins (J01F); Quinolone antibacterials (J01M); Other antibiotics in J01; and antimycotics for systemic use (J02A). According to the guidelines for clinical governance, the antibiotics were categorized into three categories: non-restricted, restricted and controlled antibiotics. In the current study, we focused on systemic agents per oral or parenteral administration route and did not include topical agents such as skin creams, ophthalmic ointments, etc.

Indicators and Definitions

The significant indicators were selected in accordance with the standardized method of rational drug use developed by the WHO and International Network (WHO, 1993). Antibiotics prescribing was analyzed using seven indicators as follows:

1)Proportion of treated with antibiotics = numbers of prescriptions or records that included at least one antibiotic/total numbers of prescriptions or records × 100%;
2)Proportion of prescribed with muti-antibiotics = numbers of prescriptions or records that included greater than or equal to two antibiotics with different ATC codes/numbers of prescriptions or records that included at least one antibiotic × 100%;
3)Proportion of prescribed with parenteral antibiotics = numbers of prescriptions or records that included at least one injection/numbers of prescriptions or records that included at least one antibiotic × 100%;
4)Average drugs expenditure per visit = total drug expenditure/numbers of prescriptions or records;
5)Average antibiotics expenditure per visit = antibiotic expenditure/numbers of prescriptions or records that included at least one antibiotic;
6)Proportion of antibiotics expenditure = antibiotic expenditure/total drug expenditure;
7)Defined daily doses (DDDs) of antibiotics = total volume of medicine consumed (i)/DDD_i (DDD_i is a unit of measure that refers to the average maintenance dosage of each medicine used for its main indication in adults daily, which was obtained from the WHO DDD index web page) (The WHO Collaborating Centre for Drug Statistics Methodology, 2019).

TABLE 1 | Patient-related associated with antibiotic prescribing in health-care institutions in nine cities, 2016–19.

Characteristics	2016	2017	2018	2019
Total antibiotic-containing prescriptions(n)	3,162,377	3,315,731	3,411,229	3,641,361
Patient gender	—	—	—	—
Male	1,602,436 (50.67)	1,681,661 (50.72)	1,735,374 (50.87)	1,846,230 (50.70)
Female	1,559,941 (49.33)	1,634,070 (49.28)	1,675,855 (49.13)	1,795,131 (49.30)
Patient age	—	—	—	—
<5 years	251,576 (7.96)	264,135 (7.97)	263,248 (7.72)	278,383 (7.65)
5–17 years	213,394 (6.75)	223,776 (6.75)	221,732 (6.50)	265,223 (7.28)
18–49 years	1,121,451 (35.46)	1,161,607 (35.03)	1,193,072 (34.97)	1,267,680 (34.81)
50–64 years	725,602 (22.94)	753,637 (22.73)	784,684 (23.00)	819,011 (22.49)
≥65 years	850,354 (26.89)	912,576 (27.52)	948,493 (27.81)	1,011,064 (27.77)
Care type	—	—	—	—
Outpatient	1,347,105 (42.60)	1,358,617 (40.97)	1,395,468 (40.91)	1,500,187 (41.20)
Emergency	318,918 (10.08)	368,358 (11.11)	392,639 (11.51)	447,119 (12.28)
Inpatient	1,496,354 (47.32)	1,588,756 (47.92)	1,623,122 (47.58)	1,694,055 (46.52)
Clinical department	—	—	—	—
Emergency internal medicine	263,085 (8.32)	307,436 (9.27)	329,825 (9.67)	374,524 (10.29)
Emergency surgery	55,833 (1.77)	60,922 (1.84)	62,814 (1.84)	72,595 (1.99)
Internal medicine	787,401 (24.90)	800,498 (24.14)	819,174 (24.01)	860,962 (23.64)
Surgery	731,658 (23.14)	768,100 (23.17)	795,327 (23.31)	829,376 (22.78)
Stomatology	100,235 (3.17)	111,921 (3.38)	113,132 (3.32)	117,115 (3.22)
Respiratory medicine	347,636 (10.99)	356,565 (10.75)	372,205 (10.91)	391,078 (10.74)
Infectious disease	81,314 (2.57)	93,313 (2.81)	96,588 (2.83)	112,126 (3.08)
Obstetrics and gynecology	212,919 (6.73)	216,299 (6.52)	211,487 (6.20)	210,909 (5.79)
Pediatrics	365,991 (11.57)	383,804 (11.58)	380,064 (11.14)	431,723 (11.86)
ICU	56,938 (1.80)	59,688 (1.80)	61,880 (1.81)	65,569 (1.80)
Other	159,367 (5.04)	157,185 (4.74)	168,733 (4.95)	175,384 (4.82)
Survey region	—	—	—	—
Beijing	380,342 (12.03)	375,927 (11.34)	380,944 (11.17)	398,830 (10.95)
Chengdu	517,728 (16.37)	530,832 (16.01)	515,394 (15.11)	533,951 (14.66)
Shenyang	194,648 (6.16)	199,419 (6.01)	205,098 (6.01)	213,693 (5.87)
Tianjin	199,067 (6.29)	233,805 (7.05)	250,094 (7.33)	269,587 (7.40)
Zhengzhou	194,049 (6.14)	204,359 (6.16)	218,265 (6.40)	230,066 (6.32)
Guangzhou	570,311 (18.03)	579,717 (17.48)	604,075 (17.71)	659,371 (18.11)
Hangzhou	430,645 (13.62)	467,655 (14.10)	519,215 (15.22)	566,629 (15.56)
Harbin	182,713 (5.78)	190,673 (5.75)	179,559 (5.26)	200,107 (5.50)
Shanghai	492,874 (15.59)	533,344 (16.09)	538,585 (15.79)	569,127 (15.63)

Unless otherwise indicated, values shown are n(%).

Statistical Analysis

The quantitative data were analyzed by mean value each quarter to assess trends in the overall number or rate of indicators during the study period. We used the Cochran-Armitage trend test was used to examine the trend of categorical variables over time and linear regression analysis was applied to test the trend of continuous variables over time. The Cochran-Armitage test emphasizes the dichotomous data of occurrence, which can be applied to evaluate the trend in the rates such as disease incidence, symptom Occurrence and etc (Bieler and Williams, 1993; Wang et al., 2017). Besides, the test was performed in one survey focused on the trend in antibiotics prescription rates in general hospitals in two provinces of China (Li et al., 2019). In this study, one of the classified variables is whether a prescription contained antibiotics or not, and the other variable is quarter. R 4.0.4 and IBM SPSS (version 22.0; Chicago, IL, United States) was used to perform the above analysis, and the p-value <0.05 was considered statistically significant.

RESULTS

General Information of the Sampled Prescriptions/Records With Antibiotics

In the nine cities selected, we collected all 89 public hospitals prescribing data for 4 years consecutively. A total of 5,601,377 outpatient prescriptions, 1,527,034 emergency prescriptions and 6,402,287 inpatient records included at least one antibiotic were obtained in the current study. Among them, 50.74% were male patients, and 49.26% were female patients. Among 13, 530, 698 sample data, 35.06% were patients between the age of 18 and 49 years old. Internal medicine (24.15%) and Surgery (23.09%) were the top two clinical departments where patients received prescriptions containing antibiotics. The details of the annual data are shown in **Table 1**.

Overall Pharmaceuticals Consumption

Trends in the usage of pharmaceuticals are shown in **Figure 2** and **Figure 3**. The average amount of all drugs prescriptions for outpatients was 251.43 ten thousand, and it increased

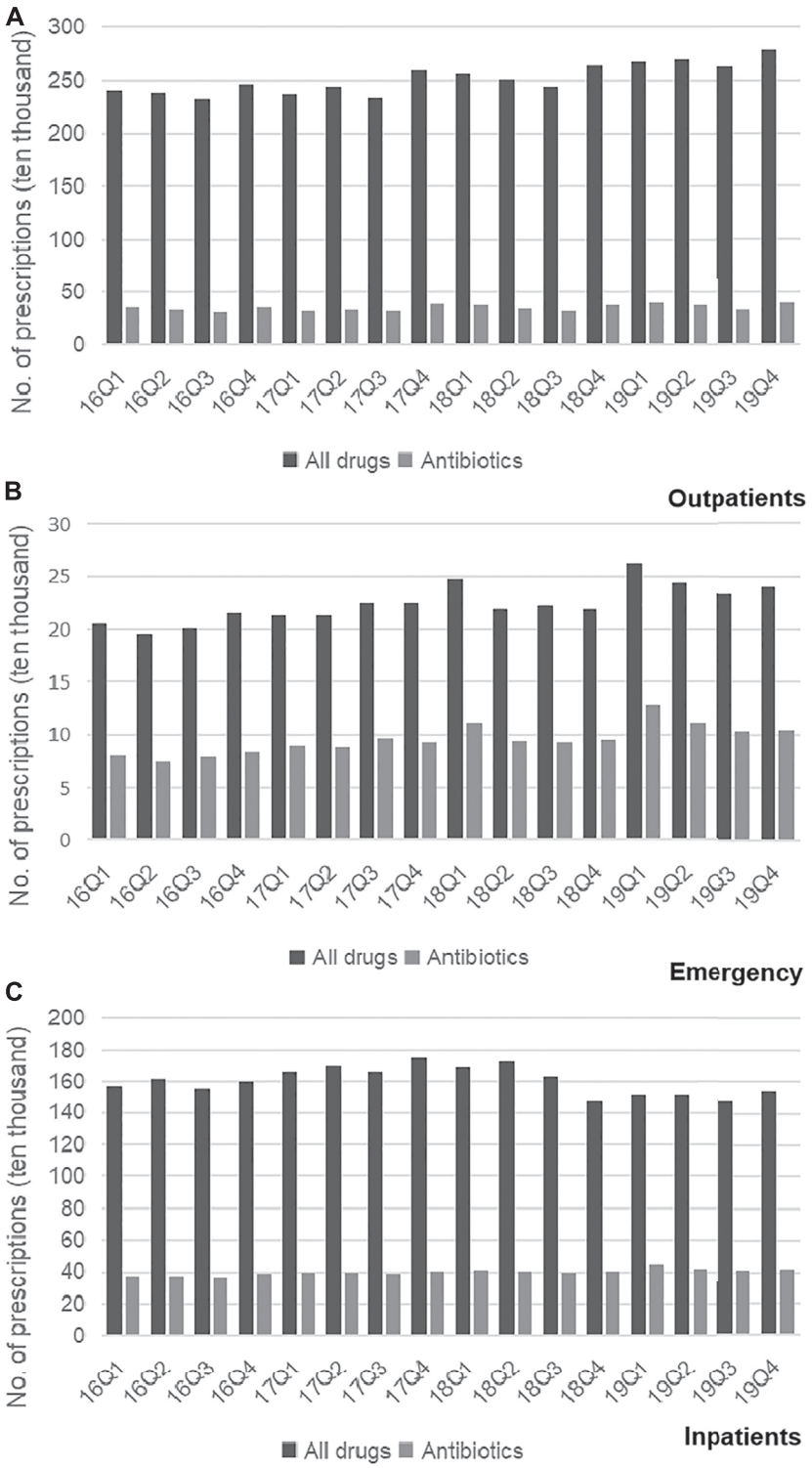

FIGURE 2 | (A) Changes in prescriptions of drugs in Outpatients, 2016–19. **(B)** Changes in prescriptions of drugs in Emergency, 2016–19. **(C)** Changes in prescriptions of drugs in Inpatients, 2016–19.

significantly throughout the study period ($p < 0.01$); the average amount of prescriptions containing antibiotics was 35.01 ten thousand, and it increased significantly throughout the study period ($p < 0.05$). In contrast, the average proportion of treated with antibiotics and proportion of prescribed with muti-antibiotics for outpatients was 13.92 and 12.21%, respectively,

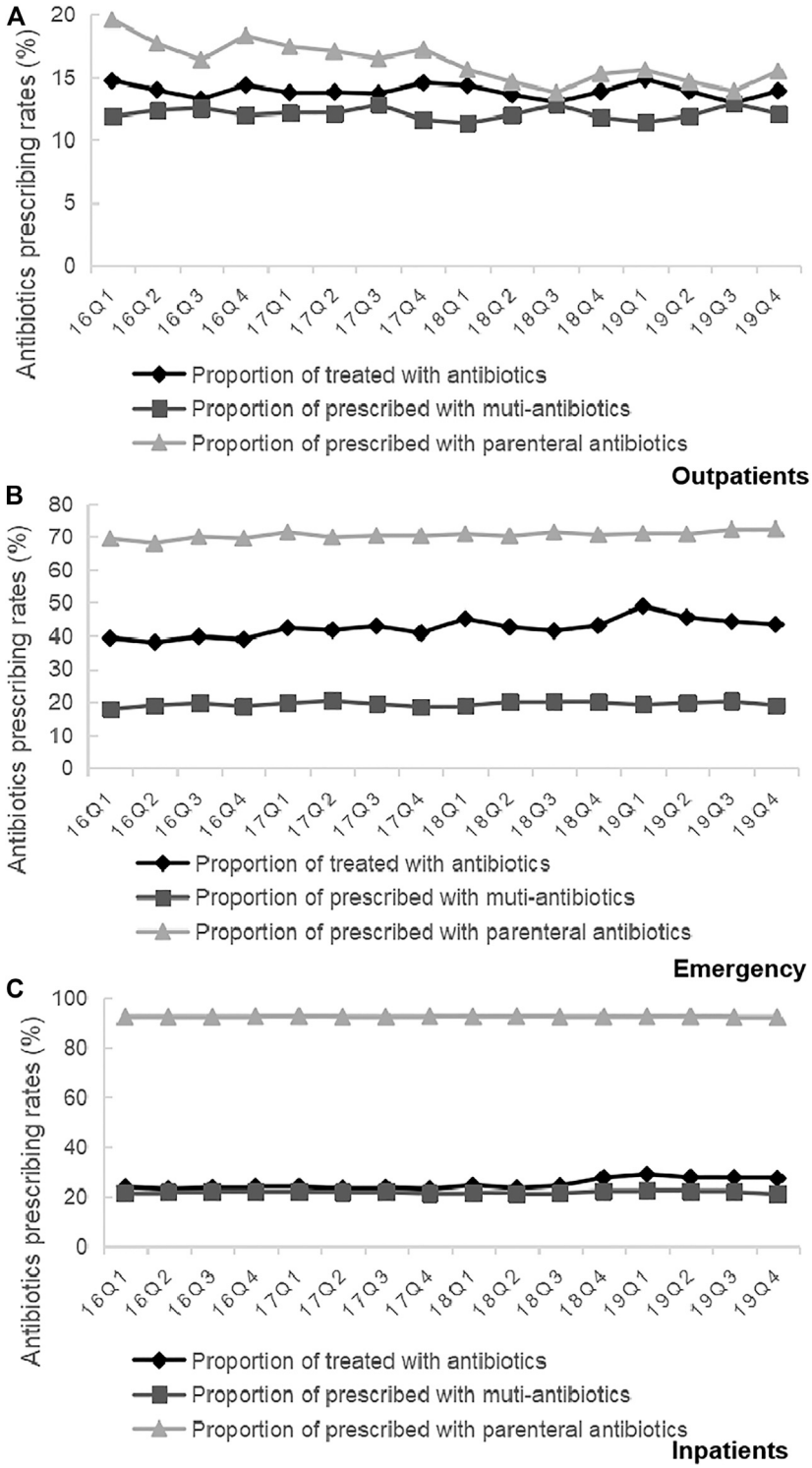

FIGURE 3 | (A) Changes in antibiotics prescribing rates in Outpatients, 2016–19. **(B)** Changes in antibiotics prescribing rates in Emergency, 2016–19. **(C)** Changes in antibiotics prescribing rates in Inpatients, 2016–19.

which dropped from 2016Q1 to 2019Q4 ($p < 0.01$). The average amount of all drugs prescriptions for patients of emergency department was 22.39 ten thousand, and it increased significantly throughout the study period ($p < 0.01$); the average amount of prescriptions containing antibiotics was 9.54 ten thousand, and it increased significantly throughout

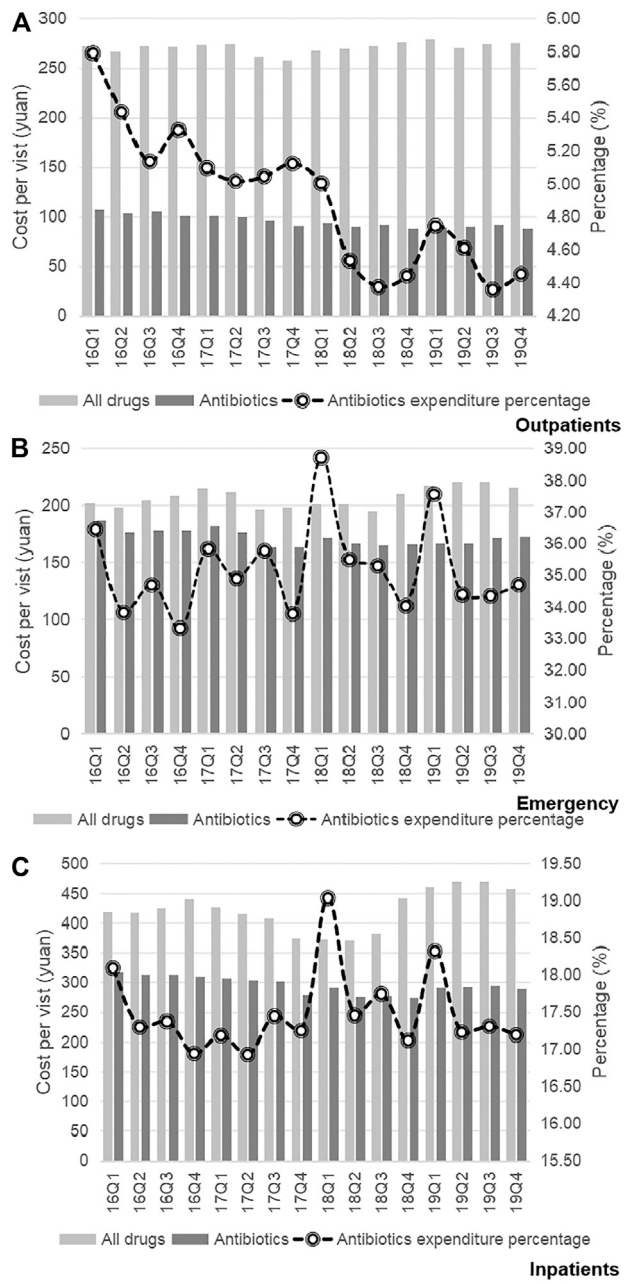

FIGURE 4 | (A) Changes in antibiotics expenditure in Outpatients, 2016–19. **(B)** Changes in antibiotics expenditure in Emergency, 2016–19. **(C)** Changes in antibiotics expenditure in Inpatients, 2016–19.

the study period ($p < 0.01$). The average proportion of treated with antibiotics and proportion of prescribed with parenteral antibiotics in emergency department was 42.44 and 70.50%, respectively, which increased from 2016Q1 to 2019Q4 ($p < 0.01$). For inpatients, the number of all drugs records was 163.24 ten thousand on average, with no significant changes during the study period ($p = 0.14$); the average amount of records containing antibiotics was 40.98 ten thousand, and it increased significantly throughout the study period ($p < 0.01$). The proportion of treated with antibiotics was 25.19% on average, showing significant uptrend over the 16 quarters ($p < 0.01$). **Figure 4** shows the changes in antibiotics expenditure. As for the expenditure of the drugs, the cost of antibiotics for outpatients increased from 201.91 yuan to 216.11 yuan ($p < 0.01$), and the proportion of antibiotics cost dropped from 5.79 to 4.45% ($p < 0.05$); but there was no significant change in the cost of all drugs ($p = 0.31$). The cost of all drugs in emergency departments increased from 201.91 yuan to 216.11 yuan ($p < 0.05$), and the cost of antibiotics cost in dropped from 187.18 yuan to 172.56 yuan ($p < 0.01$); but there was no significant change in the proportion of antibiotics cost ($p = 0.31$). For inpatients, the cost of antibiotics showed a linear downward trend over time, with a decrease of 30.09 yuan ($p < 0.01$), but no significant change was observed in the cost of all drugs ($p = 0.14$) and the proportion of antibiotics cost ($p = 0.75$).

Antibiotics Prescribing Patterns

Table 2 shows the quarterly trend in the proportion of antibiotics from 2016 to 2019 in different clinical departments. For outpatients, the proportion of treated with antibiotics in Infectious disease ($p < 0.01$) and Other ($p < 0.01$) increased over time; the proportion of treated with antibiotics in Obstetrics and gynecology ($p < 0.01$), Respiratory medicine ($p < 0.01$), Stomatology ($p < 0.01$), Internal medicine ($p < 0.01$) and Surgery ($p < 0.01$) were opposite. For patients in emergency departments, there was an uptrend in the proportion of treated with antibiotics in the Emergency internal medicine ($p < 0.01$) and Emergency surgery ($p < 0.01$). For inpatients, the proportion of treated with antibiotics in the Pediatrics ($p < 0.01$), Obstetrics and gynecology ($p < 0.01$), Infectious disease ($p < 0.01$), Respiratory medicine ($p < 0.01$), Stomatology ($p < 0.01$), Internal medicine ($p < 0.01$), Surgery ($p < 0.01$), ICU ($p < 0.01$) and Other ($p < 0.01$) increased from 2016Q1 to 2019Q4. In terms of the muti-antibiotics usage, the percentage in Internal medicine ($p < 0.01$) for outpatients increased from 2016Q1 to 2019Q4, which in contrast to the percentage in Pediatrics ($p < 0.01$), Infectious disease ($p < 0.01$), Respiratory medicine ($p < 0.01$), Surgery ($p < 0.01$) and Other ($p < 0.01$). For patients in emergency departments, there was an uptrend in the proportion of muti-antibiotics in the Emergency internal medicine ($p < 0.01$), but the proportion of muti-antibiotics in Emergency surgery was opposite ($p < 0.01$). For inpatients, the proportion of prescribed with muti-antibiotics in the Pediatrics ($p < 0.01$), Obstetrics and gynecology ($p < 0.01$), Respiratory medicine ($p < 0.01$), Internal medicine ($p < 0.01$), ICU ($p < 0.01$) and Other ($p < 0.01$) increased over time, but the proportion of prescribed with muti-antibiotics in Surgery ($p < 0.01$) was opposite (**Table 3**). In terms of the parenteral antibiotics usage, the percentage in Pediatrics ($p < 0.01$), Obstetrics and gynecology ($p < 0.01$), Respiratory medicine ($p < 0.01$), Internal medicine ($p < 0.01$) and Surgery ($p < 0.01$) for outpatients dropped from 2016Q1 to 2019Q4. For patients in emergency departments, there was an uptrend in the proportion of prescribed with parenteral antibiotics in the Emergency internal medicine ($p < 0.01$) and Emergency surgery ($p < 0.01$). For inpatients, the proportion of prescribed with parenteral antibiotics in Obstetrics and gynecology ($p < 0.01$) increased

TABLE 2 | Changes in proportion of treated with antibiotics in different clinical departments, 2016–19.

Care type	Clinical department	16Q1	16Q2	16Q3	16Q4	17Q1	17Q2	17Q3	17Q4	18Q1	18Q2	18Q3	18Q4	19Q1	19Q2	19Q3	19Q4	Z[a]
Outpatient prescriptions																		
	Pediatrics	45.96	46.08	41.8	48.97	44.58	46.92	42	47.67	43.16	47.26	42.17	47.81	44.14	46.52	43.41	46.47	−1.511
	Obstetrics and gynecology	23.25	22.47	22.21	22.34	22.82	22.64	22.46	22.37	22.2	22.35	21.8	21.26	21.37	21.35	20.85	20.77	−11.717**
	Infectious disease	21.22	18.87	20.4	16.81	20.24	20.13	24.79	18.7	24.68	19.21	20.91	17.84	26.53	22.31	23.18	19.66	10.493**
	Respiratory medicine	52.29	49.26	46.22	48.18	50.24	48.31	44.48	45.99	49.47	45.19	42.36	45.86	50.29	46.87	43.23	45.21	−19.759**
	Stomatology	53.64	52.74	53.32	54	53.16	51.93	52.51	52.67	53.76	52.35	51.43	49.76	49.65	50.68	49.35	48.96	−13.615**
	Internal medicine	6.12	5.85	5.8	5.75	5.55	5.37	5.82	5.56	5.7	5.27	5.51	5.39	5.83	5.4	5.51	5.42	−16.503**
	Surgery	24.58	24.58	24.55	23.72	23.76	23.9	23.72	22.5	22.1	21.99	22.14	21.39	21.4	21	21	20.67	−37.848**
	Other	8.92	8.42	8.51	8.86	8.18	8.15	8.95	9	10.54	9.56	9.31	9.56	11.01	9.43	8.66	9.49	16.714**
Emergency prescriptions																		
	Emergency internal medicine	40.45	38.6	40.08	39.41	43.87	42.71	44.5	41.78	46.58	43.75	42.54	42.83	49.28	45.59	44.18	42.92	33.050**
	Emergency surgery	34.11	35.81	38.61	36.66	35.39	37.66	36.89	37.02	37.77	37.91	37.41	44.62	47.53	45.62	44.98	46.29	33.521**
Inpatient records																		
	Pediatrics	39.39	39.5	38.05	43.01	41.36	40.9	38.94	41.42	39.74	41.4	39.96	52.47	51.01	52.4	49.56	51.46	47.203**
	Obstetrics and gynecology	36.28	36.55	36.32	37.06	36.17	36.42	36.14	36.27	35.51	36.35	35.94	41.11	41.11	41.29	40.58	39.74	18.951**
	Infectious disease	35.27	34.97	34.91	33.7	36.16	36.26	37.8	36.45	38.08	36.22	36.77	39.02	42.25	43.3	42.7	40.54	20.799**
	Respiratory medicine	52.49	52.34	52.69	54.33	54.96	53.31	52.45	51.43	52.91	52.12	52.74	60.11	61.32	59.27	58.58	57.73	28.821**
	Stomatology	45.82	44.76	49.58	49.61	50.74	51.36	54.7	51.8	52.28	52.4	56.83	63.4	64.98	64.86	65.47	61.51	24.124**
	Internal medicine	15.53	14.45	14.81	14.84	15.11	14.38	14.63	14.16	15.79	14.15	14.95	16.22	17.82	16.39	16.56	16.19	34.522**
	Surgery	30.89	30.95	31.8	32.13	31.39	31.22	32.14	31.04	32.28	32.76	34.23	38.61	39.08	38.81	39.04	38.03	84.067**
	ICU	27.41	26.83	27.77	28.88	28.01	30.29	31.45	30.42	31.35	30.42	31.15	36.02	38.93	36.96	36.48	35.11	36.336**
	Other	12.95	11.6	12.03	12.16	12.89	11.98	12.29	12.17	13.78	11.84	12.37	16.11	18.97	16.41	16.03	16.01	37.138**

[a]*The Cochran-Armitage trend test was applied to calculate* p *values.*

TABLE 3 | Changes in proportion of prescribed with muti-antibiotics in different clinical departments, 2016–19.

Care type	Clinical department	16Q1	16Q2	16Q3	16Q4	17Q1	17Q2	17Q3	17Q4	18Q1	18Q2	18Q3	18Q4	19Q1	19Q2	19Q3	19Q4	Z[a]
Outpatient prescriptions																		
	Pediatrics	6.15	5.94	5.76	6.33	6.25	5.78	5.23	5.39	5.56	5.33	4.98	5.78	5.92	5.92	5.98	5.33	-3.812**
	Obstetrics and gynecology	13.23	12.51	12.62	13.53	14.18	14.54	14.92	15.41	14.78	15.05	15.17	13.93	12.75	12.22	12.98	12.88	-0.706
	Infectious disease	15.37	17.74	13.78	13.89	15.15	17.4	16.21	16.34	13.35	13.79	13.93	14.81	12.14	13.95	14.7	15.02	-4.391**
	Respiratory medicine	10.01	10.43	9.31	8.45	9.04	8.52	8.76	7.28	7.24	7.52	6.98	6.12	7.22	7.59	7.67	7.7	-16.134**
	Stomatology	38.46	36.92	37.54	36.98	36.66	36.45	37.68	38.65	38.94	37.87	38.84	37.1	38.08	37.67	38.04	38.33	1.952
	Internal medicine	14.75	15.94	15.66	16.51	15.89	15.78	15.94	16.64	14.63	16.35	17.55	18.25	16.63	17.59	18.76	19.64	19.379**
	Surgery	7.48	7.58	7.38	7.28	6.59	6.99	7.1	6.16	6.1	6.65	7.12	6.31	6.15	5.88	6.19	5.86	-11.327**
	Other	8.37	9.5	10.05	8	7.99	8.63	9.43	7.76	8.07	8.55	8.67	8.01	7.68	7.98	8.59	7.96	-3.945**
Emergency prescriptions																		
	Emergency internal medicine	16.77	18.06	18.94	17.75	18.6	19.51	18.65	17.8	18.32	19.15	19.52	19.17	18.64	19.18	19.84	18.57	6.765**
	Emergency surgery	23.86	23.57	22.93	23.94	26.56	25.23	23.43	22.42	23.3	24.77	23.4	23.64	22.86	23.32	22.32	21.66	-3.624**
Inpatient records																		
	Pediatrics	17.71	17.57	19.1	19.37	18.09	17.6	17.55	18.45	16.94	17	19.34	21.13	19.52	19.28	21.37	19.9	7.345**
	Obstetrics and gynecology	28.11	28.33	29.85	29.15	29.42	29.28	28.89	28.32	27.48	28.5	28.32	31.44	30.77	30.98	29.93	29.82	4.508**
	Infectious disease	29.37	29.64	29.79	27.11	29.76	29.36	29.01	28.49	30.35	29.04	28.69	29.54	28.87	30.26	29.47	26.96	-0.788
	Respiratory medicine	25.41	25.6	26.05	25.28	26.15	25.37	25.24	24.86	27.27	25.52	25.55	26.38	28.65	27.11	27.36	26.17	6.318**
	Stomatology	25.98	30.1	25.9	27.54	28.25	29.37	25.61	29.21	26.55	27.49	26.22	30.71	30.73	29.46	26.23	26.39	0.6
	Internal medicine	17.73	18.53	18.86	18.34	18.22	18.57	18.87	17.86	17.7	17.94	18.14	18.65	19.04	18.83	18.82	18.3	2.411*
	Surgery	20.3	20.99	20.66	21.03	20.76	20.56	20.4	20.01	20	19.51	19.78	20.01	19.66	19.89	18.88	17.98	-13.240**
	ICU	35.91	34.6	36.04	36.43	37.66	36.5	37	35.59	38.63	36.18	35.67	36.61	39.24	37.43	37.41	36.09	2.700**
	Other	18.41	17.59	18.51	19.8	19.52	18.49	20.04	18.73	21.31	20.03	20.05	21.75	22.95	22.44	21.58	20.66	9.403**

[a]*The Cochran-Armitage trend test was applied to calculate* p *values.*

TABLE 4 | Changes in proportion of prescribed with parenteral antibiotics in different clinical departments, 2016–19.

Care type	Clinical department	16Q1	16Q2	16Q3	16Q4	17Q1	17Q2	17Q3	17Q4	18Q1	18Q2	18Q3	18Q4	19Q1	19Q2	19Q3	19Q4	Z[a]
Outpatient prescriptions																		
	Pediatrics	39.84	35.91	34.03	39.21	38.12	35.35	34.52	36.5	35.66	31.8	30.59	35.31	35.01	30.91	31.04	34.15	-7.345**
	Obstetrics and gynecology	9.44	9.57	8.53	8.5	8.64	8.61	7.55	7.03	6.41	6.04	6.99	6.19	5.88	5.44	5.41	4.9	-4.508**
	Infectious disease	45.83	43.17	43.49	41.06	47.15	45.69	48.07	46.94	42.85	40.44	41.28	41.72	39.51	38.11	40.6	42.3	-0.788
	Respiratory medicine	17.78	14.25	13.61	13.2	14.38	11.76	11.6	9.48	10.08	7.77	7.16	7.33	8.21	7.47	6.16	5.61	-6.318**
	Stomatology	10.11	9.98	10.35	9.27	9.74	10.8	10.96	10.55	10.57	11.98	12.61	12.2	11.28	12.63	12.21	12.19	0.6
	Internal medicine	13.32	12.84	12.69	11.92	11.72	12.15	12.54	11.04	10.27	10.25	9.85	8.88	9.99	9.95	9.91	9.66	-2.411*
	Surgery	13.86	14.06	14.19	12.69	11.76	11.83	11.59	9.87	8.39	9.24	9.53	7.94	8.05	8.54	9.12	8.12	-13.240**
	Other	7.22	5.98	6.43	5.4	6.01	6.18	6.38	4.43	9.13	8.05	7.52	6.73	8.95	8.1	7.3	7.49	9.403**
Emergency prescriptions																		
	Emergency internal medicine	67.47	65.63	68.43	67.8	69.91	68.04	68.82	67.85	69.05	68.13	69.93	68.94	69.16	68.88	70.85	70.83	11.738**
	Emergency surgery	79.85	78.57	76.52	77.76	80.38	78.49	77.65	81.4	83.04	80.41	78.18	78.52	82.13	80.6	78.18	79.64	3.384**
Inpatient records																		
	Pediatrics	95.15	94.76	94.82	95.33	95.59	95.65	95.62	95.64	95.62	95.48	95.58	95.86	95.39	95.3	94.79	95.05	0.398
	Obstetrics and gynecology	91.68	91.62	91.73	92.82	91.73	91.57	92.05	92.07	93.09	92.86	92.62	92.57	92.29	92.85	92.35	91.86	3.258**
	Infectious disease	91.1	91.19	91.42	91.56	91.77	91.36	90.7	91.3	90.66	89.91	89.08	88.8	89.24	90.05	89.61	89.61	-6.427**
	Respiratory medicine	94.58	94.6	94.66	94.72	94.56	94.37	94.63	94.67	93.97	94.03	94.02	94.34	94.43	94.38	94.07	94.23	-3.244**
	Stomatology	95.39	94.93	96.53	95.86	95.33	95.26	96.13	96.53	96.72	97.15	96.74	95.24	95.95	95.32	95.27	94.64	-0.804
	Internal medicine	90.05	89.65	89.48	89.65	89.96	89.67	89.61	89.88	89.97	89.95	89.34	89.52	89.99	89.75	89.63	89.61	-0.762
	Surgery	93.88	94.13	94.17	94.17	94.25	94.08	93.95	94.28	94.35	94.43	94.4	94.26	94.42	94.29	94	93.95	1.309
	ICU	95.83	96.46	96.07	96.1	95.94	95.87	95.77	95.95	95.64	96.12	97.03	96.53	96.31	96.38	97.07	95.37	1.595
	Other	89.82	89.21	89.88	90.03	90.15	89.29	90.07	89.63	88.86	88.25	88.37	88.32	89.16	87.78	88.21	88.36	-5.880**

[a]*The Cochran-Armitage trend test was applied to calculate* p *values.*

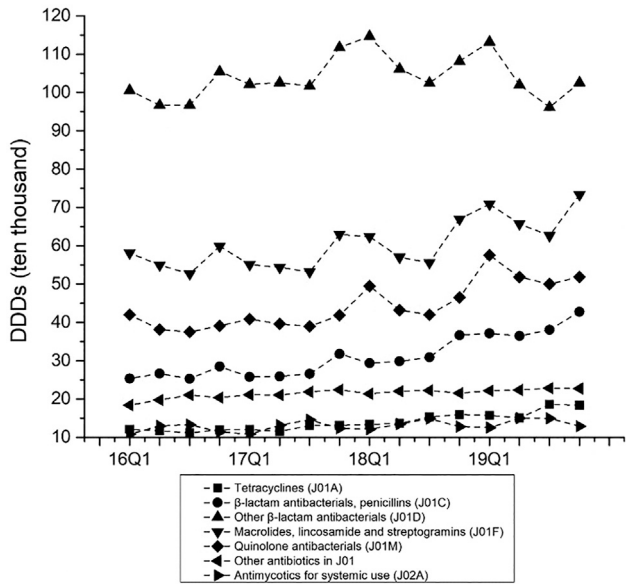

FIGURE 5 | Changes in antibiotics use by ATC classification, 2016–19.

over time; the proportion of prescribed with parenteral antibiotics in Infectious disease ($p < 0.01$), Respiratory medicine ($p < 0.01$), and Other ($p < 0.01$) (**Table 4**).

As shown in **Figure 5**, we found that the proportion of treated with antibiotics, the proportion fo prescribed with muti-antibiotics and the proportion of prescribed with parenteral antibiotics for male patients ($p < 0.01$) and female patients ($p < 0.01$) in emergency departments increased significantly from 2016Q1 to 2019Q4. There was the opposite change for outpatients and inpatients among male patients and female patients. As for the antibiotics use by age group, the overall proportion of treated with antibiotics for patients 5–17 years age ($p < 0.01$), 18–49 years age ($p < 0.01$), 50–64 years age ($p < 0.01$) and ≥65 years age ($p < 0.01$) in outpatient departments decreased over time. The overall proportion of treated with antibiotics for patients <5 years age ($p < 0.01$), 18–49 years age ($p < 0.01$), 50–64 years age ($p < 0.01$) and ≥65 years age ($p < 0.01$) in emergency departments showed an increasing trend over time, which was similar with he overall proportion of treated with antibiotics for all age group of patients in inpatient departments ($p < 0.01$) (**Table 5**). The overall proportion of prescribed with muti-antibiotics for patients 18–49 years age ($p < 0.01$) in emergency departments showed an upward trend over time, which was contrast to <5 years age ($p < 0.01$), 5–17 years age ($p < 0.01$) and ≥65 years age ($p < 0.01$). The overall proportion of prescribed with muti-antibiotics for patients <5 years age ($p < 0.01$), 5–17 years age ($p < 0.01$) and 18–49 years age ($p < 0.01$) in emergency departments increased over time. For outpatients, the overall proportion of prescribed with muti-antibiotics for patients 50–64 years age ($p < 0.01$) in emergency departments showed an downward trend over time, which was contrast to <5 years age ($p < 0.01$) and 5–17 years age ($p < 0.01$) (**Table 6**). The overall proportion of prescribed with parenteral antibiotics for all age groups of patients showed a downward trend over time, which was contrast to patients in emergency departments. No significant change was observed in the proportion of prescribed with parenteral antibiotics among all age groups in inpatient departments (**Table 7**).

Antibiotics Category Patterns

Figure 5 presents the quarterly average consumption of antibiotics classes. The consumption amounts of Tetracyclines ($p < 0.01$), β-lactam antibacterials, penicillins ($p < 0.01$), Macrolides, lincosamide and streptogramins ($p < 0.01$), Quinolone antibacterials ($p < 0.01$), Other antibiotics in J01 ($p < 0.01$), and Antimycotics for systemic use ($p = 0.03$) increased each quarter over the study period. The consumption amounts of restricted antibiotics ($p < 0.01$) and non-restricted antibiotics ($p < 0.01$) increased over time (**Figure 6**).

Figure 7 shows the proportion of antibiotics classes from 2016 to 2019. Other β-lactam antibacterials (1,663.03 ten thousand DDD, 36.08%) far exceeded other antibiotics categories. Stablely ranked first, followed by Macrolides, lincosamide and streptogramins (965.74 ten thousand DDD, 20.95%), Quinolone antibacterials (710.42 ten thousand DDD, 15.41%), and β-lactam antibacterials, penicillins (497.01 ten thousand DDD, 10.78%). In terms of individual medicines, the most frequently used antibiotics were levofloxacin (J01MA12), followed by clarithromycin (J01FA09), cefuroxime (J01DC02), amoxicillin (J01CA04) and azithromycin (J01FA10). The proportion of restricted antibiotics, non-restricted antibiotics and controlled antibiotics was 50.00, 38.41 and 9.04%, respectively (**Table 8**).

DISCUSSION

Based on the prescribing data from the CPDHP during 2016–19, this multi-center study analyzed 3,422,710 outpatient and emergency prescriptions and 26, 118, 436 inpatient records from 89 tertiary hospitals in nine cities. Results showed that the reasonable use of antibiotics in outpatient and emergency departments was superior to inpatient departments. The overall status of antibiotics use for inpatients in China rebounded during 2016–19 is noteworthy, showing a significant upward trend in treating with antibiotics. A close association exists between AMR and the amount of antibiotics used (Bell et al., 2014), which indicates a severe need to promote antibiotic consumption. Furthermore, the antibiotics use varied by different survey areas, clinical departments, patient gender, patient age and antibiotics categories.

Because the sampled 89 hospitals were from nine different administrative provinces, there may be inter-drug supply variation resulting from differences in various drug procurement policy. However, the preliminary test in our study demonstrated that there were no obvious differences in antibiotics consumption and cost in nine cities. As a consequence, differences in drug supply varieties, specifications and prices in different cities may have slight influence on the results in this

TABLE 5 | Changes in proportion of treated with antibiotics by patient characteristics, 2016–19.

Care type	Patient characteristics	16Q1	16Q2	16Q3	16Q4	17Q1	17Q2	17Q3	17Q4	18Q1	18Q2	18Q3	18Q4	19Q1	19Q2	19Q3	19Q4	Z[a]
Outpatient prescriptions																		
	Patient gender																	
	Male	14.82	14.21	13.35	14.50	13.85	13.92	13.79	14.62	14.32	13.65	13.02	13.84	14.85	13.96	13.02	13.81	-12.461**
	Female	14.64	13.81	13.21	14.27	13.71	13.73	13.64	14.52	14.39	13.57	13.02	13.87	14.84	13.84	12.91	14.02	-5.881**
	Patient age																	
	<5 years	39.71	39.14	35.16	41.97	37.07	38.54	35.22	40.91	36.55	39.82	35.59	40.83	37.66	39.24	36.30	39.99	-0.633
	5–17 years	36.85	36.43	27.86	41.43	34.20	37.80	28.13	40.81	34.09	35.46	27.12	38.93	35.25	37.08	28.25	38.70	-3.696**
	18–49 years	18.18	17.19	17.01	17.12	17.33	17.00	17.37	17.14	18.00	16.88	16.64	16.77	18.27	17.02	16.46	16.72	-10.145**
	50–64 years	10.27	9.74	9.57	9.66	9.79	9.52	9.73	9.50	10.37	9.21	9.25	9.15	10.34	9.33	9.03	9.06	-13.578**
	≥65 years	8.53	7.79	7.63	7.85	8.01	7.64	7.81	7.80	8.49	7.46	7.29	7.31	8.39	7.43	7.07	7.26	-16.611**
Emergency prescriptions																		
	Patient gender																	
	Male	38.29	37.17	39.16	37.87	41.20	40.83	42.35	39.59	44.54	41.60	40.58	42.49	47.73	44.46	43.66	42.19	31.365**
	Female	40.34	39.02	40.46	40.02	43.63	42.70	43.65	42.19	45.88	43.78	42.52	43.77	50.31	46.76	45.02	44.73	31.545**
	Patient age																	
	<5 years	31.38	36.05	30.63	34.01	34.01	36.59	33.94	36.49	28.24	33.83	30.40	32.14	38.18	39.54	40.08	32.22	2.625**
	5–17 years	45.35	44.68	41.79	46.33	47.28	44.73	40.89	46.56	43.19	43.21	37.88	43.70	45.98	45.52	44.17	45.55	-0.421
	18–49 years	45.43	44.04	45.51	45.47	47.81	46.95	47.67	46.46	50.63	47.89	47.12	49.54	54.06	50.86	50.05	50.24	27.829**
	50–64 years	35.34	33.63	35.61	34.36	37.93	37.66	39.86	36.77	42.49	38.62	37.65	39.07	46.33	42.26	40.29	39.54	23.703**
	≥65 years	34.83	32.58	35.11	34.00	39.35	38.01	39.66	36.71	42.44	39.60	38.35	39.38	45.66	41.26	39.89	38.60	26.900**
Inpatient records																		
	Patient gender																	
	Male	23.90	23.20	23.73	24.16	24.02	23.46	23.92	23.29	24.64	23.67	24.73	28.14	29.54	28.24	28.18	27.69	83.270**
	Female	23.73	22.82	23.30	23.83	23.78	22.91	23.17	22.77	24.22	23.08	23.77	26.95	28.35	27.03	26.96	26.72	63.077**
	Patient age																	
	<5 years	39.36	38.58	37.61	41.81	40.70	40.06	38.95	40.93	40.48	41.85	40.00	48.44	47.41	48.01	44.91	47.51	31.012**
	5–17 years	35.65	35.66	36.36	38.84	37.69	37.25	36.66	37.43	36.02	36.40	37.58	47.15	46.75	48.41	44.78	46.46	32.678**
	18–49 years	26.08	26.31	26.20	27.16	26.23	26.17	26.59	25.98	26.27	26.46	27.25	30.98	31.62	31.07	31.12	30.36	54.780**
	50–64 years	20.82	20.27	21.02	21.06	20.76	20.50	20.71	20.18	21.21	20.82	21.84	23.98	24.84	24.03	24.27	23.70	14.132**
	≥65 years	22.58	21.02	21.53	21.74	22.58	21.26	21.59	21.01	23.72	21.37	22.09	25.38	28.04	25.85	25.56	25.28	65.867**

[a]*The Cochran-Armitage trend test was applied to calculate* p *values.*

TABLE 6 | Changes in proportion of prescribed with muti-antibiotics by patient characteristics, 2016–19. prescribed with parenteral antibiotics (c).

Care type	Patient characteristics	16Q1	16Q2	16Q3	16Q4	17Q1	17Q2	17Q3	17Q4	18Q1	18Q2	18Q3	18Q4	19Q1	19Q2	19Q3	19Q4	Z[a]
Outpatient prescriptions																		
	Patient gender																	
	Male	11.76	12.07	12.30	11.74	11.84	11.78	12.47	11.12	11.12	11.70	12.46	11.53	11.45	11.83	12.56	11.89	-0.088
	Female	12.00	12.60	12.77	12.18	12.46	12.39	13.10	12.05	11.54	12.24	13.14	12.04	11.40	11.95	13.23	12.21	-1.651
	Patient age																	
	<5 years	5.67	5.74	5.21	5.96	5.70	5.12	4.55	5.03	5.15	5.05	4.35	5.25	5.49	5.27	5.29	4.82	-4.037**
	5–17 years	6.69	6.34	6.14	6.60	6.81	6.51	6.19	6.13	5.89	5.72	5.69	6.29	6.12	6.37	6.28	5.81	-3.273**
	18–49 years	13.90	14.30	14.60	14.49	14.21	14.49	15.09	14.61	13.39	14.44	15.53	14.73	13.63	14.28	15.64	15.41	5.609**
	50–64 years	14.44	15.29	14.80	14.83	14.68	14.58	15.34	14.42	13.27	14.71	15.28	14.26	13.36	14.70	15.56	15.36	-0.097
	≥65 years	11.81	12.52	12.00	11.55	11.73	12.04	12.30	11.49	11.21	11.72	11.76	11.20	11.51	11.64	11.50	11.54	-3.263**
Emergency prescriptions																		
	Patient gender																	
	Male	18.68	19.62	20.28	19.43	20.63	21.03	20.34	19.27	20.03	20.79	20.60	20.43	20.14	20.46	20.38	19.59	2.110*
	Female	17.09	18.52	19.13	18.16	18.89	19.94	18.57	17.97	17.94	19.38	19.84	19.45	18.39	19.26	20.20	18.61	4.162**
	Patient age																	
	<5 years	6.11	7.40	5.26	5.98	3.61	5.39	5.88	3.98	4.03	5.84	6.13	5.76	8.54	12.97	10.58	14.04	6.836**
	5–17 years	9.94	12.02	13.13	9.77	10.62	12.07	11.95	10.22	11.01	12.03	13.19	11.88	12.57	13.86	14.43	11.58	3.285**
	18–49 years	14.66	16.71	17.28	15.97	17.08	17.51	16.89	16.34	15.59	17.52	18.32	17.37	16.00	16.96	18.08	16.60	3.709**
	50–64 years	18.55	20.11	21.38	20.54	21.05	22.25	20.40	20.52	19.43	21.40	20.74	21.36	20.03	20.59	20.84	19.60	0.098
	≥65 years	24.43	25.49	25.05	24.80	25.02	26.84	25.07	23.67	24.67	25.64	24.88	25.49	24.64	25.58	24.84	23.81	-1.056
Inpatient records																		
	Patient gender																	
	Male	21.37	21.80	22.05	21.84	21.93	21.78	21.81	21.41	21.97	21.25	21.46	22.11	22.43	22.24	21.84	21.08	0.435
	Female	21.23	21.81	22.23	22.13	22.00	21.71	21.81	21.07	21.17	21.04	21.30	22.40	22.59	22.21	21.89	20.99	0.143
	Patient age																	
	<5 years	15.40	15.26	15.84	15.78	14.98	14.84	15.07	15.30	14.69	14.84	16.25	17.32	16.37	16.43	17.70	16.70	5.425**
	5–17 years	19.50	20.48	19.64	24.48	21.80	22.01	18.70	23.13	19.91	21.17	20.66	25.66	22.13	23.05	21.33	22.47	4.318**
	18–49 years	22.39	23.04	23.31	23.64	23.36	23.42	22.97	23.09	22.59	22.51	22.58	23.87	23.90	23.58	22.71	22.44	-0.17
	50–64 years	22.15	22.52	22.95	22.64	22.96	22.31	22.63	21.69	22.10	21.48	21.82	22.46	22.77	22.47	22.15	20.94	-4.258**
	≥65 years	21.01	21.43	21.68	21.04	21.39	21.16	21.65	20.42	21.79	20.92	20.95	21.41	22.35	21.89	21.63	20.71	1.739

[a]*The Cochran-Armitage trend test was applied to calculate* p *values.*

TABLE 7 | Changes in proportion of prescribed with parenteral antibiotics by patient characteristics, 2016–19.

Care type	Patient characteristics	16Q1	16Q2	16Q3	16Q4	17Q1	17Q2	17Q3	17Q4	18Q1	18Q2	18Q3	18Q4	19Q1	19Q2	19Q3	19Q4	Z[a]
Outpatient prescriptions																		
	Patient gender																	
	Male	20.97	19.01	17.68	19.79	18.53	18.43	17.93	18.67	16.80	16.15	14.97	16.52	16.73	16.04	15.29	16.99	-31.456**
	Female	18.38	16.58	15.27	17.06	16.50	15.91	15.24	16.00	14.60	13.34	12.75	14.20	14.63	13.46	12.69	14.25	-35.986**
	Patient age																	
	<5 years	38.76	34.83	31.26	37.31	36.26	33.78	32.12	34.94	35.15	30.91	27.80	33.82	34.83	28.66	28.09	31.95	-19.134**
	5–17 years	34.80	31.50	27.93	35.49	33.01	31.22	28.07	33.12	29.77	27.51	25.84	31.55	29.52	28.94	26.58	31.56	-12.961**
	18–49 years	14.64	13.67	14.08	12.75	13.10	12.98	13.48	11.51	11.26	10.64	11.01	9.62	10.62	10.43	10.78	9.80	-38.410**
	50–64 years	14.40	12.79	12.82	11.45	12.50	11.83	12.89	10.42	11.21	9.82	10.15	8.82	10.51	9.47	9.62	9.04	-27.070**
	≥65 years	14.42	12.63	12.16	11.34	12.58	11.99	12.92	10.08	11.39	10.08	10.21	9.14	10.99	10.26	10.18	9.30	-20.658**
Emergency prescriptions																		
	Patient gender																	
	Male	69.95	68.54	70.49	70.03	71.80	70.12	70.68	70.47	71.62	70.95	71.52	70.98	71.67	71.49	72.66	73.24	9.238**
	Female	68.87	67.47	69.48	68.94	71.04	69.54	69.97	69.98	70.13	69.38	71.31	70.17	70.35	70.11	71.68	71.44	7.607**
	Patient age																	
	<5 years	21.09	18.70	16.34	14.35	11.81	14.67	14.35	11.23	11.28	10.58	13.66	13.98	22.62	24.06	21.17	23.05	3.396**
	5–17 years	41.14	39.20	42.00	38.02	38.92	39.17	41.74	37.66	39.88	37.20	40.82	38.29	40.58	41.65	43.16	43.48	2.299*
	18–49 years	59.81	61.21	63.27	61.43	62.37	61.96	62.45	62.50	59.92	61.98	64.02	62.64	59.75	61.81	64.13	63.34	3.780**
	50–64 years	74.88	75.03	75.41	76.99	77.77	76.61	76.19	78.26	76.56	77.58	77.22	77.81	77.53	78.49	77.87	78.01	5.910**
	≥65 years	88.31	87.17	86.57	87.60	88.57	87.97	86.87	87.99	87.67	88.03	87.39	88.25	88.53	88.39	88.14	88.36	2.854**
Inpatient records																		
	Patient gender																	
	Male	92.74	92.82	92.84	93.02	93.01	92.87	92.75	93.13	92.93	92.95	92.80	92.88	92.96	92.88	92.66	92.67	-1.19
	Female	92.25	92.08	92.09	92.38	92.35	92.06	92.22	92.28	92.27	92.44	92.16	92.21	92.34	92.27	92.04	91.95	-0.88
	Patient age																	
	<5 years	95.00	94.58	94.89	95.36	95.67	95.12	95.29	95.39	95.30	95.42	95.90	95.71	95.43	94.98	94.87	94.73	0.41
	5–17 years	92.93	92.93	92.84	93.24	93.09	93.33	92.81	93.45	92.94	93.68	92.77	93.43	92.77	93.18	92.45	92.25	-1.343
	18–49 years	91.94	92.04	92.08	92.33	92.15	91.95	92.05	92.16	92.26	92.37	92.14	91.98	91.97	92.01	91.88	91.61	-1.924
	50–64 years	92.34	92.41	92.45	92.40	92.37	92.42	92.40	92.66	92.53	92.62	92.32	92.27	92.52	92.33	92.32	92.19	-0.88
	≥65 years	92.64	92.52	92.53	92.74	92.85	92.49	92.52	92.71	92.54	92.54	92.45	92.60	92.84	92.80	92.47	92.58	0.105

[a]*The Cochran-Armitage trend test was applied to calculate* p *values.*

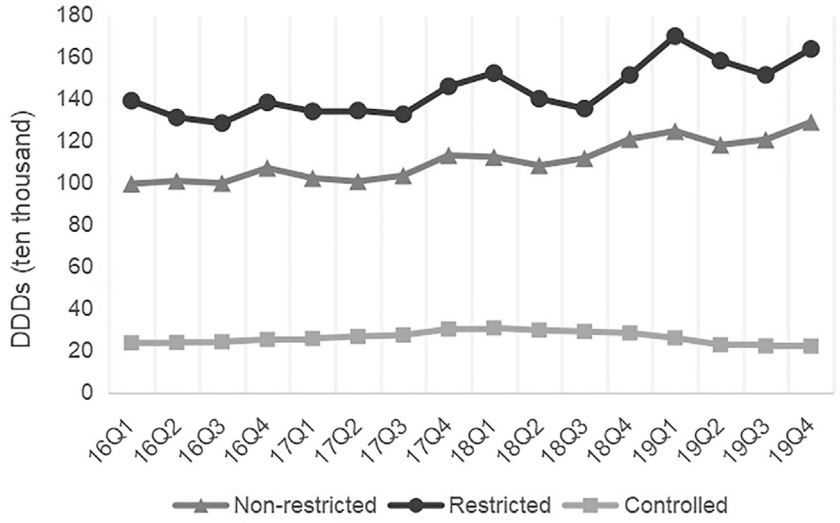

FIGURE 6 | Changes in antibiotics use by different usage levels, 2016–19.

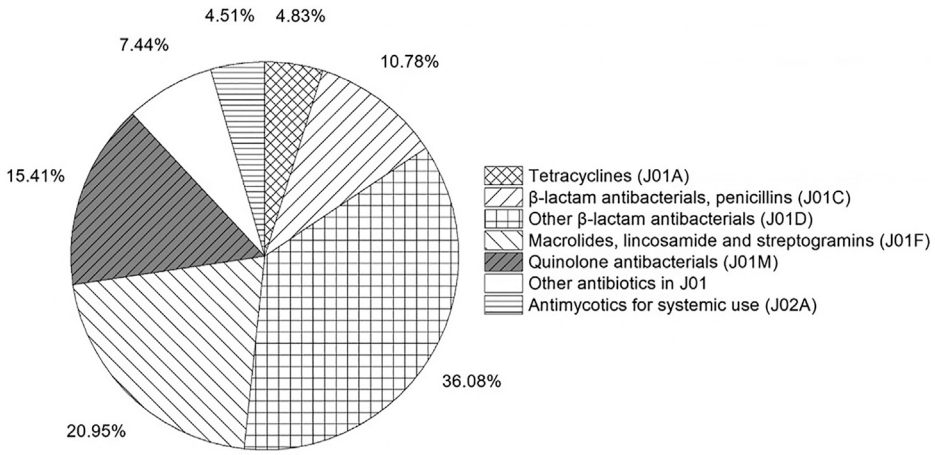

FIGURE 7 | Percentage of antibiotic classes used by ATC classification, 2016–19.

TABLE 8 | Percentage of antibiotic classes used by different usage levels, 2016–19.

Usage level	DDDs (ten thousand)	Usage ratio (%)
Restricted	2,304.87	50.00
Non-restricted	1770.74	38.41
Controlled	416.86	9.04

study. One possible reason is that most of the surveyed antibiotics were included in the National Reimbursement Drug List (NRDL) or Chinese National Essential Medicine List (NEML). The prices and dosage forms or specifications of surveyed antibiotics are set in a relatively uniform level. WHO recommended that the proportion of treated with antibiotics should not be over 30% (WHO, 2006). Except for emergency departments, the percentage use of antibiotics in the current study (16.05% for outpatients, 25.19% for inpatients) was lower than this recommended level and lower than in many other countries. For example, the average proportion of prescribed antibiotics in some developing countries was 40–50%. The proportion of prescribed antibiotics for outpatients was 60–70% in some low-and-middle-income countries such as Pakistan, Indonesia and Mozambique (WHO, 2004; WHO, 2009).

The percentage of antibiotics use under effective control may be related to a series of antimicrobial administration policies adopted by the Chinese government. China has developed Antimicrobial Stewardship Programs (ASPs) to promote antibiotics use. It's important to note that a part of the provinces has implemented restrictions of antimicrobial use

for outpatient in recent years. Previous studies also revealed that antimicrobial stewardship could reduce the consumption of antibiotics (Hou et al., 2014; Uda et al., 2019; Walker et al., 2019).

Several issues are still noteworthy even though the use of antibiotics in China has been controlled to some extent. First, the average proportion of treated with antibiotics in Stomatology was high (over 50%). More than one-third of dental patients received antibiotic combination treatment, and the percentage of prescribing parenteral antibiotics increased with time. Second, the percentage of antibiotics used for adults (>18 years) increased during the study period is noteworthy. The percentage of prescribing parenteral antibiotics for patients 5–17 years age and 18–49 years of age exceeded 50%. This finding may be related to young children's low immunity and the pressure from patients who insist on antibiotics therapy. As the primary breadwinners of a family, young people want to be cured as fast as people in most cases when they fall ill (Lin et al., 2016). Third, co-prescription of at least two antibiotics for patients ≥65 years of age and in ICU may result from the severity and complexity of those patients. Finally, this study revealed that most categories of antibiotics use increased with time.

The above finding suggests that nationwide education and prescription control of antibiotics to treat common but not severe infectious should be reinforced (Kim et al., 2019). In fact, different antibiotic classification catalogs may have different influence on the management of antibiotics use in health-care institutions. Hence, a unitary classification of antibiotic is urgently needed to make for easier administration. Besides, the medical decision-making surrounding antibiotic use is very complex and the size and types of medical care are varied, which can lead to the flexibility of implementation of ASPs in Chinese urban hospitals. Thus, the institutionalized ASPs should be implemented at the hospital level to improve the utilization of antibiotics. Individual hospitals should implement ASPs, which reflects their environment. To design effective interventions to improve antibiotics use, it is significant to understand patients' characteristics and obstacles for antibiotics prescribing from physicians or hospitals perspectives (Wang et al., 2016).

There are several limitations to our study. First, due to data collection difficulties, we cannot link antibiotics records with patients medical records or laboratory results. Second, primary health-care institutions and secondary hospitals were not considered in this study, which might not represent the situation of antibiotics consumption comprehensively. Third, our study mainly focused on the patterns of antibiotic use and expenditure. Other factors that may influence our indicators were not considered in the current study, which needed further exploration. Despite these limitations, this study still revealed crucial new information on antibiotics use and how these changes in antibiotic consumption changed in tertiary hospitals from nine cities, providing evidence for policy-makers to improve antibiotic stewardship interventions.

CONCLUSION

In conclusion, this study found that the proportion of treated with antibiotics meet the WHO standards, suggesting the ASP was effective in China. This study also demonstrated that the reasonable use of antibiotics in outpatient and emergency departments was superior to inpatient departments. The antibiotics use varied by different survey areas, clinical departments, patient gender, patient age and antibiotics categories. Further research is required to obtain more detailed information on the inappropriate use of antibiotics in China. To better facilitate proper antibiotics use, the next stage of policy should focus on improving the appropriateness of antibiotics uses at the individual level.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

XH and XL initiated the study concept; YZ and RD conducted data analysis; YZ wrote the first draft of the manuscript; YZ, YQ, RD, and XL collected data; XH led the data collection; XL, and XH contributed to the data analysis and interpretation of the data and revised the draft of the manuscript; all authors read and approved the final manuscript.

FUNDING

This study was supported by the National Natural Science Foundation of China (Grant No: 71673147, 72074123) and the CHINA MEDICAL BOARD (Grant No: 17-277).

ACKNOWLEDGMENTS

The authors are grateful for the support of the Cooperation Project Database of Hospital Prescriptions for providing the information used in this study.

REFERENCES

Bao, L., Peng, R., Wang, Y., Ma, R., Ren, X., Meng, W., et al. (2015). Significant Reduction of Antibiotic Consumption and Patients' Costs after an Action Plan in China, 2010-2014. *PLoS One* 10 (3), e0118868. doi:10.1371/journal.pone.0118868

Bell, B. G., Schellevis, F., Stobberingh, E., Goossens, H., and Pringle, M. (2014). A Systematic Review and Meta-Analysis of the Effects of Antibiotic Consumption on Antibiotic Resistance. *BMC Infect. Dis.* 14, 13. doi:10.1186/1471-2334-14-13

Bieler, G. S., and Williams, R. L. (1993). Ratio Estimates, the delta Method, and Quantal Response Tests for Increased Carcinogenicity. *Biometrics* 49 (3), 793–801. doi:10.2307/2532200

Chinese Ministry of Health (2011). Action Plan for Antimicrobial Stewardship in Clinical Application. Available at:http://www.nhc.gov.cn/xxgk/pages/viewdocument.jsp?dispatchDate=&staticUrl=/zwgkzt/wsbysj/201104/51376.shtml&wenhao=%E5%8D%AB%E5%8A%9E%E5%8C%BB%E6%94%BF%E5%8F%91%E3%80%942011%E3%80%9556%E5%8F%B7&utitle=%E5%8D%AB%E7%94%9F%E9%83%A8%E5%8A%9E%E5%85%AC%E5%8E%85%E5%85%B3%E4%BA%8E%E5%81%9A%E5%A5%BD%E5%85%A8%E5%9B%BD%E6%8A%97%E8%8F%8C%E8%8D%AF%E7%89%A9%E4%B8%B4%E5%BA%8A%E5%BA%94%E7%94%A8%E4%B8%93%E9%A1%B9%E6%95%B4%E6%B2%BB%E6%B4%BB%E5%8A%A8%E7%9A%84%E9%80%9A%E7%9F%A5&topictype=&topic=&publishedOrg=%E5%8C%BB%E6%94%BF%E5%8F%B8&indexNum=000013610/2011-04275&manuscriptId=51376 (accessed on June 4, 2020).

Dong, L., Yan, H., and Wang, D. (2008). Antibiotic Prescribing Patterns in Village Health Clinics across 10 Provinces of Western China. *J. Antimicrob. Chemother.* 62, 410–415. doi:10.1093/jac/dkn153

European Centre for Disease Prevention and Control (2018). Factsheet for Experts—Antimicrobial Resistance. Available at: https://www.ecdc.europa.eu/en/antimicrobial-resistance/facts/factsheets/experts (accessed on June 3, 2020).

Gandra, S., Barter, D. M., and Laxminarayan, R. (2014). Economic burden of Antibiotic Resistance: How Much Do We Really Know. *Clin. Microbiol. Infect.* 20, 973–980. doi:10.1111/1469-0691.12798

Hou, D., Wang, Q., Jiang, C., Tian, C., Li, H., and Ji, B. (2014). Evaluation of the Short-Term Effects of Antimicrobial Stewardship in the Intensive Care Unit at a Tertiary Hospital in China. *PloS one* 9 (7), e101447. doi:10.1371/journal.pone.0101447

Jim, O'N. (2015). Tackling Drug-Resistant Infections Globally: Final Report and Recommendations. Available at: https://amr-review.org/sites/default/files/160525_Final%20paper_with%20cover.pdf (accessed on June 3, 2020).

Kim, B., Myung, R., Lee, M. J., Kim, J., and Pai, H. (2019). Trend of Antibiotics Usage for Acute Pyelonephritis in Korea Based on National Health Insurance Data 2010-2014. *BMC Infect. Dis.* 19, 554. doi:10.1186/s12879-019-4191-0

Li, H., Yan, S., Li, D., Gong, Y., Lu, Z., and Yin, X. (2019). Trends and Patterns of Outpatient and Inpatient Antibiotic Use in China's Hospitals: Data from the Center for Antibacterial Surveillance, 2012-16. *J. Antimicrob. Chemother.* 74, 1731–1740. doi:10.1093/jac/dkz062

Lin, H., Dyar, O. J., Rosales-Klintz, S., Zhang, J., Tomson, G., Hao, M., et al. (2016). Trends and Patterns of Antibiotic Consumption in Shanghai Municipality, China: a 6 Year Surveillance with Sales Records, 2009-14. *J. Antimicrob. Chemother.* 71 (6), 1723–1729. doi:10.1093/jac/dkw013

Liu, P. (2018). Natural Selection-- History and Progress of Research and Development of Antimicrobials. *China Food Drug Adm. Mag.* 2018 (01), 76–78. (in Chinese).

National Health and Family Plan Commission of China (2016). National Action Plan to Combat Antimicrobial Resistance. Available at: http://www.gov.cn/xinwen/2016-08/25/content_5102348.htm (accessed on June 4, 2020).

The WHO Collaborating Centre for Drug Statistics Methodology (2019). ATC/DDD Index 2020. Available at: https://www.whocc.no/atc_ddd_index/ (accessed on September 2, 2020).

Uda, K., Okubo, Y., Kinoshita, N., Morisaki, N., Kasai, M., Horikoshi, Y., et al. (2019). Nationwide Survey of Indications for Oral Antimicrobial Prescription for Pediatric Patients from 2013 to 2016 in Japan. *J. Infect. Chemother.* 25 (10), 758–763. doi:10.1016/j.jiac.2019.03.004

Van Boeckel, T. P., Gandra, S., Ashok, A., Caudron, Q., Grenfell, B. T., Levin, S. A., et al. (2014). Global Antibiotic Consumption 2000 to 2010: an Analysis of National Pharmaceutical Sales Data. *Lancet Infect. Dis.* 14 (8), 742–750. doi:10.1016/S1473-3099(14)70780-7

Walker, A. J., Curtis, H. J., and Goldacre, B. (2019). Impact of Chief Medical Officer Activity on Prescribing of Antibiotics in England: an Interrupted Time Series Analysis. *J. Antimicrob. Chemother.* 74 (4), 1133–1136. doi:10.1093/jac/dky528

Wang, D. Z., Wang, C., Shen, C. F., Zhang, Y., Zhang, H., Song, G. D., et al. (2017). Comparison of Application of Cochran-Armitage Trend Test and Linear Regression Analysis for Rate Trend Analysis in Epidemiology Study. *Zhonghua Liu Xing Bing Xue Za Zhi* 38 (05), 684–687. (in Chinese). doi:10.3760/cma.j.issn.0254-6450.2017.05.026

Wang, J., Wang, P., Wang, X., Zheng, Y., and Xiao, Y. (2014). Use and Prescription of Antibiotics in Primary Health Care Settings in China. *JAMA Intern. Med.* 174 (12), 1914–1920. doi:10.1001/jamainternmed.2014.5214

Wang, Y. Y., Du, P., Huang, F., Li, D. J., Gu, J., Shen, F. M., et al. (2016). Antimicrobial Prescribing Patterns in a Large Tertiary Hospital in Shanghai, China. *Int. J. Antimicrob. Agents* 48, 666–673. doi:10.1016/j.ijantimicag.2016.09.008

WHO (1993). How to Investigate Drug Use in Health Facilities: Selected Drug Use Indicators. EDM Research Series No. 007. Available from: http://apps.who.int/medicinedocs/en/d/Js2289e/ (Accessed October 2, 2018).

WHO (2009). Medicines Use in Primary Care in Developing and Transitional Countries. Available at: http://www.who.int/entity/medicines/publications/primary_care_8April09.pdf (accessed on January 4, 2021).

WHO (2004). The World Medicines Situation 2004. Available at: http://apps.who.int/medicinedocs/pdf/s6160e/s6160e.pdf (accessed on January 4, 2021).

WHO (2006). Using Indicators to Measure Country Pharmaceutical Situations. Available at: http://apps.who.int/medicinedocs/index/assoc/s14101e/s14101e.pdf (Accessed January 4, 2021).

World Health Organization (2018). Antimicrobial Resistance. Available at: https://www.who.int/news-room/facts-in-pictures/detail/antimicrobial-resistance (accessed on June 3, 2020).

World Health Organization (2015). Global Action Plan on Antimicrobial Resistance. Available at: https://apps.who.int/iris/bitstream/handle/10665/193736/9789241509763_eng.pdf?sequence=1 (accessed on June 3, 2020).

World Health Organization (2016). WHO Methodology for A Global Programme on Surveillance of Antimicrobial Consumption. Available at: https://www.who.int/medicines/areas/rational_use/WHO_AMCsurveillance_1.0.pdf?%20ua=1 (accessed on June 4, 2020).

Xiao, Y. (2018). Antimicrobial Stewardship in China: Systems, Actions and Future Strategies. *Clin. Infect. Dis.* 67, S135–S141. doi:10.1093/cid/ciy641

Xiao, Y., and Li, L. (2016). China's National Plan to Combat Antimicrobial Resistance. *Lancet Infect. Dis.* 16, 1216–1218. doi:10.1016/S1473-3099(16)30388-7

Xiao, Y., and Li, L. (2013). Legislation of Clinical Antibiotic Use in China. *Lancet Infect. Dis.* 13, 189–191. doi:10.1016/S1473-3099(13)70011-2

Yin, J., Wu, C., Wei, X., and Sun, Q. (2018). Antibiotic Expenditure by Public Healthcare Institutions in Shandong Province in China, 2012-2016. *Front. Pharmacol.* 9, 1396. doi:10.3389/fphar.2018.01396

Yin, X., Song, F., Gong, Y., Tu, X., Wang, Y., Cao, S., et al. (2013). A Systematic Review of Antibiotic Utilization in China. *J. Antimicrob. Chemother.* 68 (11), 2445–2452. doi:10.1093/jac/dkt223

Zou, X. X., Fang, Z., Min, R., Bai, X., Zhang, Y., Xu, D., et al. (2014). Is Nationwide Special Campaign on Antibiotic Stewardship Program Effective on Ameliorating Irrational Antibiotic Use in China? Study on the Antibiotic Use of Specialized Hospitals in China in 2011-2012. *J. Huazhong Univ. Sci. Technolog Med. Sci.* 34, 456–463. doi:10.1007/s11596-014-1300-6

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Zhu, Qiao, Dai, Hu and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Public Health

BRIEF RESEARCH REPORT
published: 04 November 2021
doi: 10.3389/fpubh.2021.761017

Short Practical Regimen of Acupuncture for Melasma: A Prospective Cohort Study in a Tertiary Hospital in Thailand

Thanan Supasiri[1], *Nuntida Salakshna*[2,3]* *and Krit Pongpirul*[1,4]*

[1] Department of Preventive and Social Medicine, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand, [2] Samitivej Esthetics Institute, Samitivej Sukhumvit Hospital, Bangkok, Thailand, [3] Science Division, Mahidol University International College, Nakhon Pathom, Thailand, [4] Department of International Health and Department of Health, Behavior, and Society, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Thamil Selvee Ramasamy,
University of Malaya, Malaysia
Im Quah-Smith,
University of New South
Wales, Australia

***Correspondence:**
Nuntida Salakshna
nppn99@hotmail.com
Krit Pongpirul
doctorkrit@gmail.com

Specialty section:
This article was submitted to
Family Medicine and Primary Care,
a section of the journal
Frontiers in Public Health

Received: 19 August 2021
Accepted: 04 October 2021
Published: 04 November 2021

Citation:
Supasiri T, Salakshna N and
Pongpirul K (2021) Short Practical
Regimen of Acupuncture for Melasma:
A Prospective Cohort Study in a
Tertiary Hospital in Thailand.
Front. Public Health 9:761017.
doi: 10.3389/fpubh.2021.761017

Background: Acupuncture shows benefits for patients with melasma, although no optimal number of sessions have been determined.

Methods: The prospective observational study was conducted in melasma patients who were treated with acupuncture procedures two times a week and were evaluated after the 5th and the 10th sessions of acupuncture, with a 1-week follow-up after the last session. Participants Groups A and B received five and 10 acupuncture sessions, respectively. Melasma was assessed by using the melanin index (MI), melasma area and severity index (MASI), patient-reported improvement scores, and acupuncture-related adverse events.

Results: Out of 113 participants, 67 received five sessions of acupuncture treatment while 39 received 10 sessions. At 1 week after five sessions of acupuncture in Group A, the mean MI decreased by 28.7 (95% CI −38.5 to −18.8, $p < 0.001$), whereas the median MASI decreased by 3.4 (95% CI −6.9 to −1.2, $p < 0.001$) points. At 1 week after ten sessions of acupuncture in Group B, the mean MI decreased by 31.3 (95% CI −45 to −17.6, $p < 0.001$), whereas the median MASI decreased by 5.4 (95%CI −9.9 to −3, $p < 0.001$) points. The first five sessions of acupuncture had a higher incremental effect than the last five sessions, although there was no statistically significant difference. Twenty-nine participants reported minor side effects. Group B had a risk ratio (RR) of having adverse events 1.8 times (95% CI 1.0–3.4, $p = 0.05$) compared with Group A.

Conclusion: Short acupuncture regimens of 5–10 sessions in melasma seem to be effective and practical with minor side effects.

Keywords: acupuncture–therapy, melasma, dose–response, melasma area and severity index, Thailand

SUMMARY BOX

What is already known about this subject?

- Acupuncture has effectively improved melasma conditions at low cost with minimal side effects.
- Traditionally, one acupuncture course comprises 10 sessions for a chronic condition but dose-response evidence of acupuncture for melasma has been lacking.
- The number of acupuncture sessions performed in various randomized controlled trials varied from 21 to 60 over 4–16 weeks.

What are the new findings?

- Five sessions of acupuncture improved the melanin index (MI) by 28.7 and melasma area and severity index (MASI) by 3.4 points.
- About 10 sessions of acupuncture improved MI by 31.3 and MASI by 5.4 points, of which the improvement was mainly from the first five sessions.
- Short acupuncture regimens of 5–10 sessions in melasma seem to be effective and practical with minor side effects.

How might it impact clinical practice in the foreseeable future?

- The dose-response study has not been common but could have clinical and financial implications.

INTRODUCTION

Melasma is acquired hyperpigmentation characterized by ill-defined light to dark brownish macules and patches, typically affecting photo-exposed areas of the face (1–3). The most commonly implicated etiological factors are chronic ultraviolet exposure, female hormone stimulation, and genetic predisposition (4–6). Its appearance negatively impacts the quality of life, emotional wellbeing, and may even lead to psychiatric consequences (7, 8).

In the last decade, there was a paradigm shift in the pathophysiology of melasma which is now considered to be a more complex interplay between epidermal and dermal components including the proliferation of dermal vessels and oxidative damage by reactive oxygen species (ROS) (4, 5). Apart from conventional treatments such as hydroquinone, triple-combination cream (hydroquinone, tretinoin, and fluocinolone acetonide), and other bleaching agents, many new emerging therapies which target vascular components such as tranexamic acid or vascular targeting laser and light treatment (such as pulsed dye laser and intense pulsed light) have shown benefits in lightening the melasma area (9–12).

Acupuncture has shown benefits for patients with melasma. The effective rate of acupuncture for melasma has been reported to be around 87–96% (13–15). In Chinese medicine theory, acupuncture can promote blood and qi circulation, reducing blood stasis, rebalance and regulate various internal systems in the body to promote the metabolism of epidermal cells, as well as regulate several hormones and antioxidants thus improving the melasma (13, 16, 17). Another randomized control trial (RCT) found that after the acupuncture treatment for melasma, blood estrogen level decreased while progesterone increased, which corresponds to the association of increasing estrogen and decreasing progesterone level in melasma patients (14). Moreover, acupuncture for melasma has been shown to be effective at low cost with minimal side effects which is beneficial for this chronic recurrent skin disease. However, it requires multiple frequent treatments, which can be inconvenient for the patients. Traditionally, one course of acupuncture treatment for a chronic disease consists of 10 sessions (18), but the acupoints used and the optimal number of treatment sessions for acupuncture for melasma have been inconclusive. Therefore, this study aims to investigate and compare a short-fixed regimen of acupuncture between 5 and 10 sessions by body acupuncture with encircling and intralesional needling methods.

METHODS

Study Design and Population

This prospective observational cohort was conducted at the Acupuncture Clinic, King Chulalongkorn Memorial Hospital, Bangkok, Thailand, from September–December 2018. We recruited all the melasma patients who came to the clinic for acupuncture treatment and were willing to participate in this observational study. We excluded participants who were pregnant, had severe underlying disease or psychosis, had a bleeding tendency, or had skin infections in the acupoints area. All participants were directed to maintain their usual behavior, such as using the same frequency and amount of sunscreen protection or continuing to not use sunscreen regularly.

Acupuncture Procedure

All the patients were treated with fixed acupoints: encircling the melasma area (**Figures 1**, **2**), Fengchi (GB20), Hegu (LI4), Xuehai (SP10), Jusanli (ST36), Sanyinjiao (SP6), and Taichong (LR3). All acupoints were performed bilaterally except those encircling the melasma area. The acupoints were carefully selected from the literature review of standard teaching materials in China (18) and those frequently used in previous studies (13, 19). We used sterile disposable stainless steel needles sized.25 mm × 40 mm (EACU^{tm}, Maanshan Bond Medical Instruments Co., Ltd., China). Acupuncture was done two times a week by well-trained acupuncturists, licensed with acupuncture Master degrees certified from China. The "De Qi" sensation in every acupoint was mandatory, and the needles were retained 30 min each time. Only manual needle-stimulation was used. The participants were evaluated at every fifth session of acupuncture. Since this was an observational study, we used the usual protocol used in our acupuncture clinic. If the first five sessions of acupuncture did not yield satisfactory results and the patients were willing to continue the treatment, we continued the acupuncture treatment for five more sessions. Therefore, participants were divided into two groups: Group A received five acupuncture sessions and Group B received ten acupuncture sessions.

Outcome Measurements

The participants were evaluated at the first visit, every five sessions of acupuncture, and 1 week after their last session.

The primary outcome was the change in the melanin index (MI) measured by Mexameter® MX 18 (Courage+Khazaka electronic GmbH, Köln, Germany) on the melasma affected part of the face. The MI was calculated as the average of a triplicate measurement in the area where the melasma was darkest.

The secondary outcomes consisted of the change in the melasma area and severity index (MASI), patient-reported improvement scores, and related adverse events from acupuncture. The MASI was developed by Kimbrough-Green et al. (20) and assesses three factors—area (A) of involvement, darkness (D), and homogeneity (H)—at four locations: the forehead (f), right malar region (rm), left malar region (lm) and chin (c). The MASI score can be calculated using the following formula:

$$\text{MASI score} = 0.3A_f(D_f + H_f) + 0.3A_{rm}(D_{rm} + H_{rm}) + 0.3A_{lm}(D_{lm} + H_{lm}) + 0.1A_c(D_c + H_c)$$

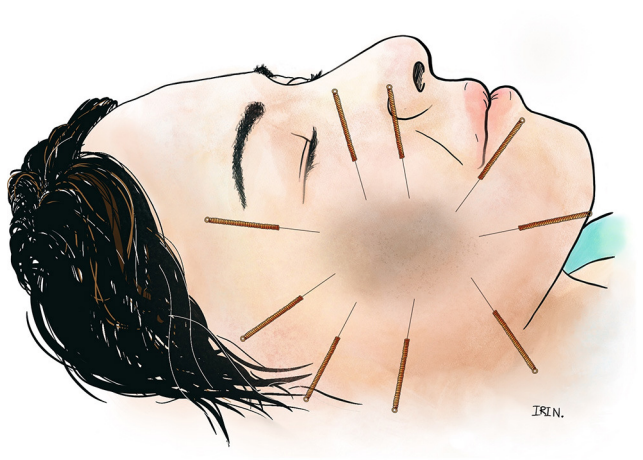

FIGURE 1 | Encircle acupuncture of a facial melasma lesion.

The area of involvement is rated from 0 to 6 while the darkness and homogeneity of pigmentation are rated from 0 to 4; thus, the MASI score ranges from 0 (none) to 48 (severe melasma).

The subjective improvement scores were measured by the 6-point Likert Scale, from−1 to +4 (-1, worsening; 0, no improvement; +1, slight improvement; +2, moderate improvement; +3, marked improvement; +4, almost cured or completely cured). The adverse events due to acupuncture were recorded by interviewing the participants and physical examinations by physicians each visit.

All participants were also photographed at every evaluation visit. The MI, MASI, and satisfactory scores were evaluated by the same evaluator who has been trained by a dermatologist.

Statistical Analysis

The baseline characteristics and data at each time point were analyzed by descriptive statistics. The data after treatment compared with the baseline in each group was analyzed by using a Student's paired *t*-test or Wilcoxon signed-rank tests as appropriate. Linear regression was used to compare the results between Group A and Group B after adjusting for unequal baselines. Linear regression was also used to see if any baseline factors are affecting the outcome. For Group B, a mixed model with the subject as a random effect was used to compare the result of the first five sessions (1st−5th) and the last five sessions (6th−10th). Pairwise comparison between the first and last five sessions was made using Scheffe's Test. If the data did not meet the assumptions to use the mixed model or linear regression, the nonparametric test was used. The data were analyzed using the Stata/MP 15.1 (StataCorp LLC, College Station, TX, USA) software package.

Trial Registration

The study was registered on August 19, 2018, before recruiting the first case, with the Thai Clinical Trials Registry, Identification Number TCTR20180824002.

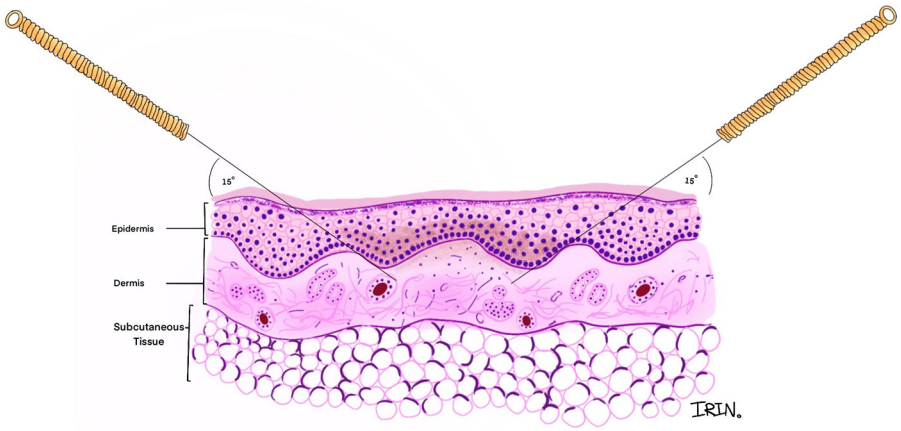

FIGURE 2 | Cross-sectional view of angle and direction of the encircle acupuncture.

Ethics Approval

The study was approved by the Institutional Review Board (IRB), Faculty of Medicine, Chulalongkorn University (IRB No.544/2559). Each participant provided written informed consent.

RESULTS

One hundred and thirteen melasma patients came for acupuncture treatment during September–December 2018. The baseline characteristics are shown in **Table 1**. The acupuncturists finished the course of treatment after five sessions for 67 participants due to the significant improvement achieved. The other 39 participants continued until 10 sessions. Seven participants did not complete at least five sessions of acupuncture and thus, were not evaluated.

Melanin Index (MI)

The average baseline MI for all the participants in the darkest area was 298.9 (SD 79.7). **Table 2** shows the baseline MI and measurements at each time-point for Groups A and B.

In Group A, after the 5th session of acupuncture, the MI decreased by 25 (95% CI −34 to −16.1, $p < 0.001$): from 287.4 (SD 64.8) to 262.3 (SD 62.5). At 1 week after completing five sessions, the mean MI decreased by 28.7 (95% CI −38.5 to −18.8, $p < 0.001$) to 258.7 (SD 61.4) from the baseline. In Group B, after the fifth session of acupuncture, the MI decreased by 22.5 (95% CI −35.5 to −9.5, $p = 0.001$): from 315.9 (SD 98.5) to 293.4 (SD 90.6). After the 10th session of acupuncture, the MI decreased by 30.3 (95% CI −42.4 to −18.2, $p < 0.001$) to 285.6 (SD 94.7) from the baseline. At 1 week after completing 10 sessions of acupuncture, the mean MI decreased by 31.3 (95% CI −45 to −17.6, $p < 0.001$) to 284.6 (SD 89.2) from the baseline.

Comparing the efficacy of the first five sessions (1st to 5th) to the last five sessions (6th to 10th) in Group B, the MI decreased by 22.5 (95%CI −35.5 to −9.5) and 7.8 (95% CI −20.5 to 4.8) respectively. However, there was no statistically significant difference between these two periods (p-value from Scheffe's Test = 0.44).

From the linear regression analysis of the post-fifth-session treatment outcomes, apart from the different baseline MI between the groups, other factors including the duration of melasma, prior acupuncture experience, and family history of melasma did not affect the treatment outcome (see **Appendix 1**).

Between the two groups at 1-week follow-up post-treatment, after adjusting for the unequal baseline, there was no significant difference between the two groups ($p = 0.67$) (refer to **Appendix 2**).

TABLE 1 | Characteristics of the participants.

	Group A: Five sessions of acupuncture treatment ($n = 67$)	Group B: Ten sessions of acupuncture treatment ($n = 39$)
Female [n (%)]	64 (95.5)	37 (94.9)
Age [mean (SD)] (years)	44.3 (9.7)	52.5 (9.7)
Duration of melasma [mean (SD)] (years)	7.4 (6.8)	8.3 (6.2)
Prior acupuncture experience [n (%)]	18 (26.9)	13 (33.3)
Family history of melasma [n (%)]	48 (71.6)	22 (56.4)

Melasma Area and Severity Index (MASI)

The baseline median MASI scores for the overall participants were 6 (2.6, 11.4). Group B had a higher baseline MASI score than Group A, as shown in **Table 3** with the details of the MASI at each time point.

In Group A, after completing five sessions of acupuncture, the median change of the MASI score was −2.9 (-5.3, −0.9) scores. Moreover, at 1 week post-fifth session, the median change of the MASI score compared with the baseline was −3.4 (−6.9, −1.2) scores. In Group B, after completing five sessions of acupuncture, the median change of the MASI score was −3.3 (−6, −0.7) scores. After completing 10 sessions of acupuncture, the median change of the MASI score compared with the baseline was −5.1 (−8.9, −2.5) scores. The median change of the MASI score between the 6th−10th sessions was −1.2 (−3, −0.6) scores. At 1 week after completing 10 sessions of acupuncture, the median change of the MASI score was −5.4 (−9.9, −3) scores.

Due to the non-normal distribution of the MASI scores, the Wilcoxon rank-sum test was used to compare the change of the MASI score between the two groups ($p = 0.08$).

The frequency distribution of patient-reported improvement scores at each time point is shown in **Table 4**. After the full treatment in both groups, 95.5% of Group A and 97.4% of Group B felt markedly improved or cured.

Related Adverse Events From Acupuncture

There were no major adverse events. Of 106 participants, 29 (27.4%) participants reported minor side effects, wherein 14 (20.9%) participants were from group A and 15 (38.5%) participants were from group B. Five patients reported two symptoms: two had hematoma and pain, two had pain and dizziness, and the last had headache and sweating. The other 24 participants had one discomfort symptom. Group B had a risk ratio (RR) of having adverse events 1.8 times (95% CI 1–3.4, p =0.05) compared with group A. All the participants were willing to continue the treatment despite the adverse events.

DISCUSSION

Acupuncture has shown clinical benefits for melasma through MI, MASI, and patient-reported improvement. Both the MI and MASI scores of the groups showed significant improvement after five sessions of acupuncture treatment. More than 95% of all the participants subjectively reported marked improvement, which

TABLE 2 | Mean melanin index (MI) at each evaluation visit.

Melanin index	Group A ($n = 67$) mean (SD)	*p*-value	Group B ($n = 39$) mean (SD)	*p*-value
Baseline	287.4 (64.8)	Reference	315.9 (98.5)	Reference
After the 5th session	262.3 (62.5)	< 0.001*	293.4 (90.6)	0.001*
One week after the 5th session	258.7 (61.4)	< 0.001*		
After the 10th session			285.6 (94.7)	< 0.001*
One week after the 10th session			284.6 (89.2)	< 0.001*

**p-value calculated using Student's paired t-test.*

TABLE 3 | Median melasma area and severity index (MASI) changes from baseline.

MASI	Group A ($n = 67$) median (IQR)	*p*-value	Group B ($n = 39$) median (IQR)	*p*-value
Baseline	5.2 (2.4, 9.0)	Reference	8.1 (4.8, 14.4)	Reference
After the 5th session	1.5 (1.2, 3.0)	< 0.001*	4.8 (1.8, 8.0)	< 0.001*
One week after the 5th session	1.2 (0.6, 2.1)	< 0.001*		
After the 10th session			1.8 (1.2, 4.6)	< 0.001*
One week after the 10th session			1.4 (1.1, 4.4)	< 0.001*

**p-value from Wilcoxon signed-rank test.*
Patient-Reported Melasma-Improvement.

TABLE 4 | Patient-reported improvement scores.

Satisfactory score	Group A ($n = 67$) n	%	Group B ($n = 39$) N	%
After the 5th session				
Worsening	-	-	-	-
No improvement	-	-	-	-
Slight improvement	-	-	2	5.7
Moderate improvement	3	4.5	4	11.4
Marked improvement	23	34.3	9	25.7
Almost or completely cured	41	61.2	20	57.1
After the 10th session				
Worsening			-	-
No improvement			-	-
Slight improvement			-	-
Moderate improvement			1	2.6
Marked improvement			7	18.0
Almost or completely cured			31	79.5

corroborates the high effective rate of acupuncture treatment for melasma from other previous studies' (13, 14).

The MASI and MI in our study were significantly reduced in both Groups A and B by 3.4 and 5.4 points, respectively for MASI, and 28.7 and 31.3, respectively for MI, at 1 week after the last session, which suggested a significant clinical magnitude of acupuncture in melasma. Recent systematic reviews and meta-analyses of oral tranexamic acid, a novel effective melasma therapy that modulates the vascular component of melasma demonstrated a decrease in MASI (1.60–1.78 points) and MI (0.69 points) (21, 22). Similarly, a meta-analysis of intense pulsed light combination therapy ranged from one to six treatment sessions in melasma reduced the MASI score only by 0.61 points (23). Furthermore, the two times per week interval of acupuncture established the shorter course of treatment compared with laser and light-based therapy. Our proposed short regimen of acupuncture is perhaps a practical and worthwhile treatment in melasma.

Since melasma is a chronic recurrent skin disease, the short practical regimen of acupuncture has many positive clinical and financial implications as in other chronic diseases. White et al. (24) mentioned the adequate dosage for acupuncture, and some meta-analysis studies proposed the adequacy criteria of at least six acupuncture sessions for low back pain (25) and chronic knee pain (26). Sangtin et al. applied a survival analysis technique in their retrospective observational data and found that 50% of stroke patients had significant improvement after eight sessions of acupuncture and reached peak improvement by 16 sessions (27). For melasma, a previous systematic review of 8 RCTs found that the number of acupuncture sessions varied from 21 to 60 over 4–16 weeks (two to seven times weekly) (19). In real practice, coming to the hospital multiple times per week for a long period is not practical both for the patients and hospital staff. In our study, the MI and MASI scores significantly decreased after 5–10 sessions with more than 95% of patients reported markedly improved or cured. Moreover, the first five sessions seem to have more magnitude of improvement than the 6th−10th sessions, although without statistically significant difference. Thus, given our setting, our protocol of 5–10 sessions is more practical than the protocol of other studies of >20 sessions (19), while

still being effective with high patient satisfaction. The melasma lesion initially assessed at the first visit was more severe than at the subsequent sessions of the treatment. This might affect the outcomes that are greater in the first five sessions compared with the 6th−10th sessions. Moreover, the renewal of the epidermis takes around 28 days which is the sum of the turnover time of each compartment (28–30). Timing for the first five acupuncture sessions is approximately equal to the shedding time of the upper part of epidermis. This may result in the obvious improvement after first five sessions.

We realized that from Chinese medicine theory, one of the important basic theories is syndrome differentiation and treatment according to each syndrome. However, we found that our short fixed regimen of acupoints yielded effective results for the majority of patients, also making the treatment easier to train and perform. The treatment was likely effective because our selected acupoint included encircling and intralesional melasma area, resulting in a local increase of blood flow (31). This correlates with the evidence from the findings of an RCT that acupuncture can increase skin blood perfusion (31). Moreover, a recent systematic review of melasma treatment targeting vascular components by tranexamic acid and laser/light therapies showed promising evidence (3, 9, 11). Apart from the blood circulation and vascular component which are the recent target treatment in melasma, the latest evidence revealed that melasma is a photoaging disorder (5, 6) which is related to oxidative damage by ROS (4). Antioxidant properties of acupuncture in melasma also correlates with promising evidence for the use of antioxidants to deal with melasma in modern medicine (13, 32). Apart from the local effects, acupoints in the rest of the body acupoint—Fengchi (GB20), Hegu (LI4), Xuehai (SP10), Jusanli (ST36), Sanyinjiao (SP6), and Taichong (LR3)—have already been established to improve commonly seen syndromes of melasma patients: liver-qi stagnation, spleen-deficiency, and kidney-yin/yang-deficiency (17, 33).

In our study, the minor yet frequent adverse events of acupuncture were pain and hematoma in the facial area. Therefore, when conducting acupuncture on the facial area, clinicians should apply pressure after taking off the needles for a longer period. Higher sessions of treatment seemed to have a higher rate of adverse events.

This observational study has several limitations. The data was collected pragmatically, based on the clinical practice protocol of our acupuncture clinic. The patients would receive at least five sessions as the standard short regimen and could decide whether they were willing to continue the treatment. Hence, our study could not be generalized in terms of a direct comparison of clinical efficacies between the two groups. However, the findings from our observations could be suggestive of the dose-response relationship of the acupuncture treatment for melasma. Some factors including duration of melasma, prior acupuncture experience, and family history of melasma did not affect the clinical outcome of the treatment according to the regression result. Also, this study had a relatively short follow-up time so further research should be conducted to see how long the effects last.

CONCLUSIONS

A short acupuncture regimen of 5–10 sessions in melasma seems to be effective and practical with minor side effects.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board (IRB), Faculty of Medicine, Chulalongkorn University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

TS, NS, and KP conceived the idea, designed the study, collected and analyzed the data, drafted, and approved the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by the Ratchadapiseksompotch Fund, Faculty of Medicine, Chulalongkorn University, Grant Number RA 060/60 and RA-MF-12/62.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Irin Lertparinyaphorn for drawing the figures, Dr. Thanapoom Rattananupong, Assoc. Prof. Somrat Lertmaharit, and Dr. Yuda Chongpison for statistical analysis advice.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2021.761017/full#supplementary-material

REFERENCES

1. Kauh YC, Zachian TF. Melasma. *Adv Exp Med Biol.* (1999) 455:491–9. doi: 10.1007/978-1-4615-4857-7_72
2. Miot LD, Miot HA, Silva MG, Marques ME. Physiopathology of melasma. *An Bras Dermatol.* (2009) 84:623–35. doi: 10.1590/S0365-05962009000600008
3. Masub N, Nguyen JK, Austin E, Jagdeo J. The vascular component of melasma: a systematic review of laboratory, diagnostic, and therapeutic evidence. *Dermatol Surg.* (2020) 46:1642–50. doi: 10.1097/DSS.0000000000002770
4. Sarkar R, Bansal A, Ailawadi P. Future therapies in melasma: What lies ahead? *Indian J Dermatol Venereol Leprol.* (2020) 86:8–17. doi: 10.4103/ijdvl.IJDVL_633_18
5. Kwon SH, Na JI, Choi JY, Park KC. Melasma: updates and perspectives. *Exp Dermatol.* (2019) 28:704–8. doi: 10.1111/exd.13844
6. Passeron T, Picardo M. Melasma, a photoaging disorder. *Pigment Cell Melanoma Res.* (2018) 31:461–5. doi: 10.1111/pcmr.12684
7. Deshpande SS, Khatu SS, Pardeshi GS, Gokhale NR. Cross-sectional study of psychiatric morbidity in patients with melasma. *Indian J Psychiatry.* (2018) 60:324–8. doi: 10.4103/psychiatry.IndianJPsychiatry_115_16
8. Handel AC, Miot LD, Miot HA. Melasma: a clinical and epidemiological review. *An Bras Dermatol.* (2014) 89:771–82. doi: 10.1590/abd1806-4841.20143063
9. Perper M, Eber AE, Fayne R, et al. Tranexamic Acid in the Treatment of Melasma: A Review of the Literature. *Am J Clin Dermatol.* (2017) 18:373–81. doi: 10.1007/s40257-017-0263-3
10. Del Rosario E, Florez-Pollack S, Zapata L. Jr, et al. Randomized, placebo-controlled, double-blind study of oral tranexamic acid in the treatment of moderate-to-severe melasma. *J Am Academy of Dermatology.* (2018) 78:363–9. doi: 10.1016/j.jaad.2017.09.053
11. Hassan AM, Elfar NN, Rizk OM, Eissa NY. Pulsed dye laser versus intense pulsed light in melasma: a split-face comparative study. *J Dermatolog Treat.* (2018) 29:725–32. doi: 10.1080/09546634.2018.1441487
12. Kong SH, Suh HS, Choi YS. Treatment of melasma with pulsed-dye laser and 1,064-nm Q-switched Nd:YAG laser: a split-face study. *Ann Dermatol.* (2018) 30:1–7. doi: 10.5021/ad.2018.30.1.1
13. Rerksuppaphol L, Charoenpong T, Rerksuppaphol S. Randomized clinical trial of facial acupuncture with or without body acupuncture for treatment of melasma. *Complement Ther Clin Pract.* (2016) 22:1–7. doi: 10.1016/j.ctcp.2015.10.004
14. Wu X, Xiang Y. The Effects of Acupuncture Combined with Auricular Acupressure in the Treatment of Chloasma. *Evid Based Complement Alternat Med.* (2018) 2018:6438458. doi: 10.1155/2018/6438458
15. Xun J. Clinical observation in 30 cases of chloasma treated by auricular point pressing and pricking. *J Tradit Chin Med.* (2003) 23:207.
16. Chen W. Fifty cases of chloasma treated by acupuncture plus intensive pulse light irradiation. *J Tradit Chin Med.* (2007) 27:265—7.
17. Feng X, Fu J, Liu F. Clinical observation on the combined use of acupuncture and herbal medicine for treatment of chloasma. *Journal of traditional Chinese medicine = Chung i tsa chih ying wen pan.* (2010) 30:15–7. doi: 10.1016/S0254-6272(10)60004-0
18. Wang Q. Zhen Jiu Zhi Liao Xue (Book in Chinese). *10th ed Beijing: Zhong Guo Zhong Yi Yao Chu Ban She.* (2009).
19. Chai Q, Fei Y, Cao H, Wang C, Tian J, Liu J. Acupuncture for melasma in women: a systematic review of randomised controlled trials. *Acupuncture in Medicine.* (2015) 33:254–61. doi: 10.1136/acupmed-2014-010633
20. Kimbrough-Green CK, Griffiths CE, Finkel LJ, et al. Topical retinoic acid (tretinoin) for melasma in black patients. *A vehicle-controlled clinical trial Archives of dermatology.* (1994) 130:727–33. doi: 10.1001/archderm.1994.01690060057005
21. Kim HJ, Moon SH, Cho SH, Lee JD, Kim HS. Efficacy and safety of tranexamic acid in melasma: a meta-analysis and systematic review. *Acta Derm Venereol.* (2017) 97:776–81. doi: 10.2340/00015555-2668
22. Zhang L, Tan WQ, Fang QQ, et al. Tranexamic acid for adults with melasma: a systematic review and meta-analysis. *Biomed Res Int.* (2018) 2018:1683414. doi: 10.1155/2018/1683414
23. Yi J, Hong T, Zeng H, et al. A meta-analysis-based assessment of intense pulsed light for treatment of melasma. *Aesthetic Plast Surg.* (2020) 44:947–52. doi: 10.1007/s00266-020-01637-x
24. White A, Cummings M, Barlas P, et al. Defining an adequate dose of acupuncture using a neurophysiological approach – a narrative review of the literature. *Acupuncture in Medicine.* (2008) 26:111. doi: 10.1136/aim.26.2.111
25. Manheimer E, White A, Berman B, Forys K, Ernst E. Meta-analysis: acupuncture for low back pain. *Ann Intern Med.* (2005) 142:651–63. doi: 10.7326/0003-4819-142-8-200504190-00014
26. White A, Foster NE, Cummings M, Barlas P. Acupuncture treatment for chronic knee pain: a systematic review. *Rheumatology (Oxford, England).* (2007) 46:384–90. doi: 10.1093/rheumatology/kel413
27. Sangtin S, Supasiri T, Tangsathitporn R, Pongpirul K. Application of survival analysis techniques to determine the optimal number of acupuncture therapy sessions for stroke patients. *Acupuncture in Medicine.* (2020) 2020:964528419883268. doi: 10.1177/0964528419883268
28. Ogawa E, Sato Y, Minagawa A, Okuyama R. Pathogenesis of psoriasis and development of treatment. *J Dermatol.* (2018) 45:264–72. doi: 10.1111/1346-8138.14139
29. Iizuka H. Epidermal turnover time. *J Dermatol Sci.* (1994) 8:215–7. doi: 10.1016/0923-1811(94)90057-4
30. Ohno K, Kobayashi Y, Uesaka M, et al. A computational model of the epidermis with the deformable dermis and its application to skin diseases. *Sci Rep.* (2021) 11:13234. doi: 10.1038/s41598-021-92540-1
31. Kim KW, Hong S, Kim HS, et al. Physiological impact of nanoporous acupuncture needles: Laser Doppler perfusion imaging in healthy volunteers. *PLoS ONE.* (2019) 14:e0226304. doi: 10.1371/journal.pone.0226304
32. Babbush KM, Babbush RA, Khachemoune A. The therapeutic use of antioxidants for melasma. *J Drugs Dermatol.* (2020) 19:788–92. doi: 10.36849/JDD.2020.5079
33. Hu J. Acupuncture treatment of chloasma. *Journal of traditional Chinese medicine = Chung i tsa chih ying wen pan.* (2008) 28:305–7. doi: 10.1016/S0254-6272(09)60017-0

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Supasiri, Salakshna and Pongpirul. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Pharmacology

SYSTEMATIC REVIEW
published: 29 November 2021
doi: 10.3389/fphar.2021.773712

Methodological Characteristics of Clinical Trials Supporting the Marketing Authorisation of Advanced Therapies in the European Union

Carolina Iglesias-Lopez[1], *Antònia Agustí*[1,2], *Antonio Vallano*[1,3]* *and Merce Obach*[3]

[1]Department of Pharmacology, Therapeutics and Toxicology, Universitat Autònoma de Barcelona, Barcelona, Spain, [2]Clinical Pharmacology Service, Vall d'Hebron University Hospital, Barcelona, Spain, [3]Medicines Department, Catalan Healthcare Service, Barcelona, Spain

Several advanced therapy medicinal products (ATMPs) have been approved in the European Union (EU). The aim of this study is to analyse the methodological features of the clinical trials (CT) that supported the marketing authorization (MA) of the approved ATMPs in the EU. A systematic review of the characteristics of pivotal CT of ATMPs approved in the EU until January 31st, 2021 was carried out. A total of 17 ATMPs were approved and 23 CT were conducted to support the MA (median, 1, range, 1–3). Of those studies, 8 (34.78%) were non-controlled and 7 (30.43%) used historical controls. Only 7 (30.4%) were placebo or active-controlled studies. Among all CT, 21 (91.3%) were open-label and 13 (56.52%) had a single-arm design. To evaluate the primary endpoint, 18 (78.26%) studies used an intermediate and single variable. The median (IQR) number of patients enrolled in the studies was 75 (22–118). To date, ATMPs' approval in the EU is mainly supported by uncontrolled, single-arm pivotal CT. Although there is a trend toward an adaptive or a life cycle approach, a switch to more robust clinical trial designs is expected to better define the benefit and the therapeutic added value of ATMPs.

Keywords: drug development, drug approval, research design, methods, clinical trials, advanced therapies, cell- and tissue-based therapy, genetic therapy

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Kurt Neumann,
Independent researcher, Kerékteleki, Hungary
Kata Mazalin,
Servier, Hungary

***Correspondence:**
Antonio Vallano
avallano@catsalut.cat

Specialty section:
This article was submitted to Drugs Outcomes Research and Policies, a section of the journal Frontiers in Pharmacology

Received: *10 September 2021*
Accepted: *28 October 2021*
Published: *29 November 2021*

Citation:
Iglesias-Lopez C, Agustí A, Vallano A and Obach M (2021) Methodological Characteristics of Clinical Trials Supporting the Marketing Authorisation of Advanced Therapies in the European Union. Front. Pharmacol. 12:773712. doi: 10.3389/fphar.2021.773712

INTRODUCTION

Advanced therapy medicinal products (ATMPs) are a medicinal class that includes gene, cell and tissue therapies. The success of ATMP development and the approval of these therapies in the European Union (EU) has been crucial to the growth of clinical research during the last few years in this field, particularly for gene therapy.

Multiple indications are being targeted, most of them being refractory and recurrent stages of a disease that lacks effective therapeutic alternatives, and a significant proportion of them affecting the paediatric population (Alamo et al., 2019). With the introduction of ATMPs that can cover unmet needs and have the potential to cure life-threatening diseases, biological therapies initiated a shift from traditional clinical development pathway to an accelerated and highly product-specific one. The adaptive pathway concept and Priority Medicines scheme (PRIME) were launched in the EU specifically to speed the access of products targeting a significant unmet medical need. Several approved ATMPs were granted a PRIME designation

and accelerated marketing authorisation application assessment during their development, allowing early access to these medicines (Iglesias-Lopez et al., 2021a).

Due to the type of target diseases, the inherent complexity of these products, and their accelerated developments, less comprehensive clinical data might be generated. These characteristics may lead to uncertainties in the benefit/risk profile for the product at the time of marketing authorization (MA). The aim of this study is to further analyse the clinical development of the current approved ATMPs'. Here, we describe the methodological features of the clinical trials that have driven ATMPs to their European approval and we compare the gene therapy trials versus the cell and tissue engineered trials.

METHODS

A systematic review of the pivotal trials' features that supported the MA of the ATMPs approved in the EU was carried out using the following approach:

1) *Search strategy*: Data collection was primarily extracted from European Public Assessment Reports on the European Medicines Agency (EMA) website (www.ema.europa.eu). The search was carried out until January 31st, 2021. In addition, a search for the main clinical trials of the approved ATMPs was conducted using ClinicalTrials.gov database and the related publications.
2) *Eligibility criteria*: Only products classified as ATMPs according to the EMA criteria (European Medicines Agency, 2010; Iglesias-Lopez et al., 2019) and authorised under centralised procedure in the EU have been considered. Combined ATMPs class, i.e., ATMP combined with a medical device, have been grouped according to the main ATMP category: gene therapy medicinal product, somatic cell therapy medicinal products or tissue engineered products. Only those trials identified or referenced as pivotal, and therefore, decisive for the MAA were analysed.
3) *Data extraction and collected variables*: The authors designed specific data extraction forms using Excel 2019 (Microsoft Corporation, Redmond, WA, United States) to collect information. For each ATMP the following variables were collected: type of ATMP, pharmacotherapeutic group, ATC code, therapeutic area (according to MeSH terms), diseases and other circumstances for its use (according to chapter's title from the international version of the ICD-10), number of assessed clinical indications and pivotal clinical trials conducted. For each pivotal clinical trial, the following variables were selected: phase, design, type of randomization, type of control, type of study blinding, number of arms, participating centres, type of hypothesis and primary endpoint, presence and type of health-related quality of life (HRQoL) endpoints, presence of pre-specified analysis, duration of the main phase of the study, pivotal trial ongoing at the time of MAA, overall number of patients that participated in the study (enrolled, on intervention arm or control arm and safety set), age and sex of population, existence or absence of previous treatments, and geographic location of the pivotal trial. To determine if the study was ongoing at the time of the submission, the MAA submission date and the final data collection date for the primary outcome measure of the pivotal clinical trial were reviewed. Standard definitions of analysis set were used to classify among intended to treat (ITT), modified ITT (mITT) and per protocol set (PP) following ICH (E9) and EMA guidelines (ICH, 1998b; European Medicnes Agency, 2007). To assign the type of hypothesis in the case of two variables being used to evaluate the primary endpoint, the most robust variable was selected, i.e., final *versus* surrogate variables.
4) *Statistical analysis*: Statistical analysis for categorical and continuous variables was made using means of proportions, mean, standard deviation (SD), median, quartiles 25 and 75 (Q25, Q75), and range (minimum and maximum). The statistical analysis was performed using SAS® 9.4 (SAS Institute Inc., Cary, NC, United States).

RESULTS

A total of 17 ATMPs have been approved in the EU (**Table 1**) and 23 main trials were conducted to support the MA for these products (median, 1, range, 1–3). The ATMPs trials by disease area, according to ICD-10 classification, included: neoplasms (7), endocrine, nutritional and metabolic diseases (2), diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism (2), diseases of the eye and adnexa (2), diseases of the nervous system (1), diseases of the musculoskeletal system and connective tissue (3), diseases of the digestive system (1). In addition, there were 6 ATMPs for rare inherited disorders and 6 for neoplasms in which 4 were indicated for haematological malignancies and 2 for solid tumours. The detailed results of this study are presented in **Table 2** by type of ATMP, in **Table 3** for gene therapy studies and in **Table 4** for cell and tissue therapy studies.

Regarding the design of the studies, 13 (56.52%) were Phase 2/3 and Phase 3 trials, 9 (39.13%) were Phase 1/2 or Phase 2 trials, and 1 (4.35%) was a retrospective study. For all types of therapies, 8 (34.78%) trials were non-controlled, 7 (30.44%) where active- or placebo-controlled, and 7 (30.43%) used an historical control as comparator. Differences were observed between gene and non-gene therapies (**Supplementary Table S1** in Supplementary material). Six (42.87%) gene therapy studies were non-controlled and 6 (42.87%) used a historical control, whilst cell and tissue therapies studies were mainly controlled (66.66%). A total of 14 (60.87%) studies were not randomized. Similarly, differences in the existence of randomization between gene and non-gene therapies studies were also observed. Most of the studies for gene products lacked randomisation (85.71%), whereas this was present in 75% of the cell therapies studies and 80% of the tissue therapies studies. A total of 21 (91.30%) were open-label studies; all gene and tissue therapy studies were open-

TABLE 1 | Approved ATMPs in the European Union and therapeutic indication.

Trade name	International non-proprietary name (INN) or common name	Pharmacotherapeutic group/ATC code	Therapeutic area (MeSH)	Chapter's title from the international version of the ICD-10
Gene therapy medicinal products				
Kymriah®	Tisagenlecleucel	Antineoplastic agents/L01XX71	Precursor B-Cell Lymphoblastic Leukemia-Lymphoma	Neoplasms
Kymriah®	Tisagenlecleucel	Antineoplastic agents/L01XX71	Lymphoma, Large-B-cell, Diffuse	Neoplasms
Yescarta®	Axicabtagene ciloleucel	Antineoplastic agents/L01XX70	Lymphoma, Large-B-cell, Diffuse	Neoplasms
Tecartus®	Autologous peripheral blood T cells CD4 and CD8 selected and CD3 and CD28 activated transduced with retroviral vector expressing anti-CD19 CD28/CD3-zeta chimeric antigen receptor and cultured	Antineoplastic agents/L01X	Lymphoma, Mantle-Cell	Neoplasms
Imlygic®	Talimogene laherparepvec	Antineoplastic agents/L01XX51	Melanoma	Neoplasms
Glybera®	Alipogene tiparvovec	Lipid modifying agents/C10AX10	Hyperlipo-proteinemia type I	Endocrine, nutritional and metabolic diseases
Strimvelis®	Autologous CD34 $^{+}$ enriched cell fraction that contains CD34 $^{+}$ cells transduced with retroviral vector that encodes for the human ADA cDNA sequence	Immunostimulants/L03	Severe combined immunodeficiency	Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
Luxturna®	Voretigene neparvovec	Ophthalmologicals, other ophthalmologicals/S01XA27	Leber congenital amaurosis Retinitis Pigmentosa	Diseases of the eye and adnexa
Zynteglo®	Betibeglogene autotemcel	Other haematological agents/ B06AX02	Beta-Thalassemia	Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
Zolgensma®	Onasemnogene abeparvovec	Other drugs for disorders of the musculoskeletal system/M09AX09	Muscular Atrophy Spinal	Diseases of the nervous system
Libmeldy®	Atidarsagene autotemcel	Other nervous system drugs/N07	Leukodystrophy, Metachromatic	Endocrine, nutritional and metabolic diseases
Somatic-cell therapy medicinal products				
Provenge®	Autologous peripheral-blood mononuclear cells activated with prostatic acid phosphatase granulocyte-macrophage colony-stimulating factor (Sipuleucel-T)	Other immunostimulants/L03AX17	Prostatic Neoplasms	Neoplasms
Zalmoxis®	Allogeneic T cells genetically modified with a retroviral vector encoding for a truncated form of the human low affinity nerve growth factor receptor (ΔLNGFR) and the herpes simplex I virus thymidine kinase (HSV-TK Mut2)	Antineoplastic agents/L01	Hematopoietic Stem Cell Transplantation Graft vs Host disease	Neoplasms Factors influencing health status and contact with health services
Alofisel®	Darvadstrocel	Immunosuppressants/L04	Rectal Fistula	Diseases of the digestive system
Tissue-engineered medicinal products				
Chondrocelect®	Characterised viable autologous cartilage cells expanded *ex vivo* expressing specific marker proteins/	Other drugs for disorders of the musculoskeletal system/M09AX02	Cartilage Diseases	Diseases of the musculoskeletal system and connective tissue
MACI®	Matrix-applied characterised autologous cultured chondrocytes	Other drugs for disorders of the musculoskeletal system/M09AX02	Fractures, Cartilage	Diseases of the musculoskeletal system and connective tissue
Spherox®	Spheroids of human autologous matrix-associated chondrocytes	Other drugs for disorders of the musculoskeletal system/M09AX02	Cartilage Diseases	Diseases of the musculoskeletal system and connective tissue
Holoclar®	*Ex vivo* expanded autologous human corneal epithelial cells containing stem cells	Ophthalmologicals/S01XA19	Stem Cell Corneal Diseases	Diseases of the eye and adnexa

label, and this was also the approach for 50% of cell products trials. However, there is a difference in the blinding evaluation of the relevant endpoints between gene and non-gene therapy studies, as such evaluation is mostly absent in the case of gene therapies (85.71%) but is present in the case of cell and tissue engineered therapies (50% for cell therapy studies and 100%

TABLE 2 | Design features of pivotal clinical trials for the approved advanced therapy medicinal products in the European Union.

ATMP clinical development		Gene therapy medicinal products	Somatic cell therapy medicinal products	Tissue engineered therapies	All types of therapies
Number of products	**N**	**10**	**3**	**4**	**17**
Number of indications per product	Mean (SD)	1.10 (0.32)	1 (0)	1 (0)	1.06 (0.24)
Total number of pivotal trials and studies	N	14	4	5	23
—	Mean (SD)	1.27 (0.65)	1.33 (0.58)	1.25 (0.5)	1.28 (0.57)
—	(min, Max)	(1, 3)	(1, 2)	(1, 2)	(1, 3)
Clinical trials	—	—	—	—	—
Phase 1	N (%)	0 (0)	0 (0)	0 (0)	0 (0)
Phase 1/2	N (%)	4 (28.57)	1 (25)	0 (0)	5 (21.74)
Phase 2	N (%)	3 (21.43)	0 (0)	1 (20)	4 (17.39)
Phase 2/3	N (%)	3 (21.43)	0 (0)	0 (0)	3 (13.04)
Phase 3	N (%)	4 (28.57)	3 (75)	3 (60)	10 (43.48)
Observational retrospective studies	N (%)	0 (0)	0 (0)	1 (20)	1 (4.35)
Randomization	—	—	—	—	—
No	N (%)	12 (85.71)	1 (25)	1 (20)	14 (60.87)
Yes 1:1	N (%)	0 (0)	1 (25)	4 (80)	5 (21.74)
Yes ≥2:1	N (%)	2 (14.29)	2 (50)	0 (0)	4 (17.39)
Control	—	—	—	—	—
Not controlled	N (%)	6 (42.87)	0 (0)	2 (40)	8 (34.78)
Placebo controlled	N (%)	0 (0)	2 (50)	0 (0)	2 (8.70)
Active controlled	N (%)	1 (7.14)	1 (25)	3 (60)	5 (21.74)
Historical control	N (%)	6 (42.87)	1 (25)	0 (0)	7 (30.43)
Other	N (%)	1 (7.14)	0 (0)	0 (0)	1 (4.35)
Blinding	—	—	—	—	—
Open label	N (%)	14 (100)	2 (50)	5 (100)	21 (91.30)
Single blind	N (%)	0 (0)	0 (0)	0 (0)	0 (0)
Double blind	N (%)	0 (0)	2 (50)	0 (0)	2 (8.70)
Blinding evaluation	—	—	—	—	—
Yes	N (%)	2 (14.28)	2 (50)	5 (100)	19 (82.61)
No	N (%)	12 (85.71)	2 (50)	0 (0)	4 (17.39)
Multicentric	—	—	—	—	—
No	N (%)	4 (28.57)	0 (0)	0 (0)	4 (17.39)
Yes	N (%)	10 (71.43)	4 (100)	5 (100)	19 (82.60)
Number of arms	—	—	—	—	—
1 arm	N (%)	11 (78.57)	1 (25)	1 (20)	13 (56.52)
2 arms	N (%)	2 (14.29)	3 (75)	3 (60)	8 (34.78)
3 arms	N (%)	1 (7.14)	0 (0)	1 (20)	2 (8.70)
Design	—	—	—	—	—
Parallel groups	N (%)	2 (14.29)	3 (75)	3 (60)	8 (34.78)
Single arm	N (%)	11 (78.57)	1 (25)	1 (20)	13 (56.52)
Other	N (%)	1 (7.14)	0 (0)	1 (20)	2 (8.70)
Main Outcomes	—	—	—	—	—
Final variable	N (%)	2 (14.28)	2 (50)	1 (20)	5 (21.74)
Intermediate variable	N (%)	12 (85.71)	2 (50)	4 (80)	18 (78.26)
Co-primary	N (%)	2 (14.28)	1 (25)	1 (20)	4 (17.39)
Composite variable	N (%)	1 (7.14)	0 (0)	0 (0)	1 (4.35)
Single variable	N (%)	11 (78.57)	3 (75)	4 (80)	18 (78.26)
Type of variable for main outcome	—	—	—	—	—
Qualitative	N (%)	13 (92.85)	3 (75)	1 (20)	17 (73.91)
Quantitative (discrete and continuous)	N (%)	2 (14.28)	1 (25)	4 (80)	7 (30.43)
Health related quality of life	—	—	—	—	—
No	N (%)	7 (50)	2 (50)	1 (20)	10 (43.48)
General questionnaires	N (%)	5 (35.71)	1 (25)	1 (20)	7 (30.43)
Specific questionnaires	N (%)	4 (28.57)	1 (25)	4 (80)	9 (39.13)
Prespecified previous analysis	—	—	—	—	—
Interim analysis	N (%)	11 (78.57	3 (75)	3 (75)	17 (73.91)
Final analysis type (primary analysis)	—	—	—	—	—
ITT	N (%)	10 (71.43)	3 (75)	5 (100)	18 (78.26)
mITT	N (%)	2 (14.28)	0 (0)	0 (0)	2 (8.69)
PP	N (%)	2 (14.28)	0 (0)	0 (0)	2 (8.69)
Hypothesis	—	—	—	—	—
Superiority	N (%)	1 (7.14)	3 (75)	1 (20)	5 (21.74)
Non-inferiority	N (%)	0 (0)	0 (0)	2 (40)	2 (8.7)

(Continued on following page)

TABLE 2 | (*Continued*) Design features of pivotal clinical trials for the approved advanced therapy medicinal products in the European Union.

ATMP clinical development		Gene therapy medicinal products	Somatic cell therapy medicinal products	Tissue engineered therapies	All types of therapies
Number of products	**N**	**10**	**3**	**4**	**17**
Other	N (%)	13 (92.85)	1 (25)	2 (40)	16 (69.56)
Mean time for the main phase (months)	Mean (SD)	11.5 (9.30)	70.50 (91.22)	24 (9.80)	35.33 (31.08)
Ongoing at the time of the MAA submission (final data for primary outcome measure)	—	—	—	—	—
Yes	N (%)	8 (57.14)	3 (75)	1 (25)	12 (57.14)
No	N (%)	6 (42.86)	1 (25)	3 (75)	10 (47.62)
Population	—	—	—	—	—
Population randomized/enrolled	N	1,065	798	543	2,406
—	Median (Q25 - Q75)	22 (18.75–106.5)	134.5 (27–437)	104 (88.50–131)	75 (22–118)
—	(min, Max)	(5, 437)	(17, 512)	(75, 144)	(5, 512)
Population on intervention arm	N	797	495	254	1,546
—	Median (Q25 - Q75)	21.5 (11.5–93.75)	68.5 (20.25–282.5)	64.5 (53.25–72.75)	41 (16.25–93.75)
—	(min, Max)	(5, 296)	(17, 341)	(52, 73)	(5, 341)
Population on control arm	N	151	416	183	750
—	Median (Q25, Q75)	75.5 (NA)	140 (105–171)	61 (50–72)	88.50 (27.5–140.5)
—	(min, Max)	(10, 141)	(105, 171)	(50–72)	(10, 171)
Population on safety set	N	933	780	439	2,152
—	Median (Q25 -Q75)	22.5 (13.5, 93.75)	128.5 (25.75–430.75)	110 (81.75–137.5)	63.5 (20–118)
—	(min, Max)	(5, 419)	(17, 506)	(75, 144)	(5, 419)
Age of adult population (years)	Mean (SD)	54.29 (9.24)	52.77 (16.67)	37.14 (5.56)	47.84 (18.45)
Age of paediatric population (years)	Mean (SD)	6.15 (8.26)	NA	NA	6.15 (8.26)
Sex	—	—	—	—	—
Female	N (%)	443 (47)	191 (30.31)	231 (42.54)	865 (37.53)
Male	N (%)	498 (53)	630 (76.73)	312 (57.45)	1,440 (62.47)
Location of the pivotal clinical trial	—	—	—	—	—
United States	N (%)	9 (64.28)	1 (25)	0 (0)	10 (43.48)
Europe	N (%)	10 (71.42)	3 (75)	5 (100)	18 (78.26)
Canada	N (%)	5 (35.71)	1 (25)	0 (0)	6 (26.09)
Others	N (%)	7 (50)	3 (75)	0 (0)	10 (43.48)
Previous treatments	—	—	—	—	—
Yes and No	N (%)	1 (7.14)	0 (0)	0 (0)	1 (4.35)
No	N (%)	3 (21.74)	0 (0)	2 (40)	5 (21.74)
Yes	N (%)	10 (65.21)	2 (50)	3 (60)	15 (65.22)

ITT: intended to treat; mITT: modified intended to treat; NA: not applicable; PP: per protocol set; Zynteglo pooled analysis (Studies HGB-204, HGB-205 and LFT-303) was counted as one pivotal study; Holoclar retrospective study was counted as a pivotal study, since it was considered the main study which lead to the Marketing Authorisation of the product; The final analysis type (primary analysis) for TK0008 study of Zalmoxis was not available; The mean time for the main phase excludes Provenge (defined as 'until disease progression or death') and TK0008 study for Zalmoxis; Age of adult population: data not available for TK0008 study for Zalmoxis; Age of paediatric population: data only available for Tecartus, Libmeldy, Kymriah and Strimvelis; Previous treatments: not applicable for Zalmoxis. For the Health related quality of life outcomes, the percentages can exceed 100% given that there might be multiple questioners for the same product (i.e., generic and disease-specific).

tissue engineered therapy studies). A total of 13 (56.52%) studies were single-arm trials and 10 (43.48%) had two or more arms. A difference in the number of arms between gene and non-gene therapy studies was also observed, where single-arm studies comprised 78.57% of total trials for gene therapy products versus the two- or three-arm designs present in 75% of cell therapy studies and 80% of tissue therapy studies. Accordingly, there are some differences in the design between gene and non-gene therapies studies, mainly in the parallel designs for cell and tissue engineered therapy studies versus single-arm designs for gene therapy studies. Of all studies analysed, 19 (82.60%) were multicentric.

Regarding the methodology used in these pivotal studies, 16 (69.56%) of the studies did not use a superiority or non-inferiority hypothesis but an alternative premise, e.g., comparison with historical controls. There is a difference between gene and non-gene therapies studies, where this type of alternative premises was mainly used for gene therapies trials (92.85%), while standard superiority or non-inferiority tests were used more frequently for cell and tissue engineered therapies trials (75 and 60%, respectively). To evaluate the primary objective, 18 (78.26%) of the trials used an intermediate and single main variable, which was mainly qualitative (73.91%). Final and quantitative variables were used in 5 (21.74%) and 7 (30.43%), respectively, which represents a smaller proportion (**Table 5**). Of these confirmatory studies, 18 (78.26%) used the intention-to-treat (ITT) principle in assessing the primary efficacy, 2 (14.28%) gene therapy trials used modified intention-to-treat (mITT) and 2 (14.28%) used per protocol set (PP). A total of 16 (69.56%) analysed studies included HRQoL questionnaires, 9 (39.13%) of

TABLE 3 | Design features of pivotal clinical trials for the approved gene therapy medicinal products in the European Union.

Gene therapies	Glybera®			Imlygic®	Strimvelis®	Yescarta®	Kymriah®		Luxturna®	Zynteglo®		Zolgensma®	Tecartus®	Libmeldy®
CHMP Positive Opinion date	Jun-23-11			Oct-22-15	Apr-01-16	Jun-28-18	Jun-29-18		Set-20-18	Apr-26-19		Mar-26-20	Oct-15-20	Oct-15-20
Authorisation status/type	Withdrawn			Authorised	Authorised	Authorised	Authorised		Authorised	Authorised		Authorised	Authorised	Authorised
Type of authorisation	Under exceptional circumstances			Standard	Standard	Standard	Standard		Standard	Conditional		Conditional	Conditional	Standard
Clinical trial Acronym	CT-AMT-011–01	CT-AMT-011–02	CT-AMT-010–01	Study 005/05	Study AD1115611/ Gene-ADA	ZUMA-1	Study B2202	Study C2201	AAV2-hRPE65v2-301/302	Studies HGB-204, HGB-205 and LFT-303	Studies HGB-207, HGB-212	Study CL-303 (STR1VE)	ZUMA-2	Study 201,222
Phase	II/III	II/III	II/III	III	I/II	I/II	II	II	III	I/II	III	III	II	I/II
Randomization	No	No	No	2:1	No	No	No	No	2:1	No	No	No	No	No
Control	Non-controlled	Non-controlled	Non-controlled	Active control	Historical control	Historical control	Historical control	Historical control	Delayed-intervention control group	Non-controlled	Non-controlled	Historical control	Non-controlled	Historical control
Blinding design	Open label	Open label	Open label	Open label	Open label	Open label	Open label	Open label	Open label	Open label	Open label	Open label	Open label	Open label
Blinding evaluation	No	No	No	No	No	No	No	No	Yes	No	No	No	Yes	No
Multicentric	Single-centre	Dual-centre	Single-centre	Multicentric	Single-centre	Multicentric	Multicentric	Multicentric	Dual-centre	Multicentre	Multicentre	Multicentre	Multicentre	Single-centre
Number of arms	Three	One	One	Two	One	One	One	One	Two	One	One	Two	One	One
Design	Parallel arms (dose range)	Single arm	Single arm	Parallel arms	Single arm	Single arm	Single arm	Single arm	Parallel arms	Single arm	Single arm	Single arm	Single arm	Single arm
Main Outcomes	Intermediate and single variable	Intermediate and single variable	Intermediate and composite variable	Intermediate and single variable	Final and single variable	Intermediate and single variable	Intermediate and single variable	Intermediate and single variable	Intermediate and single variable	Intermediate and single variable	Intermediate and single variable	Final and co-primary variable	Intermediate and single variable	Intermediate and co-primary variable
Type of variable for main outcome	Qualitative	Qualitative	Qualitative	Qualitative	Qualitative	Qualitative	Qualitative	Qualitative	Quantitative	Qualitative	Qualitative	Qualitative	Qualitative	Qualitative and Quantitative
Health related quality of life	No	General questionnaire	No	Specific questionnaire	No[a]	No	General questionnaire	General and specific questionnaires	No	No	General and specific questionnaires	No	General questionnaire	Specific questionnaire
Prespecified previous analysis	Interim analysis	Interim analysis	N/A	Interim analysis	None	Interim analysis	Interim analysis	Interim analysis	None	Interim analysis	Interim analysis	Interim analysis	Interim analysis	Interim analysis
Final analysis type (primary efficacy analysis)	ITT	ITT	ITT	ITT	ITT	mITT	PP	PP	ITT	ITT	ITT	ITT	mITT	ITT
Hypothesis	Description of efficacy of intervention	Description of efficacy of intervention	Description of efficacy of intervention	Superiority over an active control	Superiority over historical control group	Intervention compared to historical control	Description of efficacy of intervention	Intervention compared to historical control	Intervention compared non-intervention (natural history)	Description of efficacy of intervention	Description of efficacy of intervention	Superiority versus natural observation study	Description of efficacy of intervention	Superiority versus natural history cohort (or untreated sibling when available)

(Continued on following page)

TABLE 3 | (*Continued*) Design features of pivotal clinical trials for the approved gene therapy medicinal products in the European Union.

Gene therapies	Glybera®			Imlygic®	Strimvelis®	Yescarta®		Kymriah®	Luxturna®	Zynteglo®		Zolgensma®	Tecartus®	Libmeldy®
Mean time for the main phase (months)	3	3	3	12	36	12	3	12	12	12	12	14	3	24
Ongoing at the time of the MAA submission (final data for primary outcome measure)	Yes	Yes	No	No	No	Yes	Yes	Yes	No	Two studies ongoing	Yes	Yes	No	No
Population	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Population randomised/ enrolled	22	5	18	437	12	111	92	147	31	22	19	22	105	22
Population on intervention arm	14	5	8	296	12	101	75	99	21	22	10	22	92	20
Population on control arm	NA	NA	NA	141	NA	NA	NA	NA	10	NA	NA	NA	NA	NA
Population on safety set	14	5	8	419	12	101	75	99	29	23	14	22	92	20
Age of population (years)	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Mean	45.6	41.8	N/A	63.07	1.7	56.3	12	54	N/A	N/A	N/A	0.31	65	3.6
Sex														
Female	9	1	N/A	250	5	33	32	36	18	15	6	12	15	11
Male	5	4	N/A	187	7	68	43	63	13	7	5	10	77	9
Geographic region	—	—	—	—	—	—	—	—	—	—	—	—	—	—
North America	X	X	—	X	—	X	X	X	X	X	X	X	X	—
Europe	—	—	X	X	X	X	X	X	—	X	X	—	X	X
Others	—	—	—	X	X	X	X	X	—	X	X	—	—	—
Previous treatments	Yes	Yes	Yes	Yes/No	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Yes	No

ITT: intended to treat; mITT: modified intended to treat; NA: not applicable; N/A: not available; PP: per protocol set.
[a]Not at the time of the submission. The HRQoL objective applied to the long-term follow-up (4–8 years after gene therapy) only.

TABLE 4 | Design features of pivotal clinical trials for the approved cell and tissue engineered therapy medicinal products in the European Union.

	Cell therapies				Tissue therapies				
	Provenge®	**Zalmoxis®**		**Alofisel®**	**ChondroCelect®**	**Holoclar®**	**MACI®**	**Spherox®**	
CHMP Positive Opinion date	Jun-12-13	Jun-23-16		Dec-14-17	Jun-25-09	Mar-06-13	Apr-25-13	May-18-17	
Authorisation status	Withdrawn	Withdrawn		Authorised	Withdrawn	Authorised	Withdrawn	Authorised	
Type of authorisation	Standard	Conditional marketing authorisation		Standard	Standard	Conditional	Standard	Standard	
Clinical trial Acronym	9902B (IMPACT)	TK007	TK008	ADMIRE-CD	TIG/ACT/01&EXT′	HLSTM01	SUMMIT	Cod 16 HS 14	Cod 16 HS 13
Phase	III	I/II	III	III	III	Observational retrospective study	III	II	III
Randomization	2:1	No	3:1	1:1	1:1	No	1:1	1:1:1	1:1
Control	Placebo	Historical control[a]	Active treatment	Placebo	Active treatment	Non-controlled	Active treatment	Non-controlled	Active treatment
Blinding	Double-blind	Open-label	Open-label	Double-blind	Open-label	Open-label	Open-label	Open-label	Open-label
Blinding evaluation	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Multicentric	Multicentric	Multicentric	Multicentric	Multicentric	Multicentric	Dual-centre	Multicentric	Multicentric	Multicentric
Number of arms	Two	One	Two	Two	Two	One	Two	Three	Two
Design	Parallel groups	Single arm	Parallel groups	Parallel groups	Parallel groups	Retrospective case-series	Parallel groups	Single arm (three doses)	Parallel groups
Main Outcomes	Final and single variable	Intermediate and single variable	Intermediate and single variable	Final and co-primary variable	Intermediate and co-primary variable	Final and single variable	Intermediate and single variable	Intermediate and single variable	Intermediate and single variable
Type of variable for main outcome	Qualitative	Qualitative	Quantitative	Qualitative	Quantitative	Qualitative	Quantitative	Quantitative	Quantitative
Health related quality of life	No	No	General questionnaire	Specific questionnaire	Specific questionnaire	No	General and Specific questionnaire	Specific questionnaire	Specific questionnaire
Prespecified previous analysis	Interim analysis	None	Interim analysis	Interim analysis	None	NA	Interim analysis	Interim analysis	Interim analysis
Final analysis type (primary efficacy analysis)	ITT	ITT	NA	ITT	ITT	ITT	ITT	ITT	ITT
Hypothesis	Superiority over placebo	Description of efficacy of intervention	NA	Superiority over placebo	Non-inferiority vs SOC	Exploratory	Superiority over SOC	Superiority over baseline for the three dose groups	Comparison with baseline and non-inferiority/ superiority with SOC
Duration of the main phase (months)	Until disease progression or death	135	NA	6	36	NA	24	12	24
Ongoing at the time of the MAA submission (primary completion)	No	No	Yes	No	No	NA	Yes	Yes	Yes
Population	—	—	—	—	—	—	—	—	—
Population enrolled	512	57	17	212	118	NA	144	75	102
Population on intervention arm	341	30	17	107	57	104	72	73	52
Population on control arm	171	140	Not known	105	61	NA	72	NA	50
Population on Safety set	506	52	17	205	118	NA	144	75	102

(Continued on following page)

TABLE 4 | (*Continued*) Design features of pivotal clinical trials for the approved cell and tissue engineered therapy medicinal products in the European Union.

	Cell therapies				Tissue therapies				
	Provenge®	Zalmoxis®		Alofisel®	ChondroCelect®	Holoclar®	MACI®	Spherox®	
Age of population	—	—	—	—	—	—	—	—	—
Mean	71	49	N/A	38.3	33.9	46.8	34	34	37
Sex	—	—	—	—	—	—	—	—	—
Female	NA	30	N/A	161	42	24	51	53	61
Male	512	22	N/A	96	76	80	93	22	41
Geographic region	—	—	—	—	—	—	—	—	—
North America	X	—	—	—	—	—	—	—	—
Europe	—	X	X	X	X	X	X	X	X
Others	—	X	X	X	—	—	—	—	—
Previous treatments	Yes	NA	NA	Yes	Yes	Yes	No	No	No

ITT: intended to treat; NA: not applicable; N/A: not available; SOC: standard of care.
[a]*Upon assessment of the TK007 data and as only limited data from the TK008 study were available, the applicant was asked to perform a comparison of the MM-TK treated patients (TK007 and TK008 combined) with results from suitable historical controls.*

those being disease-specific. No differences were observed in the type of HRQoL questionnaires between gene and non-gene products studies, i.e., generic versus disease-specific variables.

The mean (SD) time for the main phase of the trial was 35.33 (31.08) months, approximately 1 year for the gene therapies and more than 2 years for cell and tissue engineered therapies. A total of 12 (57.14%) studies were ongoing at the time of submission, meaning that the final data collection for primary outcome measuring was not completed. Globally, 17 (73.91%) of the studies had a prespecified interim analysis, with similar proportion among the three types of ATMPs (75–78.57%).

Regarding the overall population size and location of these studies, the median (IQR 25–75) number of patients enrolled in the analysed ATMPs pivotal clinical trials was 75 (22–118). The mean ± SD age of the adult population included in these confirmatory trials was 48 ± 18.45 years old. There is no sufficient data to establish a mean ± SD age for paediatric populations. The sex distribution is higher for males (62.47%) than for women (37.53%). The analysed clinical trials were equally performed in both women and men, but the overall sex distribution was higher for males due to Provenge®'s indication, i.e., treatment of metastatic castration-resistant prostate cancer. The median (IQR 25–75) sample size in the intervention arm was 41 (16–94) patients and 63 (20–118) for the safety set. More than half of participants in these clinical trials had received previous treatments (65.22%). From the 23 pivotal studies analysed, 18 included sites located in the EU (78.26%), and 10 (43.48%) in the United States of America (US) or in other regions, such as Israel, Japan or Australia.

DISCUSSION

Clinical research on ATMPs has increased during the last few years (Alamo et al., 2019). The introduction of ATMPs and the long-term expectancy of their benefit adds a new challenge for the regulatory agencies. In the present study, we aimed to describe the most relevant methodological features of the clinical trials that have driven ATMPs to their approval. The major findings reveal that the pivotal studies of currently approved advanced therapies typically share the following characteristics: 1) they are small, open-label, non-randomised, single-arm studies without control or using historical ones, and 2) intermediate and single variables are used to evaluate the primary efficacy outcome. In addition, this type of designs is more common for gene therapies than for cell and tissue therapies.

Hanna *et al* previously reported the methodological characteristics of clinical trials assessing ATMPs in an early development phase based on clinical trials registries (Hanna et al., 2016). The results showed very similar characteristics to those found in this study such as small sample size, non-randomised trials, single-arm trials, surrogate endpoints, and adaptive designs. Coppens *et al.*, also reported that the level of scientific evidence required for the approval might differ among different regulatory agencies (Coppens et al., 2018). Elsallab *et al.* showed that clinical trials of ATMPs did not meet the same strict standards for clinical evidence that were applied to other biologicals submissions (Elsallab et al., 2020). This previously reported data, together with the results of the present study, highlight the limited clinical evidence upon which the authorisation of most ATMPs is based. Of these approved ATMPs, it was considered that eleven (64.70%) had sufficient data for a full MA, while for the remaining six products, five (29.41%) obtained a conditional approval and one (5.88%) was granted with a MA under exceptional circumstances.

The low disease prevalence, the disease severity and burden, the lack or scarce availability of disease-modifying treatments, the patient population's heterogeneity and the strong presence of paediatric patient populations comprise some of the factors that could contribute to this type of designs.

The type of target diseases has been one of the key factors that might have given more flexibility in terms of level of evidence required for the MA. Our analysis shows that these

TABLE 5 | Primary clinical variables of pivotal clinical trials for the approved ATMPs in the European Union.

Type of product	Product	Type of target disease	Intermediate (I) or final (F) variable	Primary variable description
GTMP	Kymriah (ALL)	Haematological malignancies	I	Overall remission rate, which included CR and CR with incomplete blood count recovery
	Kymriah (DLBCL)	Haematological malignancies	I	Overall response rate defined as the proportion of patients with a BOR of CR or PR, where the BOR was defined as the best disease response recorded from tisagenlecleucel until progression disease or start of new anticancer therapy
	Yescarta	Haematological malignancies	I	Objective response rate, defined as a CR or PR per the revised International Working Group Response Criteria for Malignant Lymphoma as determined by study investigators
	Tecartus	Haematological malignancies	I	Objective response rate, defined as CR or PR using central assessment per Lugano Classification
	Imlygic	Solid tumour	I	Durable response rate was defined as the percentage of participants with a CR or PR maintained continuously for at least 6 months from the time the objective response was first observed and initiating within 12 months of starting therapy as assessed by the Endpoint Assessment Committee
SCTMP	Provenge	Solid tumour	F	Overall survival defined as time from randomization to death due to any cause was analysed for the ITT population
GTMP	Glybera	Inherited monogenic diseases	I	Reduction in fasting plasma triglycerides (median of baseline vs median of week 3–12 post AMT-011) ≥ 40% Achievement of 40% reduction of median fasting triglycerides concentrations 12 weeks after treatment with AMT-011 Reduction in individual median fasting plasma triglyceride levels of ≤10 mmol/L concurrent with a low-fat diet, or 40% reduction, concurrent with a low-fat diet
	Strimvelis	Inherited monogenic diseases	F	Survival at 3 years post-gene therapy
	Luxturna		I	Subject's bilateral performance (no eye patching) on the mobility test, as measured by a change score, 1 year following vector administration as compared to a subject's Baseline bilateral mobility test performance
	Zynteglo	Inherited monogenic diseases	I	The proportion of subjects who meet the definition of transfusion independence (TI). TI is defined as a weighted average Hb ≥ 9 g/dl without any packed red blood cells transfusions for a continuous period of ≥12 months at any time during the study after drug product infusion
	Zolgensma	Inherited monogenic diseases	F/Co-primary	Proportion of patients that achieve functional independent sitting for at least 30 s at the 18 months of age study visit. It is defined by the Bayley Scales of Infant and Toddler Development (Version 3), confirmed by video recording, as a patient who sits up straight with head erect for at least 30 s Survival at 14 months of age
	Libmeldy	Inherited monogenic diseases	I/Co-primary	Total Gross Motor Function Measure score 2 years after treatment was the primary endpoint The co-primary endpoint was the ARSA activity
TEP	Chondrocelect	Condrophaties	I/Co-primary	Histomorphometry on end point biopsies at 12 months post-surgery and overall Histology Assessment on First Subscale of ICRS II Score Change from Baseline in Overall Knee Injury and Osteoarthritis Outcome Score at 12–18 Months
	MACI	Condrophaties	I	Change from Baseline to Week 104 for the Participant's Knee Injury and Osteoarthritis Outcome Score Pain and Function (Sports and Recreational Activities) Scores
	Spherox	Condrophaties	I	Change of overall Knee Injury and Osteoarthritis Outcome Score from baseline to final assessment determined for each dosage group and between the dosage groups Change of overall Knee Injury and Osteoarthritis Outcome Score from baseline to final assessment compared between intervention arm and comparator

(Continued on following page)

TABLE 5 | (*Continued*) Primary clinical variables of pivotal clinical trials for the approved ATMPs in the European Union.

Type of product	Product	Type of target disease	Intermediate (I) or final (F) variable	Primary variable description
SCTMP	Zalmoxis	Adjunctive treatment in haploidentical haematopoietic stem cell transplantation	I	Proportion of patients who achieved immune reconstitution, empirically defined a priori as an absolute $CD3^+$ cell count of 100/μL or more for two consecutive observations (and/or $CD4^+$ cells ≥50/μL and/or $CD8^+$ cells ≥50/μL) Disease-free survival measured from the date of randomization until the date of relapse (or progression), or death from any cause, whichever occurs first
	Alofisel	Complex perianal fistula(s)—Crohn's disease	F/Co-primary	Combined remission of perianal fistulising Crohn's disease and absence of collections >2 cm of the treated fistula confirmed by MRI images, at week 24. Remission was defined as clinical closure of external openings that were draining at baseline despite gentle finger compression
TEP	Holoclar	Limbal stem cell deficiency	F	Composite endpoint of the rate of patients with a successful transplantation at 12 months post-intervention, based on the co-presence of clinical signs

Intermediate variable*: a clinical endpoint such as measure of a function or of a symptom (disease-free survival, angina frequency, exercise tolerance) but is not the ultimate endpoint of the disease, such as survival or the rate of irreversible morbid events (stroke, myocardial infarction).*

Final variable: *describes a valid measure of clinical benefit due to treatment: the impact of treatment on how a patient feels, functions and survives. It is clinically relevant, sensitive (responsive to change) and is both accepted and used by physicians and patients. Clinical endpoints may be a clinical event (e.g. mortality) a composite of several events, a measure of clinical status, or health related quality of life (HRQoL) [*Ref: *EUnetHTA, 2015a, b)*. Guideline on Endpoints used for Relative Effectiveness Assessment of pharmaceuticals: Clinical endpoints. https://www.eunethta.eu/wp-content/uploads/2018/01/Clinical-endpoints.pdf*].*

ARSA, arysulfatase A enzyme; CR, complete response; GTMP, gene therapy medicinal product; ITT, intended to treat; PR, partial response; SCTMP, somatic cell therapy medicinal product; TEP, tissue engineered medicinal product.

designs are more commonly used for the development of gene therapy products, which target orphan diseases such as hematologic cancers or rare inherited monogenic disorders (40 and 60% of approved gene therapies, respectively), usually with unmet medical needs. Gene therapies were mainly authorised after conducting a single open-label study, usually non-randomised and non-controlled or using historical controls, and only few of them being Phase III studies. By contrast, tissue therapies trials consisted of Phase III studies controlled with the standard of care, and two out of three cell therapy trials conducted placebo-controlled studies. The approved tissue therapies primarily cover products for articular cartilage damage or prostate cancer, which are relatively common among the overall population and with several treatments available. Moreover, the target population might have also contributed to these alternative designs for gene therapies products, given that 60% of approved gene therapies target paediatric population, while all of the tissue and cell therapies target adults. The targeted paediatric diseases are life-threatening or with a huge impact on patients' and caregivers' quality-of-life, and randomised, controlled trials could have posed ethical concerns, as well as recruitment issues.

It is noteworthy to mention that different types of historical controls were used to compare the efficacy of the intervention: historical references from retrospective studies and retrospective databases, prospective natural history cohorts' studies, untreated sibling data and within-subject comparison between pre- and post-treatment assessments (Hassan et al., 2012; European Medicines Agency, 2021; Maude et al., 2018; Crump et al., 2017). The current EMA guideline states that orphan products are assessed according to the same standards as those for other products but considering their limitations due to low patient recruitment (European Medicnes Agency, 2006). While the same guideline states that most orphan drugs and paediatric indications submitted for regulatory approval are based on randomised controlled trials and deviation from such standards is uncommon, in the case of the current approved ATMPs, alternative approaches as historical controls were frequently used, i.e., Strimvelis®, Kymriah®, Luxturna®, Zolgensma® and Libmeldy®.

On the other hand, the line of treatment is another factor that might have justified these types of designs so far. As an example of the approved ATMPs, CAR-T therapies are indicated at least as a third-line therapy for relapsed or refractory cancer patients. The four pivotal studies conducted for these products were non-controlled, open-label, Phase II studies where the intervention arm was compared to a historical control. After the approval of the aforementioned therapies, the EMA has published recommendations on clinical considerations on CAR-T-cell product development (European Medicnes Agency, 2020), where it is stated that randomized controlled trial design should be followed even for those cases of late-stage refractory disease. It will be interesting to see how these recommendations are implemented in the near future.

Another important factor observed in the studied designs is the use of surrogate variables instead of a clinically relevant final endpoint. Intermediate endpoints can be used as a primary

endpoint for MA, especially when there is a high unmet need, when clinical events are rare/delayed in slowly progressive diseases and a very long follow-up is needed for their assessment, and for rare and/or life-threatening diseases with no therapeutic alternative available (EUnetHTA, 2015b; ICH, 1998a). In the case of all approved gene therapies that target cancer diseases, the proportion of patients with objective overall response rate (ORR) was used as the intermediate primary variable, unlike cell therapy trials that used overall survival (OS) as a final endpoint. OS analysis usually requires a large sample size, a long follow-up and should be evaluated in a randomised, control trial to avoid cofounding factors due to the switch-over of control to intervention or subsequent therapies (Gutman et al., 2013; Pazdur, 2008). However, ORR has been the most commonly used surrogate endpoint in support of accelerated/conditional approvals, but also of standard approvals, since it is directly attributable to a drug's effect, providing an accurate assessment in single-arm trials conducted in patients with refractory tumours (Food and Drug Administration, 2018a). On the other hand, for gene therapies targeting inherited monogenic diseases, biomarkers were commonly used to predict changes in the desired clinical endpoints, and at least one of the pivotal studies included HRQoL outcomes. Exceptionally for other products, a novel clinical meaningful endpoint, i.e. Luxturna® (Russell et al., 2017), or survival as a final primary outcome were used, i.e. Zolgensma® (Del Rosario et al., 2020; Cech et al., 2012).

These types of non-robust designs for new drugs in areas of high unmet medical need are mainly justified on the basis of ethical reasons, based on the potential life-saving opportunities or quality of life improvement for patients who may not survive or will progress rapidly until robust clinical data is available. On the other hand, the difficulties of conducting standard clinical developments with orphan drugs are well-recognised, and single small trials using alternative approaches have been the basis for numerous MAA in the recent years (Blin et al., 2020; Micallef and Blin, 2020; Picavet et al., 2013; Pontes et al., 2018). This regulatory flexibility sometimes comes at the cost of having a less comprehensive clinical data, and in consequence, greater uncertainty about the product's benefit-risk balance at the time of MA (Iglesias-Lopez et al., 2021b). In addition, since the introduction of the adaptive pathway concept, the shift towards accelerated clinical developments has also been associated with an intrinsic uncertainty on effectiveness and safety, which can result in promising Phase II results but an unsuccessful Phase III or post-marketing studies (Pharma Intelligence, 2019; Novartis press release, 2021a, b). This highlights the possibility for a patient to receive an early-authorised treatment without meaningful clinical benefits and with exposure to its adverse effects, missing clinical opportunities, and wasting healthcare system resources (Ermisch et al., 2016).

The speeding up access to new drugs is achieved by putting aside traditional Phase III clinical trials in favour of post-marketing evidence generation. This fact is translated into the need to perform long and extensive post-marketing studies, where the costs of evidence generation as well as the costs of therapy are likely to be transferred from the MA holder to healthcare systems (Ermisch et al., 2016; Joppi et al., 2016). It is known, that this post-authorisation commitments can be challenging due to the long-term follow-up, which may lead to delays to complete the studies, and given that patients are more reluctant to participate in a post-marketing trial with all its constraints, if the medicine is already available, above all in those cases where the trial includes randomization (Joint briefing paper, 2015).

Costly treatments with high uncertainties in regard to its benefits, translates to a complex evaluation by the Health Technology Assessment bodies (HTAb), as well as there is industry pressure for corporate pharma and its investors to ensure sustainability in drug development.

Several detailed methodological recommendations for clinical trial designs have been launched to address the shortcomings of carrying out studies in small population (Day et al., 2018; ASTERIX project, 2021; IDEAL project, 2021; Friede et al., 2018) and examples of effective use of a historical control have also been reported (Mulberg et al., 2019). Multi-arm designs and platform designs sharing where a common control is shared have been raised as a potential solution (International Rare Diseases Research Consortium, 2016; Food and Drug Administration, 2018b). Comparator data can also be taken from pragmatic trials, observational studies or registries, but ensuring its quality (EUnetHTA, 2015a). In addition, real world data plays a key role to provide sufficient therapeutic evidence for these type of therapies and efforts are being made for a better use of registries (European Medicines Agency, 2017).

Methodological and clinical guidelines for a specific medical condition is an effective manner of obtaining regulatory guidance and providing a predictable decision-making regulatory framework. Given that ATMPs are innovative and more complex than traditional pharmaceuticals or other biological drugs, some specific requirements related to the study design and methodology, study population, safety, dose selection, as well as preclinical and product controls need to be considered for the development of these therapies. The FDA has launched several guidelines for the development of ATMPs aimed at certain types of conditions based on the acquired experience of the current approved advanced therapies. These guidelines address the point of uncontrolled designs and the need of more robust study designs in order to provide proper evidence of efficacy (Food and Drug Administration, 2020a; Food and Drug Administration, 2020b; Food and Drug Administration, 2021). Although still limited, with the current experience of the approved ATMPs in the EU, the EMA has started to launch new recommendations on the types of study designs and methodologies that can support the MA more robustly (European Medicnes Agency, 2020). This fact might lead to a switch on the current trend used in clinical designs based on uncontrolled pivotal studies or with historical control comparisons to randomised-controlled trials.

The limitations of this study are the small sample size and the fact that further analysis, once more therapies are approved, is required to determine with greater accuracy

the most common clinical design and methodology for ATMPs, as well as to elucidate the potential differences between gene therapy trials versus cell and tissue therapy trials. Another limitation is that approved ATMPs have not been compared to other approved medicines. Nevertheless, this is an exhaustive study that evaluates the pivotal trials for approved ATMPs.

CONCLUSION

The results of our study show that most authorised ATMPs are based on small, open-label, uncontrolled and single-arm pivotal trials using single and intermediate variables to evaluate outcomes. ATMPs are innovative therapies that mainly target orphan diseases and high unmet medical needs. This fact has led to certain methodological. weaknesses in their pivotal clinical trials, which in turn has resulted in limited data to robustly assess the benefit/risk of the product. A gradual shift towards the production of more methodologically sound randomized-controlled trials is expected to better define the benefit and the therapeutic added value of ATMPs.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

Conception and design of the study: CI-L, AV, AA, and MO. Acquisition of data: CI-L. Analysis and interpretation of data: CI-L, AV, AA, and MO. Drafting and revising the manuscript: CI-L. Reviewed and edited the manuscript: CI-L, AA, MO, and AV. All authors have approved the final article.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2021.773712/full#supplementary-material

REFERENCES

Álamo, J. E., Timón, M., Gómez-Platero, C. G., Abad, C. D., González, M. V., et al. (2019). Clinical Trials of Advanced Therapy Investigational Medicinal Products in Spain: Preparing for the European Clinical Trials Regulation. *Cell Gene Ther. Insights* 5, 1431–1449. doi:10.18609/cgti.2019.147

ASTERIX project (2021). Advances in Small Trials dEsign for Regulatory Innovation and eXcellence. Available at: http://www.asterix-fp7.eu/(Accessed August 28, 2021).

Blin, O., Lefebvre, M. N., Rascol, O., and Micallef, J. (2020). Orphan Drug Clinical Development. *Therapie* 75, 141–147. doi:10.1016/j.therap.2020.02.004

Cech, D. J., and Martin, S. T. (2012). Chapter 5-Evaluation of Function, Activity, and Participation," in *Functional Movement Development Across the Life Span*. Elsevier, 88–104. doi:10.1016/b978-1-4160-4978-4.00005-3

Coppens, D. G. M., de Wilde, S., Guchelaar, H. J., De Bruin, M. L., Leufkens, H. G. M., Meij, P., et al. (2018). A Decade of Marketing Approval of Gene and Cell-Based Therapies in the United States, European Union and Japan: An Evaluation of Regulatory Decision-Making. *Cytotherapy* 20, 769–778. doi:10.1016/j.jcyt.2018.03.038

Crump, M., Neelapu, S. S., Farooq, U., Van Den Neste, E., Kuruvilla, J., Westin, J., et al. (2017). Outcomes in Refractory Diffuse Large B-Cell Lymphoma: Results from the International SCHOLAR-1 Study. *Blood* 130, 1800–1808. doi:10.1182/blood-2017-03-769620

Day, S., Jonker, A. H., Lau, L. P. L., Hilgers, R. D., Irony, I., Larsson, K., et al. (2018). Recommendations for the Design of Small Population Clinical Trials. *Orphanet J. Rare Dis.* 13, 195. doi:10.1186/s13023-018-0931-2

Del Rosario, C., Slevin, M., Molloy, E. J., Quigley, J., and Nixon, E. (2020). How to Use the Bayley Scales of Infant and Toddler Development. *Arch. Dis. Child. Educ. Pract. Ed.* 106, 108–112. doi:10.1136/archdischild-2020-319063

Elsallab, M., Bravery, C. A., Kurtz, A., and Abou-El-Enein, M. (2020). Mitigating Deficiencies in Evidence during Regulatory Assessments of Advanced Therapies: A Comparative Study with Other Biologicals. *Mol. Ther. Methods Clin. Dev.* 18, 269–279. doi:10.1016/J.OMTM.2020.05.035

Ermisch, M., Bucsics, A., Vella Bonanno, P., Arickx, F., Bybau, A., Bochenek, T., et al. (2016). Payers' Views of the Changes Arising through the Possible Adoption of Adaptive Pathways. *Front. Pharmacol.* 7, 305. doi:10.3389/fphar.2016.00305

EUnetHTA (2015a). Criteria for the Choice of the Most Appropriate Comparator(s) Summary of Current Policies and Best Practice Recommendations. Available at: https://eunethta.eu/wp-content/uploads/2018/03/Criteria_WP7-SG3-GL-choice_of_comparator_amend2015.pdf (Accessed April 19, 2021).

EUnetHTA (2015b). Guideline on Endpoints Used for Relative Effectiveness Assessment of Pharmaceuticals: Surrogate Endpoints. Available at: https://eunethta.eu/wp-content/uploads/2018/03/surrogate_endpoints.pdf.

European Medicines Agency (2021). *Committee for Medicinal Products for Human Use (CHMP) Assessment Report: Libmeldy*. Available at: https://www.ema.europa.eu/en/documents/assessment-report/libmeldy-epar-public-assessment-report_en.pdf.

European Medicines Agency (2017). Patient Registry Initiative- Strategy and Mandate of the Cross-Committee Task Force. Available at: www.ema.europa.eu/contact (Accessed July 23, 2021).

European Medicines Agency (2010). Reflection Paper on Classification of Advanced Therapy Medicinal Products. Available at: https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-classification-advanced-therapy-medicinal-products_en-0.pdf (Accessed March 6, 2020).

European Medicnes Agency (2006). Committe for Medicinal Products for Human Use (CHMP), Guideline on Clinial Trials in Small Populations Populations. Available at: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-trials-small-populations_en.pdf (Accessed March 28, 2021).

European Medicnes Agency (2020). Guideline on Quality, Non-clinical and Clinical Aspects of Medicinal Products Containing Genetically Modified Cells. Available at: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-quality-non-clinical-clinical-aspects-medicinal-products-containing-genetically-modified_en-0.pdf (Accessed February 2, 2021).

European Medicnes Agency (2007). *Reflection Paper on Methodological Issues in Confirmatory Clinical Trials Planned with an Adaptive Design*. London: EMEA Doc. Ref. CHMP/EWP/2459/02 http://www.emea.europa.eu (Accessed May 10, 2021).Available at:

Food and Drug Administration (2018a). Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics Guidance for Industry. Available at: https://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htmand/orhttps://www.fda.gov/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/default.htm (Accessed April 11, 2021).

Food and Drug Administration (2020a). Human Gene Therapy for Hemophilia; Guidance for Industry. Available at: https://www.fda.gov/vaccines-blood-biologics/guidance-compliance- (Accessed February 2, 2021).

Food and Drug Administration (2021). Human Gene Therapy for Neurodegenerative Diseases; Draft Guidance for Industry. Available at: https://www.fda.gov/vaccines-blood-biologics/guidance-compliance- (Accessed February 2, 2021).

Food and Drug Administration (2020b). Human Gene Therapy for Retinal Disorders; Guidance for Industry. Available at: https://www.fda.gov/vaccines-blood-biologics/guidance-compliance- (Accessed February 2, 2021).

Food and Drug Administration (2018b). Master Protocols: Efficient Clinical Trial Design Strategies to Expedite Development of Oncology Drugs and Biologics Guidance for Industry DRAFT GUIDANCE. Available at: https://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm (Accessed July 20, 2021).

Friede, T., Posch, M., Zohar, S., Alberti, C., Benda, N., Comets, E., et al. (2018). Recent Advances in Methodology for Clinical Trials in Small Populations: the InSPiRe Project. *Orphanet J. Rare Dis.* 13, 186. doi:10.1186/s13023-018-0919-y

Gutman, S. I., Piper, M., and Grant, M. D. (2013). Progression-Free Survival: What Does It Mean for Psychological Well-Being or Quality of Life? Rockville, MD: Agency for Healthcare Research and Quality (US). Available at: https://www.ncbi.nlm.nih.gov/books/NBK137763/(Accessed April 11, 2021).

Hanna, E., Rémuzat, C., Auquier, P., and Toumi, M. (2016). Advanced Therapy Medicinal Products: Current and Future Perspectives. *J. Mark Access Health Pol.* 4, 31036. doi:10.3402/jmahp.v4.31036

Hassan, A., Booth, C., Brightwell, A., Allwood, Z., Veys, P., Rao, K., et al. (2012). Outcome of Hematopoietic Stem Cell Transplantation for Adenosine Deaminase-Deficient Severe Combined Immunodeficiency. *Blood* 120, 3615–3626. doi:10.1182/blood-2011-12-396879

ICH (1998a). ICH Topic E8 General Considerations for Clinical Trials Step 5 Note for Guidance on General Considerations for Clinical Trials. Available at: http://www.emea.eu.int (Accessed April 11, 2021).

ICH (1998b). ICH Topic E9 Statistical Principles for Clinical Trials Step 5 Note for Guidance on Statistical Principles for Clinical Trials. Available at: http://www.emea.eu.int (Accessed March 27, 2021).

IDEAL project (2021). Integrated DEsign and AnaLysis of Clinical Trials in Small Population Groups. Available at: https://www.ideal.rwth-aachen.de/(Accessed August 28, 2021).

Iglesias-López, C., Agustí, A., Obach, M., and Vallano, A. (2019). Regulatory Framework for Advanced Therapy Medicinal Products in Europe and United States. *Erratum in: Front Pharmacol* 11, 766. doi:10.3389/fphar.2019.00921

Iglesias-Lopez, C., Obach, M., Vallano, A., and Agustí, A. (2021a). Comparison of Regulatory Pathways for the Approval of Advanced Therapies in the European Union and the United States. *Cytotherapy* 23, 261–274. doi:10.1016/j.jcyt.2020.11.008

Iglesias-Lopez, C., Agustí, A., Vallano, A., and Obach, M. (2021b). Current landscape of clinical development and approval of advanced therapies. *Mol Ther Methods Clin Dev* 23, 606–618. doi:10.1016/j.omtm.2021.11.003

International Rare Diseases Research Consortium (2016). Small Population Clinical Trials Task Force Workshop Report and Recommendations. Available at: https://www.irdirc.org/wp-content/uploads/2017/12/SPCT_Report.pdf (Accessed May 16, 2021).

Joint briefing paper (2015). "Adaptive Licensing" or "Adaptive Pathways": Deregulation under the Guise of Earlier Access Executive. Available at: http://english.prescrire.org/Docu/DOCSEUROPE/(Accessed April 19, 2021).

Joppi, R., Gerardi, C., Bertele', V., and Garattini, S. (2016). Letting post-marketing Bridge the Evidence gap: The Case of Orphan Drugs. *BMJ* 353, i2978. doi:10.1136/bmj.i2978

Maude, S. L., Laetsch, T. W., Buechner, J., Rives, S., Boyer, M., Bittencourt, H., et al. (2018). Tisagenlecleucel in Children and Young Adults with B-Cell Lymphoblastic Leukemia. *N. Engl. J. Med.* 378, 439–448. doi:10.1056/nejmoa1709866

Micallef, J., and Blin, O. (2020). Orphan Drug Designation in Europe: A Booster for the Research and Development of Drugs in Rare Diseases. *Therapie* 75, 133–139. doi:10.1016/j.therap.2020.02.003

Mulberg, A. E., Bucci-Rechtweg, C., Giuliano, J., Jacoby, D., Johnson, F. K., Liu, Q., et al. (2019). Regulatory Strategies for Rare Diseases under Current Global Regulatory Statutes: A Discussion with Stakeholders. *Orphanet J. Rare Dis.* 14 (1), 36. doi:10.1186/s13023-019-1017-5

Novartis press release (2021a). Novartis' Kymriah Fails to Meet Primary Goal in Non-hodgkin Lymphoma Trial. Available at: https://www.clinicaltrialsarena.com/news/novartis-kymriah-lymphoma-trial/(Accessed August 27, 2021).

Novartis press release (2021b). Novartis Kymriah® Demonstrates Consistent Efficacy and Safety Outcomes in US Patients when Used in Real-World Setting | Novartis. Available at: https://www.novartis.com/news/media-releases/novartis-kymriah-demonstrates-consistent-efficacy-and-safety-outcomes-us-patients-when-used-real-world-setting (Accessed April 18, 2021).

Pazdur, R. (2008). Endpoints for Assessing Drug Activity in Clinical Trials. *Oncologist* 13 (Suppl. 2), 19–21. doi:10.1634/theoncologist.13-S2-19

Pharma Intelligence (2019). Disappointing End for MolMed's Zalmoxis Cell Therapy in EU. Available at: https://pink.pharmaintelligence.informa.com/PS140998/Disappointing-End-For-MolMeds-Zalmoxis-Cell-Therapy-In-EU (Accessed March 8, 2020).

Picavet, E., Cassiman, D., Hollak, C. E., Maertens, J. A., and Simoens, S. (2013). Clinical Evidence for Orphan Medicinal Products-A Cause for Concern? *Orphanet J. Rare Dis.* 8, 164. doi:10.1186/1750-1172-8-164

Pontes, C., Fontanet, J. M., Vives, R., Sancho, A., Gómez-Valent, M., Ríos, J., et al. (2018). Evidence Supporting Regulatory-Decision Making on Orphan Medicinal Products Authorisation in Europe: Methodological Uncertainties. *Orphanet J. Rare Dis.* 13, 206. doi:10.1186/s13023-018-0926-z

Russell, S., Bennett, J., Wellman, J. A., Chung, D. C., Yu, Z. F., Tillman, A., et al. (2017). Efficacy and Safety of Voretigene Neparvovec (AAV2-hRPE65v2) in Patients with RPE65-Mediated Inherited Retinal Dystrophy: a Randomised, Controlled, Open-Label, Phase 3 Trial. *Lancet* 390, 849–860. doi:10.1016/S0140-6736(17)31868-8

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The findings and conclusions in this article should not be construed to represent any agency determination or policy.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Iglesias-Lopez, Agustí, Vallano and Obach. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 12 January 2022
doi: 10.3389/fphar.2021.811700

The Transition From Ambrisentan to Macitentan in Patients With Pulmonary Arterial Hypertension: A Real-word Prospective Study

*Yusi Chen[1], Jun Luo[1], Jingyuan Chen[1], Eugene Kotlyar[2], Zilu Li[1], Wenjie Chen[1] and Jiang Li[1]**

[1]Department of Cardiovascular Medicine, The Second Xiangya Hospital of Central South University, Changsha, China, [2]St. Vincent's Hospital, Sydney, NSW, Australia

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Antonio Molino,
University of Naples Federico II, Italy
Salvatore Fuschillo,
Fondazione Salvatore Maugeri
(IRCCS), Italy

***Correspondence:**
Jiang Li
lijiangcs@csu.edu.cn
orcid.org/0000-0003-4904-6635

Specialty section:
This article was submitted to
Drugs Outcomes Research and
Policies,
a section of the journal
Frontiers in Pharmacology

Received: 09 November 2021
Accepted: 16 December 2021
Published: 12 January 2022

Citation:
Chen Y, Luo J, Chen J, Kotlyar E, Li Z, Chen W and Li J (2022) The Transition From Ambrisentan to Macitentan in Patients With Pulmonary Arterial Hypertension: A Real-word Prospective Study. Front. Pharmacol. 12:811700. doi: 10.3389/fphar.2021.811700

Background: In a long-term event-driven trial, macitentan has demonstrated beneficial time to clinical worsening in patients with pulmonary arterial hypertension (PAH) and reduced PAH-related hospitalization rates compared with placebo. Macitentan is the most recently approved endothelin receptor antagonist (ERA) and is the first ERA that has shown efficacy for morbidity and mortality in PAH patients; therefore, patients and physicians may consider converting treatment from ambrisentan to macitentan. Our study evaluated the safety, efficacy, and quality of life in PAH patients transitioning from ambrisentan to macitentan.

Methods: This was a real-world, prospective study with a 12-month follow-up. PAH patients who had received stable doses of ambrisentan for over 3 months, were within the World Health Organization Functional Class II/III, and 6-min walk distance ≥ of 250 m were enrolled. The study included a screening period, followed by a transition phase, after which patients entered the long-term follow-up. Clinical data and treatment satisfaction outcomes were collected to assess and monitor the safety and efficacy of the transition. The trial was registered at the Chinese Clinical Trial Registry (www.chictr.org.cn; No. ChiCTR2000034898).

Results: One hundred and fifty-seven enrolled PAH patients completed the transition. All criteria for continuous treatment transition were met by 145 patients (92.4%). Results showed improvements in exercise capacity, cardiac function, and hemodynamics compared with baseline. During the process, 4 patients discontinued macitentan due to adverse events. There was no statistical difference in the overall incidence of adverse events before and after the transition.

Conclusion: Transition to macitentan from ambrisentan was successful and well-tolerated by PAH patients, and was associated with greater efficacy and satisfaction.

Keywords: pulmonary arterial hypertension, macitentan, ambrisentan, endothelin receptor antagonist, real-word study

INTRODUCTION

Pulmonary arterial hypertension (PAH) is a progressive pulmonary vascular disease associated with high mortality, characterized by pulmonary vascular remodeling and increased pulmonary vascular resistance (PVR) (Boucherat et al., 2017). Without timely treatment, PAH eventually leads to right ventricular failure and even death. Over the past several decades, advances in targeted therapies have significantly improved the prognosis of PAH patients. Medications targeting the endothelin pathway play an important role in the treatment of PAH. Endothelin 1 (ET-1), a potent vasoconstrictor that modulates pulmonary vascular remodeling, is overexpressed within the remodeled pulmonary arteries and results in vascular narrowing (Benza et al., 2015a; Dai et al., 2019). ET-1 binds to two receptors, endothelin type A (ET-A) and endothelin type B (ET-B) to mediate its effects. Endothelin receptor antagonists (ERA) that inhibit either ET-A or ET-A and ET-B receptors are used as monotherapy or as a component of combination regimens to treat PAH.

At present, approved ERA medications for PAH include bosentan, ambrisentan, and macitentan, all of which are available in China. Bosentan was the first available ERA; however, it has been shown to increase the incidence of dose-dependent increases in liver aminotransferase concentrations (Markova et al., 2013). More recently, two other ERAs, macitentan and ambrisentan, which have rare hepatotoxicity, have been approved for the treatment of PAH.

Ambrisentan, the only available selective ERA, is developed theoretically to inhibit the vasoconstrictor function of ET-A while not damaging the vasodilatory and clearance effects of ET-B (Hitzerd et al., 2020). The ARIES-1 (Ambrisentan in Pulmonary Arterial Hypertension, Randomized, Double-Blind, Placebo-Controlled, Multicenter, Efficacy Study) and ARIES-2 trials had shown that ambrisentan alleviates exercise intolerance and improves World Health Organization Functional Class (WHO FC) and hemodynamics in PAH patients (Galiè et al., 2008) after 12 weeks of therapy. Furthermore, the AMBITION trial demonstrated that an upfront combination of ambrisentan with tadalafil, a phosphodiesterase type 5 inhibitor (PDE-5i), was associated with a 63% reduction in risk of PAH-related hospitalization compared with monotherapy (Vachiéry et al., 2019) and delayed clinical deterioration in PAH when used initially (Hoeper et al., 2016). However, no large-scale studies to date have demonstrated that ambrisentan monotherapy results in improved long-term morbidity and mortality in PAH patients (Oudiz et al., 2009; Hoeper et al., 2016).

Macitentan is the most recently developed ERA to be approved for management of PAH, and has favorable pharmacokinetics, without necessitating monthly monitoring hepatic risk. In the early stages of research and development, macitentan exhibited better tissue penetration and receptor affinity than did bosentan and ambrisentan. In an event-driven 100 week-long trial, compared with placebo, macitentan demonstrated a significant reduction in clinical worsening in PAH, first PAH-related events, or all-cause mortality in patients with WHO FC II and III, and significantly reduced PAH-related hospitalization rates (Pulido et al., 2013). Macitentan is the first ERA that has been approved to reduce morbidity and mortality in patients with PAH. SERAPHIN found that among the 64% of PAH patients who had received background-targeted drug therapy, macitentan significantly improved the primary endpoint by 38% (Pulido et al., 2013). Furthermore, in a hemodynamic sub-study of SERAPHIN, macitentan therapy was associated with the improved cardiac index, right atrial pressure, mean pulmonary artery pressure (mPAP), PVR, and N-terminal pro-brain natriuretic peptide (NT-proBNP) levels (Galie et al., 2017). The multicenter randomized controlled trial phase III Macitentan in Eisenmenger Syndrome to Restore Exercise Capacity study revealed that NT-proBNP levels and PVR decreased in patients with Eisenmenger syndrome who received macitentan compared with those who received placebo, although it did not show superiority over placebo on the primary endpoint of change from baseline to week 16 in the 6-min walk test (Gatzoulis et al., 2019).

With the wide medications options available to physicians managing PAH patients, there is often a solid pharmacological and clinical reason for patients switching to a newer agent from ambrisentan to reduce adverse effects (AEs) or achieve greater efficacy. Additionally, ambrisentan has not been fully covered by insurance in China since 2019, and the financial burdens of the patient and the health care system may be reduced by at least 60% with macitentan therapy. Given these benefits, it is anticipated that many patients and clinicians will choose to transition from ambrisentan to macitentan therapy. However, clinical experience in switching patients from ambrisentan to macitentan is rarely reported. Although these drugs share a common mode of action, there are likely differences in individual patients' responses in terms of clinical effectiveness and tolerance. Furthermore, side effects, such as liver toxicity and drug interactions differ between the two ERA. Moreover, although there is a pharmacokinetic basis for different efficacy of macitentan, no trials have been performed to directly compare the efficacy and safety of the two ERAs.

Since physicians and patients may consider converting from ambrisentan to macitentan, this open-label, prospective study (ChiCTR2000034898), was carried out to resolve whether such transition could be reasonably and safely accomplished, the proportion of patients that successfully transition, the frequency and severity of AEs during and after transition, and various clinical parameters. The study was also aimed at providing a reference for cautious, carefully supervised, safe, and efficient therapeutic transitions within the endothelin pathway and to monitor the process of this approach.

MATERIALS AND METHODS

Study Design

This trial was a 12-month follow-up, real-world, prospective, open-label cohort study. The main objective was to evaluate the safety, tolerability, and efficacy of replacing ambrisentan with macitentan therapy in patients with PAH using a rapid switch protocol. Additionally, the impact of the change in therapy on

treatment satisfaction was assessed by the patients' quality of life (QOL) questionnaire. The trial received approval from the ethics committee at Second Xiangya Hospital of Central South University and was undertaken under the Declaration of Helsinki. All enrolled subjects provided informed consent before the entry into the study. The trial was registered at the Chinese Clinical Trial Registry (www.chictr.org.cn; No. ChiCTR2000034898).

Selection Process and Criteria

The study consisted of a screening period, followed by a transition phase (the first 1 month), after which patients proceed to long term follow-up phase (12 months). All PAH patients voluntarily tried to change the ERA to macitentan due to the unsatisfactory treatment effect of ambrisentan, the desire to obtain better drug treatment effects, or the desire to reduce the financial burden. Patients who wish to switch must also undergo a rigorous assessment by doctors to ensure they are suitable for drug transition. Patients were enrolled in the study if they met the following criteria: 1) aged over 18 years; 2) stable for more than 3 months, defined as no worsening of WHO FC within 3 months, no clinical evidence of heart failure, and diuretic dose increase less frequently than once per month; 3) PAH diagnosis was confirmed by right heart catheterization: mPAP ≥ 25 mmHg; pulmonary capillary wedge pressure ≤ 15 mmHg; PVR ≥ 3 Wood units; 4) ambrisentan used for at least 6 months before transition, and a stable dose for no less than 3 months; 5) 6-min walk distance (6 MWD) of ≥250 m; 6) function class is no worse than WHO FC III; 7)women required to use at least one method of contraception during the study; 8) idiopathic PAH (IPAH), PAH associated with connective tissue disease (CTD-PAH), or PAH due to congenital heart disease (CHD-PAH), drugs, or toxins use.

Transitioning Protocol

The decision to change to macitentan was at the discretion of the clinicians and patients during clinic visits, and the strategy for changing drugs was negotiated with the patients. During the visit, the cardiologist explained the revised dosing scheme and potential AEs and ensured that patients understood how they could contact our team. On receipt of the medication, patients were guided to take their last dose of ambrisentan at the regular time of their transition date and to take a new ERA the next day. Patients were able to contact the clinical team *via* telephone whenever they had any questions. The patient could withdraw permission and discontinue participation at any point in the study.

Outcome Measures

To evaluate the safety and efficacy of the transition, AEs and their relationship to the medication were monitored throughout the study, as well as any changes in concomitant therapies. However, to facilitate analyses, we reported the absolute number and incidence of AEs at baseline and months 1, 3, 6, and 12 for the statistical analysis stage. Serum levels of transaminases and hemoglobin were measured after transition at months 1, 3, 6, and 12.

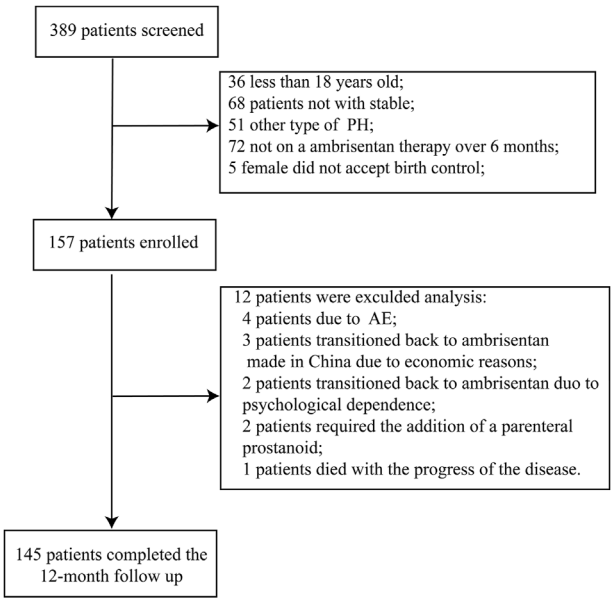

FIGURE 1 | Flow chart of patients selection.

To assess whether patients maintained clinical stability, efficacy assessment was conducted at baseline and months 6 and 12 and included a physical examination, vital signs, 6 MWD, WHO FC, NT-proBNP, echocardiography, and REVEAL 2.0 risk score calculator (Benza et al., 2015b). Absolute change in 6 MWD from baseline, and the proportion of patients with an increased (>15%), maintained (±15%), and decreased (>15%) 6 MWD from baseline were calculated. A cut-off value of 15% was selected to take account of the instability of the 6 MWD results. Echocardiographic parameters measured the diameter of the right atrium (RA) and ventricle, the diameter of the pulmonary artery (PA), the tricuspid annular plane systolic excursion (TAPSE), the systolic velocity of the tricuspid annulus (S′), the right ventricular area change rate (RVFAC), and the estimated pulmonary artery systolic pressure (sPAP). The echocardiography was elucidated by cardiologists blinded to the study procedures. Risk stratification was assessed by the REVEAL 2.0 risk score calculator. The correspondence between the low-, intermediate-, and high-risk groups, as defined by the REVEAL 2.0 calculator, were as follows: low risk = REVEAL score ≤ 6, intermediate risk = REVEAL score 7 and 8, and high risk = REVEAL score ≥ 9 (Hoeper et al., 2017). Preceding right heart catheterizations were necessitated to confirm the diagnosis of PAH and were collected as a baseline characteristic but were not used to monitor patients during the transition. QOL was assessed by a medicines questionnaire (Lopez et al., 2021) at baseline and at the final follow-up to evaluate patients' perceptions of treatment effectiveness, convenience, and satisfaction.

Statistical Analysis

Quantitative data (age, 6 MWD, NT-proBNP, and echocardiography parameters) were described as means and standard deviations (normal or approximately normal

distribution) or medians and interquartile ranges (IQR) (non-normal distribution). Normally distributed data were compared using a one-way analysis of variance for repeated measurement data, and non-normal distributions were compared using Friedman's M test. Qualitative data (AEs) and ranked data (WHO FC and REVEAL score) were described by the number and percentage. Ranked data were compared using Friedman's M test, and qualitative data were compared using Cochran's Q test. Differences with a p-value of < 0.05 were considered significant.

RESULTS

Patients' Clinical Characteristics

389 PAH patients were screened for participation in the study, 232 were excluded; 36 were less than 18 years of age; 68 were unstable; 51 had non-group 1 Pulmonary Hypertension; 72 were not on ambrisentan therapy for at least 6 months; 5 female patients were not willing to use any contraception (details in **Figure 1**). Patients were mostly female (74.5%) and WHO FC II at baseline (58.6%; details in **Table 1**). The majority of patients were CHD-PAH (60.7%), followed by IPAH (22.1%) and CTD-PAH (14.4%). Most patients (87.6%) were receiving a PDE-5i, or riociguat in addition to ambrisentan at baseline. Patients had been receiving ambrisentan for a median of 1.5 years (IQR, 1.0–2.9 years) before participating in the study. Sixty-one patients (42.1%) were receiving 10 mg once per day, whereas eighty-four patients (57.9%) were receiving 5 mg once daily.

TABLE 1 | Baseline characteristics.

Characteristic	All patients (N = 145)
Age, years[a]	32.0 (26.0, 42.0)
Sex, n (%)	
Female	108 (74.5)
Male	37 (25.5)
Duration of ambrisentan treatment before transition, years[a]	1.5 (1.0, 2.9)
mPAP, mmHg[b,c]	61.4 ± 18.6
PVR, Wood[a,c]	14.4 (8.0, 22.1)
mRAP, mmHg[b,c]	10.0 ± 5.8
mRVP, mmHg[b,c]	41.7 ± 12.4
Etiology of PAH, n (%)	
Idiopathic	32 (22.1)
Drug or toxin-induced	4 (2.8)
Associated with connective tissue disease	21 (14.4)
Associated with congenital heart disease	88 (60.7)
Ambrisentan dose, mg, QD	
5 n (%)	84 (57.9)
10 n (%)	61 (42.1)
WHO FC, n (%)	
II	85 (58.6)
III	60 (41.4)
BMI (kg/m2)[b]	21.0 ± 3.2
PAH medications in addition to taking ambrisentan, n (%)	
Monotherapy	18 (12.4)
PDE-5i	122 (84.1)
Riociguat	4 (2.8)
Riociguat and PDE-5i	1 (0.7)

[a]*Data are described as medians and interquartile ranges (non-normal distribution).*
[b]*Data are described as means and standard deviations (normal or approximately normal distribution).*
[c]*Historical data from the most recent right heart catheterization before enrollment.*
mPAP, mean pulmonary arterial pressure; PVR, pulmonary vascular resistance; mRAP, mean right atrial pressure; mRVP, mean right ventricular pressure; QD, once per day; PAH, pulmonary arterial hypertension; WHO FC, World Health Organization functional class; BMI, body mass index; PDE-5i, phosphodiesterase type 5 inhibitor.

Treatment Transition

Of the 157 enrolled patients, all patients switched safely and stably from ambrisentan to macitentan within the first month of the study. All patients were stable when they ceased taking ambrisentan. At the time of conversion, all patients were on 10 mg of macitentan once daily, and most (145 of 157) participants were on a drug regimen throughout the study and completed the whole study from January 2020 to July 2021. One hundred and forty-five patients (92.4%) met all preplanned criteria for constant transition at month 12, and 12 patients were excluded from analyses due to early discontinuation. Of the 12 excluded patients, 2 (1.3%) patients transitioned to ambrisentan of Chinese manufacture for poor financial reasons; 3 (1.9%) transitioned back to ambrisentan due to psychological dependence; 4 (2.5%) female patients suspended macitentan due to AEs; 2 (1.3%) patients required the addition of a parenteral treprostinil due to disease progression; one patient died at the 37-week follow-up. The death was likely due to the sudden cessation of subcutaneous treprostinil therapy and was not considered related to the ERA treatment.

Safety

One or more AEs occurred in 107 subjects (73.8%) during the trial. **Table 2** compared AEs that occurred in only ≥ 5% of the study population. Most AEs commenced at the transition phase. The most frequent AEs were headache, peripheral edema, and anemia, consistent with ERA class effects. Despite the high incidence, most AEs were mild to moderate. Most AE profiles were similar before and after the drug transition, the exceptions being anemia and menstrual disorder after 6 months. The number of subjects with anemia gradually increased from 10 (6.9%) at baseline to a peak of 29 (20.0%) at 6 months ($p < 0.001$) and increased to 22 (15.2%) at 12 months ($p = 0.024$).

Among the 112 transitioned female patients, 13 (15.7% of the 83 in reproductive age) patients experienced menstrual disorders, and the most common symptoms were menorrhagia and metrorrhagia. The number of subjects with menstrual disorder gradually increased from one (1.3%) to nine (11.4%) ($p < 0.001$) at month 6. Four women experienced severe anemia after 2–4 months of transition that was felt to be due to menorrhagia associated with the new therapy. The macitentan was discontinued on the patients' request and iron supplementation was instituted for at least 3 months. No other patients discontinued the macitentan due to AEs.

Clinical Outcomes

WHO FC, 6 MWD, and NT-proBNP assessments at baseline and months 6 and 12 are displayed in **Table 3**. From baseline to month 6, WHO FC stayed unchanged in 100 patients (69%) and ameliorated in 34 patients (23.4%; **Figure 2A**). Nine patients

TABLE 2 | Adverse events.

	Baseline	Month 1	Month 3	Month 6	Month 12	*p*-value
Patients with ≥ 1 AE, n (%)	67 (46.2)	73 (50.3)	73 (50.3)	65 (44.8)	64 (44.1)	0.431[ǂ]
Headache, n (%)	27 (18.6)	27 (18.6)	27 (18.6)	24 (16.6)	20 (13.8)	0.101[ǂ]
Peripheral edema, n (%)	32 (22.1)	33 (22.8)	35 (24.1)	29 (20.0)	28 (19.3)	0.467[ǂ]
Anemia, n (%)	10 (6.9)	14 (9.7)	20 (13.8)	29 (20.0)	22 (15.2)	0.004[ǂ], 0.451[¶], 0.06[§], <0.001[ɡ], 0.024[¥], 0.187[ч]
Menstrual disorder, n (%)*	1 (1.3)	1 (1.3)	5 (6.3)	9 (11.4)	4 (5.1)	<0.001[ǂ], 1.0[¶], 0.057[§], <0.001[ɡ], 0.15 [¥]
ALT, mean ± SD	18.3 ± 16.2	18.6 ± 16.2	19.7 ± 23.6	15.8 ± 12.27	15.9 ± 12.8	
AST, mean ± SD	22.9 ± 13.34	21.1 ± 13.1	26.3 ± 24.9	24.9 ± 22.6	25.1 ± 18.9	
ALT > ULN, n (%)	7 (4.8)	10 (6.9)	11 (7.6)	5 (3.4)	4 (2.8)	0.070[ǂ]
ALT > 3ULN, n (%)	1 (0.7)	0 (0)	1 (0.7)	0 (0)	0 (0)	0.558[ǂ]
AST > ULN, n (%)	13 (9.0)	10 (6.9)	8 (5.5)	16 (11)	7 (4.8)	0.051[ǂ]
AST > 3ULN, n (%)	1 (0.7)	0 (0)	2 (1.4)	3 (2.1)	1 (0.7)	0.316[ǂ]

**The proportion was calculated only in women of reproductive age (n = 79). [ǂ]Five groups (Baseline, Month 1, Month 3, Month 6, and Month 12) of data were compared. Once the difference (p < 0.05) was significant, other groups were compared with baseline. [¶]Month 1 versus Baseline, [§]Month 3 versus Baseline, [ɡ]Month 6 versus Baseline, [¥]Month 12 versus Baseline, [ч]Month 6 versus Month 12. AE: adverse event; ALT: alanine aminotransferase; AST: aspartate aminotransferase; SD: standard deviation; ULN: normalized upper limit.*

TABLE 3 | Parameters at the baseline and months 6 and 12 assessments.

	Baseline	Month 6	Month 12	*p*-value
Echocardiography				
RA (mm)[ǂ]	40.33 ± 8.54	39.35 ± 8.90	39.09 ± 8.81	0.114[§]
RV (mm)[ǂ]	41.61 ± 9.06	40.23 ± 10.44	39.31 ± 9.42	0.006[§], 0.056[ɡ], 0.003[ч]
LA (mm)[ǂ]	32.17 ± 6.94	32.69 ± 6.21	33.18 ± 6.03	0.155[§]
LVEDD (mm)[ǂ]	40.39 ± 8.04	41.15 ± 7.50	40.98 ± 7.82	0.494[§]
PA (mm)[ǂ]	31.31 ± 5.91	31.23 ± 5.86	30.97 ± 5.54	0.802[§]
sPAP (mmHg)[ǂ]	86.10 ± 26.56	74.72 ± 21.23	77.14 ± 22.69	<0.001[§], <0.001[ɡ], <0.001[ч], 0.566[¥]
TAPSE (mm)	13.98 ± 2.91[ǂ]	14.00 (12.00–17.75)[¶]	14.5 (13.00–17.75)[¶]	0.001[§], 0.226[ɡ], 0.001[ч]
S′ (cm/s)[¶]	11.00 (10.00–13.00)	13.00 (11.00–14.00)	13.00 (11.00–15.00)	<0.001[§], 0.004[ɡ], <0.001[ч], 0.102[¥]
RVFAC[ǂ]	33.12 ± 7.65%	34.59 ± 7.81%	34.61 ± 8.06%	0.023[§], 0.020[ɡ], 0.022[ч], 1.0[¥]
6 MWD (meters)[ǂ]	424.74 ± 82.34	435.92 ± 75.03	440.59 ± 77.57	0.004[§], 0.016[ɡ], 0.004[ч], 0.714[¥]
NT-proBNP (pg/ml)[¶]	443.00 (140.90–1,143.5)	253.90 (110.0–1,009.00)	234.80 (108.95–939.50)	<0.001[§], 0.045[ɡ], <0.001[ч], 0.005[¥]
S_pO_2	92.14	92.29	93.08	0.798
WHO FC, n (%)				<0.001[§], 0.046[ɡ], 0.008[ч], 0.499[¥]
I	0	9 (6.2)	13 (9.0)	
II	85 (58.6)	90 (62.1)	90 (62.1)	
III	60 (41.4)	46 (31.7)	42 (29.0)	
Quality of life[ǂ]	74.10 ± 6.74	-	79.76 ± 13.66	<0.001[§]
REVEAL score, n (%)				<0.001[§], <0.001[ɡ], <0.001[ч], 1.00[¥]
≤6	13 (9.0)	33 (22.8)	37 (25.5)	
7–8	42 (28.9)	58 (40.0)	51 (35.2)	
≥9	90 (62.1)	54 (37.2)	57 (39.3)	

[ǂ]Data are described as means and standard deviations (normal or approximately normal distribution). [¶]Data are described as medians and interquartile ranges (non-normal distribution). [§]Three groups (Baseline, Month 6, and Month 12) of data were compared. Once the difference (p < 0.05) was significant, other groups were compared with baseline. [ɡ]Month 6 versus Baseline, [ч]Month 12 versus Baseline, [¥]Month 6 versus Month 12 (only when months 6 and 12 compared with baseline both were statistically significant, the two groups would be compared.). RA: right atrium; RV: right ventricle; LA: left atrium, LVEDD: left ventricular end-diastolic dimension; PA: pulmonary artery; TAPSE: tricuspid annular plane systolic excursion; S': systolic velocity of the tricuspid annulus; RVFAC: right ventricular area change rate; 6 MWD: 6-min walking distance; NT-proBNP, N-terminal pro-brain natriuretic peptide; S_pO_2: pulse oxygen saturation; WHO FC: world health organization functional class.

ameliorated from WHO FC II to FC I, and 25 (17.2%) from WHO FC III to FC II, whereas 11 patients deteriorated from WHO FC II to FC III. From baseline to month 12, WHO FC stayed unchanged in 93 patients (64.1%) and ameliorated in 41 patients (28.3%; **Figure 2B**). Twelve patients ameliorated from WHO FC II to FC I, 28 from WHO FC III to FC II, and one from WHO FC III to FC I, whereas 11 patients declined from WHO FC II to FC III. From months 6–12, WHO FC stayed unchanged in 137 patients (94.5%), and four patients each ameliorated from WHO FC II to FC I and from WHO FC III to FC II.

The 6 MWD improved from 424.74 ± 82.34 m at baseline to 435.92 ± 75.03 m at month 6 ($p = 0.016$) and 440.59 ± 77.57 m at month 12 ($p = 0.004$). The median change in 6 MWD from months 6 to 12 was 1.00 m (IQR, −29.50 to 23.50 m; $p = 0.714$). In 101 (69.7%) patients, 6 MWD was maintained, and in 33 (22.8%) patients it was increased at month 6 (**Figure 3**). Eleven patients showed a decrease in 6 MWD from baseline at month 6. In 136 patients (93.8%), 6 MWD was either maintained or increased at month 12 (**Figure 3**). Nine patients showed a decrease in 6 MWD from baseline to month 12. Compared with month 6, only two patients showed a decrease in 6 MWD at month 12, others remained stable or increased. There was a clinically significant change in NT-proBNP levels from baseline to months 6 and 12 after

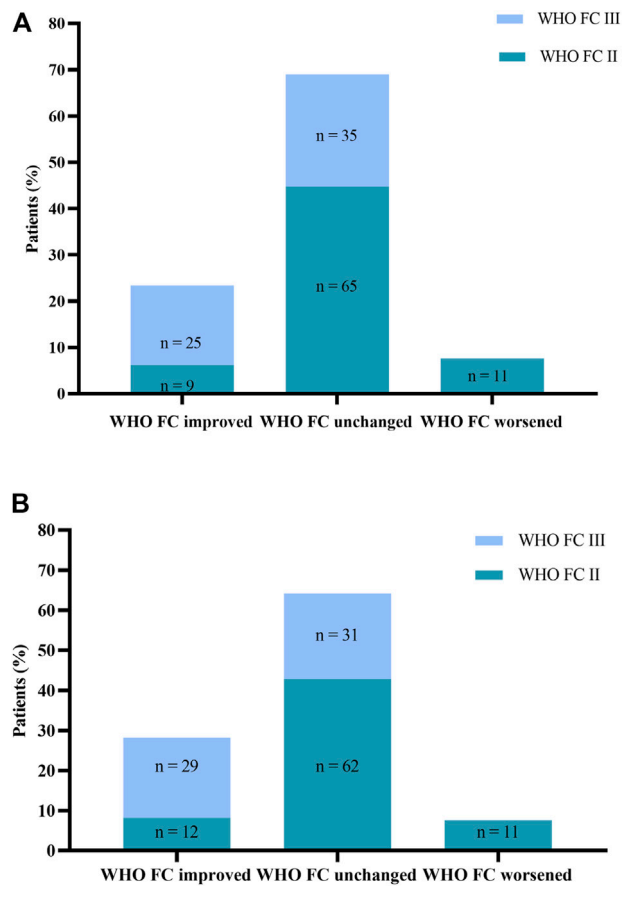

FIGURE 2 | (A) Change in World Health Organization Functional Class (WHO FC). from baseline to Month 6. **(B)** Change in WHO FC from baseline to Month 12.

the transition, the median decreased from baseline of 443.00 pg/ml (IQR, 140.90–1,143.5 pg/ml) to 253.90 pg/ml (IQR, 110.0–1,009.00 pg/ml; $p = 0.045$) and 234.80 pg/ml (IQR, 108.95–939.50 pg/ml; $p < 0.001$), respectively. The difference between month 6 and month 12 remained statistically significant ($p = 0.005$).

Regarding echocardiography parameters, no significant differences were found between diameters of the right atrium (RA), left atrium (LA), and left ventricular end-diastolic dimension (LVEDD) and pulmonary artery (PA) width. However, the basal diameter of the right ventricle (RV) was reduced from 41.61 ± 9.06 mm to 39.31 ± 9.42 mm ($p = 0.003$) at month 12, and sPAP was lowered from 86.10 ± 26.56 mmHg to 74.72 ± 21.23 mmHg ($p < 0.001$) at month 6 and to 77.14 ± 22.69 mmHg ($p < 0.001$) at month 12. TAPSE, S′, and RVFAC were used for the evaluation of RV function. TAPSE increased from 13.98 ± 2.91 mm to 14.5 (13.00–17.75) mm ($p = 0.001$) at month 12, the median change S′ from 11.00 (10.00–13.00) cm/s to 13.00 (11.00–14.00) cm/s ($p = 0.004$) at month 6 and to 13.00 (11.00–15.00) cm/s ($P < 0.001$) at month 12, and RVFAC from 33.12 ± 7.65%–34.59 ± 7.81% ($p = 0.020$) at month 6 and to 34.61 ± 8.06% ($p = 0.022$) at month 12, respectively. Of note, for those differences observed between baseline and months 6 or 12, no significant difference was found between months 6 and 12. The data on TAPSE, S' and RVFAC was only available in 75 patients as most echocardiography did not focus on the parameters of RV.

The REVEAL 2.0 calculator was used to assess risk stratification at baseline and months 6 and 12 (**Figure 4**). At baseline, there were 13 patients with REVEAL risk scores ≤ 6, 42 patients with REVEAL risk scores between 7 and 8, and 90 patients with REVEAL risk scores ≥ 9. There were 33 patients with REVEAL risk scores ≤ 6 at 6 months after transition and 37 patients 12 months after the transition. There were 58 patients with REVEAL risk scores between 7 and 8 at 6 months after transition and 51 patients 12 months after the transition. There were 54 patients with REVEAL risk scores ≥ 9 at 6 months after transition and 57 at 12 months after the transition.

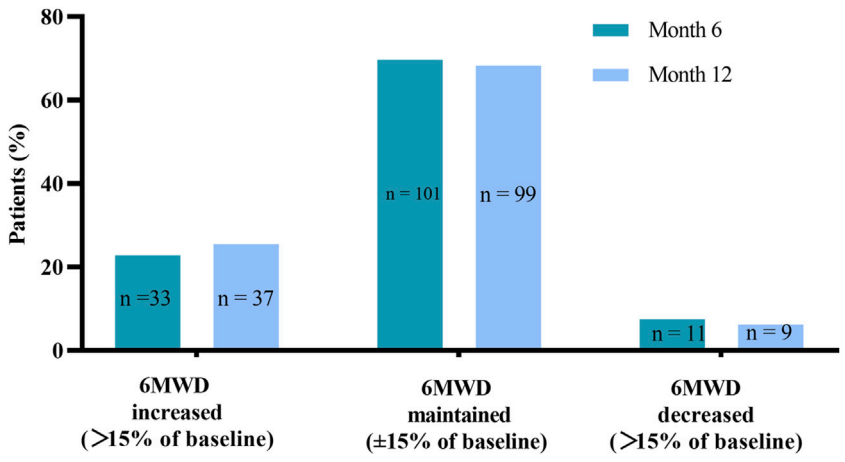

FIGURE 3 | Change in 6-min walk distance (6 MWD) from baseline to Month 6 and Month 12.

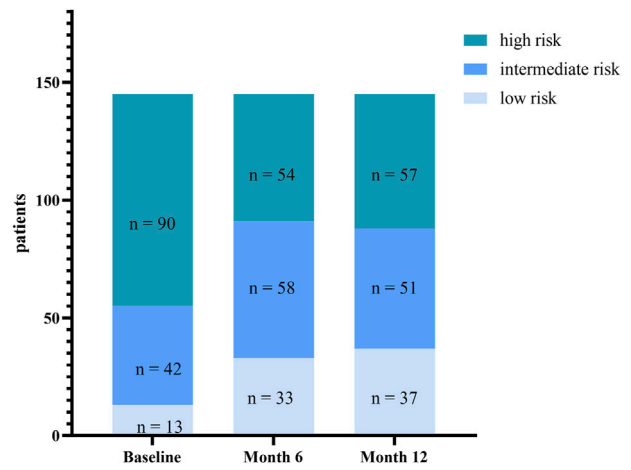

FIGURE 4 | Change in risk stratification (assessed by REVEAL score 2.0) from baseline to Month 6 and Month 12.

Satisfaction and QOL

QOL scale was collected from 157 patients to assess treatment satisfaction. The data were not available for 12 patients who discontinued macitentan before month 12. QOL questionnaire contained therapeutic effects, the convenience of taking, impact on routine activities, overall satisfaction, and AEs (Lopez et al., 2021). The QOL scale score was 74.10 ± 6.74 at baseline and 79.76 ± 13.66 at month 12 ($p < 0.001$). There were no significant differences in items of AEs and convenience between time points, but there was a significant improvement in the activities of daily living and on the overall satisfaction.

DISCUSSION

The current evidence on the safety and efficacy of drug transition from one PAH-specific therapy to another includes the transition involving prostacyclin and its analogs (Hoeper et al., 2002; Gomberg-Maitland et al., 2005; Ewert et al., 2007; Rubenfire et al., 2007; Sitbon et al., 2007; Shlobin et al., 2012; Bourge et al., 2013; Minai et al., 2013; Shapiro et al., 2013; Tamura et al., 2013; Frantz et al., 2014; Frantz et al., 2015; Fukumoto et al., 2015; Lichtblau et al., 2015; Chakinala et al., 2017; Frost et al., 2019) or PDE-5i (Tay et al., 2008; Verlinden et al., 2020), but minimal data on the switch between ERA. A study reported that the switch from bosentan or sitaxentan to ambrisentan in patients with liver abnormalities is safe and can offer clinical improvement (McGoon et al., 2009). Two studies found that switching from sitaxentan to bosentan or ambrisentan is safe and tolerable as monitored by NT-proBNP levels, echocardiography, hemodynamic parameters, and side effects (Safdar, 2011; Fox et al., 2013). An observational study in 2018 showed that most PAH patients taking bosentan showed good safety and tolerance after replacement with ambrisentan or macitentan (Dawson et al., 2018). Moreover, a retrospective study in 2019 showed that after replacing bosentan or ambrisentan with macitentan, 6 MWD was increased and right ventricular function was improved in most patients (Cadenas-Menéndez et al., 2019). However, to date, no studies have especially explored the replacement of ambrisentan with macitentan. Thus, the evidence for the safety and efficacy of the transition from ambrisentan to macitentan in PAH patients needs to be considered.

In our study of 157 patients, the transition from ambrisentan to macitentan was successfully achieved in most patients, was well-tolerated, and without deterioration of clinical status. All enrolled patients ($n = 157$) continued to receive macitentan or at least 1 month after discontinuation of ambrisentan, and 145 patients complied with all prespecified criteria for the sustained switch. After study completion, 148 patients continued with macitentan. Patients were carefully selected and assessed throughout the study to prevent disease deterioration and to monitor for AEs.

Although most patients underwent some AEs, there were no differences in the overall incidence of AEs before and after the switch, but we need to take cautious of anemia and menstrual disorder in females. When taking ambrisentan, most AEs were peripheral edema and headache. After the transition, the incidence of peripheral edema and headache was not significantly changed, although there was a trend to less headache. Anemia and menstrual disorder, however, increased significantly. The proportion of patients with anemia at month 6 was higher in our study than that in the SERAPHIN trial (Pulido et al., 2013), and the most severe anemia occurred in females of reproductive age. Such anemia may not only be directly caused by macitentan but indirectly caused by metromenorrhagia associated with macitentan. Four female patients experienced severe anemia, shown by a decline in hemoglobin to 73, 75, 78, and 83 g/L, with associated menstrual disorder. They had to discontinue macitentan and receive iron therapy. After at least three months of therapy, hemoglobin normalized, and the menstrual disorder gradually recovered. Three of the women that stopped the macitentan did not have a recurrence of anemia after re-introduction of the macitentan and they remained stable. One of the women did not wish to restart the macitentan therapy. Although the mechanism by which macitentan causes anemia is unknown, our experience suggests that anemia can be associated with menstrual disorders in women. In addition, anemia can aggravate the symptoms of palpitation, chest tightness, and shortness of breath, resulting in patients' noncompliance for prescribe. Thus, patients must monitor hemoglobin monthly. Once an abnormality is found, patients need to contact the doctors to evaluate it in time.

Regarding clinical stability, we used a strict definition and included multiple clinical parameters. Although several patients experienced a deterioration in WHO FC, 6 MWD, and NT-proBNP, most patients demonstrated stable or improved 6 MWD and WHO FC and thus, transitioned successfully. We observed 36 patients who moved from REVEAL 2.0 high-risk score to intermediate or low risk after the transition, and 33

patients maintained their condition 12 months later. Six months after the transition, the number of PAH patients (n = 91, 62.8%) in the low- or intermediate-risk groups reached a peak. Risk stratification score ameliorated, which indicates a better prognosis in our cohort of PAH patients. But our study was not professionally designed or powered to demonstrate superiority or non-inferiority over ambrisentan because of lacking a control group. The improvement observed may not be strongly reasoned to suggest that macitentan has better efficacy than ambrisentan. However, it is only suggested that patients taking macitentan do have a better effect, at least better than the baseline. Eleven patients experienced a worsening from baseline to month 6 in 6 MWD and nine patients at month 12. Eleven patients underwent deterioration in WHO FC, but none declined to WHO FC IV. The cause for the clinical worsening in a small proportion of patients was likely multifaceted and may include progressive deterioration of PAH together with the initiation of intravenous prostacyclin and the sudden withdrawal of the medication.

The hemodynamic parameters as assessed by echocardiography were comparable between pre-and post-transition, and there were no marked changes in the diameters of the RA, LA, LVEDD, and PA. However, we did observe significant decreases in the diameter of the RV and sPAP as well as increases in TAPSE, S′, and RVFAC from baseline to post-transition. These echo markers indicate improvement in RV function and may provide additional prognostic information in PAH (Augustine et al., 2018). The improvements in the mean diameter of RV and TAPSE were seen after 12 months, whereas the mean S' and RVFAC significantly improved after 6 months. The improvement after transition may be attributable to optimized drug exposure by eliminating the drug interaction of ambrisentan or the possible advantage of dual ETA and ETB receptor antagonism. Nonetheless, these echocardiographic parameters were only obtained from 75 patients (52%), which may lead to bias due to the incomplete data and should only be considered exploratory in nature, and requires further validation with additional studies. Invasive hemodynamic measurements before and after the switch would also lend further strength to our data.

Treatment with macitentan may offer several benefits over ambrisentan therapy for Chinese patients. In China, although treatment decisions are based on the PAH guidelines, they are also largely influenced by patients' financial situations (Zhang et al., 2011; Zhai et al., 2017). Both ambrisentan and macitentan are expensive in China. However, on April 1st, 2020, macitentan has been added to the list of medical insurance in Hunan province. After being reimbursed by 60–70% by the medical insurance, monthly out-of-pocket expenses of macitentan are about $187–249, while that of ambrisentan is $624 due to the withdrawal of medical insurance for the latter. Thus, ambrisentan therapy results in a substantial pill burden, which can increase non-adherence. Overall, the transition from ambrisentan to macitentan results in a cost-saving of approximately $14.5 per day. Hence, the transition enhanced overall treatment satisfaction and life happiness (QOL).

Our study is the first clinical study to demonstrate that converting from ambrisentan to macitentan can be completed safely in clinically stable patients. Limitations of this study are the open-label nature of the trial with no placebo or active comparator. In addition, the patient-reported QOL and treatment administration time data were inherently subjective, and long-term data were limited by missing data and might have been affected by completer bias of patients who may have discontinued the trial for reasons such as treatment dissatisfaction. As such, the clinical correlation of observed changes is not fully elucidated and the data should be cautiously interpreted. Despite these limitations, this study provides useful insight into the safety and efficacy of converting PAH patients from ambrisentan to macitentan therapy in a real-world clinical setting. The strengths of the study include the large size of the cohort, strict inclusion criteria, and 12-month follow-up.

CONCLUSION

In summary, we have reported the sustained transition of a carefully selected cohort from ambrisentan to macitentan in PAH patients. These data indicate that the transition is safe and compared with baseline, the switch to macitentan therapy was associated with improvements in exercise capacity, NT-pro BNP, and right ventricular function (assessed by echocardiography) in the majority of PAH patients. However, the increase in the incidence of menstrual disorder and anemia requires careful monitoring and may lead to further intervention in younger women. We found that macitentan was associated with good compliance and maybe a good alternative to ambrisentan if the switch is required due to side effects or accessibility issues.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the ethics committee at Second Xiangya Hospital of Central South University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

YC and JLi designed the study. YC, JL and JC performed the trial and analyzed the data, and YC wrote the manuscript. WC and ZL

were responsible for data acquisition. JLi was involved in the data analysis and interpretation. EK participated in critical revisions of the manuscript. All of the authors have read and approved the final manuscript.

FUNDING

The work was supported by National Natural Science Foundation of China. Grant Award Number: 81870233.

REFERENCES

Augustine, D. X., Coates-Bradshaw, L. D., Willis, J., Harkness, A., Ring, L., Grapsa, J., et al. (2018). Echocardiographic Assessment of Pulmonary Hypertension: a Guideline Protocol from the British Society of Echocardiography. *Echo Res. Pract.* 5, G11–G24. doi:10.1530/ERP-17-0071

Benza, R. L., Gomberg-Maitland, M., Demarco, T., Frost, A. E., Torbicki, A., Langleben, D., et al. (2015a). Endothelin-1 Pathway Polymorphisms and Outcomes in Pulmonary Arterial Hypertension. *Am. J. Respir. Crit. Care Med.* 192, 1345–1354. doi:10.1164/rccm.201501-0196OC

Benza, R. L., Miller, D. P., Foreman, A. J., Frost, A. E., Badesch, D. B., Benton, W. W., et al. (2015b). Prognostic Implications of Serial Risk Score Assessments in Patients with Pulmonary Arterial Hypertension: a Registry to Evaluate Early and Long-Term Pulmonary Arterial Hypertension Disease Management (REVEAL) Analysis. *J. Heart Lung Transpl.* 34, 356–361. doi:10.1016/j.healun.2014.09.016

Boucherat, O., Vitry, G., Trinh, I., Paulin, R., Provencher, S., and Bonnet, S. (2017). The Cancer Theory of Pulmonary Arterial Hypertension. *Pulm. Circ.* 7, 285–299. doi:10.1177/2045893217701438

Bourge, R. C., Tapson, V. F., Safdar, Z., Benza, R. L., Channick, R. N., Rosenzweig, E. B., et al. (2013). Rapid Transition from Inhaled Iloprost to Inhaled Treprostinil in Patients with Pulmonary Arterial Hypertension. *Cardiovasc. Ther.* 31, 38–44. doi:10.1111/1755-5922.12008

Cadenas-Menéndez, S., Álvarez Vega, P., Oterino Manzanas, A., Alonso Lecue, P., Roig Figueroa, V., Bedate Díaz, P., et al. (2019). Evolution of Patients with Pulmonary Arterial Hypertension Starting Macitentan after the Discontinuation of Other Endothelin-Receptor Antagonists: Results of a Retrospective Study. *Am. J. Cardiovasc. Drugs* 20, 481–487. doi:10.1007/s40256-019-00392-y

Chakinala, M. M., Feldman, J. P., Rischard, F., Mathier, M., Broderick, M., Leedom, N., et al. (2017). Transition from Parenteral to Oral Treprostinil in Pulmonary Arterial Hypertension. *J. Heart Lung Transpl.* 36, 193–201. doi:10.1016/j.healun.2016.06.019

Dai, Y., Chen, X., Song, X., Chen, X., Ma, W., Lin, J., et al. (2019). Immunotherapy of Endothelin-1 Receptor Type A for Pulmonary Arterial Hypertension. *J. Am. Coll. Cardiol.* 73, 2567–2580. doi:10.1016/j.jacc.2019.02.067

Dawson, A., Reddecliffe, S., Coghlan, C., Schreiber, B. E., and Coghlan, J. G. (2018). Safety and Feasibility Audit of a home-based Drug-Transitioning Approach for Patients with Pulmonary Arterial Hypertension: an Observational Study. *Eur. J. Cardiovasc. Nurs.* 17, 612–618. doi:10.1177/1474515118769534

Ewert, R., Opitz, C. F., Wensel, R., Winkler, J., Halank, M., and Felix, S. B. (2007). Continuous Intravenous Iloprost to Revert Treatment Failure of First-Line Inhaled Iloprost Therapy in Patients with Idiopathic Pulmonary Arterial Hypertension. *Clin. Res. Cardiol.* 96, 211–217. doi:10.1007/s00392-007-0490-3

Fox, B., Langleben, D., Hirsch, A. M., Schlesinger, R. D., Eisenberg, M. J., Joyal, D., et al. (2013). Hemodynamic Stability after Transitioning between Endothelin Receptor Antagonists in Patients with Pulmonary Arterial Hypertension. *Can. J. Cardiol.* 29, 672–677. doi:10.1016/j.cjca.2012.05.013

Frantz, R. P., Durst, L., Burger, C. D., Oudiz, R. J., Bourge, R. C., Franco, V., et al. (2014). Conversion from Sildenafil to Tadalafil: Results from the Sildenafil to Tadalafil in Pulmonary Arterial Hypertension (SITAR) Study. *J. Cardiovasc. Pharmacol. Ther.* 19, 550–557. doi:10.1177/1074248414528066

Frantz, R. P., Schilz, R. J., Chakinala, M. M., Badesch, D. B., Frost, A. E., Mclaughlin, V. V., et al. (2015). Hospitalization and Survival in Patients Using Epoprostenol for Injection in the PROSPECT Observational Study. *Chest* 147, 484–494. doi:10.1378/chest.14-1004

Frost, A., Janmohamed, M., Fritz, J. S., Mcconnell, J. W., Poch, D., Fortin, T. A., et al. (2019). Safety and Tolerability of Transition from Inhaled Treprostinil to Oral Selexipag in Pulmonary Arterial Hypertension: Results from the TRANSIT-1 Study. *J. Heart Lung Transpl.* 38, 43–50. doi:10.1016/j.healun.2018.09.003

Galiè, N., Jansa, P., Pulido, T., Channick, R. N., Delcroix, M., Ghofrani, H. A., et al. (2017). SERAPHIN Haemodynamic Substudy: the Effect of the Dual Endothelin Receptor Antagonist Macitentan on Haemodynamic Parameters and NT-proBNP Levels and Their Association with Disease Progression in Patients with Pulmonary Arterial Hypertension. *Eur. Heart J.* 38, 1147–1155. doi:10.1093/eurheartj/ehx025

Galiè, N., Olschewski, H., Oudiz, R. J., Torres, F., Frost, A., Ghofrani, H. A., et al. (2008). Ambrisentan for the Treatment of Pulmonary Arterial Hypertension: Results of the Ambrisentan in Pulmonary Arterial Hypertension, Randomized, Double-Blind, Placebo-Controlled, Multicenter, Efficacy (ARIES) Study 1 and 2. *Circulation* 117, 3010–3019. doi:10.1161/CIRCULATIONAHA.107.742510

Gatzoulis, M. A., Landzberg, M., Beghetti, M., Berger, R. M., Efficace, M., Gesang, S., et al. (2019). Evaluation of Macitentan in Patients with Eisenmenger Syndrome. *Circulation* 139, 51–63. doi:10.1161/CIRCULATIONAHA.118.033575

Gomberg-Maitland, M., Tapson, V. F., Benza, R. L., Mclaughlin, V. V., Krichman, A., Widlitz, A. C., et al. (2005). Transition from Intravenous Epoprostenol to Intravenous Treprostinil in Pulmonary Hypertension. *Am. J. Respir. Crit. Care Med.* 172, 1586–1589. doi:10.1164/rccm.200505-766OC

Hitzerd, E., Neuman, R. I., Broekhuizen, M., Simons, S. H. P., Schoenmakers, S., Reiss, I. K. M., et al. (2020). Transfer and Vascular Effect of Endothelin Receptor Antagonists in the Human Placenta. *Hypertension* 75, 877–884. doi:10.1161/HYPERTENSIONAHA.119.14183

Hoeper, M. M., Kramer, T., Pan, Z., Eichstaedt, C. A., Spiesshoefer, J., Benjamin, N., et al. (2017). Mortality in Pulmonary Arterial Hypertension: Prediction by the 2015 European Pulmonary Hypertension Guidelines Risk Stratification Model. *Eur. Respir. J.* 50, 1700740. doi:10.1183/13993003.00740-2017

Hoeper, M. M., Mclaughlin, V. V., Barberá, J. A., Frost, A. E., Ghofrani, H. A., Peacock, A. J., et al. (2016). Initial Combination Therapy with Ambrisentan and Tadalafil and Mortality in Patients with Pulmonary Arterial Hypertension: a Secondary Analysis of the Results from the Randomised, Controlled AMBITION Study. *Lancet Respir. Med.* 4, 894–901. doi:10.1016/S2213-2600(16)30307-1

Hoeper, M. M., Spiekerkoetter, E., Westerkamp, V., Gatzke, R., and Fabel, H. (2002). Intravenous Iloprost for Treatment Failure of Aerosolised Iloprost in Pulmonary Arterial Hypertension. *Eur. Respir. J.* 20, 339–343. doi:10.1183/09031936.02.02462001

Lichtblau, M., Harzheim, D., Ehlken, N., Marra, A., Pinado, F. P., Grünig, E., et al. (2015). Safety and Long-Term Efficacy of Transition from Sildenafil to Tadalafil Due to Side Effects in Patients with Pulmonary Arterial Hypertension. *Lung* 193, 105–112. doi:10.1007/s00408-014-9657-7

Lopez, J. L. T., Rabanales-Sotos, J., Hidalgo, M., Garcia, R. M. M., Martinez, C. L., and Abelian, G. B. (2021). Reliability and Validity of the Treatment Satisfaction with Medicines Questionnaire (SATMED-Q) in Persons with Arterial Hypertension. *Int. J. Environ. Res. Public Health* 18, 3212. doi:10.3390/ijerph18063212

Markova, S. M., De Marco, T., Bendjilali, N., Kobashigawa, E. A., Mefford, J., Sodhi, J., et al. (2013). Association of CYP2C9*2 with Bosentan-Induced Liver Injury. *Clin. Pharmacol. Ther.* 94, 678–686. doi:10.1038/clpt.2013.143

McGoon, M. D., Frost, A. E., Oudiz, R. J., Badesch, D. B., Galie, N., Olschewski, H., et al. (2009). Ambrisentan Therapy in Patients with Pulmonary Arterial Hypertension Who Discontinued Bosentan or Sitaxsentan Due to Liver Function Test Abnormalities. *Chest* 135, 122–129. doi:10.1378/chest.08-1028

Minai, O. A., Parambil, J., Dweik, R. A., Davila, G. H., Peterson, L., Rollins, K. D., et al. (2013). Impact of Switching from Epoprostenol to IV Treprostinil on Treatment Satisfaction and Quality of Life in Patients

with Pulmonary Hypertension. *Respir. Med.* 107, 458–465. doi:10.1016/j.rmed.2012.10.023

Oudiz, R. J., Galiè, N., Olschewski, H., Torres, F., Frost, A., Ghofrani, H. A., et al. (2009). Long-term Ambrisentan Therapy for the Treatment of Pulmonary Arterial Hypertension. *J. Am. Coll. Cardiol.* 54, 1971–1981. doi:10.1016/j.jacc.2009.07.033

Provencher, S., Paruchuru, P., Spezzi, A., Waterhouse, B., Gomberg-Maitland, M., and Gomberg-Maitland, M. (2015). Quality of Life, Safety and Efficacy Profile of Thermostable Flolan in Pulmonary Arterial Hypertension. *Plos One* 10, e0120657. doi:10.1371/journal.pone.0120657

Pulido, T., Adzerikho, I., Channick, R. N., Delcroix, M., Galiè, N., Ghofrani, H. A., et al. (2013). Macitentan and Morbidity and Mortality in Pulmonary Arterial Hypertension. *N. Engl. J. Med.* 369, 809–818. doi:10.1056/NEJMoa1213917

Rubenfire, M., Mclaughlin, V. V., Allen, R. P., Elliott, G., Park, M. H., Wade, M., et al. (2007). Transition from IV Epoprostenol to Subcutaneous Treprostinil in Pulmonary Arterial Hypertension: a Controlled Trial. *Chest* 132, 757–763. doi:10.1378/chest.06-2118

Safdar, Z. (2011). Effect of Transition from Sitaxsentan to Ambrisentan in Pulmonary Arterial Hypertension. *Vasc. Health Risk Manag.* 7, 119–124. doi:10.2147/VHRM.S15026

Shapiro, S., Traiger, G., Hill, W., Zhang, L., and Doran, A. K. (2013). Safety, Tolerability, and Efficacy of Overnight Switching from Sildenafil to Tadalafil in Patients with Pulmonary Arterial Hypertension. *Cardiovasc. Ther.* 31, 274–279. doi:10.1111/1755-5922.12038

Shlobin, O. A., Brown, A. W., Weir, N., Ahmad, S., Lemma, M., and Nathan, S. D. (2012). Transition of PH Patients from Sildenafil to Tadalafil: Feasibility and Practical Considerations. *Lung* 190, 573–578. doi:10.1007/s00408-012-9408-6

Sitbon, O., Manes, A., Jais, X., Pallazini, M., Humbert, M., Presotto, L., et al. (2007). Rapid Switch from Intravenous Epoprostenol to Intravenous Treprostinil in Patients with Pulmonary Arterial Hypertension. *J. Cardiovasc. Pharmacol.* 49, 1–5. doi:10.1097/FJC.0b013e31802b3184

Tamura, Y., Ono, T., Fukuda, K., Satoh, T., and Sasayama, S. (2013). Evaluation of a New Formulation of Epoprostenol Sodium in Japanese Patients with Pulmonary Arterial Hypertension (EPITOME4). *Adv. Ther.* 30, 459–471. doi:10.1007/s12325-013-0029-0

Tay, E. L., Geok-Mui, M. K., Poh-Hoon, M. C., and Yip, J. (2008). Sustained Benefit of Tadalafil in Patients with Pulmonary Arterial Hypertension with Prior Response to Sildenafil: a Case Series of 12 Patients. *Int. J. Cardiol.* 125, 416–417. doi:10.1016/j.ijcard.2007.01.046

Vachiéry, J. L., Galiè, N., Barberá, J. A., Frost, A. E., Ghofrani, H. A., Hoeper, M. M., et al. (2019). Initial Combination Therapy with Ambrisentan + tadalafil on Pulmonary Arterial Hypertension–related Hospitalization in the AMBITION Trial. *J. Heart Lung Transpl.* 38, 194–202. doi:10.1016/j.healun.2018.11.006

Verlinden, N. J., Benza, R. L., and Raina, A. (2020). Safety and Efficacy of Transitioning from the Combination of Bosentan and Sildenafil to Alternative Therapy in Patients with Pulmonary Arterial Hypertension. *Pulm. Circ.* 10, 2045894020945523. doi:10.1177/2045894020945523

Zhai, Z., Zhou, X., Zhang, S., Xie, W., Wan, J., Kuang, T., et al. (2017). The Impact and Financial burden of Pulmonary Arterial Hypertension on Patients and Caregivers: Results from a National Survey. *Medicine (Baltimore)* 96, e6783. doi:10.1097/MD.0000000000006783

Zhang, R., Dai, L. Z., Xie, W. P., Yu, Z. X., Wu, B. X., Pan, L., et al. (2011). Survival of Chinese Patients with Pulmonary Arterial Hypertension in the Modern Treatment Era. *Chest* 140, 301–309. doi:10.1378/chest.10-2327

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Chen, Luo, Chen, Kotlyar, Li, Chen and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Pharmacology

CORRECTION
published: 11 April 2022
doi: 10.3389/fphar.2022.891907

Corrigendum: The Transition From Ambrisentan to Macitentan in Patients With Pulmonary Arterial Hypertension: A Real-Word Prospective Study

*Yusi Chen [1], Jun Luo [1], Jingyuan Chen [1], Eugene Kotlyar [2], Zilu Li [1], Wenjie Chen [1] and Jiang Li [1]**

[1]Department of Cardiovascular Medicine, The Second Xiangya Hospital of Central South University, Changsha, China, [2]St. Vincent's Hospital, Sydney, NSW, Australia

Keywords: pulmonary arterial hypertension, macitentan, ambrisentan, endothelin receptor antagonist, real-word study

OPEN ACCESS

Edited and reviewed by:
Kevin Lu,
University of South Carolina,
United States

Correspondence:
Jiang Li
lijiangcs@csu.edu.cn

Specialty section:
This article was submitted to Drugs Outcomes Research and Policies, a section of the journal Frontiers in Pharmacology

Received: *08 March 2022*
Accepted: *15 March 2022*
Published: *11 April 2022*

Citation:
Chen Y, Luo J, Chen J, Kotlyar E, Li Z, Chen W and Li J (2022) Corrigendum: The Transition From Ambrisentan to Macitentan in Patients With Pulmonary Arterial Hypertension: A Real-Word Prospective Study. Front. Pharmacol. 13:891907. doi: 10.3389/fphar.2022.891907

A Corrigendum on

The Transition From Ambrisentan to Macitentan in Patients With Pulmonary Arterial Hypertension: A Real-Word Prospective Study
by Chen, Y., Luo, J., Chen, J., Kotlyar, E., Li, Z., Chen, W. and Li, J. (2022). Front Pharmacol. 12: 811700. doi:10.3389/fphar.2021.811700

In the original article, there was a mistake in **Table 1** and in the Section **Results**, **Patients' Clinical Characteristics** as published. "There were errors in the values for the 'Ambrisentan dose' as it was reported that one hundred and twenty-eight patients (88.3%) were receiving 10 mg once per day, whereas seventeen patients (11.7%) were receiving 5 mg once daily. This was corrected to sixty-one patients (42.1%) were receiving 10 mg once per day, and eighty-four patients (57.9%) were receiving 5 mg once daily." The corrected **Table 1** appears below and the text in the Section **Results**, **Patients' Clinical Characteristics** has been updated.

The authors apologize for this error and state that this does not change the scientific conclusions of the article in any way. The original article has been updated.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Chen, Luo, Chen, Kotlyar, Li, Chen and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

TABLE 1 | Baseline characteristics.

Characteristic	All patients (*N* = 145)
Age, years[a]	32.0 (26.0, 42.0)
Sex, n (%)	
Female	108 (74.5)
Male	37 (25.5)
Duration of ambrisentan treatment before transition, years[a]	1.5 (1.0, 2.9)
mPAP, mmHg[b,c]	61.4 ± 18.6
PVR, Wood[a,c]	14.4 (8.0, 22.1)
mRAP, mmHg[b,c]	10.0 ± 5.8
mRVP, mmHg[b,c]	41.7 ± 12.4
Etiology of PAH, n (%)	
Idiopathic	32 (22.1)
Drug or toxin-induced	4 (2.8)
Associated with connective tissue disease	21 (14.4)
Associated with congenital heart disease	88 (60.7)
Ambrisentan dose, mg, QD	
5 n (%)	84 (57.9)
10 n (%)	61 (42.1)
WHO FC, n (%)	
II	85 (58.6)
III	60 (41.4)
BMI (kg/m2) b	21.0 ± 3.2
PAH medications in addition to taking ambrisentan, n (%)	
Monotherapy	18 (12.4)
PDE-5i	122 (84.1)
Riociguat	4 (2.8)
Riociguat and PDE-5i	1 (0.7)

[a]*Data are described as medians and interquartile ranges (non-normal distribution).*
[b]*Data are described as means and standard deviations (normal or approximately normal distribution).*
[c]*Historical data from the most recent right heart catheterization before enrollment.*
mPAP, mean pulmonary arterial pressure; PVR, pulmonary vascular resistance; mRAP, mean right atrial pressure; mRVP, mean right ventricular pressure; QD, once per day; PAH, pulmonary arterial hypertension; WHO FC, World Health Organization functional class; BMI, body mass index; PDE-5i, phosphodiesterase type 5 inhibitor.

frontiers in Pharmacology

SYSTEMATIC REVIEW
published: 18 January 2022
doi: 10.3389/fphar.2021.789319

Comparison of Systemic Treatments for Metastatic Castration-Resistant Prostate Cancer After Docetaxel Failure: A Systematic Review and Network Meta-analysis

*Junru Chen†, Yaowen Zhang†, Xingming Zhang, Jinge Zhao, Yuchao Ni, Sha Zhu, Ben He, Jindong Dai, Zhipeng Wang, Zilin Wang, Jiayu Liang, Xudong Zhu, Pengfei Shen *, Hao Zeng * and Guangxi Sun **

Department of Urology, Institute of Urology, West China Hospital, Sichuan University, Chengdu, China

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Xiaohong Li,
University of Toledo, United States
Antonio Rozzi,
CHIC Compiégne-Noyon, France

***Correspondence:**
Pengfei Shen
cdhx510@163.com
Hao Zeng
zengh@scu.edu.cn
Guangxi Sun
sungx077@126.com

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Drugs Outcomes Research and Policies, a section of the journal Frontiers in Pharmacology

Received: 04 October 2021
Accepted: 14 December 2021
Published: 18 January 2022

Citation:
Chen J, Zhang Y, Zhang X, Zhao J, Ni Y, Zhu S, He B, Dai J, Wang Z, Wang Z, Liang J, Zhu X, Shen P, Zeng H and Sun G (2022) Comparison of Systemic Treatments for Metastatic Castration-Resistant Prostate Cancer After Docetaxel Failure: A Systematic Review and Network Meta-analysis. Front. Pharmacol. 12:789319. doi: 10.3389/fphar.2021.789319

Background: Lacking head-to-head trial, the optimal treatment for patients with metastatic castration-resistant prostate cancer (mCRPC) after docetaxel failure is unclear. This study is to compare the efficacy and safety of systemic treatments in patients who progressed after docetaxel to aid clinical decision-making.

Methods: Databases including MEDLINE, EMBASE, and the Cochrane Library were searched from inception to June 15th, 2021. The outcomes of interest include overall survival (OS), biochemical progression-free survival (bPFS), and serious adverse events (SAEs). The Cochrane risk of bias tools were used to assess study quality. Indirect comparisons of competing treatments were performed via Bayesian network meta-analysis.

Results: Five trials with 3,862 patients comparing four treatments (abiraterone, enzalutamide, cabazitaxel, and radium-223) were identified. All the four treatments were associated with improved OS and bPFS relative to best supportive care. Among them, enzalutamide (hazard ratio [HR] = 0.58, 95% credible interval [CrI]: 0.49–0.69) had the highest probability of ranking first in terms of OS, followed by cabazitaxel (HR = 0.70, 95% CrI: 0.59–0.83), radium-223 (HR = 0.71, 95% CrI: 0.56–0.90) and abiraterone (HR = 0.73, 95% CrI: 0.63–0.84). Similarly, enzalutamide (HR = 0.25, 95% CrI: 0.20–0.31) showed the greatest improvement of bPFS, followed by abiraterone (HR = 0.60, 95% CrI: 0.51–0.71) and cabazitaxel (HR = 0.75, 95% CrI: 0.63–0.89). In terms of safety, treatments ranked from the safest to the least safe were radium-223 (OR = 0.58, 95% CrI: 0.20–1.68), enzalutamide (OR = 0.80, 95% CrI: 0.28–2.29), abiraterone (OR = 0.94, 95% CrI: 0.39–2.27) and cabazitaxel (OR = 2.50, 95% CrI: 0.84–7.44).

Conclusion: For patients with mCRPC who progressed after docetaxel, enzalutamide may offer the most significant survival benefits and satisfying safety. Cabazitaxel is effective in post-docetaxel settings but associated with a high risk of SAEs. Although network meta-

analysis provides indirect comparisons and ranking probabilities, the results should be treated with caution as it cannot replace randomized direct comparison.

Systematic Review Registration: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020223040, identifier CRD42020223040.

Keywords: metastatic castration-resistant prostate cancer, network meta-analysis, docetaxel (DOC), enzalutamide (ENZ), abirateron, radium-223 (Ra) 223

1 INTRODUCTION

Prostate cancer (PCa) is the most common malignancy in men in America accounting for 26% of cancer diagnoses (Siegel et al., 2021). Although most patients with metastatic PCa are initially hormone-sensitive and controlled by androgen-deprivation therapy (ADT), diseases progression is inevitable and patients eventually develop castration-resistant diseases (Galletti et al., 2017). Metastatic castration-resistant prostate cancer (mCRPC) is highly aggressive with a median survival ranging from 17.5 to 34.7 months (Petrylak et al., 2004; Berthold et al., 2008; Beer et al., 2014; Ryan et al., 2015). Docetaxel combined with prednisone is the first to show survival benefits in patients with mCRPC and remains one of the standard first-line treatments for this setting (Cornford et al., 2021). However, most patients receiving docetaxel progress within 1 year (Petrylak et al., 2004; Berthold et al., 2008). Based on survival improvements compared to best supportive care, several regimens including abiraterone, enzalutamide, cabazitaxel, and radium-223 are recommended after docetaxel treatment failure (de Bono et al., 2010; de Bono et al., 2011; Fizazi et al., 2012; Scher et al., 2012; Parker et al., 2013; Sun et al., 2016).

Due to the lack of head-to-head trials comparing these drugs, the optimal treatment for patients with mCRPC who progress after docetaxel is unclear. As such, the current guidelines do not recommend one treatment over the others (Cornford et al., 2021; J Natl Compr Canc Netw, 2021). One preliminary analysis has attempted to compare these active treatments indirectly but was limited by a few number of included studies and the exclusion of radium-223 (Fryzek et al., 2018). In addition, the costs and treatment courses of these drugs vary widely (Pollard et al., 2017). This study aimed to compare the efficacy and safety of systemic treatments for mCRPC after upfront docetaxel failure to assist clinical practice.

2. MATERIALS AND METHODS

2.1 Search Strategy and Eligibility Criteria

The protocol of this study was developed following Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) reporting guideline (Tricco et al., 2018) and prospectively registered in PROSPERO (CRD42020223040). Bibliographic databases including MEDLINE (OVID interface), EMBASE (OVID interface), and the Cochrane Central Register of Controlled Trials were searched from inception to June 15th, 2021. The full search strategy was available in the protocol. The eligibility criteria included: 1) randomized controlled trials (RCTs) or cohorts; 2) patients who received first-line docetaxel for mCRPC and progressed; 3) interventions of interest were abiraterone, enzalutamide, cabazitaxel and radium-223; 4) comparators of interest were best supportive care (BSC) or active drugs; 5) studies reporting survival and safety outcomes. Prednisone alone and prednisone plus mitoxantrone were both regarded as BSC as a recent study showed similar survival outcomes for these two treatments (Green et al., 2015). Reviews, case reports, cohorts, study protocols, abstracts, and dose-escalation trials were excluded.

2.2 Study Selection and Data Extraction

Two investigators (JRC and YWZ) screened the titles and abstracts independently using Endnote X9. Full texts of potentially eligible studies were further evaluated to identify the final included studies. Data extraction was performed by two investigators (JRC and YWZ) and double-checked. Disagreements were reconciled by discussion or a third investigator (XMZ). The following data were extracted: study design, recruitment period, follow-up time, sample size, interventions, baseline characteristics, efficacy, and safety outcomes. For trials with multiple publications, the most recent data were extracted. The primary efficacy outcome of interest was overall survival (OS), defined as the time from randomization to death due to any cause. The secondary efficacy outcome was biochemical progression-free survival (bPFS), defined as the time from randomization to prostate-specific antigen (PSA) progression or death due to any cause, whichever occurred first. The safety outcome of interest was any serious adverse event (SAE).

2.3 Risk of Bias Assessment

Two investigators (JRC and YWZ) assessed the risk of bias of included studies independently. Disagreements were resolved through consensus. The following five domains were evaluated for RCTs according to the Cochrane framework (Sterne et al., 2019): randomization process, deviation from intended interventions, missing outcome data, measurement of the outcome, selection of reported results. The overall risk of bias of a trial was determined by the worst risk of bias in any of the domains. If multiple concerns were raised for one trial, it was judged as at high risk of bias overall.

2.4 Statistical Analysis

We performed the indirect comparisons using the Bayesian framework as recommended by the National Institute for Health and Care Excellence (NICE) of the United Kingdom (Shim et al., 2019). For survival outcomes, hazard ratios (HRs)

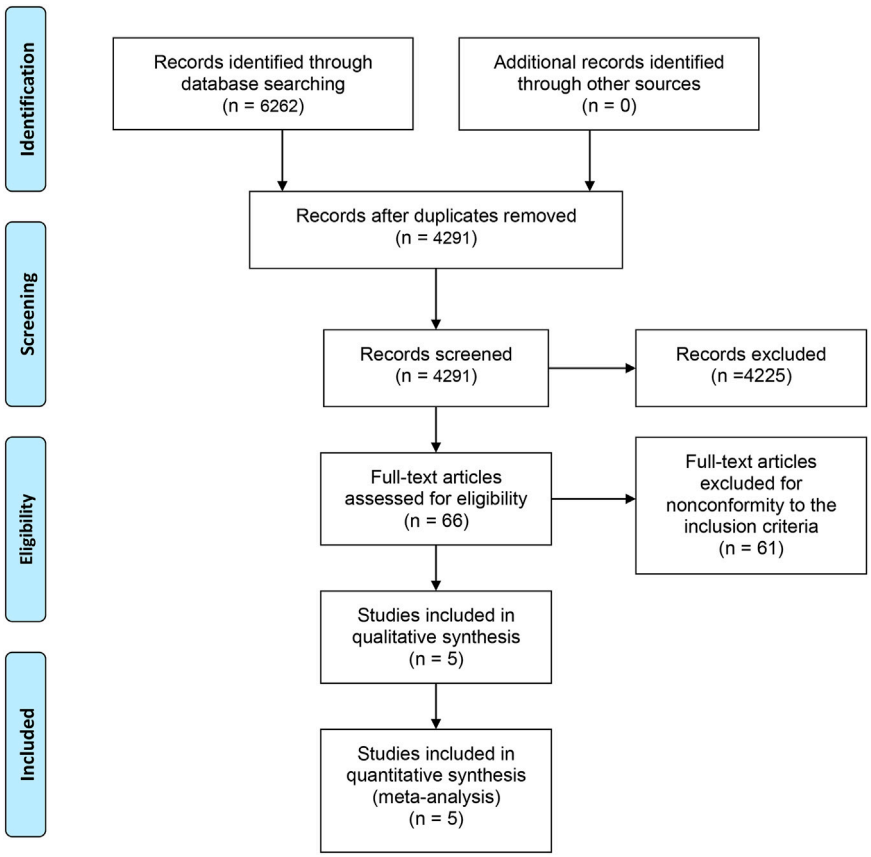

FIGURE 1 | PRISMA flowchart of study selection.

from individual trials were used to estimate overall HRs and corresponding 95% credible intervals (CrIs). The ALSYMPCA trial included patients with and without previous use of docetaxel, and only data from patients who had received docetaxel were used for analysis. For safety outcomes, the incidence of SAE in each treatment arm was used to estimate the overall odds ratios (ORs) with 95% CrIs. We fitted the consistency model and evaluated heterogeneity using I^2 statistic. A fixed-effect model was used if $I^2 > 50\%$. lMarkov chain Monte Carlo (MCMC) algorithms were applied to estimate treatment effects with 100,000 samples after a 10,000-sample burn-in. Treatment ranking probability was assessed via the surface under the cumulative ranking curve (SUCRA). SUCRA ranges from 0 to 1, with score 1 being the best (Salanti et al., 2011). Subgroup analyses were performed based on patient age, ECOG score, and the absence of visceral metastasis. All analyses were performed using the gemtc and rjags packages within R program.

3 RESULTS

3.1 Characteristics of Include Studies

A total of 6,262 studies were identified and screened via titles and abstracts. The flowchart showing the selection process is presented in **Figure 1**. Five trials involving 3,862 patients met our eligibility criteria and were included in this systematic review (de Bono et al., 2010; de Bono et al., 2011; Fizazi et al., 2012; Scher et al., 2012; Parker et al., 2013; Sun et al., 2016). All the included trials were multicenter phase 3 RCTs. The median sample size was 755 (range:214–1,199). The median follow-up duration was 13.7 months (range: 12.8–20.8). The network plot of treatment comparisons is shown in **Figure 2**.

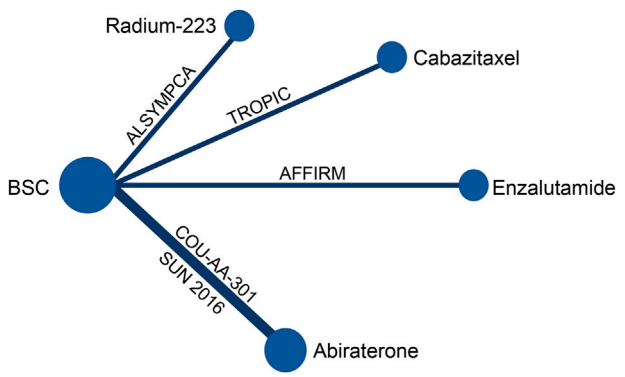

FIGURE 2 | Network plot of treatment comparisons.

TABLE 1 | Characteristics of included trials.

Trial	Study design	Recruitment period	Median follow-up (month)	Interventions	Sample size
COU-AA-301	Phase 3, multicenter, double-blind, RCT	May 2008 to Jul 2009	20.2	Abiraterone + prednisone vs. placebo + prednisone	797 vs. 398
Sun et al. (2016)	Phase 3, multicenter, double-blind, RCT	Aug 2012 to Nov 2013	12.9	Abiraterone + prednisone vs. placebo + prednisone	143 vs. 71
AFFIRM	Phase 3, multicenter, double-blind, RCT	Sep 2009 to Nov 2010	14.4	Enzalutamide vs. placebo	800 vs. 399
TROPIC	Phase 3, multicenter, open-label, RCT	Jan 2007 to Oct 2008	12.8	Cabazitaxel + prednisone vs. mitoxantrone + prednisone	378 vs. 377
ALSYMPCA	Phase 3, multicenter, open-label, RCT	Jun 2008 to Feb 2011	NA	Radium-223 vs. placebo	325[a] vs. 174[a]

Abbreviations: RCT, randomized controlled trial; NA, not available.
[a]Data of patients with previous docetaxel use.

TABLE 2 | Baseline characteristics of included patients.

Characteristic	COU-AA-301		Sun et al. (2016)		AFFIRM		TROPIC		ALSYMPCA[a]	
	Abi	Placebo	Abi	Placebo	Enza	Placebo	Caba	Placebo	Ra-223	Placebo
Median age (years)	69	69	Mean: 68.2	Mean: 67.7	69	69	68	67	71	71
Median PSA (ng/ml)	128.8	137.7	Mean: 800.0	Mean: 527.8	107.7	128.3	143.9	127.5	146	173
Gleason score<8	49%	46%	36%	28%	50%	48%	NA	NA	NA	NA
Gleason score ≥8	51%	54%	64%	72%	50%	52%	NA	NA	NA	NA
Bone metastasis	89%	90%	95%	94%	92%	92%	80%	87%	100%	100%
Lung metastasis	13%	11%	8%	13%	15%	15%	Visceral: 25%	Visceral: 25%	0%	0%
Liver metastasis	11%	8%	4%	3%	12%	9%			0%	0%
ECOG score<2	90%	89%	92%	93%	91%	92%	93%	91%	87%	87%
ECOG score≥2	10%	11%	8%	7%	9%	8%	7%	9%	13%	13%

Abbreviations: Abi, abiraterone; Caba, cabazitaxel; Enza: enzalutamide; NA, not available; Ra-223, radium-223.
[a]Data of overall population.

TABLE 3 | Risk of bias of included trials.

Trial	Randomisation process	Deviations from intended interventions	Missing outcome data	Measurement of the outcome	Selection of the reported result	Overall bias
COU-AA-301	Low	Low	Low	Low	Low	Low
Sun et al. (2016)	Low	Low	Low	Low	Low	Low
AFFIRM	Low	Low	Low	Low	Low	Low
TROPIC	Low	Some concerns	Low	Some concerns	Low	Some concerns
ALSYMPCA	Low	Low	Low	Some concerns	Some concerns	Some concerns

Two trials compared abiraterone + prednisone with prednisone + placebo (de Bono et al., 2011; Fizazi et al., 2012; Sun et al., 2016), one compared enzalutamide with placebo with no requirement of glucocorticoids (Scher et al., 2012), one compared cabazitaxel + prednisone with prednisone + mitoxantrone (de Bono et al., 2010), and one compared radium-223 with placebo (Parker et al., 2013). The ALSYMPCA trial only included patients with bone metastases and no visceral metastases (Parker et al., 2013). The characteristics of included trials and patients are summarized in **Tables 1**, **2**, respectively.

3.2 Risk of Bias

As shown in **Table 3**, the overall risk of bias was low in three trials (COU-AA-301, Sun et al., 2016 and AFFIRM) (de Bono et al., 2011; Scher et al., 2012; Sun et al., 2016), but some concerns have been raised for the other two trials (TROPIC and ALSYMPCA) (de Bono et al., 2010; Parker et al., 2013). Specifically, deviations from intended interventions might be a concern of bias for the TROPIC trial (de Bono et al., 2010), as patients and treating physicians were not masked to the treatment allocation and there was no information on whether there were deviations from

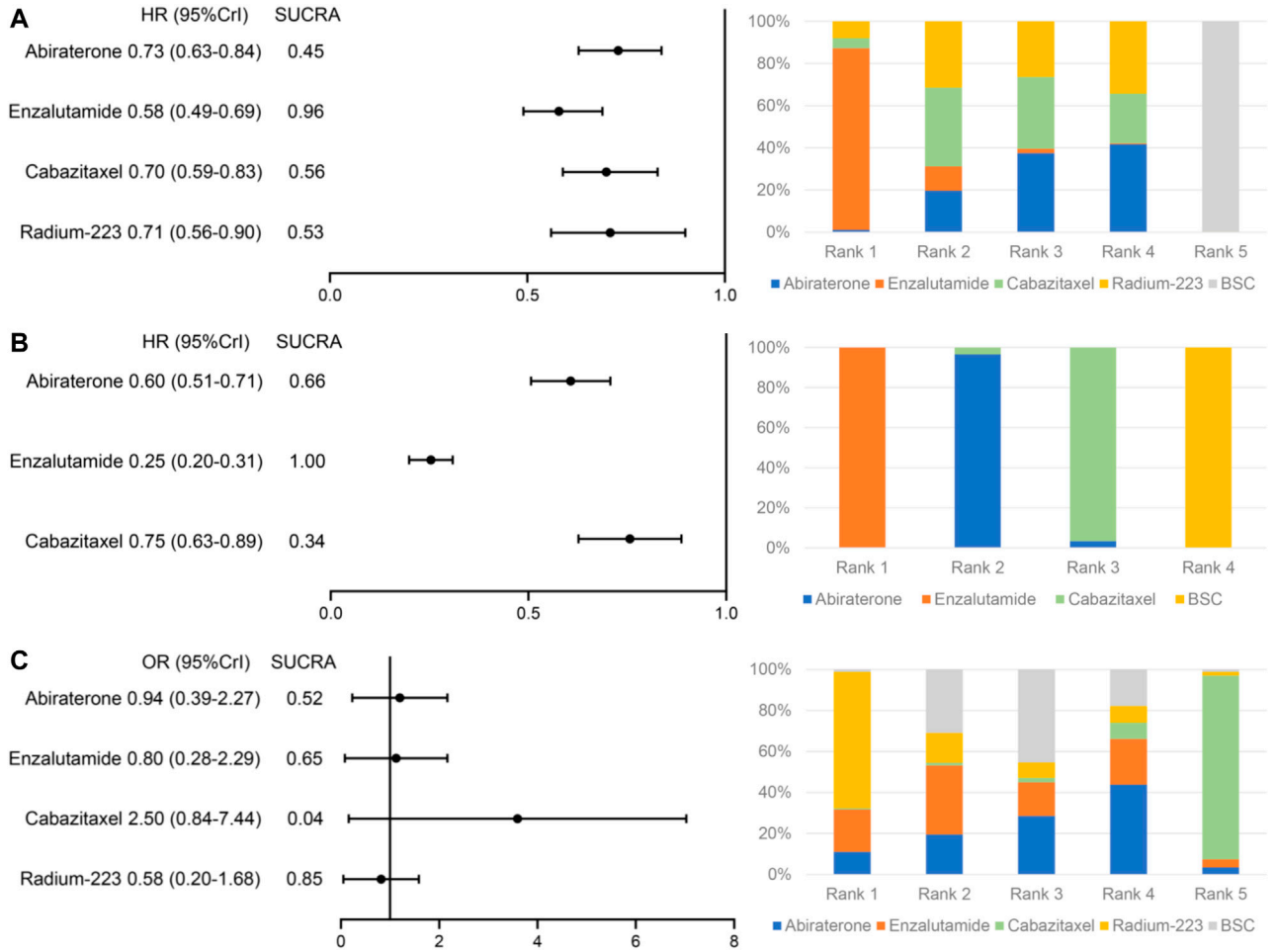

FIGURE 3 | Relative effects of systemic treatments compared to best supportive care and treatment ranking on **(A)** overall survival, **(B)** biochemical progression-free survival, and **(C)** serious adverse events.

TABLE 4 | Relative effect estimates for all pairwise treatment comparisons.

Outcomes	Comparator	Intervention				
OS, HR (95%CrI)		Abiraterone				
	Enzalutamide	1.30 (1.06–1.59)	Enzalutamide			
	Cabazitaxel	1.04 (0.83–1.30)	0.83 (0.65–1.06)	Cabazitaxel		
	Radium-223	1.05 (0.78–1.41)	0.82 (0.61–1.10)	0.99 (0.73–1.34)	Radium-223	
	BSC	0.73 (0.63–0.84)	0.58 (0.49–0.69)	0.70 (0.59–0.83)	0.71 (0.56–0.90)	BSC
bPFS, HR (95%CrI)		Abiraterone				
	Enzalutamide	2.40 (1.80–3.20)	Enzalutamide			
	Cabazitaxel	0.80 (0.63–1.02)	0.33 (0.25–0.44)	Cabazitaxel		
	BSC	0.60 (0.51–0.71)	0.25 (0.20–0.31)	0.75 (0.63–0.89)	BSC	
SAE, OR (95%CrI)		Abiraterone				
	Enzalutamide	1.10 (0.29–4.20)	Enzalutamide			
	Cabazitaxel	0.39 (0.09–1.62)	0.33 (0.07–1.51)	Cabazitaxel		
	Radium-223	1.60 (0.39–6.56)	1.40 (0.31–6.32)	1.72 (0.59–5.01)	Radium-223	
	BSC	0.94 (0.39–2.27)	0.80 (0.28–2.29)	2.50 (0.84–7.44)	0.58 (0.20–1.68)	BSC

intended intervention because of the trial context. There were concerns of bias regarding measurement of the outcome for the TROPIC and ALSYMPCA trial (de Bono et al., 2010; Parker et al., 2013). In these trials, the outcome assessors were unblinded to intervention status, which might influence the outcome assessment. Additionally, the selection of the reported result

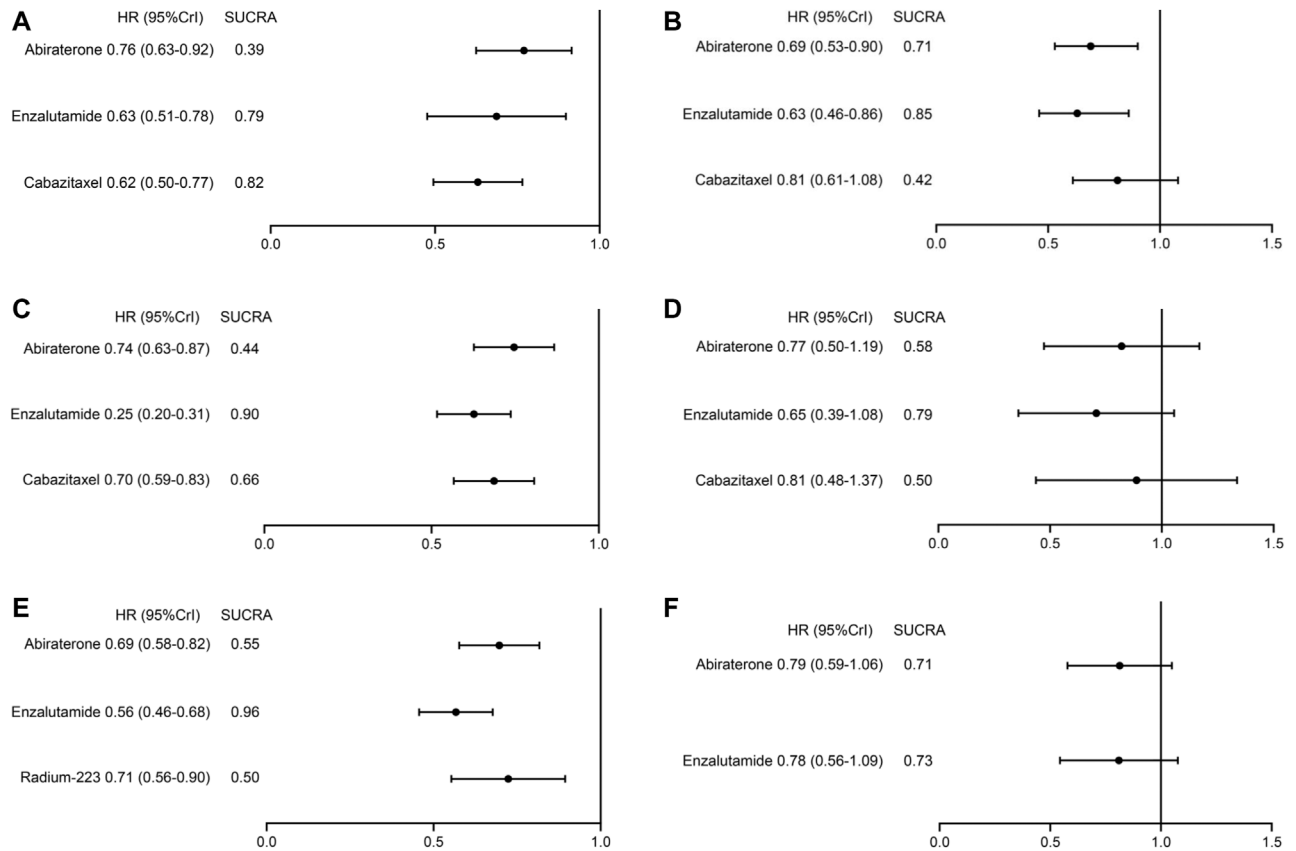

FIGURE 4 | Treatment ranking and relative effects of systemic treatments from subgroup analyses on overall survival. **(A)** patient age≥65 years, **(B)** patient age<65 years, **(C)** patients with an ECOG score<2, **(D)** patients with a ECOG score≥2, **(E)** patients without visceral metastases, **(F)** patients with visceral metastases.

raised concerns for the ALSYMPCA trial (Parker et al., 2013), as the protocol was finalized after interim analysis and unblinded outcome data were available for analysis.

3.3 Efficacy Outcomes

All the included studies reported outcomes of OS. Compared to BSC, abiraterone (HR = 0.73, 95% CrI: 0.63–0.84), enzalutamide (HR = 0.58, 95% CrI: 0.49–0.69), cabazitaxel (HR = 0.70, 95% CrI: 0.59–0.83), and radium-223 (HR = 0.71, 95% CrI: 0.56–0.90) showed significantly improved OS (**Figure 3A**). There was no significant heterogeneity ($I^2 = 0$). Furthermore, enzalutamide was associated with the highest probability of ranking first, followed by cabazitaxel, radium-223, and abiraterone, and the corresponding SUCRAs were 0.96, 0.56, 0.53, 0.45. The relative effect estimates for all pairwise treatment comparisons are summarized in **Table 4**.

The bPFS data of patients with prior docetaxel were not available in the ALSYMPCA trial. Thus, radium-223 was excluded from the treatment comparisons with respect to bPFS. The network meta-analysis showed obvious improvements in bPFS of abiraterone (HR = 0.60, 95% CrI: 0.51–0.71), enzalutamide (HR = 0.25, 95% CrI: 0.20–0.31) and cabazitaxel (HR = 0.75, 95% CrI: 0.63–0.89) relative to BSC (**Figure 3B**). No significant heterogeneity was observed (I^2 = 20%). Additionally, as shown in **Figure 3B**; **Table 4**, enzalutamide was associated with superior bPFS outcomes compared to abiraterone (HR = 0.42, 95% CrI: 0.32–0.55) and cabazitaxel (HR = 0.33, 95% CrI: 0.25–0.44) with a 100% probability of being the best treatment. The corresponding SUCRAs were 1, 0.66, and 0.34.

Subgroup analyses on OS were conducted according to patient baseline characteristics including age, ECOG score, and the presence of visceral metastasis. The results are summarized in **Figure 4**. Overall, patients with non-visceral metastasis or an ECOG score<2 were more likely to benefit from treatments. On the contrary, none of the treatments (abiraterone, enzalutamide, cabazitaxel) showed significantly improved survival compared to BSC in patients with visceral metastasis or a high ECOG score. In patients with younger age (<65 years), non-visceral metastasis, or an ECOG score<2, enzalutamide was associated with the highest probability of ranking first. However, in patients older than 65 years, cabazitaxel showed a superior ranking to other treatments.

3.4 Safety Outcomes

The data of SAEs were available in all the included studies. SAE outcomes of the COU-AA-301 and TROPIC trials were extracted from clinicaltrials.gov (NCT00638690 and NCT00417079).

Treatments ordered from the safest to the least safe based on the risk of SAEs were radium-223 (OR = 0.58, 95% CrI: 0.20–1.68), enzalutamide (OR = 0.80, 95% CrI: 0.28–2.29), abiraterone (OR = 0.94, 95% CrI: 0.39–2.27) and cabazitaxel (OR = 2.50, 95% CrI: 0. 84–7.44). The corresponding SUCRAs were 0.85, 0.65, 0.52, and 0.04 (**Figure 3**; **Table 4**).

4 DISCUSSION

Currently, there is a lack of evidence regarding the preferred treatment for patients with mCRPC after docetaxel failure. In this study, we included five eligible RCTs and conducted a Bayesian network meta-analysis to comprehensively compare and rank four systemic treatments based on their efficacy and safety. All the drugs including abiraterone, enzalutamide, cabazitaxel, and radium-223 were associated with prolonged OS and bPFS relative to BSC. Among them, enzalutamide showed relatively superior survival benefits and higher ranking than others in the overall population and most of the subgroups. For safety, radium-223 was associated with the lowest risk of SAEs, enzalutamide and abiraterone with intermediate risk of SAEs, and cabazitaxel with the highest risk of SAEs.

Compared to the previous review (Fryzek et al., 2018), our study provides several novel findings. First, radium-223 was included in the network analysis. Approximately 90% of patients with mCRPC present with bone metastases (Tannock et al., 2004; Sun et al., 2016). Radium-223 could selectively bind to bone metastases and emit high-energy alpha particles to target areas (Bruland et al., 2006). In overall comparison, radium-223 was associated with prolonged OS. Considering that the ALSYMPCA trial excluded patients with visceral metastasis, we further performed subgroup analysis for patients without visceral metastasis. Consistently, radium-223 showed improved survival compared to BSC and similar efficacy with abiraterone and enzalutamide in this setting. Second, the prior indirect analysis used PFS as one of the efficacy outcomes, while the definitions of PFS were different cross trials (Fryzek et al., 2018). In this study, we compared the bPFS outcomes, which were measured more consistently in the included trials.

Moreover, patients with mCRPC are clinically heterogeneous, we performed subgroup analyses based on patient baseline characteristics to inform decision-making in specific populations. Across treatments, better efficacy was observed in patients with better performance status (ECOG<2) and in patients without visceral metastasis. Across patient subgroups, enzalutamide had the highest probability of being the best treatment for the most time. Patient age seemed to have little impact on the efficacy of enzalutamide and abiraterone, while older (≥65 years) patients might benefit more from cabazitaxel treatment. A recent study suggested abiraterone plus apalutamide combination therapy was associated with improved survival outcomes compared to abiraterone alone in older mCRPC patients who were chemotherapy-naïve (Saad et al., 2021). Thus, AR-targeted combination therapy may also be a potential option in post-chemotherapy setting. In terms of patients with higher ECOG scores, all treatments failed to yield significant survival improvements. It is possible that these patients may be less tolerant to AEs and have higher probability of treatment discontinuation (Honecker et al., 2018). For patients with visceral metastasis, neither abiraterone nor enzalutamide was associated with significantly prolonged survival relative to BSC, which was consistent with the situation in the chemotherapy-naive mCRPC setting (Beer et al., 2014; Ryan et al., 2015). Emerging evidence has suggested that variant histology subtypes and molecular aberrations in visceral metastasis were associated with treatment resistance (Pouessel et al., 2007; Akfirat et al., 2013). In the TROPIC trial, cabazitaxel was not analyzed in this subgroup, as patients were divided by measurable disease rather than visceral disease. However, a recent retrospective study demonstrated visceral metastasis was also an independent predictor for poor survival in patients treated with cabazitaxel (Kosaka et al., 2018). Furthermore, in the CARD trial, no difference in survival was observed between cabazitaxel and abiraterone/enzalutamide in heavily treated patients harboring visceral metastasis (de Wit et al., 2019). Thus, visceral metastasis remains a challenge in the treatment of mCRPC and future trials are needed to further address this issue.

In treatment selection, safety should also be taken sufficiently into consideration. Among the four treatments, radium-223 had the lowest risk of SAEs, followed by enzalutamide and abiraterone. Cabazitaxel was associated with the highest risk of SAEs. Specifically, hematological AEs were extremely common in patients with cabazitaxel. Combined with the survival outcomes, enzalutamide seemed to have the optimal efficacy and safety profile among the four systemic treatments.

Our findings should be interpreted with caution as this review is not devoid of limitations. First, the BSC for patients in the control arm varied across trials. Prednisone was used as BSC in most trials, while the TROPIC trial administered mitoxantrone plus prednisone as control. Although mitoxantrone plus prednisone showed comparative effectiveness with prednisone alone in mCRPC after docetaxel failure, the inconsistent comparators could still bias the results against cabazitaxel (Green et al., 2015). Second, with a follow-up duration of 12.9 months, median survival was not reached in either treatment arm in the Sun et al. (2016) trial. However, the relative effects and ranking probability of abiraterone remained nearly unchanged in the sensitivity analysis in which this trial was excluded (data not shown). Third, our results are limited by the potential bias of the TROPIC and ALSYMPCA trials, which also reflects the opportunities for future trials to improve study design. Finally, in the era of precision oncology, several biomarkers have emerged to predict treatment response, such as androgen receptor splice variant 7 (AR-V7) (Scher et al., 2018; Armstrong et al., 2019). A recent study showed that patients with nuclear-localized AR-V7 protein in circulating tumor cells might have longer survival if treated with taxane chemotherapy than AR signaling inhibitors. However, the relevant information was not feasible in the included trials.

5 CONCLUSION

This interactive network meta-analysis provides the best current evidence on the efficacy and safety profiles of multiple second-line treatments after docetaxel failure in patients with mCRPC. Our findings demonstrate that enzalutamide may provide optimal efficacy and a relatively low risk of SAEs. Cabazitaxel is also effective in post-docetaxel settings but associated with a high risk of SAEs. This study offers important implications for patients and clinicians. However, the results should be used with caution due to the inherent biases across the comparisons. Further head-to-head trials are needed to confirm our findings.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

JC and YZ had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. JC, PS, HZ, and GS contributed to conception and design of the study. JC, YZ, XMZ, JD, ZPW, and XDZ contributed to acquisition, analysis, or interpretation of data. JC, YZ, ZLW, JL, and YN contributed to Statistical analysis. JC, YZ, XMZ, and GS contributed to drafting of the manuscript. PS, HZ, GS, JZ, SZ, and BH contributed to critical revision of the manuscript for important intellectual content. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This study was supported by grants from National Natural Science Foundation of China (NSFC 81974398 and 81902577), and 1.3.5 project for disciplines of excellence, West China Hospital, Sichuan University.

REFERENCES

Akfirat, C., Zhang, X., Ventura, A., Berel, D., Colangelo, M. E., Miranti, C. K., et al. (2013). Tumour Cell Survival Mechanisms in Lethal Metastatic Prostate Cancer Differ between Bone and Soft Tissue Metastases. *J. Pathol.* 230 (3), 291–297. doi:10.1002/path.4180

Armstrong, A. J., Halabi, S., Luo, J., Nanus, D. M., Giannakakou, P., Szmulewitz, R. Z., et al. (2019). Prospective Multicenter Validation of Androgen Receptor Splice Variant 7 and Hormone Therapy Resistance in High-Risk Castration-Resistant Prostate Cancer: The PROPHECY Study. *J. Clin. Oncol.* 37 (13), 1120–1129. doi:10.1200/JCO.18.01731

Beer, T. M., Armstrong, A. J., Rathkopf, D. E., Loriot, Y., Sternberg, C. N., Higano, C. S., et al. (2014). Enzalutamide in Metastatic Prostate Cancer before Chemotherapy. *N. Engl. J. Med.* 371 (5), 424–433. doi:10.1056/NEJMoa1405095

Berthold, D. R., Pond, G. R., Soban, F., de Wit, R., Eisenberger, M., and Tannock, I. F. (2008). Docetaxel Plus Prednisone or Mitoxantrone Plus Prednisone for Advanced Prostate Cancer: Updated Survival in the TAX 327 Study. *J. Clin. Oncol.* 26 (2), 242–245. doi:10.1200/JCO.2007.12.4008

Bruland, Ø. S., Nilsson, S., Fisher, D. R., and Larsen, R. H. (2006). High-linear Energy Transfer Irradiation Targeted to Skeletal Metastases by the Alpha-Emitter 223Ra: Adjuvant or Alternative to Conventional Modalities? *Clin. Cancer Res.* 12 (20 Pt 2), 6250s–6257s. doi:10.1158/1078-0432.CCR-06-0841

Cornford, P., van den Bergh, R. C. N., Briers, E., Van den Broeck, T., Cumberbatch, M. G., De Santis, M., et al. (2021). EAU-EANM-ESTRO-ESUR-SIOG Guidelines on Prostate Cancer. Part II-2020 Update: Treatment of Relapsing and Metastatic Prostate Cancer. *Eur. Urol.* 79 (2), 263–282. doi:10.1016/j.eururo.2020.09.046

de Bono, J. S., Logothetis, C. J., Molina, A., Fizazi, K., North, S., Chu, L., et al. (2011). Abiraterone and Increased Survival in Metastatic Prostate Cancer. *N. Engl. J. Med.* 364 (21), 1995–2005. doi:10.1056/NEJMoa1014618

de Bono, J. S., Oudard, S., Ozguroglu, M., Hansen, S., Machiels, J. P., Kocak, I., et al. (2010). Prednisone Plus Cabazitaxel or Mitoxantrone for Metastatic Castration-Resistant Prostate Cancer Progressing after Docetaxel Treatment: A Randomised Open-Label Trial. *Lancet* 376 (9747), 1147–1154. doi:10.1016/S0140-6736(10)61389-X

de Wit, R., de Bono, J., Sternberg, C. N., Fizazi, K., Tombal, B., Wülfing, C., et al. (2019). Cabazitaxel versus Abiraterone or Enzalutamide in Metastatic Prostate Cancer. *N. Engl. J. Med.* 381 (26), 2506–2518. doi:10.1056/NEJMoa1911206

Fizazi, K., Scher, H. I., Molina, A., Logothetis, C. J., Chi, K. N., Jones, R. J., et al. (2012). Abiraterone Acetate for Treatment of Metastatic Castration-Resistant Prostate Cancer: Final Overall Survival Analysis of the COU-AA-301 Randomised, Double-Blind, Placebo-Controlled Phase 3 Study. *Lancet Oncol.* 13 (10), 983–992. doi:10.1016/S1470-2045(12)70379-0

Fryzek, J. P., Reichert, H., Summers, N., Townes, L., Deuson, R., Alexander, D. D., et al. (2018). Indirect Treatment Comparison of Cabazitaxel for Patients with Metastatic Castrate-Resistant Prostate Cancer Who Have Been Previously Treated with a Docetaxel-Containing Regimen. *PloS one* 13 (4), e0195790. doi:10.1371/journal.pone.0195790

Galletti, G., Leach, B. I., Lam, L., and Tagawa, S. T. (2017). Mechanisms of Resistance to Systemic Therapy in Metastatic Castration-Resistant Prostate Cancer. *Cancer Treat. Rev.* 57, 16–27. doi:10.1016/j.ctrv.2017.04.008

Green, A. K., Corty, R. W., Wood, W. A., Meeneghan, M., Reeder-Hayes, K. E., Basch, E., et al. (2015). Comparative Effectiveness of Mitoxantrone Plus Prednisone versus Prednisone Alone in Metastatic Castrate-Resistant Prostate Cancer after Docetaxel Failure. *Oncologist* 20 (5), 516–522. doi:10.1634/theoncologist.2014-0432

Honecker, F., Wedding, U., Kallischnigg, G., Schroeder, A., Klier, J., Frangenheim, T., et al. (2018). Risk Factors for Unplanned Discontinuation of Scheduled Treatment in Elderly Patients with Castration-Resistant Prostate Cancer: Results of the IBuTu Study. *J. Cancer Res. Clin. Oncol.* 144 (3), 571–577. doi:10.1007/s00432-017-2577-1

J Natl Compr Canc Netw (2021) Prostate Cancer: Version 2.2021. Available at: https://www.nccn.org/professionals/physician_gls/pdf/prostate.pdf (Accessed July 10, 2021).

Kosaka, T., Hongo, H., Mizuno, R., and Oya, M. (2018). Risk Stratification of Castration-Resistant Prostate Cancer Patients Treated with Cabazitaxel. *Mol. Clin. Oncol.* 9 (6), 683–688. doi:10.3892/mco.2018.1724

Parker, C., Nilsson, S., Heinrich, D., Helle, S. I., O'Sullivan, J. M., Fosså, S. D., et al. (2013). Alpha Emitter Radium-223 and Survival in Metastatic Prostate Cancer. *N. Engl. J. Med.* 369 (3), 213–223. doi:10.1056/NEJMoa1213755

Petrylak, D. P., Tangen, C. M., Hussain, M. H., Lara, P. N., Jr, Jones, J. A., Taplin, M. E., et al. (2004). Docetaxel and Estramustine Compared with Mitoxantrone and Prednisone for Advanced Refractory Prostate Cancer. *N. Engl. J. Med.* 351 (15), 1513–1520. doi:10.1056/NEJMoa041318

Pollard, M. E., Moskowitz, A. J., Diefenbach, M. A., and Hall, S. J. (2017). Cost-effectiveness Analysis of Treatments for Metastatic Castration Resistant Prostate Cancer. *Asian J. Urol.* 4 (1), 37–43. doi:10.1016/j.ajur.2016.11.005

Pouessel, D., Gallet, B., Bibeau, F., Avancès, C., Iborra, F., Sénesse, P., et al. (2007). Liver Metastases in Prostate Carcinoma: Clinical Characteristics and Outcome. *BJU Int.* 99 (4), 807–811. doi:10.1111/j.1464-410X.2006.06663.x

Ryan, C. J., Smith, M. R., Fizazi, K., Saad, F., Mulders, P. F., Sternberg, C. N., et al. (2015). Abiraterone Acetate Plus Prednisone versus Placebo Plus Prednisone in Chemotherapy-Naive Men with Metastatic Castration-Resistant Prostate

Cancer (COU-AA-302): Final Overall Survival Analysis of a Randomised, Double-Blind, Placebo-Controlled Phase 3 Study. *Lancet Oncol.* 16 (2), 152–160. doi:10.1016/S1470-2045(14)71205-7

Saad, F., Efstathiou, E., Attard, G., Flaig, T. W., Franke, F., Goodman, O. B., et al. (2021). Apalutamide Plus Abiraterone Acetate and Prednisone versus Placebo Plus Abiraterone and Prednisone in Metastatic, Castration-Resistant Prostate Cancer (ACIS): a Randomised, Placebo-Controlled, Double-Blind, Multinational, Phase 3 Study. *Lancet OncolAdvance Online Publ.* S1470-2045 (21), 00402–2. doi:10.1016/S1470-2045(21)00402-2

Salanti, G., Ades, A. E., and Ioannidis, J. P. (2011). Graphical Methods and Numerical Summaries for Presenting Results from Multiple-Treatment Meta-Analysis: an Overview and Tutorial. *J. Clin. Epidemiol.* 64 (2), 163–171. doi:10.1016/j.jclinepi.2010.03.016

Scher, H. I., Fizazi, K., Saad, F., Taplin, M. E., Sternberg, C. N., Miller, K., et al. (2012). Increased Survival with Enzalutamide in Prostate Cancer after Chemotherapy. *N. Engl. J. Med.* 367 (13), 1187–1197. doi:10.1056/NEJMoa1207506

Scher, H. I., Graf, R. P., Schreiber, N. A., Jayaram, A., Winquist, E., McLaughlin, B., et al. (2018). Assessment of the Validity of Nuclear-Localized Androgen Receptor Splice Variant 7 in Circulating Tumor Cells as a Predictive Biomarker for Castration-Resistant Prostate Cancer. *JAMA Oncol.* 4 (9), 1179–1186. doi:10.1001/jamaoncol.2018.1621

Shim, S. R., Kim, S. J., Lee, J., and Rücker, G. (2019). Network Meta-Analysis: Application and Practice Using R Software. *Epidemiol. Health* 41, e2019013. doi:10.4178/epih.e2019013

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2021). Cancer Statistics, 2021. *CA A. Cancer J. Clin.* 71 (1), 7–33. doi:10.3322/caac.21654

Sterne, J. A. C., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., et al. (2019). RoB 2: a Revised Tool for Assessing Risk of Bias in Randomised Trials. *BMJ* 366, l4898. doi:10.1136/bmj.l4898

Sun, Y., Zou, Q., Sun, Z., Li, C., Du, C., Chen, Z., et al. (2016). Abiraterone Acetate for Metastatic Castration-Resistant Prostate Cancer after Docetaxel Failure: A Randomized, Double-Blind, Placebo-Controlled Phase 3 Bridging Study. *Int. J. Urol.* 23 (5), 404–411. doi:10.1111/iju.13051

Tannock, I. F., de Wit, R., Berry, W. R., Horti, J., Pluzanska, A., Chi, K. N., et al. (2004). Docetaxel Plus Prednisone or Mitoxantrone Plus Prednisone for Advanced Prostate Cancer. *N. Engl. J. Med.* 351 (15), 1502–1512. doi:10.1056/NEJMoa040720

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., et al. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann. Intern. Med.* 169 (7), 467–473. doi:10.7326/M18-0850

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Chen, Zhang, Zhang, Zhao, Ni, Zhu, He, Dai, Wang, Wang, Liang, Zhu, Shen, Zeng and Sun. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 25 January 2022
doi: 10.3389/fphar.2022.802123

Methylxanthine Treatment in Patients Hospitalized for Acute Exacerbation of Chronic Obstructive Pulmonary Disease in China: A Real-World Study Using Propensity Score Matching Analysis

Zijie Zhan[1], Yiming Ma[1], Ke Huang[2,3,4], Chen Liang[5], Xihua Mao[5], Yaowen Zhang[5], Xiaoxia Ren[2,3,4], Jieping Lei[2,3,4], Yan Chen[1], Ting Yang[2,3,4]* and Chen Wang[2,3,4,6]*

[1]Department of Pulmonary and Critical Care Medicine, The Second Xiangya Hospital, Central South University, Changsha, China, [2]Department of Pulmonary and Critical Care Medicine, China-Japan Friendship Hospital, Beijing, China, [3]National Clinical Research Center for Respiratory Disease, Beijing, China, [4]Institute of Respiratory Medicine, Chinese Academy of Medical Science, Beijing, China, [5]Chinese Alliance for Respiratory Diseases in Primary Care, Beijing, China, [6]Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

OPEN ACCESS

Edited by:
Brian Godman,
University of Strathclyde,
United Kingdom

Reviewed by:
Antonio Molino,
University of Naples Federico II, Italy
Yongchun Shen,
Sichuan University, China

***Correspondence:**
Yan Chen
chenyan99727@csu.edu.cn
Ting Yang
zryyyangting@163.com

Specialty section:
This article was submitted to Drugs Outcomes Research and Policies, a section of the journal Frontiers in Pharmacology

Received: *26 October 2021*
Accepted: *06 January 2022*
Published: *25 January 2022*

Citation:
Zhan Z, Ma Y, Huang K, Liang C, Mao X, Zhang Y, Ren X, Lei J, Chen Y, Yang T and Wang C (2022) Methylxanthine Treatment in Patients Hospitalized for Acute Exacerbation of Chronic Obstructive Pulmonary Disease in China: A Real-World Study Using Propensity Score Matching Analysis. Front. Pharmacol. 13:802123. doi: 10.3389/fphar.2022.802123

Background: Although medical guidelines discourage the use of methylxanthines in patients with acute exacerbation of chronic obstructive pulmonary disease (AECOPD), they are still widely used in clinical practice. This study investigated the real-world use of methylxanthines in the management of AECOPD.

Methods: Patient data from the Acute exacerbation of Chronic obstructive pulmonary disease Using REgistry data (ACURE, NCT02657525) study database were screened. Enrolled patients were divided into treatment and control groups. Propensity score (PS) matching and Cox regression analyses were used to minimize confounding factors and determine the association between methylxanthine treatment and the length of stay (LOS).

Results: Among the 2088 eligible patients, 1,563 (74.9%) were in the methylxanthine treatment group. Patients treated with methylxanthines had more severe respiratory symptoms and worse lung function than those in the control group. Doxophylline was the most commonly used methylxanthine in both secondary and tertiary hospitals. After PS matching, 966 patients were equally divided into two groups. The LOS of patients in the two groups was similar [median: 8 days, interquartile range (IQR): 7–11 days, $p = 0.730$]. Patients in the treatment group (median: 8, IQR: 4–12) had a more significant decrease in the COPD Assessment Test score from admission to discharge than those in the control group (median: 6, IQR: 2–10, $p < 0.001$). Among all matched patients, the LOS was not significantly associated with methylxanthine treatment [adjusted hazard ratio (HR): 1.02, 95% confidence intervals (CIs): 0.89–1.16]. However, in the subgroup analysis, methylxanthines were significantly associated with a short LOS in patients with blood eosinophil count >4% (adjusted HR: 1.56, 95% CIs: 1.12–2.17).

Conclusion: This study revealed that methylxanthines, especially doxophylline, are widely used in China. Methylxanthines were effective in improving symptoms in AECOPD patients. Higher blood eosinophil count may be associated with a better efficacy of methylxanthine treatment.

Keywords: chronic obstructive pulmonary disease, acute exacerbation, hospitalization, methylxanthine, length of stay

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is characterized by chronic respiratory symptoms and persistently inadequate airflow. COPD is highly prevalent worldwide and is associated with high disability and mortality according to the Global Burden of Disease study, making it a major public health challenge. In 2017, 299.4 million COPD cases and 3.2 million COPD-related deaths were reported worldwide (GBD 2017 Causes of Death Collaborators, 2018; GBD 2017 Disease and Injury Incidence and Prevalence Collaborators, 2018). In China, 99.9 million cases were reported, and COPD was the third leading cause of death in 2015 (Zhou et al., 2016; Wang et al., 2018). Acute exacerbation of COPD (AECOPD) is a significant event in the course of the disease, and develops in patients approximately 0.5–3.5 times annually. Further, AECOPD is associated with an accelerated decline of lung function, worse quality of life, and increased risk of mortality (Seemungal et al., 1998; Donaldson et al., 2002; Cai et al., 2014; Müllerova et al., 2015). Therefore, effective pharmacological treatments are required to minimize the impact of this condition on patients' health and well-being. Methylxanthines, such as theophylline, have been used in the treatment of COPD and asthma for more than 80 years. They are known to promote bronchodilation and enhance inspiratory muscle function (Barnes, 2013). In addition, theophylline has an anti-inflammatory effect and can reverse corticosteroid resistance (Ito et al., 2002; Cosio et al., 2004). However, previous randomized control trials on methylxanthine treatment for AECOPD have indicated frequent side effects and drug interactions (Barr et al., 2003; Duffy et al., 2005). Based on these results, the Global Initiative for Chronic Obstructive Lung Disease (GOLD) does not recommend the use of methylxanthines in AECOPD patients. Similarly, the Chinese expert consensus recommended the use of methylxanthines only in selected patients who have an insufficient response to short-acting bronchodilators or in some severe AECOPD patients (Cai et al., 2014; Global, 2019).

However, methylxanthines, especially doxophylline, which is a methylxanthine derivative, are still widely used for AECOPD in actual clinical settings in China. It has been reported that doxophylline has better efficacy and safety than other types of methylxanthines, such as theophylline and aminophylline (Cazzola et al., 2018). However, the current use of methylxanthines in China and their efficacy and safety in real-world settings have not been described.

In this study, we first described important aspects of current usage of methylxanthines in hospitalized AECOPD patients in China and subsequently analyzed the efficacy of methylxanthine use in such settings.

METHODS

Data Source

Data for this multicenter retrospective cohort study were obtained from the Acute exacerbation of Chronic obstructive pulmonary disease Using REgistry (ACURE, NCT02657525) database (Pei et al., 2020). The ACURE study is an ongoing, nationwide, multicenter, prospective study that is investigating the demographic characteristics, clinical features, diagnosis, treatments, disease prognosis, and economic costs associated with hospitalized patients with exacerbations of COPD in real-world settings. Additional details of the ACURE study have been described elsewhere (Pei et al., 2020). The study discussed herein was approved by the Ethics Committee of the China-Japan Friendship Hospital (2015-88) and was conducted in accordance with the ethical standards stated in the Helsinki Declaration.

Patient Selection

We used data from the ACURE database starting from January 2018 to December 2019 for patients with AECOPD hospitalized at 163 secondary or tertiary hospital sites. The inclusion criteria for patients were the following: patients who 1) were not less than 40 years old; 2) had spirometry test results and met the diagnostic criteria according to the 2020 GOLD report (Global, 2019), which is a ratio of the post-bronchodilator forced expiratory volume in one second (FEV1) and forced vital capacity (FVC) of less than 0.7; and 3) were hospitalized due to AECOPD. Patients were excluded if they: 1) refused or withdrew informed consent; 2) had enrolled in other interventional studies; 3) had a diagnosis of pneumonia or chronic respiratory disease other than COPD, including asthma, lung cancer, bronchiectasis, and pulmonary fibrosis; and 4) were allergic to methylxanthines.

Measurements and Outcomes

We extracted baseline characteristics of enrolled patients including demographic data, hospital sites of admission, smoking history, body mass index (BMI), AECOPD symptoms at admission, comorbid disease (cardiovascular disease, cerebrovascular disease, and diabetes) history, scores of the modified British Medical Research Council (mMRC) questionnaire, and COPD Assessment Test (CAT) questionnaire at admission, COPD severity, hospitalization history in the previous year, spirometry test results, and blood eosinophil counts. Dyspnea in AECOPD patients was assessed using the mMRC questionnaire, and other symptoms were assessed using the PEACE questionnaire from the Effect of Carbocisteine on Acute Exacerbation of chronic obstructive

Pulmonary disease study (Zheng et al., 2008). The severity of COPD was assessed according to the GOLD staging system given in the 2020 GOLD report (Global, 2019). A history of pharmacological treatments undertaken during hospitalization, including the use of methylxanthines, bronchodilators, and corticosteroids, were also extracted. We divided the enrolled patients into treatment and control groups based on the use of methylxanthines.

The primary outcome was the length of stay (LOS), and the secondary outcomes were a change in the CAT score from admission to discharge, total direct costs, intensive care unit (ICU) admission and mortality during hospitalization, the CAT score at the 30-days follow-up visit, all-cause readmissions, and AECOPD-related readmissions within 30 days after discharge. Total direct costs were converted to United States dollars using the average exchange rate in 2019 (1 United States dollar was equivalent to 6.90 Yuan) because most patients were enrolled in this year.

Statistical Analysis

Continuous variables with normal distributions have been presented as a mean with standard deviation, whereas non-normally distributed continuous variables have been presented a median and interquartile range (IQR). Categorical variables have been presented as frequency and percentage. To compare variables in the two groups, Student's *t*-test or the Wilcoxon rank sum test were used for continuous variables and the Chi-square test or Fisher's exact test were used for categorical variables.

To reduce the effect of confounding factors in this non-experimental study, a propensity score (PS) matching was performed, using the "MathIt" package in R software (Brookhart et al., 2013; Zhang, 2017). We calculated the PS using multivariable logistic regression. Demographic characteristics, including the age, sex, and other baseline variables with a p-value <0.1 in univariable analysis were included in the model. In addition, we used the "nearest" method and a caliper equal to 0.05 to conduct the PS match with a 1:1 ratio.

We used the Cox proportional hazards model to estimate the hazard ratio (HR) between methylxanthine treatment and the LOS. The outcome of the event was defined as hospital discharge. Both univariate and multivariate analyses were performed. All factors with a p-value <0.05, following a univariable analysis, were included in the multivariable analysis. HRs > 1 represented early discharge and a short LOS, while HRs < 1 represented late discharge and a long LOS.

In addition, we performed a subgroup analysis to examine the association between methylxanthine treatment and the LOS in subsets of the population. Patients were divided into subgroups according to the age, sex, smoking history, severity of AECOPD symptoms (cough, sputum, dyspnea, and wheezing), CAT score, GOLD stage, hospitalization in the previous year, and blood eosinophil counts. Cough symptoms were considered to be light if patients had no cough or only morning cough, whereas severe cough symptoms were considered to be episodes of cough during the day or nearly continuous cough.

Statistical analyses were performed using the R software (version 4.0.2). All statistical tests were two-sided tests, with significance set at $p < 0.05$.

RESULTS

Baseline Characteristics

A total of 2088 patients met all the eligibility criteria (**Figure 1**). Among them, 247 patients were from secondary hospitals and 1841 patients were from tertiary hospitals. The median age of the included patients was 70 (IQR: 64–76) years, and 1,679 (80.4%) patients were male. A total of 589 (28.2%) patients were current smokers, 913 (43.7%) were former smokers, and 586 (28.1%) were non-smokers. Among the included patients, 1,563 (74.9%) patients received methylxanthine treatment during hospitalization (treatment group) and 525 (25.1%) patients did not receive methylxanthine treatment (control group).

Demographic characteristics and clinical factors at admission of all enrolled patients before and after PS matching are given in **Table 1**, and treatment course during hospitalization has been summarized in **Table 2**.

The groups were imbalanced for the smoking history, AECOPD symptoms, including cough, wheezing, and dyspnea, comorbid cardiovascular disease, post-dose FEV1/FVC, GOLD stage, and use of bronchodilators, antibiotics and corticosteroids. Patients who received methylxanthines seemed to have severe AECOPD; they had severe cough ($p = 0.002$), dyspnea ($p < 0.001$), and wheezing ($p = 0.010$) at admission. Patients in the treatment group had a history of smoking ($p = 0.029$) and comorbid cardiovascular disease ($p = 0.034$). Spirometry test results showed that patients in the treatment group had a severe deterioration of lung function, as characterized by the FEV1/FVC ($p < 0.001$) and GOLD stage ($p < 0.001$). The CAT score in the treatment group (median: 20, IQR: 15–25) was slightly higher than that in the control group (median: 19, IQR: 15–23); however, the difference was statistically insignificant ($p = 0.075$). Among all AECOPD patients, 1,562 (74.8%) received either short-acting muscarinic antagonists (SAMA) or short-acting beta-antagonists (SABA). Further, 1,233 of the 1,563 (78.9%) patients in the treatment group received either SAMA or SABA; this proportion was significantly higher than that in the control group (329, 62.7%) ($p < 0.001$). Nearly half of the enrolled AECOPD patients received SABA during hospitalization, and there was no significant difference between groups; however, more patients in the treatment group (591, 37.8%) received SAMA than in the control group (135, 25.7%) ($p < 0.001$). Corticosteroids were commonly used for patients in the treatment group, regardless of the administration route. Methylxanthines were commonly used in both secondary and tertiary hospitals, and doxophylline was the most popular methylxanthine drug (63.49% in secondary hospitals and 57.92% in tertiary hospitals; **Figure 2**). The type of methylxanthine use was not significantly different between the two hospital levels ($p = 0.226$).

After PS matching, a total of 966 patients were divided equally into two groups, and all baseline characteristics were balanced. A total of 772 (79.9%) patients in the matched cohort were male, and the median age was 70 years (IQR: 64–76).

Clinical Outcomes During Hospitalization

After PS matching, the median LOS for both the treatment and control groups was approximately 8 days (IQR: 7–11), without

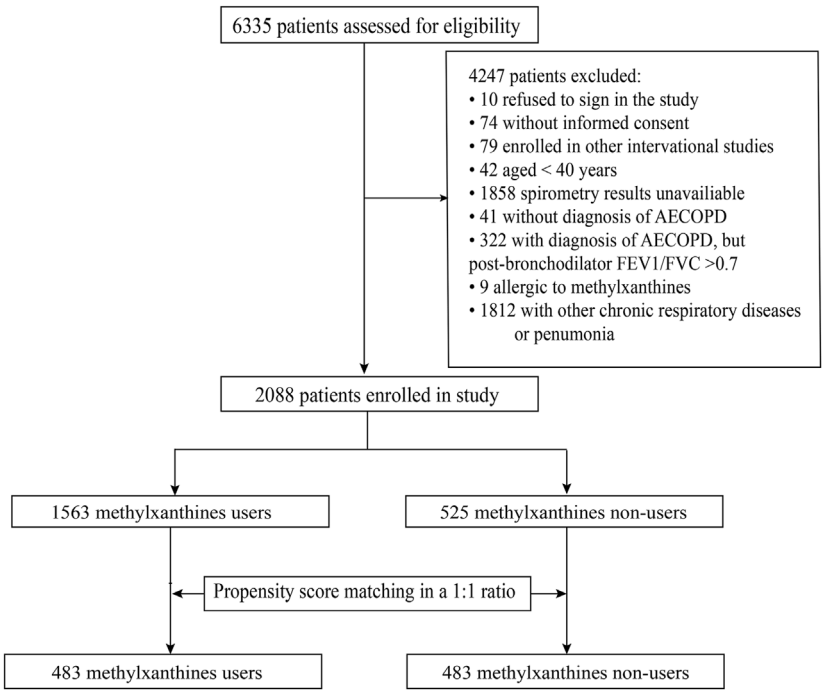

FIGURE 1 | The flowchart of patient enrollment process. AECOPD, acute exacerbations of chronic obstructive pulmonary disease, FEV1/FVC, forced expiratory volume in one second/forced vital capacity.

any significant difference (p = 0.730). However, a significant decrease was observed in the CAT score between the groups (treatment: median, 8; IQR: 4–12 *vs.* control: median: 6, IQR: 2–10; $p < 0.001$). The CAT score at discharge was also lower in the treatment group (median: 10, IQR: 8-14) than the control group (median 12, IQR: 8–16) ($p < 0.001$). There was no significant difference in the total cost, ICU admission, and in-hospital mortality between the groups (**Table 3**).

In the univariable Cox regression analysis, the use of methylxanthines was not significantly associated with the LOS (HR 0.97, 95% CI: 0.86–1.10, p = 0.644). After adjusting for the age, sex, BMI, sputum history, wheezing history, mMRC score, CAT score, comorbid cardiovascular and cerebrovascular diseases, history of hospitalization in the previous year, post-dose FEV1/FVC, use of antibiotics, and use of oral and transvenous corticosteroids, the use of methylxanthines was not significantly associated with the LOS (adjusted HR 1.02, 95% CI: 0.89–1.16, p = 0.783) (**Table 4**).

Clinical Outcomes up Till the 30-days Follow-up Visit

Among the 2088 enrolled patients, 30-days follow-up data was available for 1,246 (59.7%) patients after discharge. Baseline characteristics were similar between patients with and without 30-days follow-up data (**Supplementary Table S1**). After PS matching, 5 of 273 (1.8%) patients were readmitted to the hospital within 30 days with any cause in the treatment group. Only 1 of 300 (0.3%) patients were readmitted to the hospital within 30 days with any cause in the control group, and there was no significant difference between these groups (p = 0.107). Four of 273 (1.5%) patients were readmitted to the hospital because of AECOPD within 30 days in the treatment group, and there were no cases in the control group that had 30-days AECOPD-related readmission. However, there was no significant difference in AECOPD-related readmissions between the two groups (p = 0.051). The CAT score of the treatment group (median, 10, IQR: 8–14) at 30 days after discharge was still significantly lower than that of the control group (median, 12, IQR: 8–17, p = 0.006).

Subgroup Analysis

In the subgroup analysis of the matched cohort for LOS, the association between the use of methylxanthines and LOS remained insignificant in AECOPD patients across different categories of the age, sex, smoking history, severity of AECOPD symptoms, CAT score, GOLD stage, and history of hospitalization in the previous year. The use of methylxanthines was significantly associated with a short LOS only in patients with blood eosinophil counts >4% (adjusted HR: 1.56, 95% CI: 1.12–2.17) (**Figure 3**).

DISCUSSION

To the best of our knowledge, this is the first multicenter study that describes methylxanthine use in China and investigates the efficacy of methylxanthine treatment for hospitalized AECOPD

TABLE 1 | Baseline characteristics of hospitalized acute exacerbation of chronic obstructive pulmonary disease patients at admission.

Variables	Unmatched groups		*p*-Value	Matched groups		*p*-Value
	Treatment group (*n* = 1,563)	Control group (*n* = 525)		Treatment group (*n* = 483)	Control group (*n* = 483)	
Age (year)	69 (63–76)	70 (64–76)	0.245	70 (64–76)	70 (64–76)	0.953
Sex, male	1,262 (80.7)	417 (79.3)	0.512	385 (79.7)	387 (80.1)	0.872
Race, han	1,483 (94.9)	500 (95.2)	0.746	460 (95.2)	461 (95.4)	0.879
Smoking history			0.029*			0.718
Current smokers	451 (28.9)	138 (26.3)	-	123 (25.5)	133 (27.5)	-
Former smokers	697 (44.6)	216 (41.1)	-	206 (42.7)	205 (42.4)	-
Non smokers	415 (26.6)	171 (32.6)	-	154 (31.9)	145 (30.0)	-
BMI (kg/m^2)	22.0 (19.5–24.4)	22.0 (19.4–24.7)	0.473	22.0 (19.6–24.6)	22.2 (19.3–24.8)	0.618
Symptoms						
Cough			0.002*			0.933
None	27 (1.7)	16 (3.0)	-	11 (2.3)	14 (2.9)	-
Morning only	576 (36.9)	233 (44.4)	-	207 (42.9)	205 (42.4)	-
Episodes during day	767 (49.1)	223 (42.5)	-	219 (45.3)	216 (44.7)	-
Nearly continuous	193 (9.2)	53 (10.1)	-	46 (9.5)	48 (9.9)	-
Sputum quantity			0.542			0.638
<10 ml	137 (8.8)	47 (9.0)	-	51 (10.6)	44 (9.1)	-
10–50 ml	789 (50.5)	278 (53.0)	-	251 (52.0)	255 (52.8)	-
50–100 ml	560 (35.8)	181 (34.5)	-	158 (32.7)	167 (34.6)	-
>100 ml	77 (3.7)	19 (3.6)	-	23 (4.8)	17 (3.5)	-
Wheezing	1,347 (86.2)	428 (81.5)	0.010*	405 (83.9)	398 (82.4)	0.548
Dyspnea (mMRC score)			<0.001*			0.720
0	13 (0.8)	14 (2.7)		7 (1.4)	10 (2.1)	-
1	205 (13.1)	81 (15.4)		64 (13.3)	74 (15.3)	-
2	477 (30.5)	163 (31.0)		166 (34.4)	150 (31.1)	-
3	581 (37.2)	205 (39.0)		185 (38.3)	188 (38.9)	-
4	287 (18.4)	62 (11.8)		61 (12.6)	61 (12.6)	-
CAT score	20 (15-25)	19 (15-23)	0.075	20 (15-24)	19 (15-23)	0.420
Comorbidities						
Cardiovascular disease	777 (49.7)	233 (44.4)	0.034*	212 (43.9)	218 (45.1)	0.698
Cerebrovascular disease	69 (4.4)	34 (6.5)	0.059	24 (5.0)	25 (5.2)	0.883
Diabetes	145 (9.3)	46 (8.8)	0.723	43 (8.9)	43 (8.9)	1.000
Previous year admission			0.079			0.914
None	769 (49.2)	272 (51.8)	-	251 (52.0)	244 (50.5)	-
Once	384 (24.6)	135 (25.7)	-	119 (24.6)	127 (26.3)	-
Twice	257 (16.4)	86 (16.4)	-	79 (16.4)	81 (16.8)	-
More than twice	153 (9.8)	32 (6.1)	-	34 (7.0)	31 (6.4)	-
Spirometry						
Post-dose FEV1 (l)	0.96 (0.72–1.34)	0.99 (0.74–1.45)	0.058	0.98 (0.73–1.38)	0.99 (0.74–1.44)	0.616
Post-dose FEV1/FVC	0.49 (0.41–0.58)	0.52 (0.43–0.60)	<0.001*	0.50 (0.42–0.59)	0.52 (0.43–0.60)	0.239
GOLD stage			<0.001*			0.634
I (mild)	98 (6.3)	63 (12.0)	-	47 (9.7)	47 (9.7)	-
II (moderate)	436 (27.9)	140 (26.7)	-	120 (24.8)	136 (28.2)	-
III (severe)	649 (41.5)	213 (40.6)	-	213 (44.1)	196 (40.6)	-
IV (very severe)	380 (24.3)	109 (20.8)	-	103 (21.3)	104 (21.5)	-
Eosinophil count			0.770			0.754
Eos ≤ 2% n (%)	912 (58.8)	305 (59.6)	-	282 (58.8)	282 (59.7)	-
Eos > 2% n (%)	638 (41.2)	207 (40.4)	-	198 (41.2)	190 (40.3)	-

Data are presented as n (%) or median (interquartile range). Abbreviations: BMI, body mass index; CAT, COPD assessment test; Eos, eosinophil; FEV1, forced expiratory volume in one second; FVC, forced vital capacity; mMRC, modified british medical research council.

patients. In our study, we found that a large number of hospitalized AECOPD patients around China received methylxanthine treatment. Among the commonly used methylxanthine drugs, doxophylline was the most common in both secondary and tertiary hospitals. Patients receiving add-on methylxanthine treatment tend to have severe disease as indicated by severe symptoms, spirometry tests, and comorbid diseases. We performed a 1:1 PS matching analysis to estimate the efficacy of methylxanthines. The matched cohort analysis showed significant symptomatic relief in patients who received methylxanthines during hospitalization. However, we did not find significant differences in the LOS, total cost of hospitalization, ICU admission, and in-hospital mortality. To investigate the subgroup of AECOPD patients who could benefit from methylxanthine treatment, we performed a subgroup analysis between methylxanthine treatment and the LOS in a subset of patients across categories of variables of interest. Interestingly, we found that patients

TABLE 2 | In-hospital treatment of acute exacerbation of chronic obstructive pulmonary disease patients with or without methylxanthines.

	Before PSM		*p*-Value	After PSM		*p*-Value
	Treatment group (*n* = 1,563)	Control group (*n* = 525)		Treatment group (*n* = 483)	Control group (*n* = 483)	
Short-bronchodilators						
SAMA	591 (37.8)	135 (25.7)	<0.001*	119 (24.6)	132 (27.3)	0.340
SABA	839 (53.7)	265 (50.5)	0.203	257 (53.2)	252 (52.2)	0.747
Antibiotics	1,419 (90.8)	416 (79.2)	<0.001*	421 (87.2)	400 (82.8)	0.059
Corticosteroids						
Inhaled CS	280 (17.9)	69 (13.1)	0.011 *	69 (14.3)	68 (14.1)	0.926
Oral CS	60 (3.8)	10 (1.9)	0.033 *	13 (2.7)	9 (1.9)	0.388
Transvenous CS	500 (32.0)	119 (22.7)	<0.001 *	121 (25.1)	118 (24.4)	0.823
Nebulized CS	965 (61.7)	227 (43.2)	<0.001 *	212 (43.9)	221 (45.8)	0.560
Methylxanthines						
Doxophylline	1,264 (80.9)	NA	-	384 (79.5)	NA	-
Aminophylline	153 (9.8)	NA	-	42 (8.7)	NA	-
Diprophylline	147 (9.4)	NA	-	46 (9.5)	NA	-
Theophylline	69 (4.4)	NA	-	26 (5.4)	NA	-

Data are presented as n (%) or median (interquartile range). Abbreviations: CS, corticosteroid; PSM, propensity score matching; SAMA, short-acting muscarinic antagonist; SABA, short-acting beta-antagonists; NA, not applicable.

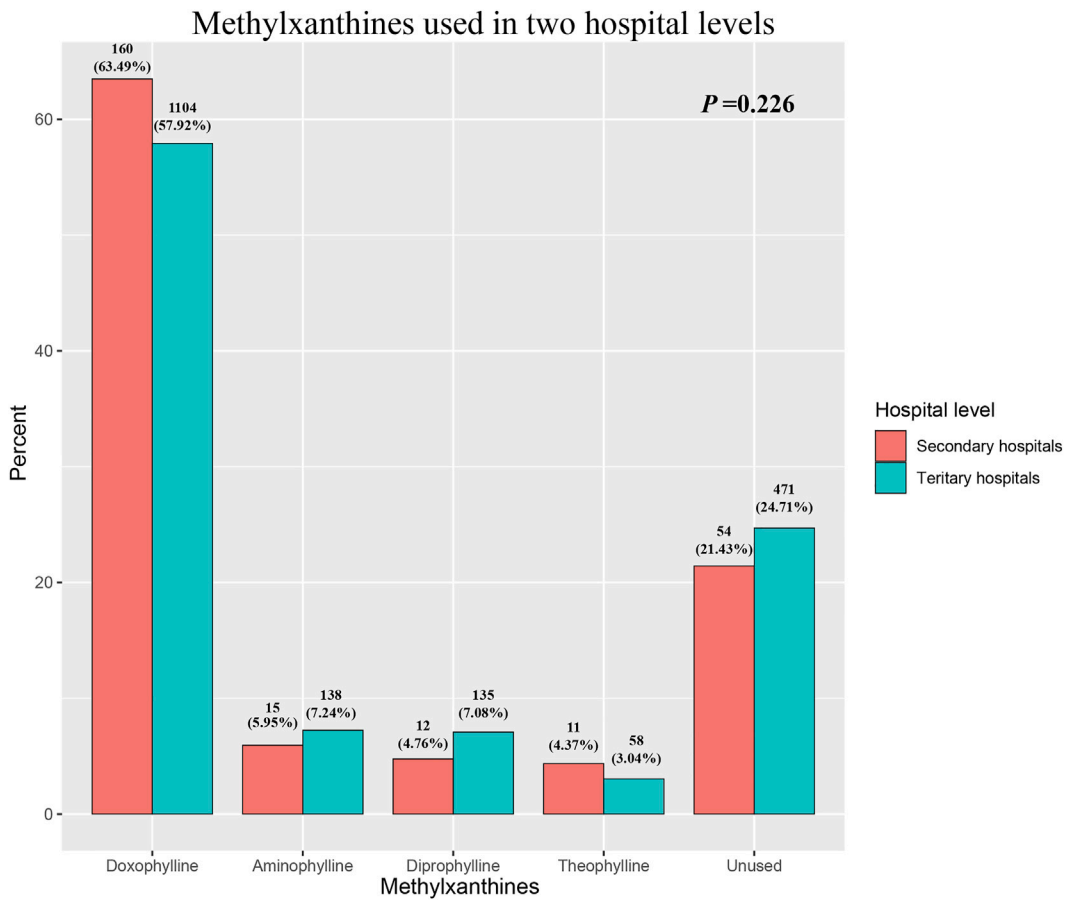

FIGURE 2 | The distribution of methylxanthines used in two hospital levels.

with higher blood eosinophil counts may be associated with a shorter LOS and could benefit from methylxanthine treatment.

The methylxanthine family of drugs includes theophylline, aminophylline, diprophylline, and doxophylline. Theophylline is a classic type of methylxanthine, and several molecular

TABLE 3 | Outcomes of hospitalized AECOPD patients treated with or without methylxanthines.

	Before PSM		*p*-Value	After PSM		*p*-Value
	Treatment group (*n* = 1,563)	Control group (*n* = 525)		Treatment group (*n* = 483)	Control group (*n* = 483)	
Length of stay, days	9 (7–12)	8 (7–11)	0.061	8 (7–11)	8 (7–11)	0.730
CAT score at discharge	11 (8-15)	12 (8-16)	<0.001*	10 (8-14)	12 (8-16)	<0.001*
Change in CAT score	8 (4–12)	6 (2–10)	<0.001*	8 (4–12)	6 (2–10)	<0.001*
Total cost, US dollars	1,323 (971–1850)	1,259 (985–1791)	0.227	1,267 (927–1738)	1,262 (994–1793)	0.280
Oxygen therapy	1,232 (78.8)	390 (74.3)	0.031*	376 (77.8)	367 (76.0)	0.492
NPPV	51 (3.3)	6 (1.1)	0.010*	7 (1.4)	6 (1.2)	0.780
IPPV	1 (0.1)	0 (0)	1.000	0	0	1.000
ICU admission	11 (0.7)	3 (0.6)	1.000	2 (0.4)	3 (0.6)	1.000
In-hospital mortality	2 (0.1)	0	1.000	1 (0.2)	0	1.000
30-days follow-up available	921 (58.9)	325 (61.9)	0.229	273 (56.5)	300 (62.1)	0.077
30-days CAT score	11 (8–15)	12 (8–17)	0.012*	10 (8–14)	12 (8–17)	0.006*
30-days all-cause readmission	34 (3.7)	2 (0.6)	0.003*	5 (1.8)	1 (0.3)	0.108
30-days AECOPD readmission	24 (2.6)	1 (0.3)	0.009*	4 (1.5)	0	0.051

Data are presented as n (%) or median (interquartile range). Abbreviations: AECOPD, acute exacerbation of chronic obstructive pulmonary disease; CAT, COPD assessment test; ICU, intensive care unit; IPPV, invasive positive pressure ventilation; NPPV, non-invasive positive pressure ventilation; PSM, propensity score matching.

mechanisms have been proposed for its therapeutic effect. It may inhibit phosphodiesterase (PDE) non-selectively, be antagonistic to adenosine receptors, and activate certain histone acetylases (Barnes, 2013). Recently, researchers have investigated the potential use of low-dose theophylline to reverse corticosteroid resistance (Cosio et al., 2004). Several randomized control trials using theophylline as an add-on therapy to corticosteroids for stable COPD patients have been conducted. However, the efficacy of low-dose theophylline add-on therapy is limited (Devereux et al., 2018; Devereux et al., 2019; Jenkins et al., 2020). As for management of acute exacerbations, an early meta-analysis revealed that treatment with theophylline or aminophylline did not improve the FEV1 during hospitalization and LOS. Inversely, some side effects such as nausea and vomiting were observed in patients in the treatment group (Barr et al., 2003). As a result, the GOLD report did not recommend theophylline or aminophylline for the management of AECOPD patients. Consistent with guidelines, our study found limited use of theophylline or aminophylline at both secondary and tertiary hospitals.

Doxophylline is another xanthine derivative with anti-inflammatory and bronchodilator effects. The molecular mechanisms of doxophylline are partially different from those of theophylline. It has decreased affinity toward adenosine A1 and A2 receptors, which may contribute to its better safety profile (Matera et al., 2017; Cazzola and Matera, 2020). Studies comparing the safety profile of doxophylline with other types of methylxanthines revealed fewer side effects than those with theophylline and aminophylline in COPD patients. In addition, it seems that doxophylline has significant efficacy in improving lung function, similar to the therapeutic effects of theophylline and aminophylline (Cazzola et al., 2018). Furthermore, although doxophylline has no significant effect on activation of certain histone acetylases, a recent study using a combination of doxophylline and dexamethasone in a murine model of lung inflammation indicated a corticosteroid-sparing effect at a low dose (Riffo-Vasquez et al., 2018). Thus, because of its good safety profile and promising efficacy, doxophylline was the most common drug used among the hospitalized AECOPD patients receiving methylxanthine treatment in our study. Similar to our study findings, methylxanthines in COPD patients are also widely used in Korea. In 2013, 63.4% of patients with COPD prescribed methylxanthines in Korea. However, the researchers did not analyze the type of methylxanthine and the disease stage for COPD patients (Lee et al., 2017).

In our study, we also designed a PS-matched cohort to investigate the efficacy of methylxanthines in real-world conditions. Unlike the results of a previous meta-analysis that indicated a non-significant change in symptom scores for theophylline and aminophylline treatment in AECOPD patients (Barr et al., 2003), we found significant symptomatic relief in patients who were administered methylxanthine treatment, as assessed by the CAT score. However, these patients did not show a significant reduction in the LOS. Interestingly, further subgroup analysis revealed a significant reduction in LOS for patients with blood eosinophil count >4%.

Eosinophils were recently regarded as an essential biomarker in COPD (Bafadhel et al., 2017). High eosinophils either in blood, sputum or airway were independently associated with frequent exacerbation, worse quality of life, and mortality (Hospers et al., 1999; Hospers et al., 2000; Hastie et al., 2017). It was argued that patients with high eosinophils share some common features with asthma, such as more reversibility to bronchodilators, increased FENO, and better treatment response to corticosteroids (Barnes, 2019). There were several well-designed randomized control trials investigating the role of eosinophils in corticosteroid treatment. For example, when comparing with dual bronchodilator therapy, a randomized clinical trial showed inhaled triple therapy adding beclomethasone dipropionate was associated with a significant reduction of moderate to severe exacerbation in patients with blood eosinophils >2% (Papi et al., 2018). Another hoc analysis of three randomized trials also indicated the exacerbation rate was significantly reduced in the group with blood eosinophils >100 cells/ul,

TABLE 4 | Cox proportional hazards regression model for length of stay.

Variables	Univariable analysis		Multivariable analysis	
	HR (95% CIs)	*p* Value	aHR (95% CIs)	*p* Value
Use of methylxanthines	0.97 (0.86–1.10)	0.644	1.02 (0.89–1.16)	0.783
Age, years	0.99 (0.98–1.00)	0.003*	0.99 (0.99–1.00)	0.155
Male (*vs.* female)	0.81 (0.69–0.95)	0.009*	0.80 (0.68–0.94)	0.008*
Race, han (*vs.* others)	0.87 (0.64–1.17)	0.358	-	-
Smoking history (*vs.* never smokers)	-	-	-	-
Current smokers	0.94 (0.80–1.11)	0.477	-	-
Former smokers	0.87 (0.80–1.01)	0.068	-	-
BMI, kg/m^2	1.00 (1.00–1.00)	<0.001*	1.00 (1.00–1.00)	<0.001*
Cough (*vs.* none)	-	-	-	-
Morning only	0.82 (0.55–1.23)	0.332	-	
Episodes during day	0.70 (0.47–1.05)	0.088	-	
Nearly continuous	0.67 (0.43–1.04)	0.074	-	
Wheezing (yes *vs.* no)	0.79 (0.67–0.94)	0.008*	0.95 (0.79–1.14)	0.567
Sputum (*vs.* < 10 ml)	-	-	-	-
10–50 ml	0.89 (0.71–1.11)	0.288	0.92 (0.73–1.16)	0.486
50–100 ml	0.77 (0.61–0.97)	0.024*	0.84 (0.66–1.08)	0.178
>100 ml	0.80 (0.55–1.16)	0.244	0.89 (0.61–1.31)	0.568
mMRC score	0.88 (0.82–0.94)	<0.001*	0.92 (0.85–1.00)	0.043*
CAT score	0.98 (0.97–0.99)	<0.001*	1.00 (0.99–1.01)	0.681
Comorbidities	-	-	-	-
Cardiovascular disease (yes *vs.* no)	0.72 (0.64–0.82)	<0.001*	0.77 (0.67–0.88)	<0.001*
Cerebrovascular disease (yes *vs.* no)	0.64 (0.48–0.86)	0.003*	0.68 (0.51–0.92)	0.011*
Previous year admission (*vs.* None)	-	-	-	-
Once	0.86 (0.74–1.00)	0.050*	0.90 (0.77–1.05)	0.170
Twice	0.85 (0.71–1.02)	0.084	0.99 (0.82–1.19)	0.920
More than twice	0.81 (0.62–1.05)	0.106	0.93 (0.72–1.22)	0.612
Post-dose FEV1, l	1.02 (0.99–1.04)	0.149	-	-
Post-dose FEV1/FVC	1.91 (1.09–3.37)	0.025*	1.93 (1.06–3.54)	0.032*
Use of antibiotics	0.68 (0.56-0.81)	<0.001*	0.69 (0.57–0.84)	<0.001*
Eosinophil percent (≥2% *vs.* <2%)	1.03 (0.90–1.17)	0.696	-	-
Use of SAMA	0.93 (0.80–1.08)	0.332	-	-
Use of SABA	0.95 (0.84–1.08)	0.451	-	-
Use of inhaled CS	0.89 (0.75–1.07)	0.226	-	-
Use of oral CS	0.61 (0.40–0.95)	0.027*	0.66 (0.43–1.04)	0.073
Use of transvenous CS	0.77 (0.67–0.89)	0.001*	0.87 (0.75–1.02)	0.085
Use of nebulized CS	1.01 (0.89–1.14)	0.921	-	-

Variables increase the length of stay if HR < 1, whereas variables reduce the length of stay if HR > 1. Abbreviations: aHR, adjusted hazard ratio; BMI, body mass index; CIs, confidence intervals; CS, corticosteroid; CAT, COPD assessment test; Eos, eosinophil; FEV1, forced expiratory volume in one second; FVC, forced vital capacity; HR, hazard ratio; mMRC, modified british medical research council; SAMA, short-acting muscarinic antagonist; SABA, short-acting beta-antagonists.

while the efficacy increased at higher blood eosinophil level (Bafadhel et al., 2018). A meta-analysis including 5 randomized control trials indicated 17% reduction in exacerbation rate in patients with blood eosinophils >2% (Cheng, 2018).

Corticosteroids were widely used potent anti-inflammatory agents. They have been reported to induce eosinophil apoptosis through inhibiting the effects of IL-5 and GM-CSF on eosinophil survival and production of cytokines from cytokine-producing cells, such as T cells by means of apoptosis. Interestingly, theophylline, one typical type of methylxanthine, showed some addictive effect on induction of eosinophil apoptosis by functioning as a PDE inhibitor to increase intracellular cyclic adenosine monophosphate on activated eosinophils, the latter can induce apoptosis in eosinophils with some cytokines, such as IL-5 and GM-CSF (Hallsworth et al., 1992; Ohta and Yamashita, 1999). Even though there was no previous studies which investigated the role of methylxanthines in eosinophilic COPD, there was one study about the association between roflumilast, a selective PDE4 inhibitor, and eosinophils which revealed an improved efficacy in high blood eosinophil groups (Martinez et al., 2018). It was also reported that treatment with roflumilast reduced sputum neutrophil and eosinophil numbers by targeting PDE4 (Grootendorst et al., 2007). Thus, it was reasonable to infer that methylxanthines may have a better efficacy in patients with higher blood eosinophils.

An important concern regarding methylxanthine use is the risk of adverse effects. However, the ACURE study lacked data on side effects experienced by patients. Therefore, we could not analyze the side effects that occurred in real-world conditions. There were also other limitations to our study. First, follow-up data after 30 days was not available for all patients in our study. However, we compared the baseline characteristics between those with and those without follow-up data, and we did not find a significant difference, which reduced potential selective bias. Second, since the ACURE study is ongoing we did not have

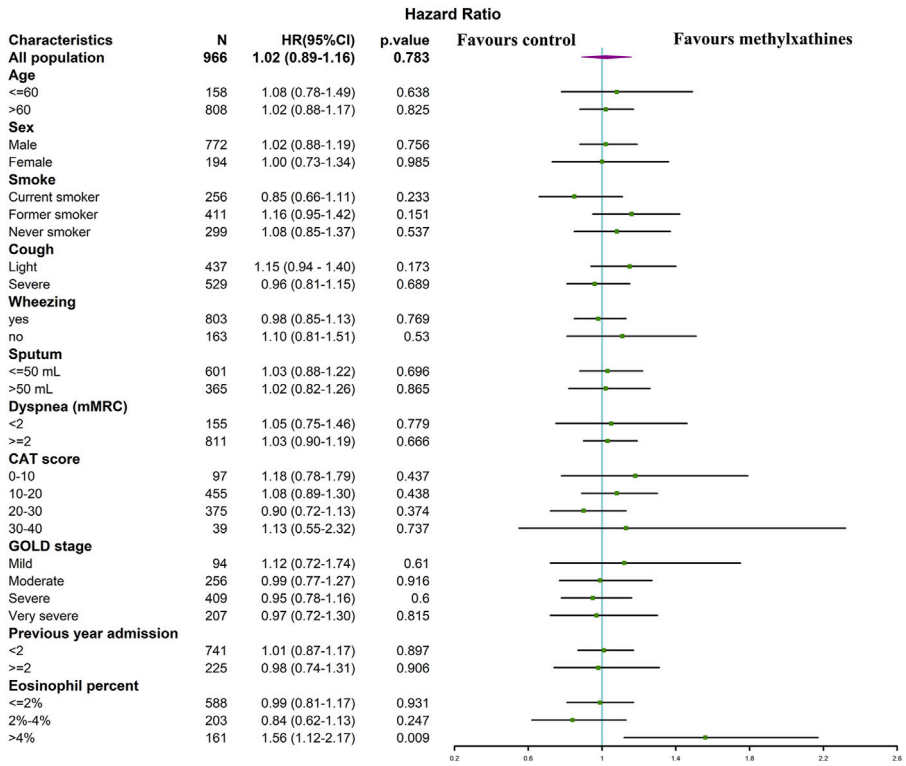

FIGURE 3 | The association between LOS and methylxanthines use in subgroups of matched cohort. Variables increase the length of stay if HR < 1, whereas variables reduce the length of stay if HR > 1.

long-term follow-up data. This information will be available in the future and then analyzed. Third, the number of cases with 30-days readmission was limited, and the impact of methylxanthines on short-term readmission needs further study. Finally, for some sub-group analyses, the number of enrolled patients was limited, and further studies for particular AECOPD patients should be conducted.

CONCLUSION

This study revealed that methylxanthines, especially doxophylline, are widely used in China for the treatment of AECOPD. Methylxanthines were found to be effective in improving AECOPD symptoms. Higher blood eosinophil levels may be associated with better efficacy of methylxanthine treatment. Further studies are needed to determine the association between eosinophil count and the efficacy of methylxanthines in a large population.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Materials**, further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of the China-Japan Friendship Hospital (2015-88). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Concept and design of the study by ZZ, CL, KH, TY, YC, and CW. Data collection and management of the study by ZZ, YM, CL, KH, XM, YZ, XR, JL, YC, TY, and CW. Statistical anlysis by ZZ, CL, XM, and YZ. Drafting of the manuscript by ZZ. All authors contributed to the article and approved the submitted version.

FUNDING

This project was supported by CAMS Innovation Fund for Medical Sciences (CIFMS) (2020-I2M-2-008), Major Program of National Natural Science Foundation of China (82090010, 82090011), Respiratory Disease Clinical Research Public Welfare Program of China Soong Ching Ling Foundation (2018MZFC-032), and National Key R&D Program of China (2018YFC1315100).

ACKNOWLEDGMENTS

We would like to thank all participants in the ACURE study that provided consent for the data used in this clinical study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2022.802123/full#supplementary-material

REFERENCES

Bafadhel, M., Pavord, I. D., and Russell, R. E. K. (2017). Eosinophils in COPD: Just Another Biomarker?. *Lancet Respir. Med.* 5, 747–759. doi:10.1016/S2213-2600(17)30217-5

Bafadhel, M., Peterson, S., De Blas, M. A., Calverley, P. M., Rennard, S. I., Richter, K., et al. (2018). Predictors of Exacerbation Risk and Response to Budesonide in Patients with Chronic Obstructive Pulmonary Disease: a post-hoc Analysis of Three Randomised Trials. *Lancet Respir. Med.* 6, 117–126. doi:10.1016/S2213-2600(18)30006-7

Barnes, P. J. (2019). Inflammatory Endotypes in COPD. *Allergy* 74, 1249–1256. doi:10.1111/all.13760

Barnes, P. J. (2013). Theophylline. *Am. J. Respir. Crit. Care Med.* 188 (8), 901–906. doi:10.1164/rccm.201302-0388pp

Barr, R. G., Rowe, B. H., and Camargo, C. A., Jr (2003). Methylxanthines for Exacerbations of Chronic Obstructive Pulmonary Disease: Meta-Analysis of Randomised Trials. *BMJ* 327 (7416), 643. doi:10.1136/bmj.327.7416.643

Brookhart, M. A., Wyss, R., Layton, J. B., and Stürmer, T. (2013). Propensity Score Methods for Confounding Control in Nonexperimental Research. *Circ. Cardiovasc. Qual. Outcomes* 6 (5), 604–611. doi:10.1161/CIRCOUTCOMES.113.000359

Cai, B. Q., Cai, S. X., Chen, R. C., Cui, L. Y., Feng, Y. L., Gu, Y. T., et al. (2014). Expert Consensus on Acute Exacerbation of Chronic Obstructive Pulmonary Disease in the People's Republic of China. *Int. J. Chron. Obstruct Pulmon Dis.* 9, 381–395. doi:10.2147/COPD.S58454

Cazzola, M., Calzetta, L., Barnes, P. J., Criner, G. J., Martinez, F. J., Papi, A., et al. (2018). Efficacy and Safety Profile of Xanthines in COPD: a Network Meta-Analysis. *Eur. Respir. Rev.* 27 (148). doi:10.1183/16000617.0010-2018

Cazzola, M., and Matera, M. G. (2020). The Effect of Doxofylline in Asthma and COPD. *Respir. Med.* 164, 105904. doi:10.1016/j.rmed.2020.105904

Cheng, S. L. (2018). Blood Eosinophils and Inhaled Corticosteroids in Patients with COPD: Systematic Review and Meta-Analysis. *Int. J. Chron. Obstruct Pulmon Dis.* 13, 2775–2784. doi:10.2147/COPD.S175017

Cosio, B. G., Tsaprouni, L., Ito, K., Jazrawi, E., Adcock, I. M., and Barnes, P. J. (2004). Theophylline Restores Histone Deacetylase Activity and Steroid Responses in COPD Macrophages. *J. Exp. Med.* 200 (5), 689–695. doi:10.1084/jem.20040416

Devereux, G., Cotton, S., Fielding, S., McMeekin, N., Barnes, P. J., Briggs, A., et al. (2018). Effect of Theophylline as Adjunct to Inhaled Corticosteroids on Exacerbations in Patients with COPD: A Randomized Clinical Trial. *JAMA* 320 (15), 1548–1559. doi:10.1001/jama.2018.14432

Devereux, G., Cotton, S., Fielding, S., McMeekin, N., Barnes, P. J., Briggs, A., et al. (2019). Low-dose Oral Theophylline Combined with Inhaled Corticosteroids for People with Chronic Obstructive Pulmonary Disease and High Risk of Exacerbations: a RCT. *Health Technol. Assess.* 23 (37), 1–146. doi:10.3310/hta23370

Donaldson, G. C., Seemungal, T. A., Bhowmik, A., and Wedzicha, J. A. (2002). Relationship between Exacerbation Frequency and Lung Function Decline in Chronic Obstructive Pulmonary Disease. *Thorax* 57 (10), 847–852. doi:10.1136/thorax.57.10.847

Duffy, N., Walker, P., Diamantea, F., Calverley, P. M., and Davies, L. (2005). Intravenous Aminophylline in Patients Admitted to Hospital with Non-acidotic Exacerbations of Chronic Obstructive Pulmonary Disease: a Prospective Randomised Controlled Trial. *Thorax* 60 (9), 713–717. doi:10.1136/thx.2004.036046

Global (2019) Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease 2019 Report. Available from: http://goldcopd.org/wp-content/uploads/2019/11/GOLD-2020- REPORT-ver1.1wms.pdf. Accessed Dec 6, 2019.

GBD 2017 Causes of Death Collaborators (2018). Global, Regional, and National Age-sex-specific Mortality for 282 Causes of Death in 195 Countries and Territories, 1980-2017: a Systematic Analysis for the Global Burden of Disease Study 2017. *Lancet* 392 (10159), 1736–1788. doi:10.1016/S0140-6736(18)32203-7

GBD 2017 Disease and Injury Incidence and Prevalence Collaborators (2018). Global, Regional, and National Incidence, Prevalence, and Years Lived with Disability for 354 Diseases and Injuries for 195 Countries and Territories, 1990-2017: a Systematic Analysis for the Global Burden of Disease Study 2017. *Lancet* 392 (10159), 1789–1858. doi:10.1016/S0140-6736(18)32279-7

Grootendorst, D. C., Gauw, S. A., Verhoosel, R. M., Sterk, P. J., Hospers, J. J., Bredenbröker, D., et al. (2007). Reduction in Sputum Neutrophil and Eosinophil Numbers by the PDE4 Inhibitor Roflumilast in Patients with COPD. *Thorax* 62, 1081–1087. doi:10.1136/thx.2006.075937

Hallsworth, M. P., Litchfield, T. M., and Lee, T. H. (1992). Glucocorticoids Inhibit Granulocyte-Macrophage colony-stimulating Factor-1 and Interleukin-5 Enhanced *In Vitro* Survival of Human Eosinophils. *Immunology* 75, 382–385.

Hastie, A. T., Martinez, F. J., Curtis, J. L., Doerschuk, C. M., Hansel, N. N., Christenson, S., et al. (2017). Association of Sputum and Blood Eosinophil Concentrations with Clinical Measures of COPD Severity: an Analysis of the SPIROMICS Cohort. *Lancet Respir. Med.* 5, 956–967. doi:10.1016/S2213-2600(17)30432-0

Hospers, J. J., Rijcken, B., Schouten, J. P., Postma, D. S., and Weiss, S. T. (1999). Eosinophilia and Positive Skin Tests Predict Cardiovascular Mortality in a General Population Sample Followed for 30 Years. *Am. J. Epidemiol.* 150, 482–491. doi:10.1093/oxfordjournals.aje.a010037

Hospers, J. J., Schouten, J. P., Weiss, S. T., Postma, D. S., and Rijcken, B. (2000). Eosinophilia Is Associated with Increased All-Cause Mortality after a Follow-Up of 30 Years in a General Population Sample. *Epidemiology* 11, 261–268. doi:10.1097/00001648-200005000-00006

Ito, K., Lim, S., Caramori, G., Cosio, B., Chung, K. F., Adcock, I. M., et al. (2002). A Molecular Mechanism of Action of Theophylline: Induction of Histone Deacetylase Activity to Decrease Inflammatory Gene Expression. *Proc. Natl. Acad. Sci. U S A.* 99 (13), 8921–8926. doi:10.1073/pnas.132556899

Jenkins, C. R., Wen, F.-Q., Martin, A., Barnes, P. J., Celli, B., Zhong, N.-S., et al. (2020). The Effect of Low Dose Corticosteroids and Theophylline on the Risk of Acute Exacerbations of COPD. The TASCS Randomised Controlled Trial. *Eur. Respir. J.* 57 (6), 2003338. doi:10.1183/13993003.03338-2020

Lee, J., Lee, J. H., Kim, J. A., and Rhee, C. K. (2017). Trend of Cost and Utilization of COPD Medication in Korea. *Int. J. Chron. Obstruct Pulmon Dis.* 12, 27–33. doi:10.2147/COPD.S121687

Martinez, F. J., Rabe, K. F., Calverley, P. M. A., Fabbri, L. M., Sethi, S., Pizzichini, E., et al. (2018). Determinants of Response to Roflumilast in Severe Chronic Obstructive Pulmonary Disease. Pooled Analysis of Two Randomized Trials. *Am. J. Respir. Crit. Care Med.* 198, 1268–1278. doi:10.1164/rccm.201712-2493OC

Matera, M. G., Page, C., and Cazzola, M. (2017). Doxofylline Is Not Just Another Theophylline!. *Int. J. Chron. Obstruct Pulmon Dis.* 12, 3487–3493. doi:10.2147/COPD.S150887

Müllerova, H., Maselli, D. J., Locantore, N., Vestbo, J., Hurst, J. R., Wedzicha, J. A., et al. (2015). Hospitalized Exacerbations of COPD: Risk Factors and Outcomes in the ECLIPSE Cohort. *Chest* 147 (4), 999–1007. doi:10.1378/chest.14-065

Ohta, K., and Yamashita, N. (1999). Apoptosis of Eosinophils and Lymphocytes in Allergic Inflammation. *J. Allergy Clin. Immunol.* 104, 14–21. doi:10.1016/s0091-6749(99)70107-7

Papi, A., Vestbo, J., Fabbri, L., Corradi, M., Prunier, H., Cohuet, G., et al. (2018). Extrafine Inhaled Triple Therapy versus Dual Bronchodilator Therapy in Chronic Obstructive Pulmonary Disease (TRIBUTE): a Double-Blind, Parallel Group, Randomised Controlled Trial. *Lancet* 391, 1076–1084. doi:10.1016/S0140-6736(18)30206-X

Pei, Z., Sun, Y., Wang, S., Chen, Y., Yang, T., Huang, K., et al. (2020). Estimating Mortality Among Inpatients with Acute Exacerbation of Chronic Obstructive Pulmonary Disease Using Registry Data. *NPJ Prim. Care Respir. Med.* 30 (1), 28. doi:10.1038/s41533-020-0186-y

Riffo-Vasquez, Y., Venkatasamy, R., and Page, C. P. (2018). Steroid Sparing Effects of Doxofylline. *Pulm. Pharmacol. Ther.* 48, 1–4. doi:10.1016/j.pupt.2017.10.008

Seemungal, T. A., Donaldson, G. C., Paul, E. A., Bestall, J. C., Jeffries, D. J., and Wedzicha, J. A. (1998). Effect of Exacerbation on Quality of Life in Patients with Chronic Obstructive Pulmonary Disease. *Am. J. Respir. Crit. Care Med.* 157 (5 Pt 1), 1418–1422. doi:10.1164/ajrccm.157.5.9709032

Wang, C., Xu, J., Yang, L., Xu, Y., Zhang, X., Bai, C., et al. (2018). Prevalence and Risk Factors of Chronic Obstructive Pulmonary Disease in China (The China Pulmonary Health [CPH] Study): a National Cross-Sectional Study. *Lancet* 391 (10131), 1706–1717. doi:10.1016/S0140-6736(18)30841-9

Zhang, Z. (2017). Propensity Score Method: a Non-parametric Technique to Reduce Model Dependence. *Ann. Transl Med.* 5 (1), 7. doi:10.21037/atm.2016.08.57

Zheng, J. P., Kang, J., Huang, S. G., Chen, P., Yao, W. Z., Yang, L., et al. (2008). Effect of Carbocisteine on Acute Exacerbation of Chronic Obstructive Pulmonary Disease (PEACE Study): a Randomised Placebo-Controlled Study. *Lancet* 371 (9629), 2013–2018. doi:10.1016/S0140-6736(08)60869-7

Zhou, M., Wang, H., Zhu, J., Chen, W., Wang, L., Liu, S., et al. (2016). Cause-specific Mortality for 240 Causes in China during 1990-2013: a Systematic Subnational Analysis for the Global Burden of Disease Study 2013. *Lancet* 387 (10015), 251–272. doi:10.1016/S0140-6736(15)00551-6

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Zhan, Ma, Huang, Liang, Mao, Zhang, Ren, Lei, Chen, Yang and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 08 February 2022
doi: 10.3389/fphar.2022.808440

Effect of 0.01% Atropine on Accommodation in Myopic Teenagers

Huixia Li[1†], Liying Zhang[2†], Hong Tian[3†], Song Zhang[4†], Xueyan Zhang[1†], Han Zhang[1†], Yujing Chen[2], Wenping Qi[5], Xiaoying Wu[5], Hongmei Jiang[6], Hailong Yang[1], Yajun Yang[5], Lei Liu[7,8] and Guisen Zhang[1]**

[1]Department of Retina, Inner Mongolia Chaoju Eye Hospital, Hohhot, China, [2]Department of Cataract, Ulanqab Chaoju Eye Hospital, Ulanqab, China, [3]Department of Cataract, Baotou Kunlun Chaoju Eye Hospital, Baotou, China, [4]Department of Graduate School, China Medical University, Shenyang, China, [5]Department of Cataract, Baotou Chaoju Eye Hospital, Baotou, China, [6]Department of Myopia, Chifeng Chaoju Eye Hospital, Chifeng, China, [7]Department of Ophthalmology, Guangdong Eye Institute, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, School of Medicine, South China University of Technology, Guangzhou, China, [8]Department of Ophthalmology, The First Affiliated Hospital of China Medical University, Shenyang, China

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Mario Damiano Toro,
Medical University of Lublin, Poland
Marjana Novič,
National Institute of Chemistry,
Slovenia

***Correspondence:**
Lei Liu
liulejiao@163.com
Guisen Zhang
zhangguisen76@sohu.com

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Drugs Outcomes Research and Policies, a section of the journal Frontiers in Pharmacology

Received: 03 November 2021
Accepted: 03 January 2022
Published: 08 February 2022

Citation:
Li H, Zhang L, Tian H, Zhang S, Zhang X, Zhang H, Chen Y, Qi W, Wu X, Jiang H, Yang H, Yang Y, Liu L and Zhang G (2022) Effect of 0.01% Atropine on Accommodation in Myopic Teenagers. Front. Pharmacol. 13:808440. doi: 10.3389/fphar.2022.808440

Purpose: The purpose of the study is to evaluate the effects of 0.01% atropine eye drops on accommodative system parameters among teenagers with low myopia.

Methods: Ninety-five myopic teenagers [39 boys (8.69 ± 2.473) and 56 girls (8.54 ± 2.054) aged 5–17 years] with no history of eye disease were enrolled. Biometric and accommodative system parameters were evaluated before and at 1 week, 1 month, 3 months, and 6 months of 0.01% atropine eye drop instillation.

Results: Participants without accommodative demand at 6 months demonstrated insignificant changes after the atropine instillation (all $p > 0.05$). Nevertheless, there were significant differences in accommodative sensitivity, accommodative amplitude, accommodative responsiveness, and negative relative accommodation (NRA) at 3 months compared with baseline after atropine instillation (all $p < 0.05$). Except spherical equivalent refraction, cornea thickness, intraocular pressure, and axial length were stable after the 0.01% atropine instillation (all $p > 0.05$).

Conclusion: Morphologically, current measurements suggested that 0.01% atropine had favorable reduction of accommodation for childhood low myopia over a half-year period.

Keywords: accommodation, low-concentration atropine, myopia, ocular biometric components, axial length, spherical equivalent

INTRODUCTION

Myopia has become a significant global public health and socioeconomic problem (Xue et al., 2018). It has been predicted that 49.8% of the global population is predicted to be myopic, and 9.8% of the global population is predicted to be high myopic in 2050 (Holden et al., 2016). Most developed countries, especially among East Asia, have been currently faced with high prevalence of myopia as well as high myopia (Morgan et al., 2012). Several studies report that the occurrence and progression of myopia are related to accommodation (Holden et al., 2014; Jonas et al., 2021). Previously, atropine can be used for cycloplegia before optometry, and the treatment of amblyopia, iridocyclitis, and malignant glaucoma (Chen et al., 2021; Pineles et al., 2021). In recent years, this "old drugs" with new application was used to control myopia (Saw et al., 2002; Chia et al., 2016; Hieda et al., 2021). However, the mechanism of its use to

control myopia is still unclear. The efficacy of atropine in controlling myopia is positively correlated with concentration (Chia et al., 2012; Chia et al., 2014). The higher the concentration, the better the myopia controlling effects, while the more obvious the adverse reactions caused by this drug (Pineles et al., 2017). Recent studies have shown that low-concentration atropine (0.01%) not only retains the effect of myopia control but also greatly reduces the incidence of adverse effects (Chia et al., 2016). However, long-term follow-up studies on the effects of atropine on accommodative parameters have not been evaluated comprehensively. Therefore, the current study used a nonrandomized before and after controlled design to observe the effects of 0.01% atropine on accommodative system parameters.

METHODS

Subjects

Teenagers with low myopia were continuously enrolled from April 2019 to March 2021. This study followed the guidelines outlined in the Declaration of Helsinki and was approved by the Institutional Review Board (IRB) and research ethics committee of Hohhot Chaoju Eye Hospital. All parents/guardians of participating teenagers received a full explanation of the procedures and provided written informed consent. This clinical study registration number is ChiCTR2000034981.

Inclusion Criteria

None of the participants had a history of eye disease, except for refractive errors, and all had a best-corrected visual acuity of 20/20 or better, with mydriatic spherical equivalent (SE) refractive error of −1.50 D to −0.75 D, astigmatism less than −1.00 D, and with normal intraocular pressure (<21 mmHg).

Exclusion Criteria

We excluded subjects who were with any history of ocular disease, previous ocular surgery, or taking systemic medications that could affect accommodation.

Refractive Status Check

All included subjects were treated with 1% atropine for cycloplegia, and a computerized automatic refractor (RC-4000, Tomey, Japan) and a retinoscope (YZ24, Suzhou Liuliu Vision Technology Co., Ltd., Suzhou, China) were used for objective flexion. For the detection of the optical state, the optometrist used a comprehensive refractor (DK-500, Japan Topcon Company) to perform subjective optometry to measure the final diopter the next day.

All included subjects underwent comprehensive ophthalmological examinations including best corrected visual acuity (BCVA), full optometric, noncontact tonometry, biometrics and slit lamp microscope, fundus, and accommodative function measurements before and at 1 week, 1 month, 3 months, and 6 months after 0.01% atropine eye-drop use. An experienced technician in each of the hospital carried out all the examinations. All subjects underwent low-concentration atropine (0.01%) eye drop instillation once a day.

Accommodative Amplitude Measurement

Refractive errors were corrected with the best refractive compensation in place during the accommodative tests if teenagers were needed. The near vision card (Near Chart) at a distance of 40 cm was set according to "minus lens method," so that the examinee could see the upper line of the optotype clearly. Minus and plus lenses with 0.25 D steps were increased subsequently until the first sustained blur.

Accommodative Sensitivity Measurement

The participant wears a telescopic correction lens with ±2.00 D inversion shot, and looked at the Near Chart at 40 cm (the upper line of the best visual acuity mark), and first place +2.00 D in front of both eyes. It was reported immediately when the reader read clearly, and immediately reversed to −2.00 D, until the font was clear and then reversed again. Finally, the numbers of cycles completed within 1 min were recorded.

Accommodative Response

The optometrist used a comprehensive refractometer (VT-10, TOPCON, Japan) on the basis of long-distance correction, and placed a 0.50 D cross column in front of the eyes of the participant. After that, the participant should fix the negative axis of the cross cylinder to 90° and the positive axis to 180°, respectively. Then the patient was asked to look at the fused cross cylinder through the cross cylinder to test the visual target. If the participant replied that the horizontal line was clearer than the vertical line, it reflected that the participant should gradually add a 0.25 D lens positive lens in front of the participant's eyes until the subject could see the two lines as clearly as well. If the participant replied that the vertical line was clearer, it reflected advanced adjustment. The adjustment would be made by gradually adding a negative lens with a 0.25 D lens in front of the participant's eyes, gradually, one by one, until the participant could see the two lines equally clearly.

Negative Relative Accommodation and Positive Relative Accommodation

Measurements were set at a distance of 40 cm on the basis of distance correction of both eyes. For negative relative accommodation (NRA), the positive lens with +0.25 D was added one at a time and fixed in front of the eyes of the participant until the target was blurred to the subject. The total number of added lenses recorded is the NRA value. For positive relative accommodation (PRA), the negative lens was added one at a time in front of the eyes of the participant with −0.25 D until the target was blurred. The number of negatives added last is the value of PRA.

Accommodation Convergence/ Accommodation Ratio

Subjective binocular detection was first evaluated at a distance of 6 m, and then at a distance of 40 cm. A phoropter was used to measure horizontal occlusion, placing a 6 Δ base-up separation prism in front of the left eye to deviate the vertical oculomotor

TABLE 1 | Baseline biometric parameters of both eyes before instillation of 0.01% atropine eye drops.

	Right eye		Left eye		*p*-Value
	Median	**Range**	**Median**	**Range**	
SE (D)	–0.75	[–2.75, 0.25]	–0.75	[–2.25, 1.875]	0.133
Cornea thickness (μm)	537	[464, 617]	541	[463, 614]	0.825
IOP (mmHg)	17	[10, 24]	17	[10, 25]	0.849
AL (mm)	23.66	[22.13, 27.89]	23.62	[22.02, 25.84]	0.454
Accommodative sensitivity (cpm/min)	6	[0, 15]	6.00	[0, 13]	0.990
Accommodative amplitude (D)	11.5	[0, 18]	12.00	[0, 20]	0.896
Accommodative responsiveness (D)	0.25	[–0.5, 1]	0.25	[–0.5, 1]	0.990
NRA (D)	2	[0.5, 3]	2.00	[0.5, 3]	0.990
PRA (D)	–2	[–5, 1.75]	–2	[–5, 1.75]	0.990
AC/A	3	[0, 6]	3.00	[0, 6]	0.990

Note. PRA, positive relative accommodation; NRA, negative relative accommodation; AC/A, accommodative convergence/accommodation ratio; IOP, intraocular pressure; AL, axial length; SE, spherical equivalent.

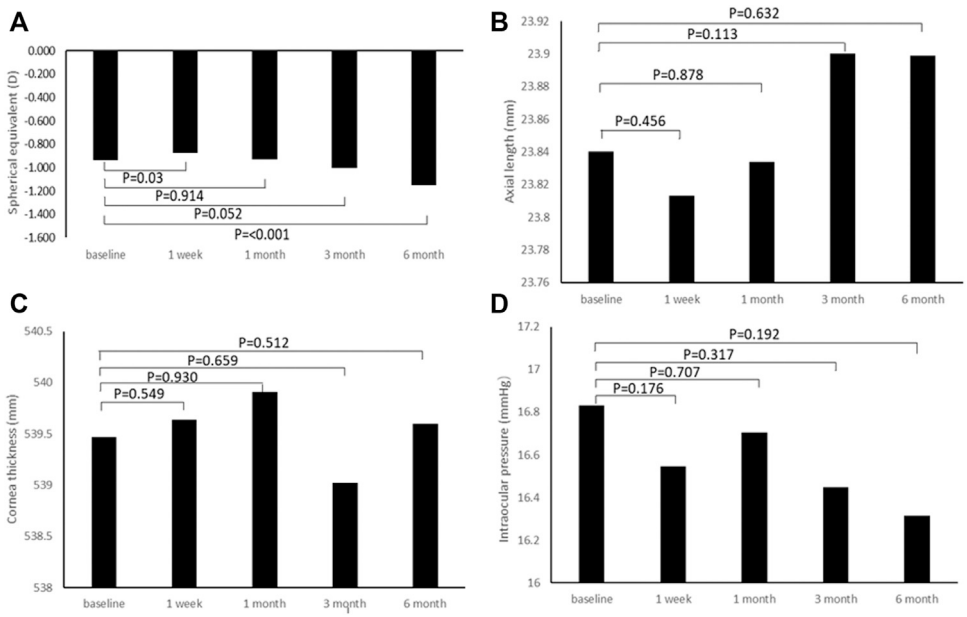

FIGURE 1 | Change in biometric parameters after the instillation of 0.01% atropine eye drops. **(A)** Spherical equivalent. **(B)** Axial length. **(C)** Cornea thickness. **(D)** Intraocular pressure.

nerve, and then using a 12 Δ base-in separation prism to neutralize the horizontal oculomotor deviation in the front of the right eye until it induces diplopia. Then the number of prisms was reduced until the subject could just recover from the diplopia image. Three hidden slope measurements were obtained for each subject, and the average value was taken. The accommodation convergence/accommodation (AC/A) ratio was calculated using the heterophoria method.

Statistical Methods

The SPSS 22.0 statistical software (SPSS Inc., Chicago, IL, USA) was used to analyze the data. Categorical variables were summarized by percentage. Numerical data were expressed as the mean ± standard deviation (SD) or median/interquartile range (IQR). Mann–Whitney U-test was used to test for difference between the right and left eyes. The nonparametric Wilcoxon matched-pairs signed-rank test was used to identify the difference among baseline and four different follow-ups (Tan et al., 2021). A *p*-value less than 0.05 was identified as statistically significant.

RESULTS

Baseline Characteristics of the Participants

Ninety-five teenagers [39 boys (8.69 ± 2.473) and 56 girls (8.54 ± 2.054)] all underwent 0.01% atropine treatment, and no severe complications were observed during the follow-up period. The mean age of the participants was 8.60 ± 2.22 (range 5–17) years. **Table 1** presents the baseline biometric and accommodation

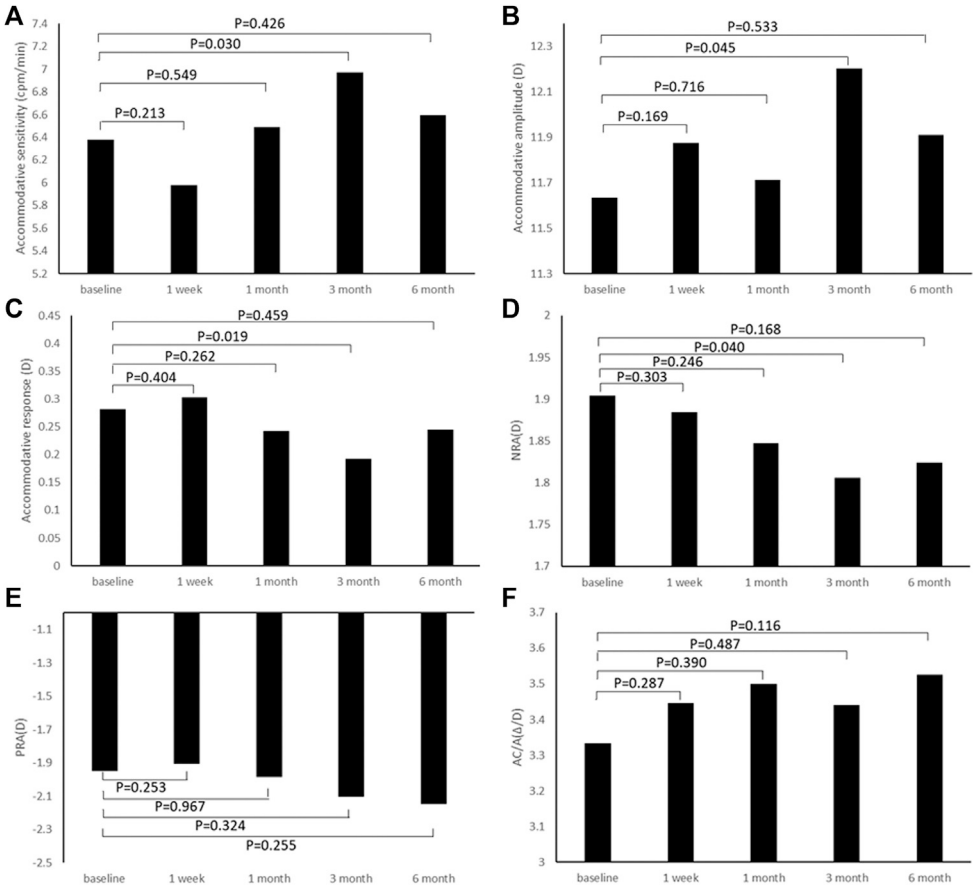

FIGURE 2 | Change in accommodative system parameters after the instillation of 0.01% atropine eye drops. **(A)** Accommodative sensitivity. **(B)** Accommodative amplitude. **(C)** Accommodative response. **(D)** NRA, negative relative accommodation. **(E)** PRA, positive relative accommodation. **(F)** AC/A, accommodative convergence/accommodation ratio.

system parameters of both eyes before the instillation of 0.01% atropine eye drops. The cycloplegic refraction of the right eye ranged from −2.75 D to 0.25 D, and that of the left eye ranged from −2.25 D to 1.875 D; thus, there was no significant difference in refraction between the right and left eyes (p = 0.133). Furthermore, there was also no significant difference in cornea thickness, intraocular pressure, and axial length as well as accommodation parameters between the right and left eyes (all $p > 0.05$). Therefore, we present the findings of only the right eye.

Effects of 0.01% Atropine on Refraction and Biometric Parameters

Figure 1A presents a comparison of SE in the relaxed state before and 1 week, 1, 3, and 6 months after instillation of 0.01% atropine eye drops. The SE was improved 1 week after atropine instillation than before in the relaxed state (p = 0.03). In contrast, there was no significant difference in the RE 1 and 3 months after atropine instillation compared with that before atropine instillation in the relaxed state, while it was significant in the RE 6 months after atropine instillation compared with before atropine instillation. **Figures 1B–D** show that there are no significant differences in AL, cornea thickness, and intraocular pressure (IOP) before and after atropine instillation (all $p > 0.05$).

Effects of 0.01% Atropine on Accommodative System Parameters

Figure 2A shows the accommodative sensitivity 1 week, 1 month, 3 months, and 6 months after the instillation of 0.01% atropine eye drops. Except for the 3 months, accommodative sensitivity, accommodative amplitude, accommodative responsiveness, and NRA did not change significantly at the three follow-ups after the instillation of 0.01% atropine eye drops compared with baseline, respectively (all $p > 0.05$, **Figures 2B–D**). However, the differences in positive relative accommodation (PRA) and AC/A value were insignificant before and after the instillation of 0.01% atropine eye drops (all $p > 0.05$, **Figures 2E,F**).

DISCUSSION

In order to prevent the outbreak of coronavirus disease 2019 (COVID-19), lockdown has caused many students to stay indoors

in many countries, which may promote myopic occurrence and progression. With the increased near work and insufficient outdoor activities during lockdown, approximately 80% of outpatient visits in the pediatric ophthalmology department in March and April 2020 were for refractive errors (RE), of which 79% were for myopia (Jayadev et al., 2020). Recently, a prospective cross-sectional study, using school-based photoscreenings involving 123,535 Chinese children aged 6–13 years during the COVID-19 home confinement, has reported a significant myopic shift for children aged 6–8 years of approximately −0.3 D (Wang et al., 2021). These findings are consistent with the outcomes of another large-scale intervention research with 12-month follow-up on myopia development among 1,001,749 school children of 1,305 elementary and high schools. Here, myopia increased approximately 1.5 times from −0.23 D during the pre-COVID-19 to −0.343 D during the post-COVID-19 period (Toro et al., 2021). Moreover, even if the data are still controversial, besides myopic progression, accommodative dysfunction in children may be also on the rise due to digital device usage. Accommodation spasm and sudden-onset esotropia are important consequences.

To some extent, better vision is associated with better accommodative function (Chen et al., 2019). Previous studies have shown that low-concentration atropine effectively inhibits the progression of myopia and axial elongation (Azuara-Blanco et al., 2020; Polling et al., 2020; Walline et al., 2020). However, the effects of 0.01% atropine on accommodative system parameters are still unclear. Our prospective nonrandomized single-arm study revealed that 0.01% atropine had no effect on accommodative system parameters before and 6 months after instillation. Interestingly, accommodative sensitivity, accommodative amplitude, accommodative responsiveness, and NRA changed significantly at the 3 month follow-up after the instillation of 0.01% atropine eye drops compared with baseline, respectively.

To date, there are a few reports evaluating the effects of low-concentration atropine on the accommodative system. Previously meta-analysis revealed that patients receiving 0.01% atropine showed no significant changes in accommodative amplitude (WMD, −0.45 D; 95% CI = −1.80, 0.90; p = 0.51) over a 1-year period (Tsai et al., 2021), which is consistent with our findings. However, in another study, Caucasian children receiving a concentration of 0.05 and 0.01% atropine revealed that accommodation was decreased by −4.2 ± 3.8 D in 0.05% atropine-treated eyes, whereas 0.01% atropine induced hypoaccommodation of −0.05 ± 2.5 D ($p < 0.01$) (Joachimsen et al., 2021). There is a nonlinear dose–response relationship between atropine and accommodative amplitude (Tran et al., 2021). Herein, future study is needed to determine the concentration that provides maximal efficacy with tolerable effects on the accommodative system.

In the present study, compared with baseline, a significant reduction in SE after 0.01% atropine eye drop instillation was observed at 1 week. Although the changes in AL were not significant compared with baseline at 1 week, the decrease in AL was observed. This trend was consistent with the findings of Ye et al. (2021). These findings suggest that the improved effect of 0.01% atropine eye drop instillation on SE might be affected by the shortened AL. Recently, it has been reported that one of the effects of atropine is that it stimulates choroidal thickening (Chiang et al., 2020; Mathis et al., 2021; Ye et al., 2021). However, our findings were specifically created to have opposing effects at 6 months after atropine eye drop instillation.

Limitations of the present study should be acknowledged. Clinically, it is difficult to make adolescents receive 0.01% atropine intervention for 12 months, and our observation duration is 6 months. Herein, the follow-up period of the current study was not long enough to judge about the accommodative system parameter change. We did not have a healthy control group to observe the natural process of disease, which was a big limitation. Finally, we did not include the teenagers with moderate or high myopia because of the relatively small number of participants. Therefore, further randomized controlled trial should be performed to provide more evidence on the accommodation function after the instillation of atropine eye drops.

CONCLUSION

Generally, our study observed the effects of 0.01% atropine eye drops on the accommodative parameters, and it is necessary to observe accommodative parameters at 3 months after the 0.01% atropine eye drop instillation. Our findings provided information for its clinical application for controlling the development of myopia in teenagers.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Inner Mongolia Chaoju Eye Hospital Medical Science Research Ethics Committee. Written informed consent to participate in this study was provided by the legal guardian/next of kin of the participants.

AUTHOR CONTRIBUTIONS

YY, LL, SZ, and GZ participated in the design of the study, performed the statistical analysis, and wrote the first draft. LL conceived the study, participated in its design and coordination, and helped in drafting the manuscript. HL performed the investigations. HL, LL, and GZ interpreted and discussed the results. All authors read and approved the final manuscript.

FUNDING

The work was supported by the Baotou City Science and Technology Program (grant No: wjwkj2020100), Grassland

Elite Project (No. CYYC10089), 2021 Inner Mongolia Autonomous Region Medical and Health Technology Plan Project, and Hohhot City Applied Technology Research and Development Project (No. 2021-Ji-02).

ACKNOWLEDGMENTS

We thank all the teenagers and their parents and teachers who participated in this study.

REFERENCES

Azuara-Blanco, A., Logan, N., Strang, N., Saunders, K., Allen, P. M., Weir, R., et al. (2020). Low-dose (0.01%) Atropine Eye-Drops to Reduce Progression of Myopia in Children: a Multicentre Placebo-Controlled Randomised Trial in the UK (CHAMP-UK)-study Protocol. *Br. J. Ophthalmol.* 104 (7), 950–955. doi:10.1136/bjophthalmol-2019-314819

Chen, D. X., He, X. G., and Xu, X. (2021). The Safety of Atropine for Myopia Prevention and Control. *Zhonghua Yan Ke Za Zhi* 57 (4), 299–304. doi:10.3760/cma.j.cn112142-20200622-00413

Chen, M., Long, Q., Gu, H., and Hong, J. (2019). Accommodation Changes after Visian Implantable Collamer Lens with central Hole for High Myopia: A STROBE-Compliant Article. *Medicine (Baltimore)* 98 (28), e16434. doi:10.1097/MD.0000000000016434

Chia, A., Chua, W. H., Cheung, Y. B., Wong, W. L., Lingham, A., Fong, A., et al. (2012). Atropine for the Treatment of Childhood Myopia: Safety and Efficacy of 0.5%, 0.1%, and 0.01% Doses (Atropine for the Treatment of Myopia 2). *Ophthalmology* 119 (2), 347–354. doi:10.1016/j.ophtha.2011.07.031

Chia, A., Chua, W. H., Wen, L., Fong, A., Goon, Y. Y., and Tan, D. (2014). Atropine for the Treatment of Childhood Myopia: Changes after Stopping Atropine 0.01%, 0.1% and 0.5%. *Am. J. Ophthalmol.* 157 (2), 451–e1 e451. doi:10.1016/j.ajo.2013.09.020

Chia, A., Lu, Q. S., and Tan, D. (2016). Five-Year Clinical Trial on Atropine for the Treatment of Myopia 2: Myopia Control with Atropine 0.01% Eyedrops. *Ophthalmology* 123 (2), 391–399. doi:10.1016/j.ophtha.2015.07.004

Chiang, S. T., Turnbull, P. R. K., and Phillips, J. R. (2020). Additive Effect of Atropine Eye Drops and Short-Term Retinal Defocus on Choroidal Thickness in Children with Myopia. *Sci. Rep.* 10 (1), 18310. doi:10.1038/s41598-020-75342-9

Hieda, O., Hiraoka, T., Fujikado, T., Ishiko, S., Hasebe, S., Torii, H., et al. (2021). Efficacy and Safety of 0.01% Atropine for Prevention of Childhood Myopia in a 2-year Randomized Placebo-Controlled Study. *Jpn. J. Ophthalmol.* 65 (3), 315–325. doi:10.1007/s10384-021-00822-y

Holden, B., Sankaridurg, P., Smith, E., Aller, T., Jong, M., and He, M. (2014). Myopia, an Underrated Global challenge to Vision: where the Current Data Takes Us on Myopia Control. *Eye (Lond)* 28 (2), 142–146. doi:10.1038/eye.2013.256

Holden, B. A., Fricke, T. R., Wilson, D. A., Jong, M., Naidoo, K. S., Sankaridurg, P., et al. (2016). Global Prevalence of Myopia and High Myopia and Temporal Trends from 2000 through 2050. *Ophthalmology* 123 (5), 1036–1042. doi:10.1016/j.ophtha.2016.01.006

Jayadev, C., Sarbajna, P., and Vinekar, A. (2020). Commentary: Impact of the COVID-19 Pandemic on Digital Eye Strain in Children. *Indian J. Ophthalmol.* 68 (11), 2383–2384. doi:10.4103/ijo.IJO_3028_20

Joachimsen, L., Farassat, N., Bleul, T., Böhringer, D., Lagrèze, W. A., and Reich, M. (2021). Side Effects of Topical Atropine 0.05% Compared to 0.01% for Myopia Control in German School Children: a Pilot Study. *Int. Ophthalmol.* 41 (6), 2001–2008. doi:10.1007/s10792-021-01755-8

Jonas, J. B., Ang, M., Cho, P., Guggenheim, J. A., He, M. G., Jong, M., et al. (2021). IMI Prevention of Myopia and its Progression. *Invest. Ophthalmol. Vis. Sci.* 62 (5), 6. doi:10.1167/iovs.62.5.6

Mathis, U., Feldkaemper, M. P., and Schaeffel, F. (2021). Effects of Single and Repeated Intravitreal Applications of Atropine on Choroidal Thickness in Alert Chickens. *Ophthalmic Res.* 64 (4), 664–674. doi:10.1159/000515755

Morgan, I. G., Ohno-Matsui, K., and Saw, S. M. (2012). Myopia. *Lancet* 379 (9827), 1739–1748. doi:10.1016/S0140-6736(12)60272-4

Pineles, S. L., Kraker, R. T., VanderVeen, D. K., Hutchinson, A. K., Galvin, J. A., Wilson, L. B., et al. (2017). Atropine for the Prevention of Myopia Progression in Children: A Report by the American Academy of Ophthalmology. *Ophthalmology* 124 (12), 1857–1866. doi:10.1016/j.ophtha.2017.05.032

Pineles, S. L., Repka, M. X., Yu, F., Velez, F. G., Perez, C., Sim, D., et al. (2021). The Use of Atropine for Treatment of Amblyopia Using the OptumLabs Data Warehouse. *J. AAPOS* 25 (2), 107–e1 e101. doi:10.1016/j.jaapos.2020.11.003

Polling, J. R., Tan, E., Driessen, S., Loudon, S. E., Wong, H. L., van der Schans, A., et al. (2020). A 3-year Follow-Up Study of Atropine Treatment for Progressive Myopia in Europeans. *Eye (Lond)* 34 (11), 2020–2028. doi:10.1038/s41433-020-1122-7

Saw, S. M., Shih-Yen, E. C., Koh, A., and Tan, D. (2002). Interventions to Retard Myopia Progression in Children: an Evidence-Based Update. *Ophthalmology* 109 (3), 415–443. discussion 422-414; quiz 425-416, 443. doi:10.1016/s0161-6420(01)00972-1

Tan, S. Y. X., Chia, A., Sugianto, R., Tong, H. E., Ang, I. Y. H., Shek, L. P., et al. (2021). Validation of a Web-Based, Time-Use Application to Assess Children's School Meal Intakes: My E-Diary for Activities and Lifestyle (MEDAL). *Nutrients* 13 (11). doi:10.3390/nu13113790

Toro, M. D., Bremond-Gignac, D., Brézin, A. P., Cummings, A. B., Kemer, O. E., Kermani, O., et al. (2021). COVID-19 Outbreak and Increased Risk of Amblyopia and Epidemic Myopia: Insights from EUROCOVCAT Group. *Eur. J. Ophthalmol.* 11206721211053175, 11206721211053175. doi:10.1177/11206721211053175

Tran, H. D. M., Sankaridurg, P., Naduvilath, T., Ha, T. T. X., Tran, T. D., Jong, M., et al. (2021). A Meta-Analysis Assessing Change in Pupillary Diameter, Accommodative Amplitude, and Efficacy of Atropine for Myopia Control. *Asia Pac. J. Ophthalmol. (Phila)* Publish Ahead of Print (5), 450–460. doi:10.1097/APO.0000000000000414

Tsai, H. R., Chen, T. L., Wang, J. H., Huang, H. K., and Chiu, C. J. (2021). Is 0.01% Atropine an Effective and Safe Treatment for Myopic Children? A Systemic Review and Meta-Analysis. *J. Clin. Med.* 10 (17). doi:10.3390/jcm10173766

Walline, J. J., Lindsley, K. B., Vedula, S. S., Cotter, S. A., Mutti, D. O., Ng, S. M., et al. (2020). Interventions to Slow Progression of Myopia in Children. *Cochrane Database Syst. Rev.* 1, CD004916. doi:10.1002/14651858.CD004916.pub4

Wang, J., Li, Y., Musch, D. C., Wei, N., Qi, X., Ding, G., et al. (2021). Progression of Myopia in School-Aged Children after COVID-19 Home Confinement. *JAMA Ophthalmol.* 139 (3), 293–300. doi:10.1001/jamaophthalmol.2020.6239

Xue, A., Zheng, L., Tan, G., Wu, S., Wu, Y., Cheng, L., et al. (2018). Genipin-Crosslinked Donor Sclera for Posterior Scleral Contraction/Reinforcement to Fight Progressive Myopia. *Invest. Ophthalmol. Vis. Sci.* 59 (8), 3564–3573. doi:10.1167/iovs.17-23707

Ye, L., Li, S., Shi, Y., Yin, Y., He, J., Zhu, J., et al. (2021). Comparisons of Atropine versus Cyclopentolate Cycloplegia in Myopic Children. *Clin. Exp. Optom.* 104 (2), 143–150. doi:10.1111/cxo.13128

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Li, Zhang, Tian, Zhang, Zhang, Zhang, Chen, Qi, Wu, Jiang, Yang, Yang, Liu and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Medicine

SYSTEMATIC REVIEW
published: 18 February 2022
doi: 10.3389/fmed.2022.780956

Efficacy and Safety of Collagenase Clostridium Histolyticum in the Treatment of Peyronie's Disease: An Evidence-Based Analysis

Dehong Cao[1†], Jinze Li[1†], You Lu[2†], Yin Huang[1], Bo Chen[1], Zeyu Chen[1], Yinzhi Shen[3], Liangren Liu[1] and Qiang Wei[1*]*

[1] Department of Urology, Institute of Urology, West China Hospital, Sichuan University, Chengdu, China, [2] Department of Child Healthcare, West China Second University Hospital, Sichuan University, Chengdu, China, [3] West China School of Medicine, Sichuan University, Chengdu, China

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Yunxiang Li,
North Sichuan Medical College, China
ZhengJu Ren,
Chongqing Medical University, China

***Correspondence:**
Qiang Wei
weiqiang339@126.com
Liangren Liu
liuliangren@scu.edu.cn

[†] These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Family Medicine and Primary Care, a section of the journal Frontiers in Medicine

Received: 22 September 2021
Accepted: 11 January 2022
Published: 18 February 2022

Citation:
Cao D, Li J, Lu Y, Huang Y, Chen B, Chen Z, Shen Y, Liu L and Wei Q (2022) Efficacy and Safety of Collagenase Clostridium Histolyticum in the Treatment of Peyronie's Disease: An Evidence-Based Analysis. Front. Med. 9:780956. doi: 10.3389/fmed.2022.780956

Background: Peyronie's disease (PD) is a chronic wound healing disorder, mainly involving tunica albuginea. Collagenase Clostridium Histolyticum (CCH) has shown its effectiveness in treating PD, but its efficacy and safety remain controversial, which propelled us to conduct the first evidence-based research on this topic.

Methods: We searched the Web of Science, PubMed, Embase, and ClinicalTrials.gov for related randomized controlled trials (RCTs). A systematic review and meta-analysis were performed to compare the penile curvature deformity (PCD), Peyronie's Disease Questionnaire peyronie's disease symptom bother (PDSB), penile pain score, total treatment-related adverse events (TAEs), and specific adverse events, including penile pain, penile edema, injection site pain, and contusion. Cochrane Collaboration's tool and Review Manager 5.3.0 version were applied, respectively, to evaluate the study quality and heterogeneity.

Results: Four articles (five RCTs) with 1,227 patients were finally included in the meta-analysis. The results revealed that CCH had excellent efficacy in relieving PCD (weighted mean difference [WMD]: −318.77, $p < 0.001$) and PDSB (WMD: −1.20, $p < 0.001$) compared to the placebo group, but there was no difference in the penile pain score (WMD: −0.64, $P = 0.39$) between the two groups. Furthermore, the incidence of TAEs in the CCH group was higher [odds ratio (OR): 12.86, $p < 0.001$].

Conclusions: The current evidence suggests that CCH has a significant effect on treating PD. Considering that all these adverse events are acceptable and curable, CCH could slow the disease progression in the acute phase or act as a substitute for patients unable or unwilling to undergo surgery. However, the conclusion could not be certainly drawn until RCTs with a larger scale proved it.

Keywords: Peyronie's disease (PD), Collagenase Clostridium Histolyticum (CCH), sexual function, efficacy, safety, meta-analysis

INTRODUCTION

Peyronie's disease (PD), also known as penile fibrous cavernositis, is a benign chronic disease characterized by the formation of fibroids in the tunica albuginea (1), leading to plaque formation, penile malformation, penile pain, sexual dysfunction, and mental disorders (2). Recent epidemiological surveys have shown that the period of the PD onset is usually 40 to 70 years old, and its prevalence is 3 to 9%. However, due to its low recognition, the prevalence of PD may actually be higher (3–5).

At present, clinical treatment of PD is mainly concentrated on the acute phase of the disease, aiming to prevent disease progression and penile malformation. For patients in the acute phase, non-surgical treatment is the primary choice, and only patients who are in the stable phase or with serious problems are treated with surgery (1, 6). Non-surgical treatments, such as oral drug therapy, intralesional local injection therapy, iontophoresis, and vacuum mechanical traction therapy, have received extensive attention (7–9). However, there is no strong evidence to prove their therapeutic advantages (10).

Local injection therapy for lesions has been accepted recently, because of its rapid drug delivery, high local drug concentration, and ease of clinical operation (7, 11). Currently, drugs for the local injection are mainly CCH, verapamil, and interferon-α2b, of which CCH is the most widely used (9, 12, 13). Previous clinical trials focused on the effectiveness of CCH (14–16), but most of them were reviews or case reports, lacking evidence-based medical evidence, and there are still some clinicians who are concerned about its safety and efficacy. Therefore, we are the first to use meta-analysis to explore the efficacy and safety of CCH in the treatment of PD all over the world.

METHODS

This present study was performed following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) criteria (17).

Search Strategy

The following search string terms: "Peyronie's," "Peyronie's disease," "penile curvature," "PD," "Collagenase Clostridium Histolyticum," "Xiafle," "Xiapex," and "CCH" were used to systematically search the Web of Science, PubMed, Embase, and CinicalTrials.gov date to May 2021 for randomized controlled trials (RCTs) that compared CCH with placebo for PD. The search language was limited to English. In addition, the reference lists of all eligible studies were reviewed manually.

Study Selection Criteria

If correlative studies suffice to meet all the following criteria, they will be included in this study: (1) RCTs or pseudo-RCTs, (2) Males who were 18 years of age or older, had a regular heterosexual partner, and were clinically diagnosed with PD; (3) These RCTs investigated the effect of CCH in patients with PD and compared it with placebo or blank controls; and (4) The study provided at least one indicator of outcomes that can be analyzed. On the contrary, studies were excluded if: (1) The study data could not be obtained; (2) Studies that combined CCH with other treatments were excluded; (3) Animal experiments, reviews, letters, editorial comments, pediatric articles, case reports, or conference abstracts; and (4) Unpublished articles and non-English articles.

Data Extraction

After scanning the title, abstract, and full text, two analysts (Cao and Shen) selected the literature in strict accordance with the inclusion criteria and then extracted the data according to the pre-designed table for cross-checking. Any argument on this topic was arbitrated by a third researcher. The extracted data included the first author, year of publication, type of study design, interventions, total number and age of subjects, the follow-up period, and outcome indicators. The following result outcomes were extracted: penile curvature deformity (PCD), Peyronie's Disease Questionnaire Peyronie's disease symptom bother (PDSB), penile pain score, and total treatment-related adverse events (TAEs). Additionally, several common complications, such as penile pain, penile edema, injection site pain, and contusion, were also included.

Study Quality Assessment

The quality assessment was based on methodological quality assessment criteria recommended by the Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (16). "High risk" stands for the high risk of bias, "low risk" stands for the low risk of bias, and "unclear risk" stands for the absence of adequate information to conduct the bias evaluation. All differences were solved by a third researcher.

Statistical Analysis

RevMan 5.3.0 (Cochrane Collaboration, Oxford, UK) was used for meta-analysis. Weighted mean difference (WMD) and odds ratio (OR) were used as the effect indexes for continuous and dichotomous data, respectively, and P values and 95% CI were given for both. Heterogeneity between studies was judged by Cochran's Q and I^2 statistics. When there was a statistical homogeneity between studies ($p > 0.1$ and $I^2 < 50\%$), the fixed effect model was introduced for analysis; otherwise, the random effect model was applied. For all statistical consequences, $p < 0.05$ was considered statistically significant. Sensitivity analysis was performed by excluding one or more studies that led to heterogeneity.

RESULTS

Description of Studies

A total of 226 related articles were obtained by a preliminary examination. Two hundred duplicates and unrelated studies were removed, and 26 studies were left. After reading the full text, four articles (five studies) were finally included for meta-analysis (**Figure 1**) (18–21), and 1,227 patients with PD were involved. Of these, 815 and 412 patients accepted CCH and placebo, respectively. In one of these articles, there were two studies, and they came from different experimental centers, so we named these two studies Gelbard a and Gelbard b (20). In addition,

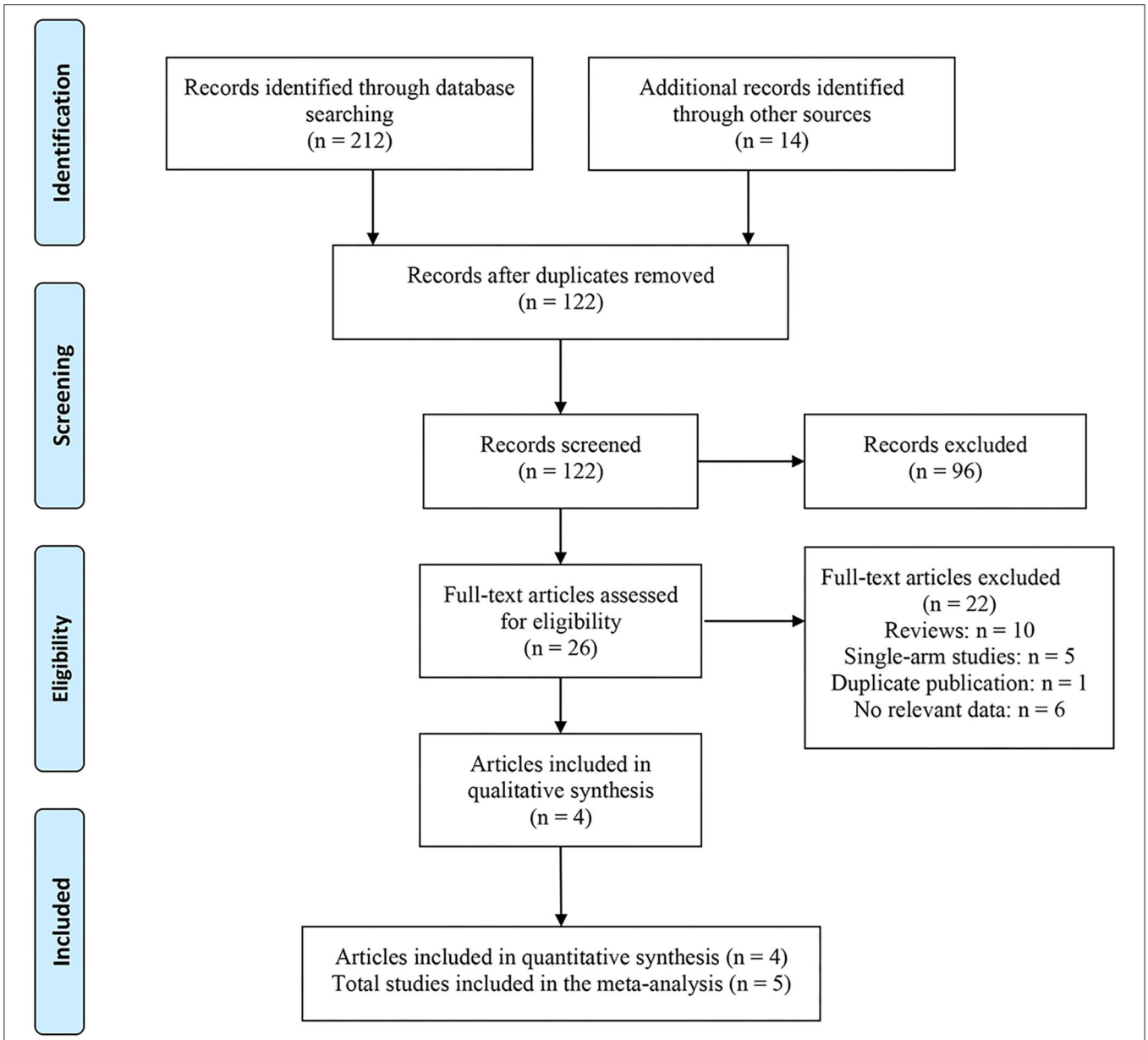

FIGURE 1 | The literature screening process.

we also looked for partial data available at the Cochrane Central Register of Controlled Trials according to the ClinicalTrials.gov Identifier of the three RCTs. The basic information and baseline characteristics of the incorporated studies are shown in **Table 1**, and the methodological quality evaluation of RCTs is exposed in **Figure 2**.

Penile Curvature Deformity

The data of mean percent change in PCD were reported in 4 studies involving 956 patients with PD (18–20). The combined results displayed a significant improvement in the CCH group compared with the placebo group (fixed-effects model; WMD: −318.77; 95% CI: −22.58 to −14.96; $p < 0.001$; $I^2 = 38\%$; **Figure 3A**).

Peyronie's Disease Symptom Bother

We extracted the data for this indicator from 4 studies (18–20), which contained a total of 945 patients with PD with 632 patients in the CCH group and 313 patients in the placebo group. Results of the heterogeneity test revealed no significant difference between the two groups ($P = 0.64$; $I^2 = 0\%$; **Figure 3B**), and a fixed-effects model was introduced. The comprehensive analysis demonstrated that the PDSB in the CCH group was lower than that in the placebo group (WMD: −1.20; 95% CI: −1.69 to −0.72; $p < 0.001$).

TABLE 1 | Basic information and characteristics of studies.

Authors	Year	Design	LOE	CCH/Placebo		Intervention	Follow-up	Outcome measures
				Patients (N)	Age (years)[a]			
Gelbard et al. (18)	1993	RCT	1b	22/27	NA	6,000–14,000 unit in 3–7 injections	3 months	TAEs
Gelbard et al. (19)	2012	RCT	1b	111/36	56.9 (7.8)/55.4 (7.0)	1–3 cycles[b] with a interval of 6 weeks	36 weeks	PCD, PDSB, penile pain score, TAEs, penile pain, penile edema, injection site pain and contusion
Gelbard a (20)	2013	RCT	1b	277/140	57.9 (8.2)/58.2 (8.9)	1–4 cycles[b] with a interval of 6 weeks	52 weeks	PCD, PDSB, penile pain score, TAEs, penile pain, penile edema, injection site pain and contusion
Gelbard b (20)	2013	RCT	1b	274/141	57.3 (8.8)/57.6 (7.5)	1–4 cycles[b] with a interval of 6 weeks	52 weeks	PCD, PDSB, penile pain score, TAEs, penile pain, penile edema, injection site pain and contusion
Lipshultz et al. (21)	2015	RCT	1b	131/68	NA	1–4 cycles[b] with a interval of 6 weeks	52 weeks	PCD, PDSB

RCT, randomized controlled trial; LOE, Level of evidence; CCH, Collagenase Clostridium Histolyticum; TAEs, treatment-related adverse events; PCD, penile curvature deformity; PDSB, Peyronie's disease questionnaire symptom bother; NA, not available. [a]Data are presented by median (standard deviation). [b]Each treatment cycle consisted of 2 intralesional injections of 0.58 mg CCH or placebo, with an interval of approximately 24–72 h between each injection. Approximately 24–72 h following the second injection of each treatment cycle, subjects in the CCH and placebo groups underwent penile plaque modeling.

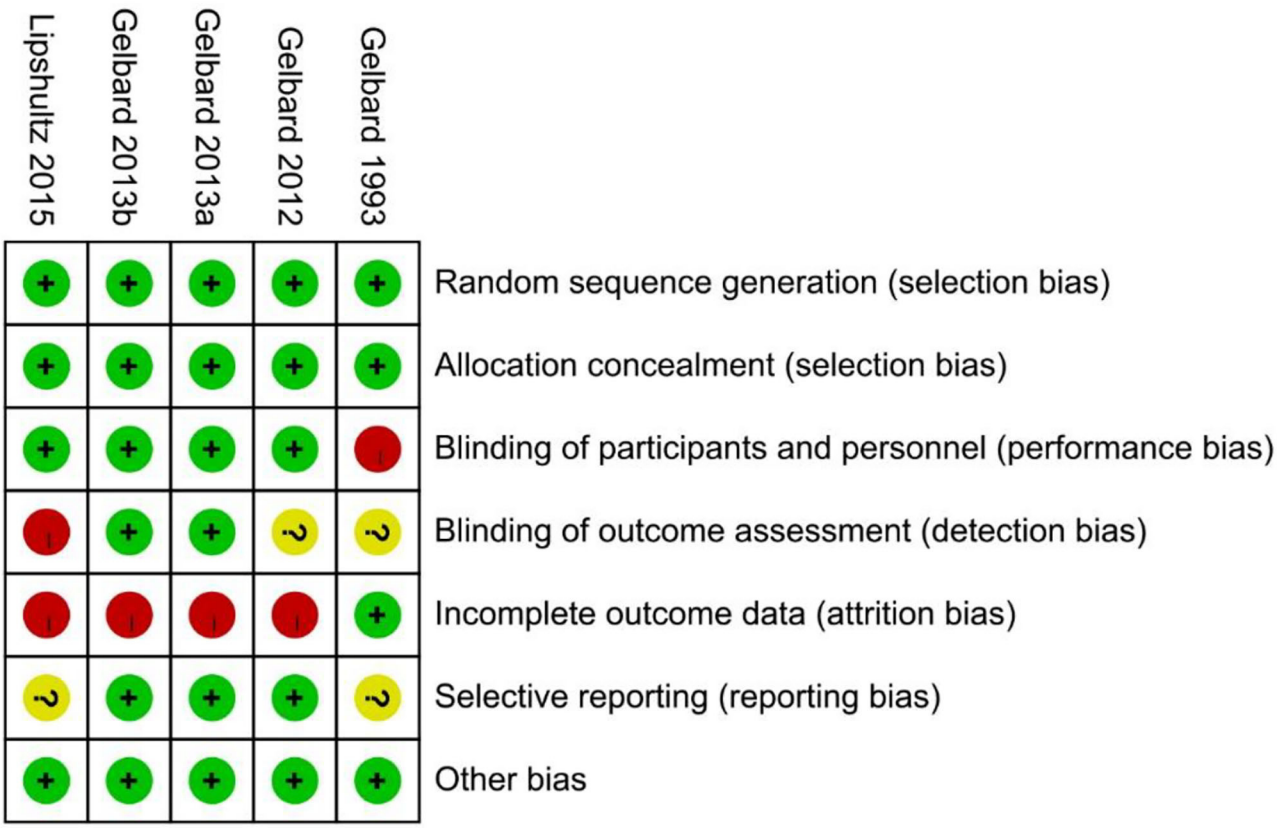

FIGURE 2 | Quality evaluation of included randomized controlled trials (RCTs).

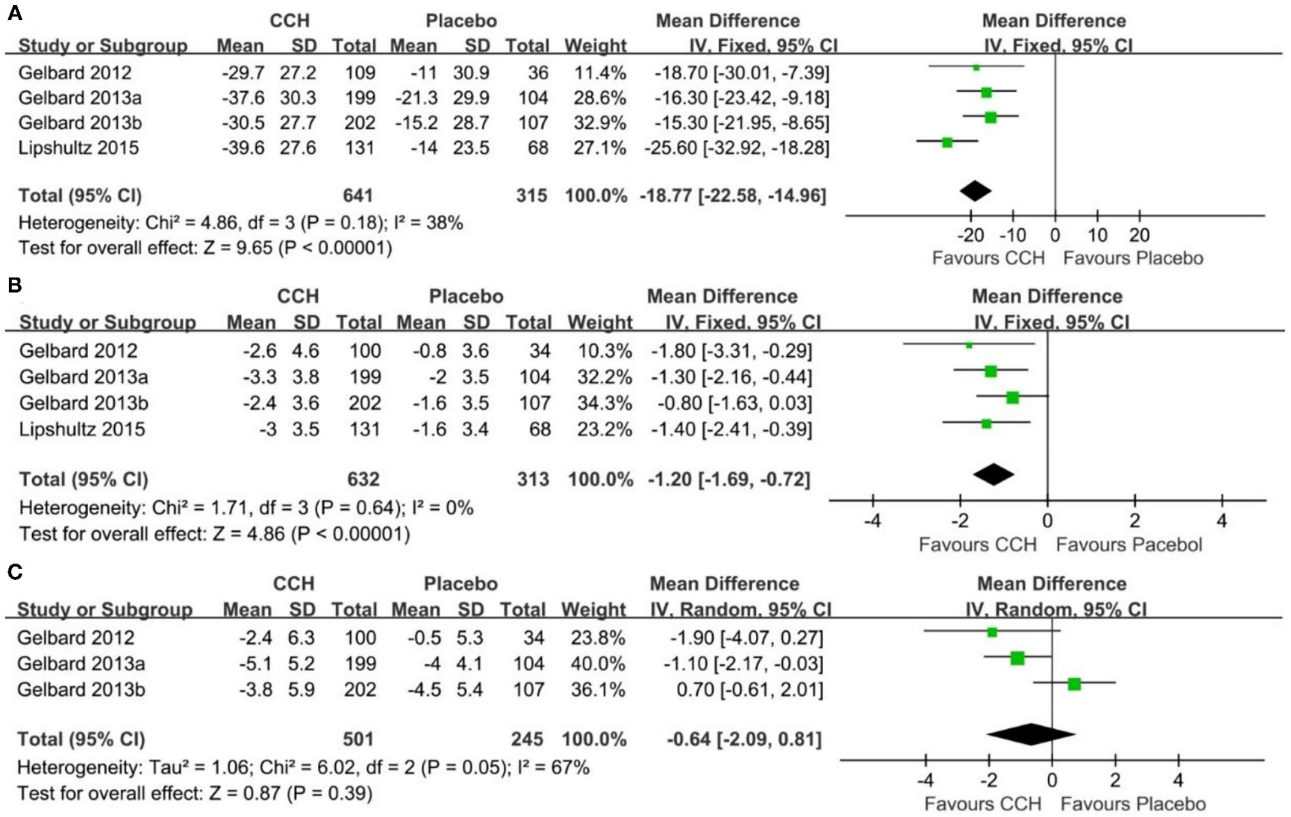

FIGURE 3 | Comparison between Collagenase Clostridium Histolyticum (CCH) and placebo in efficacy. **(A)** Penile curvature deformity, **(B)** Peyronie's disease symptom bothers, **(C)** penile pain scores.

Penile Pain Score

For the penile pain score, 3 included studies reported this outcome (19, 20). Heterogeneity test results indicated a high heterogeneity ($P = 0.05$, $I^2 = 67\%$) between the CCH group and the placebo group. The combined results showed no significant difference in the penile pain score between the CCH group and the placebo group (WMD: −0.64; 95% CI: −2.09 to 0.81; $P = 0.39$; **Figure 3C**).

Adverse Events

There were four studies recording the adverse event rate (18–20), and the merged results indicated that the TAEs of the CCH group were significantly higher than that of the placebo group (OR: 12.86; 95% CI: 9.17 to 18.04; $p < 0.001$; $I^2 = 0\%$; **Figure 4**). In addition, we conducted a meta-analysis of several common complications. In the final statistical analysis, the number of penile pain event in the CCH group was significantly higher, compared with the placebo group (OR: 8.87; 95% CI: 5.43 to 14.50; $p < 0.001$), and others showed similar results (penile edema: OR = 26.86, $p < 0.001$; injection site pain: OR = 7.91, $p < 0.001$; contusion: OR = 14.60, $p < 0.001$; **Table 2**).

DISCUSSION

The incidence of PD increases with the age, and it has been reported that 90% of patients with PD are between 50 and 59 years old (22). As a chronic benign disease, PD could cause erectile dysfunction, which seriously affects the life quality. Unfortunately, there is, currently, no ideal treatment for PD and, thus, made it a hot topic. The CCH, a mixture of AUX-I and AUX-II Clostridium collagenase with high selective hydrolysis activity on Type I and Type III collagen, can directly destroy collagen-based plaques without damaging surrounding elastic tissue and vascular smooth muscle. Gelbard et al. (18) first reported the application of CCH in PD, and they found that the therapeutic effect of CCH was superior to placebo in terms of plaque size and penile deformity. Subsequently, substantial evidence has shown that CCH is an effective non-surgical treatment for PD (13). However, the safety and efficacy of CCH in PD remain controversial. We, therefore, conducted a meta-analysis of the current data to provide a higher level of evidence. As far as we know, our study is the first meta-analysis worldwide that systematically illustrates the efficacy and safety of CCH in the treatment of PD.

The PCD is caused by the formation of fibrous plaques in the tunica albuginea, which may cause great trouble for patients with

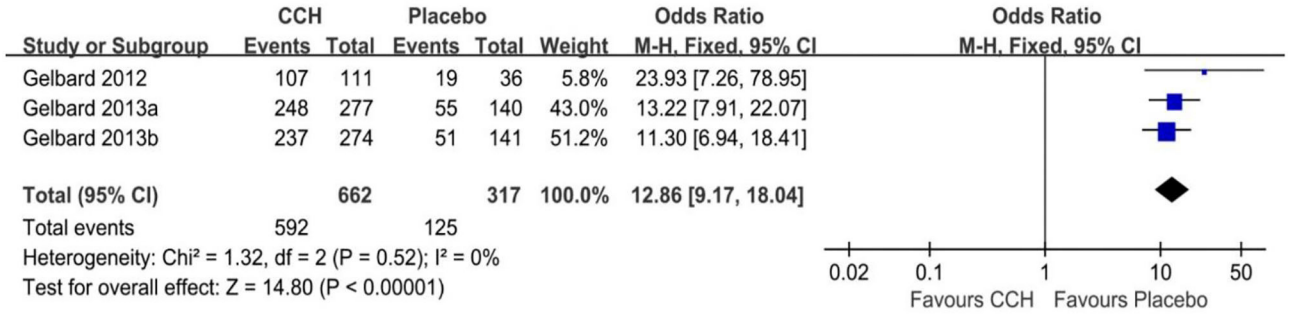

FIGURE 4 | Comparison between CCH and placebo in adverse events.

TABLE 2 | Comparison of common complications between CCH group and placebo group.

Outcomes	No. of studies	No. of patients (CCH/Placebo)	*P* value	OR (95% CI)	Heterogeneity			
					Chi^2	df	*P*	I^2 (%)
Penile pain	4	684/344	<0.001	8.87 [5.43, 14.50]	0.00	2	1.00	0
Penile edema	4	684/344	<0.001	26.86 [6.63, 108.80]	0.82	2	0.66	0
Injection site pain	4	684/344	<0.001	7.91 [4.38, 14.30]	0.38	2	0.83	0
Contusion	4	684/344	<0.001	14.60 [4.13, 51.68]	1.05	2	0.59	0

CCH, Collagenase Clostridium Histolyticum; OR, odds ratio.

PD in penetrative intercourse. A clinical trial found that, after two CCH injections, 69 patients had a significant improvement in mean percent change for PCD, and the average reduction in curvature is 23° (23). Evidence synthesis in our meta-analysis also indicated that CCH can significantly reduce PCD in patients with PD compared with placebo ($p < 0.001$), which was consistent with the results of the previous studies (21, 24, 25). Also, the man who is most likely to get PCD improvement is supposed to have curvature between 30° and 60°, longer duration, an international Index for Erectile Function score >17, no calcification, and a set to receive all standard cycles (26). On the other side, however, the practical significance of this improvement remains doubtful and requires further pieces of research, as none of the included studies resulted in a final mean curve <30°, which has been deemed unlikely to inhibit intercourse. Furthermore, for patients with severe curves (>90°), extended curves are recommended to receive surgical management to obtain a superior result (9).

The PD is characterized by penile abnormalities, coital pain, and impaired sexual function, followed by secondary psychological problems. The PDSB and the penile pain score focus on these issues and reflect subjective indicators. Our results confirmed the conclusion from previous studies that CCH can alleviate the PDSB of patients ($p < 0.001$). Increased tissue compliance and positive psychological suggestions for receiving non-surgical therapy could account for this phenomenon (9). In addition, one researcher (27) considered that the reduction of PCD, penile shortening, and pain during sexual intercourse can alleviate PD-related bothers. However, compared with placebo, CCH was not effective in reducing the penile pain score ($P = 0.39$), which may be associated with its side effect, inducing penile pain and bruising (28).

Safety is one of the most important indicators of all doctors' and patients' concerns about CCH treatment. Several studies have examined this topic, but the results were not the same. Tsambarlis et al. believed that patients with PD had a higher incidence of TAEs after receiving CCH, and 80% of the patients expressed a certain degree of dissatisfaction during treatment after therapy (29). On the contrary, several Phase 3 clinical trials have proven that patients treated with CCH have higher TAEs, but they were all acceptable, which were similar to our results. Moreover, the most common complications, such as penile pain, penile edema, injection site pain, and contusion, were slight and could recover without intervention. The occurrence of TAEs is also related to doctors' technical and postoperative care and, therefore, the clinical application should always be treated with caution. It is worth mentioning that corporal rupture, a serious adverse event, must be carefully dealt with to prevent its occurrence, as 34% of the providers encountered at least one case, and 67% of them managed the rupture surgically (30).

This study is the world's first meta-analysis of the efficacy and safety of CCH, and, thus, some limitations could not be avoided. First, although more reliable than the previous form, PDSB and the penile pain score were subjective indicators and could be biased by psychologic status in patients. Second, only 4 studies were included in our study, resulting in inadequate statistical confidence, and 3 of them shared the

same first author. However, all studies were randomized controlled trials with a high level of evidence and very low heterogeneity in all results, which could make up for this shortcoming.

CONCLUSION

The present meta-analysis demonstrated that CCH had a significant effect on treating PD. This method could relieve bothering symptoms and provide improvement in PCD and sexual function. Its adverse event rate was higher but acceptable. However, more RCTs with a larger sample size are needed to confirm our findings.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

LL and QW contributed to conception and design. DC and JL contributed to the acquisition of data and critical revision of the manuscript for important intellectual content. DC, JL, YL, YH, BC, ZC, and YS contributed to the analysis and interpretation of data. DC and JL contributed to the drafting of the manuscript. QW contributed to supervision. All authors read and agreed to the published version of the manuscript.

FUNDING

This study was funded by the National Natural Science Foundation of China (Grant No. 82000721), Post-Doctor Research Project, West China Hospital, Sichuan University (Grant No. 2019HXBH089), Health Commission of Sichuan province (Grant No. 20PJ036), and Programs from the Department of Science and Technology of Sichuan Province (Grant No. 2020YJ0054).

REFERENCES

1. Yafi FA, Pinsky MR, Sangkum P, Hellstrom WJ. Therapeutic advances in the treatment of Peyronie's disease. *Andrology.* (2015) 3:650–60. doi: 10.1111/andr.12058
2. Hughes WM, Natale C, Hellstrom WJG. The management of penile fracture: a review of the literature with special consideration for patients undergoing collagenase clostridium histolyticum injection therapy. *Curr Urol Rep.* (2021) 22:13. doi: 10.1007/s11934-020-01025-7
3. Schwarzer U, Sommer F, Klotz T, Braun M, Reifenrath B, Engelmann U. The prevalence of Peyronie's disease: results of a large survey. *BJU Int.* (2001) 88:727–30. doi: 10.1046/j.1464-4096.2001.02436.x
4. Mulhall JP, Creech SD, Boorjian SA, Ghaly S, Kim ED, Moty A, et al. Subjective and objective analysis of the prevalence of Peyronie's disease in a population of men presenting for prostate cancer screening. *J Urol.* (2004) 171:2350–3. doi: 10.1097/01.ju.0000127744.18878.f1
5. Dibenedetti DB, Nguyen D, Zografos L, Ziemiecki R, Zhou X. A population-based study of Peyronie's disease: prevalence and treatment patterns in the United States. *Adv Urol.* (2011) 2011:282503. doi: 10.1155/2011/282503
6. Chung E, Ralph D, Kagioglu A, Garaffa G, Shamsodini A, Bivalacqua T, et al. Evidence-based management guidelines on Peyronie's disease. *J Sex Med.* (2016) 13:905–23. doi: 10.1016/j.jsxm.2016.04.062
7. Capoccia E, Levine LA. Contemporary review of Peyronie's disease treatment. *Curr Urol Rep.* (2018) 19:51. doi: 10.1007/s11934-018-0800-5
8. Russo GI, Milenkovic U, Hellstrom W, Levine LA, Ralph D, Albersen M. Clinical efficacy of injection and mechanical therapy for Peyronie's disease: a systematic review of the literature. *Eur Urol.* (2018) 74:767–81. doi: 10.1016/j.eururo.2018.07.005
9. Tsambarlis P, Levine LA. Nonsurgical management of Peyronie's disease. *Nat Rev Urol.* (2018) 16:172–86. doi: 10.1038/s41585-018-0117-7
10. Sherer BA, Warrior K, Levine LA. 2013-2014 updates in Peyronie's disease management. *Curr Urol Rep.* (2014) 15:459. doi: 10.1007/s11934-014-0459-5
11. Levine LA. Peyronie's disease: a contemporary review of non-surgical treatment. *Arab J Urol.* (2013) 11:278–83. doi: 10.1016/j.aju.2013.03.008
12. Randhawa K, Shukla CJ. Non-invasive treatment in the management of Peyronie's disease. *Ther Adv Urol.* (2019) 11:1756287218823671. doi: 10.1177/1756287218823671
13. Abdel Raheem A, Johnson M, Ralph D, Garaffa G. Collagenase clostridium histolyticum: a novel medical treatment for Peyronie's disease. *Minerva Urol Nefrol.* (2018) 70:380–5. doi: 10.23736/S0393-2249.18.03118-1
14. Egui Rojo MA, Moncada Iribarren I, Carballido Rodriguez J, Martinez-Salamanca JI. Experience in the use of collagenase clostridium histolyticum in the management of Peyronie's disease: current data and future prospects. *Ther Adv Urol.* (2014) 6:192–7. doi: 10.1177/1756287214537331
15. Carson CC, Sadeghi-Nejad H, Tursi JP, Smith TM, Kaufman GJ, Gilbert K, et al. Analysis of the clinical safety of intralesional injection of collagenase Clostridium histolyticum (CCH) for adults with Peyronie's disease (PD). *BJU Int.* (2015) 116:815–22. doi: 10.1111/bju.13120
16. Cocci A, Russo GI, Briganti A, Salonia A, Cacciamani G, Capece M, et al. Predictors of treatment success after collagenase Clostridium histolyticum injection for Peyronie's disease: development of a nomogram from a multicentre single-arm, non-placebo controlled clinical study. *BJU Int.* (2018) 122:680–7. doi: 10.1111/bju.14410
17. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Bmj.* (2021) 372:n71. doi: 10.1136/bmj.n71
18. Gelbard MK, James K, Riach P, Dorey F. Collagenase versus placebo in the treatment of Peyronie's disease: a double-blind study. *J Urol.* (1993) 149:56–8. doi: 10.1016/S0022-5347(17)35998-0
19. Gelbard M, Lipshultz LI, Tursi J, Smith T, Kaufman G, Levine LA. Phase 2b study of the clinical efficacy and safety of collagenase Clostridium histolyticum in patients with Peyronie disease. *J Urol.* (2012) 187:2268–74. doi: 10.1016/j.juro.2012.01.032
20. Gelbard M, Goldstein I, Hellstrom WJ, McMahon CG, Smith T, Tursi J, et al. 3rd: Clinical efficacy, safety and tolerability of collagenase clostridium histolyticum for the treatment of peyronie disease in 2 large double-blind, randomized, placebo controlled phase 3 studies. *J Urol.* (2013) 190:199–207. doi: 10.1016/j.juro.2013.01.087
21. Lipshultz LI, Goldstein I, Seftel AD, Kaufman GJ, Smith TM, Tursi JP, et al. Clinical efficacy of collagenase Clostridium histolyticum in the treatment of Peyronie's disease by subgroup: results from two large, double-blind, randomized, placebo-controlled, phase III studies. *BJU Int.* (2015) 116:650–6. doi: 10.1111/bju.13096
22. Herati AS, Pastuszak AW. The genetic basis of peyronie disease: a review. *Sex Med Rev.* (2016) 4:85–94. doi: 10.1016/j.sxmr.2015.10.002
23. Ziegelmann MJ, Viers BR, McAlvany KL, Bailey GC, Savage JB, Trost LW. Restoration of penile function and patient satisfaction with intralesional collagenase clostridium histolyticum injection for Peyronie's disease. *J Urol.* (2016) 195:1051–6. doi: 10.1016/j.juro.2015.10.065
24. Hellstrom WJG, Gittleman M, Tan RBW, Tursi JP, Smith T, Kaufman GJ, et al. Pd22-02 meaningful change in peyronie's disease following treatment with collagenase clostridium histolyticum: results from two large double-blind, randomized, placebo-controlled phase 3 studies. *J Urol.* (2014) 12:125–6. doi: 10.1016/j.juro.2014.02.1853

25. Hellstrom WJ, Feldman RA, Coyne KS, Kaufman GJ, Smith TM, Tursi JP, et al. Self-report and clinical response to peyronie's disease treatment: Peyronie's disease questionnaire results from 2 large double-blind, randomized, placebo-controlled phase 3 studies. *Urology.* (2015) 86:291–8. doi: 10.1016/j.urology.2015.04.047
26. Masterson TA, Rezk A, Ramasamy R. Characteristics predictive of response to collagenase clostridium histolyticum for Peyronie's disease: a review of the literature. *World J Urol.* (2020) 38:279–85. doi: 10.1007/s00345-019-02850-3
27. Serefoglu EC, Smith TM, Kaufman GJ, Liu G, Yafi FA, Hellstrom WJG. Factors associated with erectile dysfunction and the Peyronie's disease questionnaire in patients with peyronie disease. *Urology.* (2017) 107:155–60. doi: 10.1016/j.urology.2017.05.029
28. Gaffney CD, Kashanian JA. Peyronie disease. *Jama.* (2020) 324:2566. doi: 10.1001/jama.2020.14819
29. Tsambarlis PN, Yong R, Levine LA. Limited success with clostridium collagenase histolyticum following FDA approval for the treatment of Peyronie's disease. *Int J Impot Res.* (2019) 31:15–9. doi: 10.1038/s41443-018-0063-1
30. Yafi FA, Anaissie J, Zurawin J, Sikka SC, Hellstrom WJ. Results of SMSNA survey regarding complications following intralesional injection therapy with collagenase clostridium histolyticum for Peyronie's disease. *J Sex Med.* (2016) 13:684–9. doi: 10.1016/j.jsxm.2016.02.105

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Cao, Li, Lu, Huang, Chen, Chen, Shen, Liu and Wei. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 28 February 2022
doi: 10.3389/fphar.2022.832078

The Impact of a Multifaceted Pharmacist-Led Antimicrobial Stewardship Program on Antibiotic Use: Evidence From a Quasi-Experimental Study in the Department of Vascular and Interventional Radiology in a Chinese Tertiary Hospital

Jinhui Xu[1†], Jian Huang[2†], YanXia Yu[1†], Dayong Zhou[2], Ying Wang[3], Sudong Xue[1], Erning Shang[1], Jiantong Sun[1], Xinyuan Ding[1], Lu Shi[1], Lufen Duan[1], Lian Tang[1], Qin Zhou[1]* and Xin Li[4,5]**

[1]Department of Pharmacy, The Affiliated Suzhou Hospital of Nanjing Medical University, Suzhou, China, [2]Department of Vascular and Interventional Radiology, The Affiliated Suzhou Hospital of Nanjing Medical University, Suzhou, China, [3]Department of Education and Training, The First Affiliated Hospital of SooChow University, SuZhou, China, [4]Department of Clinical Pharmacy, School of Pharmacy, Nanjing Medical University, Nanjing, China, [5]Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing, China

Objective: The objective of the study was to assess the impact of multifaceted clinical pharmacist-led antimicrobial stewardship (AMS) program on the rational use of antibiotics for patients who receive vascular and interventional radiology therapies.

Methods: A quasi-experimental retrospective intervention design with a comparison group was applied to the practice of antibiotic use in the department of vascular and interventional radiology in a Chinese tertiary hospital. We used difference-in-differences (DID) analysis to compare outcomes before and after the AMS intervention between the intervention group and control group, to determine whether intervention would lead to changes in irrationality of antibiotic prescribing, antibiotic utilization, cost of antibiotics, and length of hospital stay.

Results: The DID results showed that the intervention group was associated with a reduction in the average consumption of antibiotics ($p = 0.017$) and cost of antibiotics ($p = 0.006$) and cost per defined daily dose (DDD) ($p = 0.000$). There were no significant differences in the mean change of total costs and length of stay between the two groups ($p > 0.05$). The average inappropriate score of perioperative antimicrobial prophylaxis in the intervention group declined by 0.23, while it decreased by 0.02 in the control group [0.21 (95% CI, −0.271 to −0.143); $p = 0.000$]. The average inappropriate score of non-surgical antimicrobial prophylaxis in the intervention group declined by 0.14, while it increased by 0.02 in the control group [0.16 (95% CI, −0.288 to −0.035); $p = 0.010$]. The average

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Yan Song,
Shandong Academy of Medical Sciences, China
Yifan Zheng,
The First Affiliated Hospital of Sun Yat-sen University, China

***Correspondence:**
Xin Li
xinli@njmu.edu.cn
Qin Zhou
zhouyx1223@yeah.net
Lian Tang
tanglian716@aliyun.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Drugs Outcomes Research and Policies,
a section of the journal Frontiers in Pharmacology

Received: 09 December 2021
Accepted: 13 January 2022
Published: 28 February 2022

Citation:
Xu J, Huang J, Yu Y, Zhou D, Wang Y, Xue S, Shang E, Sun J, Ding X, Shi L, Duan L, Tang L, Zhou Q and Li X (2022) The Impact of a Multifaceted Pharmacist-Led Antimicrobial Stewardship Program on Antibiotic Use: Evidence From a Quasi-Experimental Study in the Department of Vascular and Interventional Radiology in a Chinese Tertiary Hospital. Front. Pharmacol. 13:832078. doi: 10.3389/fphar.2022.832078

inappropriate score of the therapeutic use of antibiotics in the intervention group declined by 0.07, while it decreased by 0.01 in the control group [0.06 (95% CI, −0.115 to −0.022); p = 0.003].

Conclusions: This study provides evidence that implementation of AMS interventions was associated with a marked reduction of antibiotic use, cost of antibiotics, and irrationality of antibiotic prescribing in China.

Keywords: antibiotics use, antimicrobial stewardship, pharmacist, difference-in-differences, vascular and interventional radiology

INTRODUCTION

By making lethal infections readily treatable, antibiotics have led the development of modern medicine. Prompt initiation of appropriate antibiotic therapy to treat infectious diseases reduces morbidity and save lives. Moreover, antibiotics also enables multiple advances in the area of medical technology, such as organ transplants and cancer chemotherapy (Beyar-Katz et al., 2017; Versporten et al., 2018). However, about 30%–50% of all antibiotics prescribed in hospitals are either unnecessary or suboptimal worldwide (Denny et al., 2019; Rowe and Linder, 2019). In recent years, antimicrobial resistance (AMR) has been recognized as a growing public health problem of huge concern to countries around the world (Huemer et al., 2020). Inappropriate antibiotic use contributes greatly to the development of AMR (Bohmer et al., 2021), leading to ineffective therapy, increase in adverse reactions, and escalation in health care costs. In order to optimize antibiotic use in medical care, many different antimicrobial stewardship (AMS) initiatives have been conducted in some developed countries recently. Several studies have shown that AMS can help physicians increase infection cure rates, improve treatment outcomes, and reduce antibiotic resistance by improving antibiotic prescribing (Dellit et al., 2007; Rice, 2018).

According to the Centers for Disease Control and Prevention (CDC) of the United States, core elements of hospital AMS programs are as follows: 1) Hospital leadership commitment. AMS programs should be supported from the senior leadership of the hospital, especially the director of pharmacy. 2) Pharmacy expertise. AMS programs should appoint a physician or pharmacist as the coleader to improve antibiotic use. 3) Accountability. As a leader or coleaders, a physician or pharmacist should be responsible for program management. 4) Interventions. The prospective audit and feedback or preauthorization should be implemented to improve antibiotic use. 5) Education. AMS should educate clinicians, pharmacists, nurses, and patients about adverse reactions caused by antibiotics, AMR, and optimal prescribing (CDC, 2019).

A recent study found that 25% (21%–29%) of antibiotic prescriptions in the United States ambulatory care clinics were inappropriate in 2015 (Ray et al., 2019). The rates of inappropriate prescriptions were highest for otitis externa (67.3%) and upper respiratory tract infection (38.7%) in the United Kingdom general practice between January 2010 and June 2015 (Nowakowska et al., 2019). Due to its huge antibiotic consumption (Zhao et al., 2018), strong economic incentives (Lin et al., 2020), and limited knowledge about antibiotics (Liu et al., 2019) for overprescribing, inappropriate antibiotic use in China is common. The proportion of inappropriate antibiotic prescriptions among outpatients and inpatients was 60.6% and 75.4% in primary health care settings in China from 2009 through 2011, respectively (Wang et al., 2014). A recent survey conducted on the use of antibiotics in 52 hospitals of China reported that the proportion of antibiotic prescription in surgical inpatients in 2016 was 69.03% (Li et al., 2019), which did not meet the goal of not more than 60% set by the National Health Commission of the People's Republic of China (National Health Commission of China, 2011). This report also showed that the percentage of antibiotic use in inpatients for acute upper respiratory infections and fever or cough symptoms all exceeded 55%, which was unnecessary for these infectious diseases mostly caused by viruses (WHO, 2011).

Thus, to promote rational use of antibiotics, it is necessary to implement AMS programs in hospitals. In 2011, AMS policy was formally implemented by the National Health Authority in medical institutions nationwide. Successful implementation of AMS requires a multidisciplinary organizational approach, with pharmacists as the core members that deliver the interventions to optimize antimicrobial use.

Several studies provided evidence that pharmacist-led interventions can improve adherence to guidelines and reduced the duration of antimicrobial therapy with substantial cost savings for patients (Mahmoudi et al., 2019; Monmaturapoj et al., 2021). Furthermore, implementation of pharmacy-based interventions in AMS programs can reduce antibiotic consumption, length of hospital day, and 30-day mortality (Xin et al., 2019; Du et al., 2020).

Previous studies assessed the effectiveness of AMS intervention on surgical antibiotic prophylaxis (Zhang et al., 2014; Abubakar et al., 2019), upper or lower respiratory tract infection (Shen et al., 2011; Foolad et al., 2018), nonspecified indications (Li et al., 2017; May et al., 2021) in those priority departments, such as the department of urology, department of obstetrics and gynecology, department of respiratory medicine, intensive care unit, and emergency department, given high antimicrobial prescribing for prophylaxis or treatment and concerns with inappropriate antibiotic use. However, the previous studies pay little attention to some emerging departments with innovative technology; for instance, as the number and breadth of vascular and interventional radiology

procedures grow, it is critical to promote rational use of antibiotics in the department of vascular and interventional radiology to avoid serious infectious complications. So far, only one study on the use of antibiotics for perioperative interventional surgery was retrieved. Yang et al. (2017) conducted a clinical pharmacist-led guidance team intervention for the rationality of prophylactic antibiotic usage for eight interventional procedures in a Chinese tertiary teaching hospital. However, it did not focus on a specific department, and the intervention included mainly with prescription reviews and administrative approach. Moreover, it was only limited to perioperative antimicrobial prophylaxis, rather than comprehensive intervention involving antibiotic prophylaxis and antibiotic therapy in surgical wards. The research design of the study was that of a before-and-after intervention trial to evaluate the effectiveness.

To date, there is no evidence to demonstrate the effectiveness of pharmacist-led AMS interventions on the rational use of antibiotics in the department of vascular and interventional radiology in a Chinese tertiary hospital. Therefore, the aim of this study was to improve the rational antibiotics use in the department of vascular and interventional radiology, and assess the value of pharmacist-led AMS programs during the study period. For this purpose, we used a quasi-experimental study design with a difference-in-differences (DID) analysis comparing outcomes before and after the AMS intervention between intervention group and control group, to determine whether intervention would lead to changes in the following outcomes: irrationality of antibiotic prescribing, antibiotic utilization, cost of antibiotics, and length of hospital stay.

MATERIALS AND METHODS

Study Setting

This was a quasi-experimental retrospective study, with a comparison group and pre- and post-intervention measures in two independent vascular and interventional radiology wards of a tertiary teaching hospital. The two independent wards were randomized into control and intervention groups. In the intervention group, a clinical pharmacist-led AMS was implemented. While in the control group, treatment strategies were performed by the physicians and nurses without pharmacist involvement.

To ensure comparability of the statistical data, inclusion and exclusion criteria were established. Diseases involving the circulatory system (such as superficial thrombophlebitis of the lower extremities, aortic aneurysm or dissection, etc.), digestive system (such as biliary obstruction, cirrhosis with esophageal and gastric variceal bleeding, etc.), respiratory system (bronchiectasis, pneumonia, etc.), genitourinary system (such as urinary tract infection, kidney failure, etc.), tumor (liver malignant tumor), injury, poisoning, certain other consequences of external causes (such as complications of procedures, injury of kidney, etc.), the blood and blood-forming organs, and certain disorders involving the immune mechanism, skin, and subcutaneous tissue, and congenital malformation (such as cellulitis, polycystic kidney, hypersplenism, etc.) were enrolled. Patients were excluded if they were transferred from other medical departments, were transferred to other medical departments for further treatment, and ages were younger than 18 years or data were missing. Accordingly, in our retrospective study, phases were divided into two stages as follows: pre-intervention period (March 2018 to October 2018) and post-intervention period (March 2019 to October 2019).

Ethical Consideration

The retrospective study complied with the Declaration of Helsinki, and ethical approval was obtained from the Ethical Committee of Nanjing Medical University (grant number: ethical review 2020103).

Pharmacist-Led Antimicrobial Stewardship Interventions

Based on the core elements of AMS (CDC, 2019), the multifaceted interventions included: 1) Establishment of criteria: provided practitioners with a standardized approach to the safe and effective use of antibiotic drugs based on global and national guidelines. 2) Daily ward round: participated in physicians' ward rounds at working days only and made recommendations on the antimicrobial prescribing. 3) Regular review of medical orders: communicated with the physicians in cases of inappropriate drug use and made suggestions to determine the optimal use of antibiotics. 4) Routine education and training: trained the medical staff on rational use of antibiotics and correct microbiology specimen submission. 5) Consultations and case discussions: conducted regular multidisciplinary consultations and case seminars for difficult cases to discuss diagnosis and modify the regimen.

Data Collection and Outcomes

The primary outcome was the score of inappropriate antimicrobial use between pre-intervention and post-intervention period, which was evaluated by the multidisciplinary team. The team consisted of an infectious disease specialist, a ward physician, and a clinical pharmacist, who were blinded to the patients' allocation status. The appropriateness of antimicrobial use was assessed by the three specialists, respectively. If the score was the same, the result would be adopted, when there was any disagreement, final decision would be made after a discussion.

Secondary outcomes included antimicrobial consumption, antimicrobial cost, the average cost per defined daily dose (DDD), the total cost of hospitalization, and length of hospital stay per patient during the periods. Antimicrobial consumption was defined as the defined daily doses (DDDs) of antibiotics used per patient that was applied in previous studies of antibiotic utilization (Yang et al., 2017; Abubakar et al., 2019). The average cost per DDD was calculated by dividing the antimicrobial cost by DDDs of antibiotics used per patient.

Clinical data including patients' baseline characteristics, antibiotic usage, cost, length of hospital stay, and clinical outcome were extracted from the hospital information system

(HIS). The HIS recorded the clinical outcome of the discharged patients, which were classified as cure, improvement, failure, and death. Interventions identified by the pharmacist-led AMS program and the accepted recommendations were collected within the electronic medical record. The use of antibiotics was stratified into three indication categories: perioperative antimicrobial prophylaxis, non-surgical antimicrobial prophylaxis, and therapeutic use of antibiotics. The specialists evaluated the phenomena of inappropriate drug use using a scoring system that consisted of different items regarding the three categories. Each item was assigned one point if an inappropriate phenomenon was identified. At the same time, the score was multiplied by the weight of each irrational medication indicator, which was developed by 30 experts' consultation using the Saaty 1–9 scaling method (Saaty, 1990).

Statistical Analysis

In this study, the variations in an index of the two groups before and after intervention are calculated to reflect the net effect of clinical pharmacist-led AMS strategy. A DID methodology (Wooldrige, 2002; Dimick and Ryan, 2014) was used to compare pre- and post-intervention changes between the AMS intervention group and the control group. The differences lie in the cluster-level summaries of the two groups. Comparisons were carried out using a mixed effect linear model with repeated measures using group, time, and group × time as fixed effects, while group is a random effect nested within the study group. The descriptive statistics of cost, DDDs, score, and days are presented as means (by time and by group), and mean changes by group. The DID methodology allows the researchers to control confounding influences of independent variables, as well as imbalance between groups in dependent variables, which are due to chances of imperfect randomization. For each group, the incremental effect of the pharmacist-led AMS intervention was computed as the differences between the pre- and post-implementation of intervention, and then the incremental effects were compared between the intervention and control groups to estimate the net effect of the intervention. The model that was used for this DID regression is written as:

$$Y_{it} = \alpha_0 + \beta \times time_{it} \times group_{it} + \gamma \times X_{it} + \varepsilon_{it}$$

where i indicates the vascular and interventional radiology ward, and t indicates the time. Y represents the main outcomes, either score of inappropriate antimicrobial use, antimicrobial consumption, cost, or length of hospital stay; $time_{it}$ is the time dummy variable, where it was measured at the pre-intervention period as 0 and post-intervention period as 1; $group_{it}$ is the treatment dummy variable, where the intervention group is 1, while the control group is 0; $time_{it} \times group_{it}$ is an interaction term between the time dummy variable and treatment dummy variable, and its coefficient β captured the average effects of the pharmacist-led AMS strategy on the outcome indicators. X_{it} is a series of covariates, including age, sex, clinical outcome, and diagnosis. ε_{it} is a residual error. Each outcome was regressed on the main effects for intervention (1 = intervention group; 0 = control group) and time period (1 = post-intervention, 0 = pre-intervention). In this procedure, the DID model was used to predict the expected value at baseline and follow-up in the intervention and control groups with all other fixed effects set to their mean values.

Continuous variables that were normally distributed were presented as the mean ± standard deviation, and group differences were compared using an independent t-test. Non-normally distributed continuous variables were described in the median and 25th and 75th interquartile range (Q25, Q75) and were assessed using Mann–Whitney U-test. Categorical variables were presented as frequencies and percentages and assessed using the Chi-squared test or Fisher's exact test. All statistical analyses were conducted using SPSS Statistics 23.0 (IBM Corp., Armonk, NY, USA), and differences with $p < 0.05$ were considered statistically significant.

RESULTS

Patient Characteristics

A total of 1,026 patients were enrolled in the study, 514 for the control group and 512 for the intervention group. There were no significant differences in age, gender, number of diagnoses, and clinical outcome of the admitted patients between the two groups (**Table 1**).

The Effects of Antimicrobial Stewardship Program on the use of Antibacterial Drugs

As shown in **Figure 1**, there were no differences between the two groups in the trends of DDDs per patient, cost of antibiotics, and average cost per DDD during the pre-intervention period ($p = 0.588$, $p = 0.285$, and $p = 0.876$), while during the post-intervention period, significant differences in the trends between the two groups were observed ($p = 0.025$, $p = 0.014$, and $p = 0.000$, respectively, for DDDs per patient, cost of antibiotics, and average cost per DDD). The effects of the intervention on antibiotic consumption, costs, and length of hospital stay in the control and intervention groups between the pre-intervention and post-intervention period are shown in **Tables 2** and **3**. The DID results suggested that the intervention group was associated with a reduction in the average consumption of antibiotics [1.84 DDDs (95% CI, −3.433 to −0.337); $p = 0.017$], cost of antibiotics [598.19 RMB (95% CI, −1,018.558 to −173.554); $p = 0.006$], and cost per DDD [70.5 RMB (95% CI, −101.086 to −39.915); $p = 0.000$]. There were no significant differences in mean change of total costs and length of stay between the two groups.

Construction of the Scoring System Involving Weights of Different Items

Table 4 presents the detailed results of the scoring system for perioperative antimicrobial prophylaxis, non-surgical antimicrobial prophylaxis, and therapeutic use of antibiotics. Among the items of the scoring system for perioperative antimicrobial prophylaxis and non-surgical antimicrobial prophylaxis, the weight of indication is the highest. On the

TABLE 1 | Comparison of the characteristics of patients in the control group and intervention group.

Characteristic	Control group (*n* = 514)	Intervention group (*n* = 512)	*p*-Value
Age, year	67 (57,75)	67 (58,75)	0.607[a]
Male, n (%)	332 (64.59%)	323 (63.08%)	0.616
Diagnosis, n (%)			
Circulatory system	181 (35.21%)	161 (31.44%)	0.200
Digestive system	139 (27.04%)	161 (31.44%)	0.121
Respiratory system	81 (15.76%)	69 (13.48%)	0.301
Genitourinary system	60 (11.67%)	66 (12.89%)	0.552
Tumor	31 (6.03%)	29 (5.66%)	0.802
Injury, poisoning, and certain other consequences of external causes	20 (3.89%)	23 (4.49%)	0.631
Others[b]	14 (2.72%)	19 (3.71%)	0.370
Multiple diagnosis	12 (2.33%)	16 (3.12%)	0.437
Clinical outcome			0.993
Cure	504	502	
Improvement	10	10	
Failure	0	0	
Death	0	0	

Note. Values are presented as median (interquartile range) or n (%). All p*-values are calculated using Chi-square unless otherwise noted.*
[a]*Mann–Whitney U-test.*
[b]*Others: the blood and blood-forming organs and certain disorders involving the immune mechanism; skin and subcutaneous tissue; and congenital malformation.*

other hand, timing of the initial dose and duration is assigned a lower weight. Regarding therapeutic use of antibiotics, the weight of indication is the highest, followed by choice, and dosage in its scoring system.

Figure 2 shows the detailed category analysis for irrational use of antibacterial drugs. The highest inappropriate score of items for the intervention group was antibiotic choice (0.12), antibiotic choice (0.15), and indication (0.11) in perioperative prophylaxis, non-surgical prophylaxis, and treatment during the pre-intervention period, respectively. After the intervention, inappropriate score of the corresponding items declined by 0.1, 0.06, and 0.05, respectively. Considering the antibiotic choice, as reported in **Supplementary Table S1**, the most frequently used antibiotics for perioperative prophylaxis in the intervention group before the intervention was cefathiamidine (37.80%). However, the use of cefathiamidine dropped to 1.90% during the post-intervention period. Meanwhile, cefazolin became the most frequently prescribed antibiotics, which increased from 2.44% to 50.48%. In terms of non-surgical prophylaxis, there was an increase in ceftriaxone use from 0 to 37.04% with a decrease in cefmetazole from 53.85% to 18.52% in the intervention group after the intervention.

Evaluation of the Rationality of Antibiotic use

The inappropriate rate of antibiotic prescriptions, average inappropriate score for both groups, and associated parameters of DID analysis between the groups in perioperative antimicrobial prophylaxis, non-surgical antimicrobial prophylaxis, and therapeutic use of antibiotics during the pre-intervention and post-intervention period are shown in **Tables 5** and **6**. For the intervention group, inappropriate rate of antibiotic prescriptions in perioperative prophylaxis, non-surgical prophylaxis, and treatment decreased by 48.10%, 40.59%, and 29.25%, respectively. For the control group, inappropriate rate of antibiotic prescriptions in perioperative prophylaxis and non-surgical prophylaxis increased by 4.35% and 2.77%, respectively, while a 2.49% decrease in inappropriate rate of antibiotic prescriptions regarding treatment. During the post-intervention period, the average inappropriate score of perioperative antimicrobial prophylaxis in the intervention group declined by 0.23, while it decreased by 0.02 in the control group [0.21 (95% CI, −0.271 to −0.143); $p = 0.000$]. The average inappropriate score of non-surgical antimicrobial prophylaxis in the intervention group declined by 0.14, while it increased by 0.02 in the control group [0.16 (95% CI, −0.288 to −0.035); $p = 0.010$]. In terms of therapeutic use of antibiotics, the average inappropriate score of the intervention group declined by 0.07, while it decreased by 0.01 in the control group [0.06 (95% CI, −0.115 to −0.022); $p = 0.003$].

Pharmacists' Recommendations Made in the Post-intervention Period

Over the 8-month post-implementation of a pharmacist-led AMS program, a total of 445 interventions were recommended, of which 321 (72.13%) were from medical order review led by clinical pharmacists (**Table 7**). Overall, 445 interventions were made with an average acceptance rate of 80.00% (356/445). The highest proportion of intervention types in perioperative antimicrobial prophylaxis was stopping due to prolonged duration (34.97%) with an acceptance rate of 66.67%. Among the recommendations for non-surgical prophylaxis, the most frequent type was initiating a preferred agent for indication, which had an acceptance rate of 47.62%. The most common intervention types for antibiotic treatment were discontinuation of antibiotics without indication (23.58%), followed by dose or frequency adjustment (19.10%), and de-escalation (13.82%). The acceptance rates were 67.24%, 91.49%, and 88.24%, respectively.

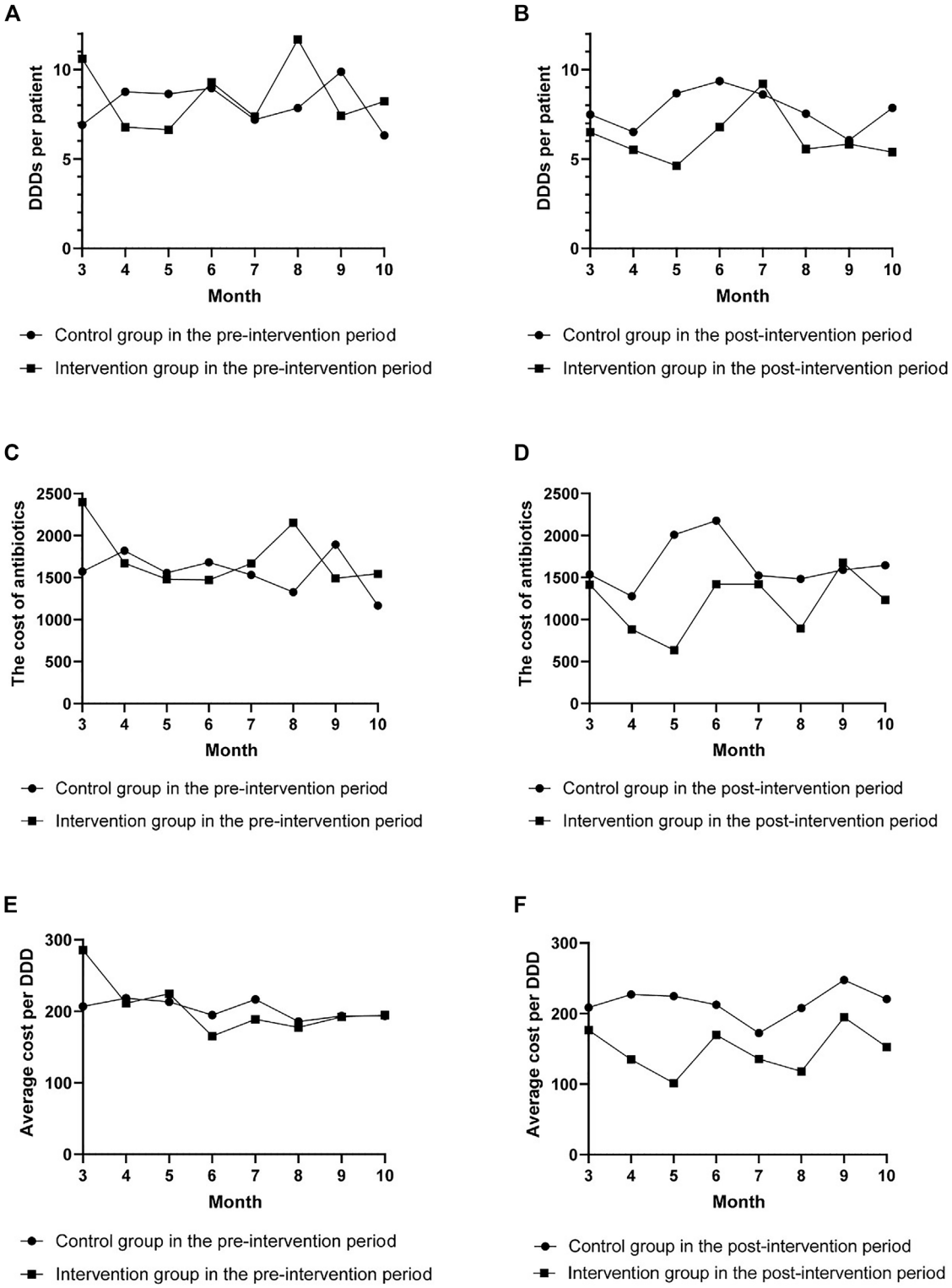

FIGURE 1 | The change in defined daily doses (DDDs) per patient, cost of antibiotics, and average cost per DDD per month. The change in DDDs per patient **(A)**, cost of antibiotics **(C)**, and average cost per DDD **(E)** per month in the pre-intervention period. The change in DDDs per patient **(B)**, cost of antibiotics **(D)**, and average cost per DDD **(F)** per month in the post-intervention period.

TABLE 2 | The mean antibiotic consumption, costs, and length of hospital stay in the control and intervention groups between the pre-intervention and post-intervention periods.

Outcomes	Pre-intervention		Post-intervention	
	Control group	Intervention group	Control group	Intervention group
DDDs per patient (DDDs)	7.93	8.29	7.67	6.19
The cost of antibiotics (RMB)	1546.99	1706.27	1624.51	1185.60
Total cost (RMB)	34529.55	27342.66	30792.38	30366.09
Average cost per DDD (RMB)	202.45	204.07	216.01	147.13
Length of stay (days)	10.85	11.30	10.92	12.38

TABLE 3 | Difference-in-differences (DID) results of the net effects of antimicrobial stewardship (AMS) program on antibiotic consumption, costs, and length of hospital stay.

Outcomes	Coefficient	*p*-Value	95% CI	Robust SE
DDDs per patient	−1.885	0.017	(−3.433, 0.337)	0.789
The cost of antibiotics	−596.056	0.006	(−1018.558, −173.554)	215.305
Total cost	2320.662	0.476	(−4065.250, 8706.575)	3254.235
Average cost per DDD	−70.501	0.000	(−101.086, −39.915)	15.586
Length of stay	0.813	0.375	(−0.986, 2.611)	0.916

Note. DID, difference-in-differences.

TABLE 4 | The scoring system involving weights of different items.

Items	Perioperative antimicrobial prophylaxis	Non-surgical antimicrobial prophylaxis	Therapeutic use of antibiotics
Indication	0.41	0.57	0.37
Choice	0.16	0.21	0.23
Timing of the initial dose	0.07	NA	NA
Dosage and dosing interval	0.11	0.13	NA
Dosage	NA	NA	0.12
Dosing frequency	NA	NA	0.10
Duration	0.25	0.09	0.06
Intravenous-to-oral conversion	NA	NA	0.05
Combination of antibiotics	NA	NA	0.07

Note. NA, not available.

DISCUSSION

To our knowledge, this is the first study to evaluate the impact of a pharmacist-led AMS program in the department of vascular and interventional radiology. The DID analysis indicated that the multifaceted pharmacist-led AMS intervention was associated with a significant decline in the antibiotic use, cost of antibiotics, and irrationality of antibiotic prescribing. Comparison of pre- and post-intervention changes between the intervention group and the control group, rather than merely levels before and after the intervention in the intervention group, necessarily adjusted for the possibility that the outcome was changing without the benefit of the intervention (Gerber et al., 2013).

There was a significant reduction in antibiotic utilization after the interventions in our study. This decline was attributed to reduction of unnecessary antimicrobial prescribing and increase in compliance with duration of antibiotic use. This observation was consistent with studies conducted in hospital and clinic settings (Hwang and Kim, 2018; Lee et al., 2018; Abubakar et al., 2019; Haseeb et al., 2020). The current study also found that the implementation of antibiotic stewardship interventions led to cost savings of antibiotics, which was similar to that found in prior studies (Butt et al., 2019; Xin et al., 2019). Some studies reported that the length of hospital stay significantly was shortened after the intervention of a pharmacist-directed antimicrobial stewardship (Mahmoudi et al., 2019; Xin et al., 2019; Arensman et al., 2020; Mahrous et al., 2020), while some studies indicated that the intervention had no effect on the length of stay (Wenzler et al., 2017; Abubakar et al., 2019; Monmaturapoj et al., 2021), which may be related to differences in the study design and analytical methods.

Inappropriate antibiotic prescribing was evaluated using a scoring system consisting of different items regarding three categories. Compared with the scoring system constructed by Shen et al. (2011), the relative weight values of each item were determined on the basis of utilizing the Saaty1–9 scale method, which can quantitatively judge the importance of each item on the inappropriate antibiotic prescribing. The result implies that

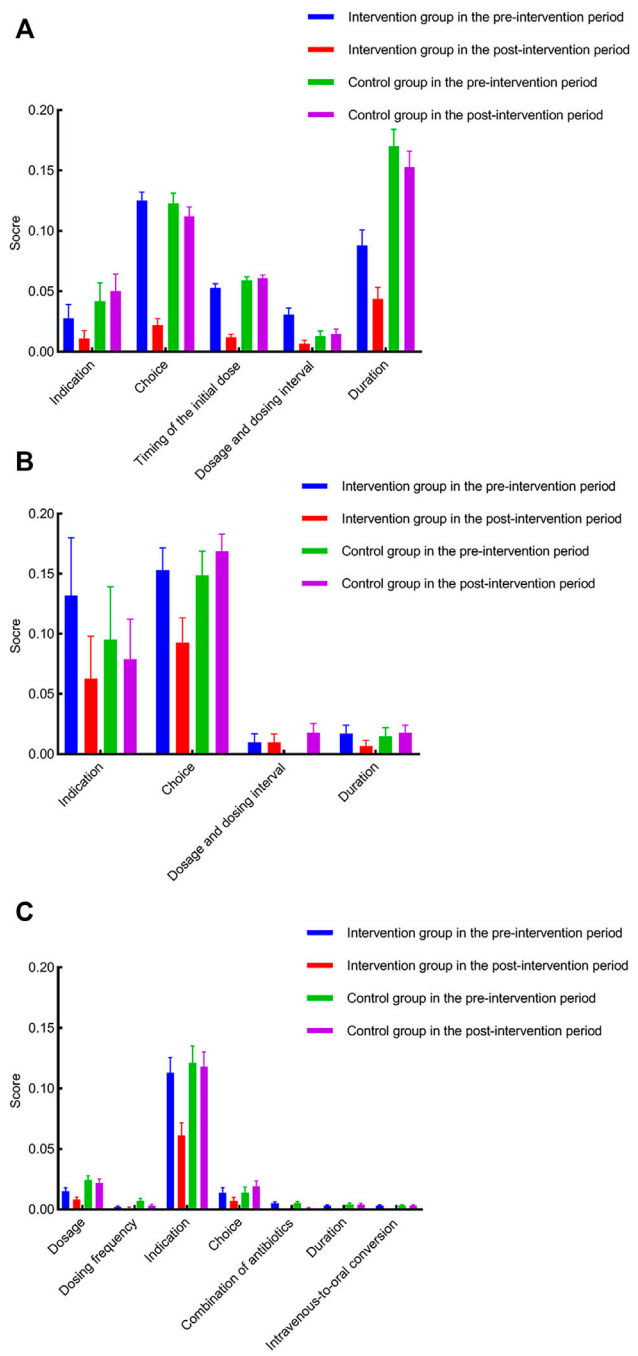

FIGURE 2 | Category analysis for irrational use of antibacterial drugs. **(A)** inappropriate score of antimicrobial prophylaxis with respect to corresponding items for both groups during the pre- and post-intervention periods; **(B)** inappropriate score of non-surgical antimicrobial prophylaxis with respect to corresponding items for both groups during the pre- and post-intervention periods; **(C)** inappropriate score of antimicrobial treatment with respect to corresponding items for both groups during the pre- and post-intervention periods.

indication had the largest impact for irrationality of antibiotics, regardless of antibiotic use for prophylaxis or treatment. Several studies showed that intervention has a significant impact on the appropriateness of antibiotic use after the implementation of the AMS program (Brink et al., 2017; Foolad et al., 2018; Butt et al., 2019; Calò et al., 2021; Jantarathaneewat et al., 2021). Our study was consistent with these findings, as we observed that the irrationality of antimicrobial prescriptions decreased in perioperative antimicrobial prophylaxis, non-surgical antimicrobial prophylaxis, and therapeutic use of antibiotics.

Inappropriateness of antimicrobial use in perioperative prophylaxis significantly decreased, especially in first-choice antibiotics, which was similar to that found in a previous study (Yang et al., 2017). Regarding the choice of antimicrobial prophylaxis, cefazolin definitely became the first choice in the post-intervention period with a concomitant decrease in cefathiamidine. Indeed, cefazolin was recommended as the first-line antibiotic according to the evidence-based trials (National Health Commission of China, 2015). Previous studies have reported that implementation of antibiotic stewardship program in surgeries was associated with reduction in the prescription of broad-spectrum and expensive antibiotics (Dona et al., 2019; Mahmoudi et al., 2019).

Inappropriateness score of antimicrobial prescribing for non-surgical prophylaxis decreased, manifested by improvement in correct antibiotic choice. The current study observed an increased use of ceftriaxone instead of cefoxitin for cirrhotic patients with esophageal and gastric variceal bleeding. Cirrhotic patients presenting with acute variceal bleeding have a high risk of developing bacterial infections; therefore, initiation of prophylactic antibiotic treatment at the time of admission is necessary (Lee et al., 2014). According to the guidelines (Korean Association for the Study of the, 2020), short-term antibiotic prophylaxis with intravenous ceftriaxone is recommended in cirrhotic patients with variceal bleeding.

A significant reduction in inappropriate prescribing for therapeutic use of antibiotics was also found after the intervention, manifested by a reduction in inappropriate prescribing in conditions for which antibiotics were generally not indicated, such as superficial thrombophlebitis of the lower extremities. Other antimicrobial stewardship efforts also have demonstrated improvements in unnecessary antibiotic prescribing for viral acute respiratory infection and asymptomatic bacteriuria (Meeker et al., 2016; Akpan et al., 2020; Calò et al., 2021; May et al., 2021). Antibiotic prescribing for these conditions where antimicrobials are not indicated puts them at unnecessary risk for adverse events and should be a target for quality improvement in interventional radiology wards.

In the intervention period, the acceptance rate of interventions was 80.00%, which was similar to that described in other studies at hospitals (Li et al., 2017; Hurst et al., 2019). The highest proportion of intervention types for surgical prophylaxis during the intervention period was stopping antimicrobial treatment, but the acceptance rate was the lowest. The reason is as follows: first, according to the Clinical Practice Guidelines for antibiotic prophylaxis during vascular and interventional radiology procedures (Chehab et al., 2018), a single preprocedural dose of cefazolin is an effective regimen for arterial endografts in the management of aortic aneurysm or

TABLE 5 | Inappropriate rate of antibiotic prescriptions and average score for both groups during the pre-intervention and post-intervention periods.

Categories	Groups	Pre-intervention		Post-intervention	
		Inappropriate rate (%)	Average score	Inappropriate rate (%)	Average score
Perioperative antimicrobial prophylaxis[a]	Control group	95.65 (66/69)	0.41	100.00 (90/90)	0.39
	Intervention group	89.77 (79/88)	0.33	41.67 (45/108)	0.10
Non-surgical antimicrobial prophylaxis	Control group	91.67 (22/24)	0.26	94.44 (34/36)	0.28
	Intervention group	96.15 (25/26)	0.31	55.56 (15/27)	0.17
Therapeutic use of antibiotics	Control group	66.67 (100/150)	0.18	64.18 (129/201)	0.17
	Intervention group	57.06 (105/184)	0.15	27.81 (47/169)	0.08

Note. [a]Antimicrobial use in some patients involving both prophylaxis and treatment.

TABLE 6 | DID analysis of average score for both groups during the pre-intervention and post-intervention periods.

Outcomes	Coefficient	*p*-Value	95% CI	Robust SE
Perioperative antimicrobial prophylaxis	−0.207	0.000	(−0.271, −0.143)	0.033
Non-surgical antimicrobial prophylaxis	−0.164	0.010	(−0.288, −0.035)	0.040
Therapeutic use of antibiotics	−0.069	0.003	(−0.115, −0.022)	0.024

Note. DID, difference-in-differences.

TABLE 7 | Source, initiator, type, and acceptance rate of antimicrobial stewardship interventions.

Interventions	Perioperative antimicrobial prophylaxis	Non-surgical antimicrobial prophylaxis	Therapeutic use of antibiotics
Source of interventions		Number (%)	
Ward round consultation	18 (11.04%)	13 (36.11%)	54 (21.95%)
WeChat consultation	6 (3.68%)	2 (5.55%)	10 (4.07%)
Medical order review	139 (85.28%)	21 (58.33%)	161 (65.45%)
Multidisciplinary consultation	NA	NA	21 (8.54%)
Initiator of interventions		Number (%)	
Clinician	24 (14.72%)	15 (41.67%)	85 (34.55%)
Clinical pharmacist	139 (85.28%)	21 (58.33%)	161 (65.45%)
Type and acceptance of interventions		Accepted% (number accepted/number)	
Initiate preferred agent for indication	85.29% (29/34)	47.62% (10/21)	90.00% (9/10)
Dose/frequency adjustment	82.14% (23/28)	50.00% (2/4)	91.49% (43/47)
Drug timing adjustment	79.55% (35/44)	NA	NA
Stop due to prolonged duration	66.67% (38/57)	75.00% (3/4)	85.71% (6/7)
Discontinuation of antibiotics without indication	NA	57.14% (4/7)	67.24% (39/58)
Stop combination of antimicrobial treatment	NA	NA	87.50% (7/8)
Change antibiotic due to toxicity, allergy, etc	NA	NA	87.50% (7/8)
Escalation	NA	NA	100.00% (16/16)
De-escalation	NA	NA	88.24% (30/34)
Sequential intravenous-to-oral therapy	NA	NA	100% (11/11)
Imaging/laboratory/etiological examination	NA	NA	93.10% (27/29)
Therapeutic drug monitoring	NA	NA	94.44% (17/18)
Total	76.69% (125/163)	52.78% (19/36)	86.18% (212/246)

Note NA, not available.

dissection. Chinese Guidelines also suggest a single dose for prophylaxis (National Health Commission of China, 2015). However, the aortic stent graft surgery involves important viscera, an endograft infection that carries high morbidity and mortality rates (as high as 27%) (Cernohorsky et al., 2011). Therefore, doctors are unwilling to use a single dose of antibiotics. Second, postoperative inflammatory response syndrome may occur after aortic stent graft procedures and hepatic artery chemoembolization, which is characterized by fever, leukocytosis, and elevated C-reactive protein (CRP) and other inflammatory markers (Blackburn and West, 2016; Daye and Walker, 2018). Sometimes it is difficult for doctors to distinguish it from infection, so they dare not easily stop the antibiotics. Compared with perioperative antimicrobial prophylaxis, non-surgical antimicrobial prophylaxis had a lower acceptance rate for intervention, which was mainly manifested in the recommendation of using preferred evidence-based drug. That is probably because clinicians in the surgical ward might pay more attention to perioperative antimicrobial prophylaxis that resulted in change of

prescribing behavior. In terms of therapeutic use of antibiotics, the lowest acceptance rate of intervention type was discontinuation of therapy without indication. For example, superficial thrombophlebitis is characterized by the appearance of tenderness, pain, redness, and swelling in the affected veins, which refers to an inflammation of superficial veins. Nonsteroidal anti-inflammatory drugs are suggested to reduce inflammation rather than antibiotics (Di Nisio et al., 2018). However, the strained relationship between doctors and patients in China has made doctors hesitant to not use antibiotics.

Our study has several strengths. Although the AMS program was performed in a surgical ward, a comprehensive intervention implemented not only focused on perioperative antimicrobial prophylaxis but also on non-surgical antimicrobial prophylaxis and therapeutic use of antibiotics. Notably, a scoring system was developed for assessment of rationality of antimicrobial prescriptions in antimicrobial prophylaxis and treatment based on consultation of multidisciplinary experts experienced in the management of infectious diseases. Previous studies usually emphasized on the analysis of change trend in antimicrobial utilization and cost savings after the implementation of AMS, whereas a comprehensive evaluation of antimicrobial prescriptive appropriateness was conducted as well as antimicrobial consumption and cost in our study.

There were also some limitations. First, it was a nonrandomized, retrospective pre–post analysis that carry a risk of selection bias. However, a quasi-experimental design with a DID methodology is often used to estimate the net effect of the intervention. Second, the present study investigated only the short-term effects of a pharmacist-led AMS intervention in the department of vascular and interventional radiology. It is unclear whether these effects would sustain after the interventions had finished. Third, this study was limited by its sample size. Finally, this was a single center study conducted in a tertiary teaching hospital; hence, the possibility of generalizing our results to other hospitals was limited. Future study with a larger sample size and a more rigorous design over an extended period needs to be confirmed. In spite of the limitations, this study has demonstrated that implementation of pharmacist-led AMS program significantly decreased the antibiotic use, cost of antibiotics, and irrationality of antibiotic prescribing in the department of vascular and interventional radiology.

CONCLUSION

This study provides important evidence from a pre-post intervention study that implementation of AMS interventions was associated with a marked reduction of antibiotic use, cost of antibiotics, and irrationality of antibiotic prescribing in the department of vascular and interventional radiology in a Chinese hospital. Our data indicate that departments that are not a priority for antimicrobial administration also need attention. It is essential for surgeons to be aware of the importance of adherence to guidelines. The findings of our study indicate that the beneficial AMS intervention led by clinical pharmacists can help to promote the rational use of antibiotics and decrease the antibiotic utilization.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The retrospective study complied with the Declaration of Helsinki, and ethical approval was obtained from the Ethical Committee of Nanjing Medical University (grant number: ethical review 2020103).

AUTHOR CONTRIBUTIONS

JX and XL initiated the study concept. YW conducted the data analysis. JX and JH wrote the first draft of the manuscript. XD, LS, and LD collected the data. YY led the data collection. DZ contributed to the data analysis and interpretation of the data. SX, ES, and JS were responsible for coordinating all research activities. LT and XL revised the draft of the manuscript. QZ supervised the study. All authors read and approved the final manuscript.

FUNDING

This study was supported by the National Natural Science Foundation of China (Grant No. 71673147), the China Medical Board (Grant No. 17-277), Suzhou Science and Technology Development Plan Project (No. SKJYD2021170) and Wu Jieping Medical Foundation (No. 320.6750.2020-12-69).

ACKNOWLEDGMENTS

The authors are grateful for the support of the information department in the Affiliated Suzhou Hospital of Nanjing Medical University for providing the original data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2022.832078/full#supplementary-material

REFERENCES

Abubakar, U., Syed Sulaiman, S. A., and Adesiyun, A. G. (2019). Impact of Pharmacist-Led Antibiotic Stewardship Interventions on Compliance with Surgical Antibiotic Prophylaxis in Obstetric and Gynecologic Surgeries in Nigeria. *PLoS One.* 14 (3), e0213395. doi:10.1371/journal.pone.0213395

Akpan, M. R., Isemin, N. U., Udoh, A. E., and Ashiru-Oredope, D. (2020). Implementation of Antimicrobial Stewardship Programmes in African Countries: a Systematic Literature Review. *J. Glob. Antimicrob. Resist.* 22, 317–324. doi:10.1016/j.jgar.2020.03.009

Arensman, K., Dela-Pena, J., Miller, J. L., LaChance, E., Beganovic, M., Anderson, M., et al. (2020). Impact of Mandatory Infectious Diseases Consultation and Real-Time Antimicrobial Stewardship Pharmacist Intervention on *Staphylococcus aureus* Bacteremia Bundle Adherence. *Open Forum Infect. Dis.* 7 (6), ofaa184. doi:10.1093/ofid/ofaa184

Beyar-Katz, O., Dickstein, Y., Borok, S., Vidal, L., Leibovici, L., and Paul, M. (2017). Empirical Antibiotics Targeting Gram-Positive Bacteria for the Treatment of Febrile Neutropenic Patients with Cancer. *Cochrane Database Syst. Rev.* 6, CD003914. doi:10.1002/14651858.CD003914.pub4

Blackburn, H., and West, S. (2016). Management of Postembolization Syndrome Following Hepatic Transarterial Chemoembolization for Primary or Metastatic Liver Cancer. *Cancer Nurs.* 39 (5), E1–E18. doi:10.1097/NCC.0000000000000302

Böhmer, F., Hornung, A., Burmeister, U., Köchling, A., Altiner, A., Lang, H., et al. (2021). Factors, Perceptions and Beliefs Associated With Inappropriate Antibiotic Prescribing in German Primary Dental Care: A Qualitative Study. *Antibiotics (Basel).* 10 (8), 987. doi:10.3390/antibiotics10080987

Brink, A. J., Messina, A. P., Feldman, C., Richards, G. A., van den Bergh, D., and Netcare Antimicrobial Stewardship Study, A. (2017). From Guidelines to Practice: a Pharmacist-Driven Prospective Audit and Feedback Improvement Model for Peri-Operative Antibiotic Prophylaxis in 34 South African Hospitals. *J. Antimicrob. Chemother.* 72 (4), 1227–1234. doi:10.1093/jac/dkw523

Butt, S. Z., Ahmad, M., Saeed, H., Saleem, Z., and Javaid, Z. (2019). Post-surgical Antibiotic Prophylaxis: Impact of Pharmacist's Educational Intervention on Appropriate Use of Antibiotics. *J. Infect. Public Health.* 12 (6), 854–860. doi:10.1016/j.jiph.2019.05.015

Calò, F., Onorato, L., Macera, M., Di Caprio, G., Monari, C., Russo, A., et al. (2021). Impact of an Education-Based Antimicrobial Stewardship Program on the Appropriateness of Antibiotic Prescribing: Results of a Multicenter Observational Study. *Antibiotics.* 10 (3), 314. doi:10.3390/antibiotics10030314

CDC (2019). The Core Elements of Hospital Antibiotic Stewardship Programs. [Online]. Available at: https://www.cdc.gov/antibiotic-use/core-elements/hospital.html (Accessed on June 5, 2020).

Cernohorsky, P., Reijnen, M. M., Tielliu, I. F., van Sterkenburg, S. M., van den Dungen, J. J., and Zeebregts, C. J. (2011). The Relevance of Aortic Endograft Prosthetic Infection. *J. Vasc. Surg.* 54 (2), 327–333. doi:10.1016/j.jvs.2010.12.067

Chehab, M. A., Thakor, A. S., Tulin-Silver, S., Connolly, B. L., Cahill, A. M., Ward, T. J., et al. (2018). Adult and Pediatric Antibiotic Prophylaxis during Vascular and IR Procedures: A Society of Interventional Radiology Practice Parameter Update Endorsed by the Cardiovascular and Interventional Radiological Society of Europe and the Canadian Association for Interventional Radiology. *J. Vasc. Interv. Radiol.* 29 (11), 1483–e2. doi:10.1016/j.jvir.2018.06.007

Daye, D., and Walker, T. G. (2018). Complications of Endovascular Aneurysm Repair of the Thoracic and Abdominal Aorta: Evaluation and Management. *Cardiovasc. Diagn. Ther.* 8 (Suppl. 1), S138–S156. doi:10.21037/cdt.2017.09.17

Dellit, T. H., Owens, R. C., McGowan, J. E., Jr., Gerding, D. N., Weinstein, R. A., Burke, J. P., et al. (2007). Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America Guidelines for Developing an Institutional Program to Enhance Antimicrobial Stewardship. *Clin. Infect. Dis.* 44 (2), 159–177. doi:10.1086/510393

Denny, K. J., Gartside, J. G., Alcorn, K., Cross, J. W., Maloney, S., and Keijzers, G. (2019). Appropriateness of Antibiotic Prescribing in the Emergency Department. *J. Antimicrob. Chemother.* 74 (2), 515–520. doi:10.1093/jac/dky447

Di Nisio, M., Wichers, I., and Middeldorp, S. (2018). Treatment of Lower Extremity Superficial Thrombophlebitis. *JAMA.* 320 (22), 2367–2368. doi:10.1001/jama.2018.16623

Dimick, J. B., and Ryan, A. M. (2014). Methods for Evaluating Changes in Health Care Policy: the Difference-In-Differences Approach. *Jama.* 312 (22), 2401–2402. doi:10.1001/jama.2014.16153

Donà, D., Luise, D., La Pergola, E., Montemezzo, G., Frigo, A., Lundin, R., et al. (2019). Effects of an Antimicrobial Stewardship Intervention on Perioperative Antibiotic Prophylaxis in Pediatrics. *Antimicrob. Resist. Infect. Control.* 8, 13. doi:10.1186/s13756-019-0464-z

Du, Y., Li, J., Wang, X., Peng, X., Wang, X., He, W., et al. (2020). Impact of a Multifaceted Pharmacist-Led Intervention on Antimicrobial Stewardship in a Gastroenterology Ward: A Segmented Regression Analysis. *Front. Pharmacol.* 11, 442. doi:10.3389/fphar.2020.00442

Foolad, F., Huang, A. M., Nguyen, C. T., Colyer, L., Lim, M., Grieger, J., et al. (2018). A Multicentre Stewardship Initiative to Decrease Excessive Duration of Antibiotic Therapy for the Treatment of Community-Acquired Pneumonia. *J. Antimicrob. Chemother.* 73 (5), 1402–1407. doi:10.1093/jac/dky021

Gerber, J. S., Prasad, P. A., Fiks, A. G., Localio, A. R., Grundmeier, R. W., Bell, L. M., et al. (2013). Effect of an Outpatient Antimicrobial Stewardship Intervention on Broad-Spectrum Antibiotic Prescribing by Primary Care Pediatricians: a Randomized Trial. *JAMA.* 309 (22), 2345–2352. doi:10.1001/jama.2013.6287

Haseeb, A., Faidah, H. S., Al-Gethamy, M., Iqbal, M. S., Barnawi, A. M., Elahe, S. S., et al. (2020). Evaluation of a Multidisciplinary Antimicrobial Stewardship Program in a Saudi Critical Care Unit: A Quasi-Experimental Study. *Front. Pharmacol.* 11, 570238. doi:10.3389/fphar.2020.570238

Huemer, M., Mairpady Shambat, S., Brugger, S. D., and Zinkernagel, A. S. (2020). Antibiotic Resistance and Persistence-Implications for Human Health and Treatment Perspectives. *EMBO Rep.* 21 (12), e51034. doi:10.15252/embr.202051034

Hurst, A. L., Child, J., and Parker, S. K. (2019). Intervention and Acceptance Rates Support Handshake-Stewardship Strategy. *J. Pediatr. Infect Dis Soc.* 8 (2), 162–165. doi:10.1093/jpids/piy054

Hwang, H., and Kim, B. (2018). Impact of an Infectious Diseases Specialist-Led Antimicrobial Stewardship Programmes on Antibiotic Use and Antimicrobial Resistance in a Large Korean Hospital. *Sci. Rep.* 8 (1), 14757. doi:10.1038/s41598-018-33201-8

Jantarathaneewat, K., Apisarnthanarak, A., Limvorapitak, W., Weber, D. J., and Montakantikul, P. (2021). Pharmacist-Driven Antibiotic Stewardship Program in Febrile Neutropenic Patients: A Single Site Prospective Study in Thailand. *Antibiotics (Basel).* 10 (4), 456. doi:10.3390/antibiotics10040456

Korean Association for the Study of the, L. (2020). KASL Clinical Practice Guidelines for Liver Cirrhosis: Varices, Hepatic Encephalopathy, and Related Complications. *Clin. Mol. Hepatol.* 26 (2), 83–127. doi:10.3350/cmh.2019.0010n

Lee, C. F., Cowling, B. J., Feng, S., Aso, H., Wu, P., Fukuda, K., et al. (2018). Impact of Antibiotic Stewardship Programmes in Asia: a Systematic Review and Meta-Analysis. *J. Antimicrob. Chemother.* 73 (4), 844–851. doi:10.1093/jac/dkx492

Lee, Y. Y., Tee, H. P., and Mahadeva, S. (2014). Role of Prophylactic Antibiotics in Cirrhotic Patients with Variceal Bleeding. *World J. Gastroenterol.* 20 (7), 1790–1796. doi:10.3748/wjg.v20.i7.1790

Li, H., Yan, S., Li, D., Gong, Y., Lu, Z., and Yin, X. (2019). Trends and Patterns of Outpatient and Inpatient Antibiotic Use in China's Hospitals: Data from the Center for Antibacterial Surveillance, 2012-16. *J. Antimicrob. Chemother.* 74 (6), 1731–1740. doi:10.1093/jac/dkz062

Li, Z., Cheng, B., Zhang, K., Xie, G., Wang, Y., Hou, J., et al. (2017). Pharmacist-driven Antimicrobial Stewardship in Intensive Care Units in East China: A Multicenter Prospective Cohort Study. *Am. J. Infect. Control.* 45 (9), 983–989. doi:10.1016/j.ajic.2017.02.021

Lin, L., Sun, R., Yao, T., Zhou, X., and Harbarth, S. (2020). Factors Influencing Inappropriate Use of Antibiotics in Outpatient and Community Settings in China: a Mixed-Methods Systematic Review. *BMJ Glob. Health.* 5 (11), e003599. doi:10.1136/bmjgh-2020-003599

Liu, C., Liu, C., Wang, D., and Zhang, X. (2019). Knowledge, Attitudes and Intentions to Prescribe Antibiotics: A Structural Equation Modeling Study of Primary Care Institutions in Hubei, China. *Int. J. Environ. Res. Public Health.* 16 (13), 2385. doi:10.3390/ijerph16132385

Mahmoudi, L., Ghouchani, M., Mahi-Birjand, M., Bananzadeh, A., and Akbari, A. (2019). Optimizing Compliance with Surgical Antimicrobial Prophylaxis Guidelines in Patients Undergoing Gastrointestinal Surgery at a Referral Teaching Hospital in Southern Iran: Clinical and Economic Impact. *Infect. Drug Resist.* 12, 2437–2444. doi:10.2147/IDR.S212728

Mahrous, A. J., Thabit, A. K., Elarabi, S., and Fleisher, J. (2020). Clinical Impact of Pharmacist-Directed Antimicrobial Stewardship Guidance Following Blood Culture Rapid Diagnostic Testing. *J. Hosp. Infect.* 106 (3), 436–446. doi:10.1016/j.jhin.2020.09.010

May, L., Martín Quirós, A., Ten Oever, J., Hoogerwerf, J., Schoffelen, T., and Schouten, J. (2021). Antimicrobial Stewardship in the Emergency Department: Characteristics and Evidence for Effectiveness of Interventions. *Clin. Microbiol. Infect.* 27 (2), 204–209. doi:10.1016/j.cmi.2020.10.028

Meeker, D., Linder, J.A., Fox, C.R., Friedberg, M.W., Persell, S.D., Goldstein, N.J., et al. (2016). Effect of Behavioral Interventions on Inappropriate Antibiotic Prescribing Among Primary Care Practices: A Randomized Clinical Trial. *Jama* 315(6), 562–570. doi:10.1001/jama.2016.0275

Monmaturapoj, T., Scott, J., Smith, P., Abutheraa, N., and Watson, M. C. (2021). Pharmacist-led Education-Based Antimicrobial Stewardship Interventions and Their Effect on Antimicrobial Use in Hospital Inpatients: a Systematic Review and Narrative Synthesis. *J. Hosp. Infect.* 115, 93–116. doi:10.1016/j.jhin.2021.06.003

National Health Commission of the People's Republic of China (2011). Notice Regarding National Special Measure Scheme on Clinical Use of Antibiotics [in Chinese]. [Online]. Available at: http://www.nhc.gov.cn/cms-search/xxgk/getManuscriptXxgk.htm?id=51376 (Accessed July 4, 2021).

National Health Commission of the People's Republic of China (2015). Guiding Principles for Clinical Application of Antibiotics. Available at: http://www.nhc.gov.cn/yzygj/s3593/201508/c18e1014de6c45ed9f6f9d592b43db42.shtml (Accessed July 5, 2021).

Nowakowska, M., van Staa, T., Mölter, A., Ashcroft, D. M., Tsang, J. Y., White, A., et al. (2019). Antibiotic Choice in UK General Practice: Rates and Drivers of Potentially Inappropriate Antibiotic Prescribing. *J. Antimicrob. Chemother.* 74 (11), 3371–3378. doi:10.1093/jac/dkz345

Ray, M. J., Tallman, G. B., Bearden, D. T., Elman, M. R., and McGregor, J. C. (2019). Antibiotic Prescribing without Documented Indication in Ambulatory Care Clinics: National Cross Sectional Study. *Bmj.* 367, l6461. doi:10.1136/bmj.l6461

Rice, L. B. (2018). Antimicrobial Stewardship and Antimicrobial Resistance. *Med. Clin. North. Am.* 102 (5), 805–818. doi:10.1016/j.mcna.2018.04.004

Rowe, T. A., and Linder, J. A. (2019). Novel Approaches to Decrease Inappropriate Ambulatory Antibiotic Use. *Expert Rev. Anti Infect. Ther.* 17 (7), 511–521. doi:10.1080/14787210.2019.1635455

Saaty, T. L. (1990). How to Make a Decision: The Analytic Hierarchy Process. *Eur. J. Oper. Res.* 48 (1), 9–26. doi:10.1016/0377-2217(90)90057-i

Shen, J., Sun, Q., Zhou, X., Wei, Y., Qi, Y., Zhu, J., et al. (2011). Pharmacist Interventions on Antibiotic Use in Inpatients with Respiratory Tract Infections in a Chinese Hospital. *Int. J. Clin. Pharm.* 33 (6), 929–933. doi:10.1007/s11096-011-9577-z

Versporten, A., Zarb, P., Caniaux, I., Gros, M. F., Drapier, N., Miller, M., et al. (2018). Antimicrobial Consumption and Resistance in Adult Hospital Inpatients in 53 Countries: Results of an Internet-Based Global Point Prevalence Survey. *Lancet Glob. Health.* 6 (6), e619–e629. doi:10.1016/S2214-109X(18)30186-4

Wang, J., Wang, P., Wang, X., Zheng, Y., and Xiao, Y. (2014). Use and Prescription of Antibiotics in Primary Health Care Settings in China. *JAMA Intern. Med.* 174 (12), 1914–1920. doi:10.1001/jamainternmed.2014.5214

Wenzler, E., Wang, F., Goff, D. A., Prier, B., Mellett, J., Mangino, J. E., et al. (2017). An Automated, Pharmacist-Driven Initiative Improves Quality of Care for *Staphylococcus aureus* Bacteremia. *Clin. Infect. Dis.* 65 (2), 194–200. doi:10.1093/cid/cix315

WHO (2011). The World Medicines Situation 2011: Rational Use of Medicines. [Online]. Available at: https://www.who.int/medicines/areas/policy/world_medicines_situation/WMS_ch14_wRational.pdf (Accessed January 4, 2022).

Wooldrige, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge: MIT press.

Xin, C., Xia, Z., and Li, G. (2019). The Impact of Pharmaceutical Interventions on The Use of Carbapenems in A Chinese Hospital: A Pre-Post Study. *Infect. Drug Resist.* 12, 3567–3573. doi:10.2147/IDR.S229009

Yang, P., Jiang, S. P., and Lu, X. Y. (2017). Effectiveness of Continuous Improvement by a Clinical Pharmacist-Led Guidance Team on the Prophylactic Antibiotics Usage Rationality in Intervention Procedure at a Chinese Tertiary Teaching Hospital. *Ther. Clin. Risk Manag.* 13, 469–476. doi:10.2147/TCRM.S131937

Zhang, H. X., Li, X., Huo, H. Q., Liang, P., Zhang, J. P., and Ge, W. H. (2014). Pharmacist Interventions for Prophylactic Antibiotic Use in Urological Inpatients Undergoing Clean or Clean-Contaminated Operations in a Chinese Hospital. *PLoS One.* 9 (2), e88971. doi:10.1371/journal.pone.0088971

Zhao, W., Wang, B., and Yu, G. (2018). Antibiotic Resistance Genes in China: Occurrence, Risk, and Correlation Among Different Parameters. *Environ. Sci. Pollut. Res. Int.* 25 (22), 21467–21482. doi:10.1007/s11356-018-2507-z

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claims that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Xu, Huang, Yu, Zhou, Wang, Xue, Shang, Sun, Ding, Shi, Duan, Tang, Zhou and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Pharmacology

SYSTEMATIC REVIEW
published: 02 March 2022
doi: 10.3389/fphar.2022.768980

Neoplasm Risk in Patients With Rheumatoid Arthritis Treated With Fostamatinib: A Systematic Review and Meta-analysis

Yuehong Chen, Huan Liu, Yunru Tian, Zhongling Luo, Geng Yin and Qibing Xie**

Department of Rheumatology and Immunology, West China Hospital, Sichuan University, Chengdu, China

Objective: This study aimed to assess neoplasm risk in patients with rheumatoid arthritis (RA) treated with fostamatinib.

Methods: Studies were collected from electronic databases of OVID Medline, OVID EMBASE, Cochrane Central Register of Controlled Trials, and Web of Science. We included studies that reported neoplasms in patients with RA treated with fostamatinib. Study selection was repeated by two reviewers based on the study selection criteria. Data were collected and methodological quality assessment was performed. Data were pooled using the Peto odds ratio (OR) with a 95% confidence interval (CI). Subgroup analyses of the fostamatinib dose, trial duration, neoplasm nature, and neoplasm-originating systems were conducted. A funnel plot was used to estimate publication bias, and sensitivity analysis was performed to test the robustness of the results.

Results: Seven trials involving 4,971 participants showing low to moderate risk of bias were included. Compared with the placebo, fostamatinib use was not associated with the risks of overall neoplasms (Peto OR = 2.62, 95%CI 0.97–7.10), malignant neoplasms (Peto OR = 3.08, 95%CI 0.96–9.91), or benign neoplasms (Peto OR = 1.71, 95%CI 0.26–11.36). Nevertheless, compared with the placebo, a longer duration of fostamatinib use had a higher risk of malignant neoplasms (Peto OR = 4.49, 95%CI 1.03–19.60) at 52 weeks. As for malignant neoplasms in the digestive system, lower doses of fostamatinib reduced the neoplasm risk (100 mg bid vs 150 mg qd: Peto OR = 0.06, 95%CI 0.01–0.59). Sensitivity analysis showed no significant differences in the effective trends, and no publication bias was found.

Conclusion: Fostamatinib is not associated with the risks of overall neoplasms as compared to placebo. Nevertheless, a longer duration of fostamatinib use may be associated with a risk of malignant neoplasms and higher doses of fostamatinib may increase malignant neoplasms in the digestive system. Further well-planned cohort studies with a larger study population are needed to elucidate these outcomes.

Systematic Review Registration: PROSPERO (CRD42020202121).

Keywords: fostamatinib, rheumatoid arthritis, neoplasm, meta-analysis, systematic review

OPEN ACCESS

Edited by:
Tin Wui Wong,
Universiti Teknologi MARA Puncak Alam, Malaysia

Reviewed by:
Ettore Silvagni,
University of Ferrara, Italy
Kurt Neumann,
Independent researcher, Kerékteleki, Hungary

****Correspondence:***
Geng Yin
yingeng1975@163.com
Qibing Xie
xieqibing1971@163.com

Specialty section:
This article was submitted to Drugs Outcomes Research and Policies, a section of the journal Frontiers in Pharmacology

Received: *01 September 2021*
Accepted: *07 February 2022*
Published: *02 March 2022*

Citation:
Chen Y, Liu H, Tian Y, Luo Z, Yin G and Xie Q (2022) Neoplasm Risk in Patients With Rheumatoid Arthritis Treated With Fostamatinib: A Systematic Review and Meta-analysis. Front. Pharmacol. 13:768980. doi: 10.3389/fphar.2022.768980

1 INTRODUCTION

Rheumatoid arthritis (RA) is an autoimmune disease characterized by persistent synovitis formation, systematic inflammation, and autoantibodies presence, leading to bone and cartilage damage if not treated appropriately (Scott et al., 2010; Smolen et al., 2016). The risk factors for RA include susceptibility genes, environmental factors (e.g., oral microbiome, smoking, periodontitis, and microbiome), epigenetic modifications, and posttranslational modifications (e.g., methylation, acetylation, and citrullination). RA is associated with several comorbidities, such as cancer and cardiovascular diseases, which are risk factors for increased mortality (Dougados, 2016; Widdifield et al., 2018).

Compared with the general population, patients with RA are associated with increased risks of overall malignant neoplasms [standardized incidence ratio (SIR) = 1.05, 95% confidence interval (CI) (1.01–1.09)], lymphoma [2.08, 95%CI (1.80–2.39)], lung cancer [1.63, 95%CI (1.43–1.87)], and malignant melanoma [1.23, 95%CI (1.01–1.49)] (Smitten et al., 2008; Simon et al., 2015; De Cock and Hyrich, 2018). The high inflammatory activity of RA is a major risk factor for developing lymphomas (Baecklund et al., 2006). Compared with low disease activity, moderate and high inflammatory disease activity increase the risk of developing lymphomas by 8-fold (odds ratio [OR] = 7.7, 95%CI 4.8–12.3) and 70-fold (OR = 71.3, 95%CI 24.1–211.4), respectively. Therefore, achieving remission of disease activity or maintaining low disease activity is the treatment target (Smolen et al., 2016).

Spleen tyrosine kinase (Syk), a non-receptor protein tyrosine located in the cytoplasm, plays a fundamental role in the activation of the B-cell receptor, which is necessary for B-cell development, proliferation, and survival. Thus, pharmacological targeting of Syk is effective in affecting the signal transduction of B-cell receptors, leading to cell apoptosis and inhibition of the activation and migration of B-cells, which are therapeutic targets for B-cell dominant diseases, such as chronic lymphocytic leukemia, B-cell malignancies, and autoimmune disorders (Buchner et al., 2010; Hoellenriegel et al., 2012). Fostamatinib is a Syk inhibitor, and R406 is the active metabolite of fostamatinib, which has been reported to effectively treat RA (Kunwar et al., 2016; Kang et al., 2019). Nevertheless, fostamatinib is reported to have an increased risk of infection, diarrhea, hypertension, neutropenia, and hypertransaminasaemia (Salgado et al., 2014; Kunwar et al., 2016; Kang et al., 2019; Chen et al., 2021). Whether the use of fostamatinib is associated with an increased risk of malignancy remains unknown. Therefore, we performed a systematic review and meta-analysis by including all the available evidence to assess neoplasm risk in patients with RA treated with fostamatinib.

2 METHODS

2.1 Setting

This systematic review and meta-analysis was conducted to investigate the neoplasm risk in patients with RA treated with fostamatinib and reported based on the preferred reporting items for systematic reviews and meta-analyses guidelines (Moher et al., 2009). This study was registered in PROSPERO with registration number CRD42020202121.

2.2 Eligibility Criteria

The inclusion criteria were set based on the PICO principle: P (patients) were RA patients; I (intervention) was fostamatinib, regardless of the dose and usage; C (comparison) was placebo, other treatment, or different doses of fostamatinib; O (outcomes) were neoplasms regardless of the neoplasm nature (malignant or benign); in addition, we also included s (study designs) for randomized controlled trials (RCTs), cohort studies, or case-control studies.

A study was excluded if it was a duplicate, commentary, conference abstract, and or did not have relevant outcomes.

2.3 Search Strategy

Electronic databases of OVID Medline, OVID EMBASE, Web of Science, and Cochrane Central Register of Controlled Trials were searched on 3 July 2020, using both MeSH terms and key words without language limitations. Search terms included "rheumatoid arthritis" and "fostamatinib". Detailed search strategy can be found in the supplemental file or the published study (Chen et al., 2021). Manual searches of reference lists of included studies and clnicaltrials.gov was also performed to identify potentially eligible studies.

2.4 Study Selection

Study selection was performed independently and repeated by a pair of reviewers (CYH, LH, and TYR), which was managed by Microsoft Office Access 2013. After preliminary screening by titles and abstracts, full texts were read based on the study selection criteria. Reference lists of included studies and published reviews and the clinicaltrials.gov website were manually checked. Any disagreement was resolved via discussion or judged by a third reviewer, if necessary.

2.5 Data Extraction

Data was collected independently and repeated by two authors (CYH, TYR, and LZL) on trial registration number, publication date or release date, trial duration-from the trial beginning to the time assessing neoplasm incidence, treatment information, number of neoplasms, and number of participants. Any disagreement was resolved via discussion or judged by a third reviewer, if necessary.

2.6 Methodological Quality Assessment

The Cochrane Collaboration tool was used to assess the risk of bias of included RCTs (JPT and So, 2011). Which focused on the items of random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, and selective reporting. For each item, if the answer was yes and correctly described, the assessment was low risk; if the answer was yes but lacked detailed description, the assessment was unclear; if the answer was yes but with the inappropriate method or if the method was not

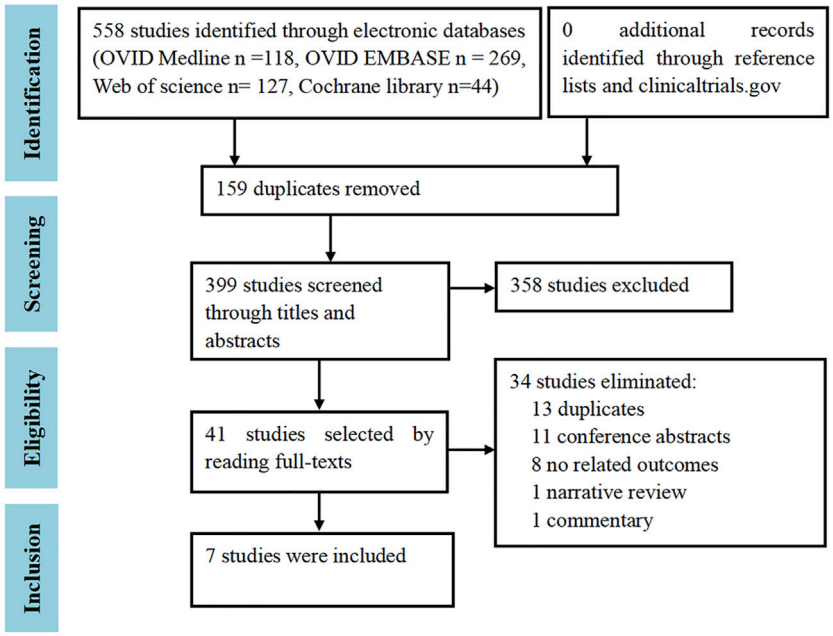

FIGURE 1 | Study selection flowchart. Adapted from Chen et al. (2021).

performed, the assessment was high risk. The risk of bias of the included studies was judged based on overall evidence. Methodological quality assessment was performed by two reviewers (CYH and LH), and any disagreement was resolved via discussion or judged by a third reviewer, if necessary. Quality assessment for cohort and case-control studies was not described here, as no such study design was included.

2.7 Data Analyses

RevMan software (version 5.1.3) was used to analyze the data. The effect size of the meta-analysis was estimated using the Peto OR with 95% CIs, considering the very low event. I^2 and heterogeneity *p*-value at the level of 0.1 was used to assess the clinical diversity, and I^2 had values of 25, 50, and 75% indicating low, moderate, and high heterogeneity, respectively, as recommended by the Cochrane Handbook (Higgins et al., 2003). Pre-set subgroup analyses by fostamatinib dose, trial duration, neoplasm nature, and neoplasm-originating systems were conducted. Sensitivity analysis using the Mantel-Haenszel random effect model was performed to test the robustness of the results. Publication bias was assessed using funnel plots.

Classification of neoplasm-originating sites and systems was based on the 10th version of the International Classification of Diseases (ICD10) (WHO), including malignant neoplasms (C00-C97), benign neoplasms (D10-D36), bone and articular cartilage (C40-C41), ill-defined, secondary, and unspecified sites (C76-C80), digestive organs (C15-C26), and urinary tract (C64-C68).

2.8 Patient and Public Involvement

As this study is a systematic review and meta-analysis, no ethical concerns or patients were involved.

3 RESULTS

3.1 Study Selection

A total of 558 citations were identified from the OVID Medline (*n* = 118), OVID EMBASE (*n* = 269), Web of Science (*n* = 127), and Cochrane Library (*n* = 44). After excluding duplicates (*n* = 159), irrelevant studies (*n* = 358) screened by titles and abstracts, and studies without relevant outcomes (*n* = 34) by reading the full texts, finally, seven trials (Evaluation of Effectivene, 1197a; Evaluation of Effectivene, 1197b; Evaluation of Effectivene, 1197c; Evaluation of Efficacy an, 1264; Treatment of Arthritis Wi, 2633; Efficacy and Safety Study; Evaluation of Long-term S, 1242) that enrolled a total of 4,971 RA patients were included (**Figure 1**). No additional studies were included by the manual checking. No cohort or case-control studies were included.

3.2 Characteristics of Included Studies

The trial duration was a median (range) of 26 (12–109) weeks. The number of participants varied from 189 to 1912, with a median of 457. About 82.7% (4,113/4,971) of the participants were women. The age was around 50 years, ranging from 50 to 53 years. The common dosage of fostamatinib was 100 mg twice daily or 150 mg once daily, which were taken orally (**Table 1**).

3.3 Methodological Quality

Only one study correctly reported methods of random sequence generation, three studies correctly reported allocation concealment, and all trials performed blinding of participants, personnel, and outcome assessment, reported

TABLE 1 | Characteristics of included RCTs.

Included RCTs	Year of results first posted	Trial duration	NO. of participants (female/male)	Age (years, mean ± SD)	Treatments (no. of participants)
NCT01197521 (Evaluation of Effectivene, 1197a)	2014	52 weeks	918 (770/148)	52 ± 12.0	Fostamatinib 100 mg bid (310) Fostamatinib 100 mg bid (4 weeks) then 150 mg qd (304) placebo (24 weeks) then fostamatinib 100 mg bid (304)
NCT01197534 (Evaluation of Effectivene, 1197b)	2014	52 weeks	908 (742/166)	53 ± 11.9	Fostamatinib 100 mg bid (308) Fostamatinib 100 mg bid (4 weeks) then 150 mg qd (298) placebo (24 weeks) then fostamatinib 100 mg bid (302)
NCT01197755 (Evaluation of Effectivene, 1197c)	2014	24 weeks	322 (261/61)	53 ± 12.3	Fostamatinib 100 mg bid (105) Fostamatinib 100 mg bid (4 weeks) then 150 mg qd (108) placebo (109)
NCT01264770 (Evaluation of Efficacy an, 1264)	2014	24 weeks	265 (210/55)	50 ± 11.8	Fostamatinib 100 mg bid (54) Fostamatinib 100 mg bid (4 weeks) then 100 mg qd (57) Fostamatinib 100 mg bid (4 weeks) then 150 mg qd (48) placebo (6 weeks) then fostamatinib 100 mg bid (27) placebo (6 weeks) then fostamatinib 100 mg bid (4 weeks) then 150 mg qd (25) Adalimumab 40 mg every 2 weeks (54)
NCT00326339 (Treatment of Arthritis Wi, 2633)	2008	12 weeks	189 (164/25)	52.1 (20–75), median	Fostamatinib 50 mg bid (46) Fostamatinib 100 mg bid (49) Fostamatinib 150 mg bid (47) placebo (47)
NCT00665925 (Efficacy and Safety Study)	2016	26 weeks	457 (390/67)	52.5 ± 12.8	Fostamatinib 100 mg bid (152) Fostamatinib 150 mg qd (152) placebo (153)
NCT01242514 (Evaluation of Long-term S, 1242)	2014	109 weeks	1912 (1576/336)	53 ± 11.8	Fostamatinib 100 mg qd (212) Fostamatinib 100 mg bid (1343) Fostamatinib 150 mg qd (357)

Bid = twice a day; qd = once a day.

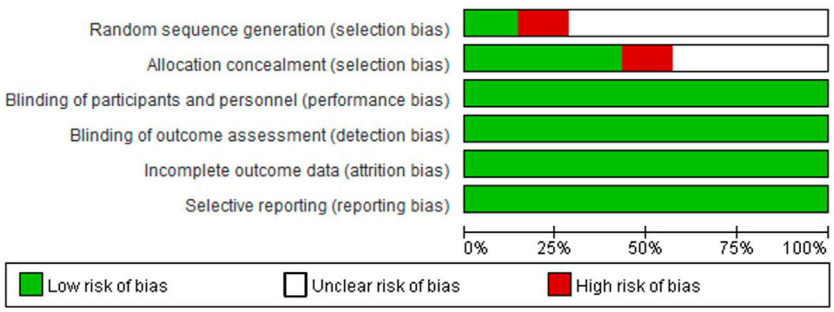

FIGURE 2 | Risk of bias graph. Adapted from Chen et al. (2021).

complete outcome data, and did not selectively report outcome data (**Figures 2**, **3**). Overall, the risk of bias in the included trials was low to moderate.

3.4 Sensitivity Analysis and Publication Bias

Different statistical methods were used, the Mantel-Haenszel random effect model vs. Peto fixed-effect model, to conduct sensitivity analysis to test the robustness of the results, which did not change the effect direction (data not shown). Publication bias, taking the data from overall neoplasms in patients treated with fostamatinib vs. placebo for an example, was assessed using a funnel plot, and the results showed that the funnel plot was symmetrical; thus, publication bias was not likely to occur (**Figure 4**).

3.5 Main Outcomes

3.5.1 Overall Neoplasms

In comparing fostamatinib vs. placebo, a total of six trials reported 16 neoplasms in 2,038 RA patients treated with fostamatinib, which did not increase the neoplasm risk compared with the placebo of two neoplasms in 967 RA patients (Peto OR = 2.62, 95%CI 0.97–7.10, I^2 = 0, 16/2038 vs. 2/967) (**Table 2**; **Figure 5**). Longer durations of fostamatinib use also did not increase the neoplasm risk without clinical diversity

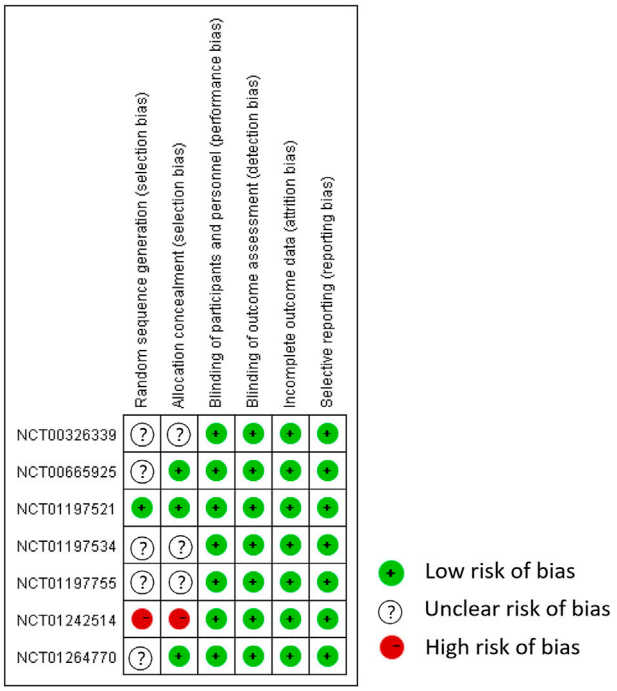

FIGURE 3 | Risk of bias summary. Adapted from Chen et al. (2021).

across the studies (Peto OR = 3.78, 95%CI 0.04–352.61, 1/142 vs. 0/47 for 12 weeks; Peto OR = 0.75, 95%CI 0.06–9.58, 2/372 vs. 1/161 for 24 weeks; Peto OR = 4.53, 95%CI 0.41–50.05, 3/304 vs. 0/153 for 26 weeks; Peto OR = 2.98, 95%CI 0.85–10.49, 10/1220 vs. 1/606 for 52 weeks) (**Table 2**; **Figure 5**).

For fostamatinib dosages of 100 mg bid vs. 150 mg qd, seven trials reported 25 neoplasms in 2,954 RA patients treated with 100 mg of fostamatinib twice daily; 100 mg fostamatinib twice daily did not increase the neoplasm risk compared with 150 mg fostamatinib use once daily for 12 neoplasms in 1,339 RA patients (Peto OR = 0.81, 95%CI 0.39–1.69, and I^2 = 35). Similarly, there was no difference in the neoplasm risk between fostamatinib dosages of 100 mg twice daily and 150 mg once daily, with longer durations of fostamatinib use (**Table 2**; **Supplementary Figure S1**).

For fostamatinib dosages of 100 mg bid vs. 100 mg qd, two studies were concerned with the neoplasm risk in RA patients treated with these different usage times of fostamatinib. A dosage timing of twice daily did not have a higher neoplasm risk compared with 100 mg of fostamatinib once daily (Peto OR = 3.22, 95%CI 0.77–13.52, 16/1424 vs. 0/269) (**Table 2**; **Supplementary Figure S2**).

3.5.2 Malignant Neoplasms

When comparing fostamatinib and placebo, six trials reported 12 malignant neoplasms in 2038 RA patients treated with fostamatinib, which did not increase the neoplasm risk compared with the placebo in one malignant neoplasm in 967 RA patients (Peto OR = 3.08, 95%CI 0.96–9.91, and I^2 = 6) (**Table 2**; **Supplementary Figure S3**). Nevertheless, the longer use of fostamatinib had a higher malignant neoplasm risk than the placebo (Peto OR = 4.49, 95%CI 1.03–19.60, 8/1220 vs. 0/606 for 52 weeks, I^2 = 0) (**Table 2**, **Supplementary Figure S3**).

When comparing different doses of fostamatinib, compared with 150 mg of fostamatinib once daily, 100 mg of fostamatinib twice daily did not have a higher malignant neoplasm risk (Peto OR = 1.09, 95%CI 0.45–2.60, I^2 = 39, 19/2954 vs. 7/1339, n = 7) (**Table 2**, **Supplementary Figure S4**). In comparison with 100 mg

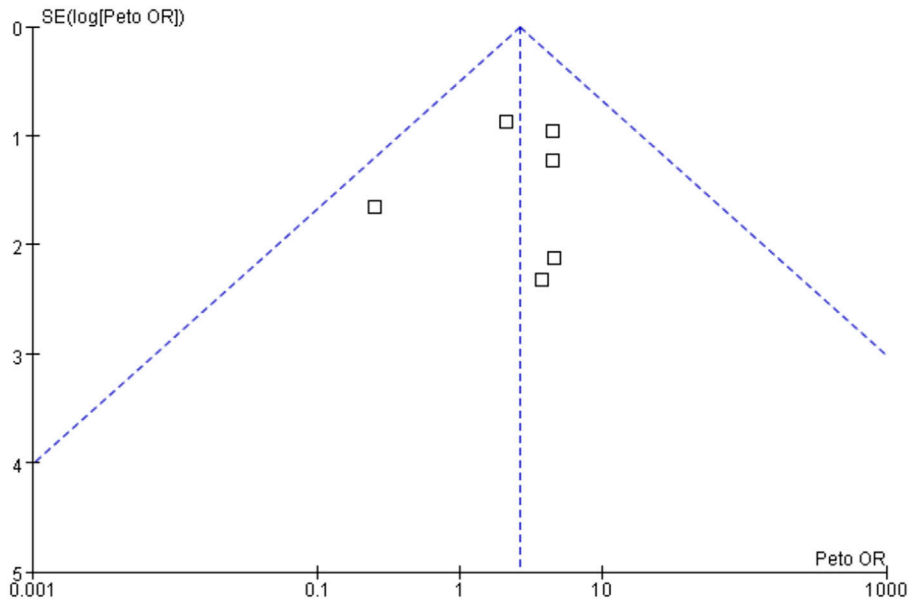

FIGURE 4 | Publication bias based on overall neoplasms in patients treated with fostamatinib versus placebo.

TABLE 2 | Pooled data of neoplasms in RA patients.

Comparisons	No of study	Fostamatinib		Comparator		Heterogeneity		Peto OR, 95% CI
		No of cancer	No of participants	No of cancer	No of participants	I2	P	
Overall neoplasms								
Fostamatinib vs PBO	6	16	2038	2	967	0	0.75	2.62 [0.97, 7.10]
12 weeks	1	1	142	0	47	—	—	3.78 [0.04, 352.61]
24 weeks	2	2	372	1	161	14	0.28	0.75 [0.06, 9.58]
26 weeks	1	3	304	0	153	—	—	4.53 [0.41, 50.05]
52 weeks	2	10	1220	1	606	0	0.56	2.98 [0.85, 10.49]
Fostamatinib 100 mg bid vs. 150 mg qd	7	25	2954	12	1339	35	0.17	0.81 [0.39, 1.69]
12 weeks	1	0	49	0	47	—	—	—
24 weeks	2	0	186	2	181	0	0.96	0.13 [0.01, 2.08]
26 weeks	1	2	152	1	152	—	—	1.96 [0.20, 18.95]
52 weeks	2	7	1224	4	602	82	0.02	0.86 [0.24, 3.02]
109 weeks	1	16	1343	5	357	—	—	0.84 [0.29, 2.42]
Fostamatinib 100 mg bid vs. 100 mg qd	2	16	1424	0	269	—	—	3.22 [0.77, 13.52]
Malignant neoplasms								
fostamatinib vs PBO	6	12	2038	1	967	6	0.38	3.08 [0.96, 9.91]
12 weeks	1	1	142	0	47	—	—	3.78 [0.04, 352.61]
24 weeks	2	1	372	1	161	68	0.08	0.36 [0.02, 7.76]
26 weeks	1	2	304	0	153	—	—	4.51 [0.24, 85.33]
52 weeks	2	8	1220	0	606	0	1	**4.49 [1.03, 19.60]**
Fostamatinib 100 mg bid vs. 150 mg qd	7	19	2954	7	1339	39	0.16	1.09 [0.45, 2.60]
12 weeks	1	0	49	0	47	—	—	—
24 weeks	2	0	186	1	181	—	—	0.14 [0.00, 7.02]
26 weeks	1	2	152	0	152	—	—	7.44 [0.46, 119.46]
52 weeks	2	6	1224	3	602	72	0.06	0.98 [0.24, 3.96]
109 weeks	1	11	1343	3	357	—	—	0.97 [0.27, 3.54]
Fostamatinib 100 mg bid vs. 100 mg qd	2	11	1424	0	269	—	—	3.21 [0.57, 18.05]
fostamatinib vs Aadamumab	1	1	159	0	54	—	—	1.03 [0.04, 25.70]
Malignant neoplasms by system								
Bone and articular cartilage								
fostamatinib vs PBO	2	1	765	1	354	68	0.08	0.36 [0.02, 7.67]
ill-defined, secondary and unspecified sites								
fostamatinib vs PBO	3	2	1222	1	759	74	0.05	1.58 [0.16, 15.90]
fostamatinib 100 mg bid vs. 150 mg qd	4	6	2719	0	1111	0	0.99	4.64 [0.80, 26.88]
Digestive organs								
fostamatinib 100 mg bid vs. 150 mg qd	2	1	1957	3	661	0	0.92	**0.06 [0.01, 0.59]**
urinary tract								
fostamatinib vs PBO	4	4	1273	0	613	0	1	4.33 [0.52, 36.07]
fostamatinib 100 mg bid vs. 150 mg qd	5	2	2263	2	968	50	0.11	0.60 [0.08, 4.81]
Benign neoplasms								
Fostamatinib vs PBO	6	4	2038	1	967	0	0.74	1.71 [0.26, 11.36]
12 weeks	1	0	142	0	47	—	—	—
24 weeks	2	1	372	0	161	—	—	3.77 [0.04, 356.08]
26 weeks	1	1	304	0	153	—	—	4.50 [0.07, 286.14]
52 weeks	2	2	1220	1	606	0	0.39	0.99 [0.09, 11.00]
Fostamatinib 100 mg bid vs. 150 mg qd	7	6	2954	5	1339	0	0.53	0.41 [0.10, 1.57]
12 weeks	1	0	49	0	47	—	—	—

(Continued on following page)

TABLE 2 | (*Continued*) Pooled data of neoplasms in RA patients.

Comparisons	No of study	Fostamatinib		Comparator		Heterogeneity		Peto OR, 95% CI
		No of cancer	No of participants	No of cancer	No of participants	I2	P	
24 weeks	2	0	186	1	181	–	–	0.12 [0.00, 6.15]
26 weeks	1	0	152	1	152	–	–	0.14 [0.00, 6.82]
52 weeks	2	1	1224	1	602	55	0.13	0.46 [0.02, 8.82]
109 weeks	1	5	1343	2	357	–	–	0.63 [0.10, 3.91]
fostamatinib 100 mg bid vs. 100 mg qd	2	5	1424	0	269	–	–	3.19 [0.25, 41.20]

Bold values indicate statistical significance.

of fostamatinib once daily, 100 mg of fostamatinib twice daily did not increase the malignant neoplasm risk (Peto OR = 3.21, 95%CI 0.57–18.05, and 11/1424 vs. 0/269, and n = 2) (**Table 2**, **Supplementary Figure S5**).

Regarding fostamatinib vs. Adamumab, one trial examined the differences in the malignant neoplasm risk, and the use of fostamatinib did not have a higher malignant neoplasm risk compared to adamumab (Peto OR = 1.03, 95%CI 0.04–25.70, and 1/159 vs. 0/54) (**Table 2**, **Supplementary Figure S6**).

3.5.3 Malignant Neoplasms by System

For the bone and articular cartilage system, one malignant neoplasm was reported in 765 RA patients treated with fostamatinib, which was comparable to that in RA patients treated with placebo (Peto OR = 0.36, 95%CI 0.02–7.67, 1/765 vs. 1/354, I^2 = 68, and n = 2) (**Table 2**; **Supplementary Figure S7**).

For ill-defined, secondary, and unspecified sites, four studies reported malignant neoplasms. Compared to placebo, fostamatinib did not have a higher malignant neoplasm risk than placebo (Peto OR = 1.58, 95%CI 0.16–15.90, 2/1222 vs. 1/759, I^2 = 74, and n = 3) (**Table 2**; **Supplementary Figure S8**). A dosage of 100 mg of fostamatinib twice daily did not increase the risk of malignant neoplasms compared to 150 mg of fostamatinib once daily (Peto OR = 4.64, 95%CI 0.80–26.88, 6/2719 vs. 0/1111, I^2 = 0, and n = 4) (**Table 2**; **Supplementary Figure S9**).

For the digestive organs, two trials reported malignant neoplasms. RA patients who used fostamatinib 100 mg twice daily had a lower risk of malignant neoplasms than those who used 150 mg of fostamatinib once daily (Peto OR = 0.06, 95%CI 0.01–0.59, 1/1957 vs. 3/661, I^2 = 0) (**Table 2**; **Supplementary Figure S10**).

For the urinary tract, there were five trials focused on malignant neoplasms. The use of fostamatinib did not have a higher risk of malignant neoplasms than the placebo (Peto OR = 4.33, 95%CI 0.52–36.07, 4/1273 vs. 0/613, I^2 = 0, and n = 4) (**Table 2**; **Supplementary Figure S11**). RA patients who used fostamatinib 100 mg twice daily did not have an increased risk of malignant neoplasms compared to those who used 150 mg of fostamatinib once daily (Peto OR = 0.60, 95%CI 0.08–4.81, 2/2263 vs. 2/968, I^2 = 50, and n = 5) (**Table 2**; **Supplementary Figure S12**).

3.5.4 Benign Neoplasms

Compared to placebo, the use of fostamatinib did not have a higher benign neoplasm risk (Peto OR = 1.71, 95%CI 0.26–11.36, I^2 = 0, 4/2038 vs. 1/967, and n = 6) (**Table 2**, **Supplementary Figure S13**). A longer duration of fostamatinib use did not increase the risk of benign neoplasms (**Table 2**; **Supplementary Figure S13**). Compared to 150 mg of fostamatinib once daily, 100 mg of fostamatinib twice daily did not elevate the risk of benign neoplasms (Peto OR = 0.41, 95%CI 0.10–1.57, I^2 = 0, 6/2954 vs. 5/1339, and n = 7) (**Table 2**; **Supplementary Figure S14**). Similarly, long-term use of 100 mg of fostamatinib twice daily did not increase the risk of benign neoplasms (**Table 2**; **Supplementary Figure S14**).

In total, two studies reported a benign neoplasm risk in RA patients treated with different dosage times of fostamatinib. For

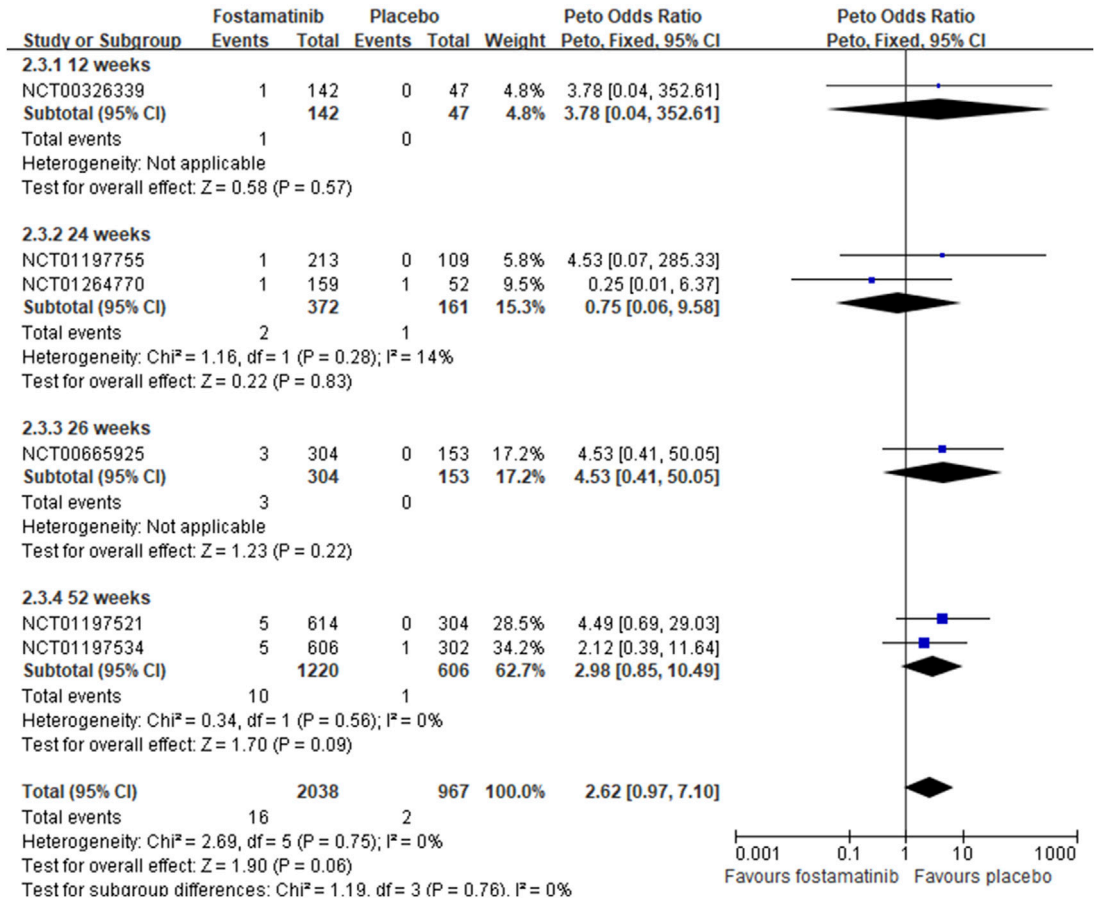

FIGURE 5 | Forest plot of overall neoplasms based on fostamatinib versus placebo.

100 mg of fostamatinib, a dosage time of twice daily did not increase benign neoplasm risk compared to once daily (Peto OR = 3.19, 95%CI 0.25–41.20, 5/1424 vs. 0/269) (**Table 2**; **Supplementary Figure S15**).

4 DISCUSSION

4.1 Main Findings

Our results indicate that fostamatinib was not associated with the risks of overall neoplasms as compared to placebo (16 cases in 2038 participants vs 2 cases in 967 participants), whereas, use of fostamatinib might be related to increased malignant neoplasm risk: longer duration of fostamatinib use might be correlated with an increased risk of malignant neoplasms, 8 cases in 1220 participants vs 0 cases in 606 participants at 52 weeks, and higher dose of fostamatinib may increase the malignant neoplasm risk in the digestive system, 3 cases in 661 participants who used 150 mg once a day vs 1 case in 1957 participants taking 100 mg twice a day.

Askling et al. reported that the highest crude incidence rates of all malignancies excluding non-melanoma skin cancer, solid malignancies, all skin cancers, and malignant lymphomas in RA patients treated with fostamatinib were 1.36, 1.47, 0.74, and 0.10 per 100 person-years, respectively (Askling et al., 2016), which were similar in RA patients with other treatments. The incidence of solid cancer in RA patients treated with tumor necrosis factor α inhibitors (TNFi) and conventional synthetic disease-modifying anti-rheumatic drugs (csDMARDs) was 0.81 and 1.17 per 100 person-years, respectively, and no difference was found between the two groups after adjusting for baseline characteristics (hazard ratio [HR] = 0.83, 95%CI 0.64–1.07) (Mercer et al., 2015).

Although the high inflammatory activity of RA has been reported to be associated with increased lymphoma (Baecklund et al., 2006), Simon et al. suggested that some types of malignant tumors are related to treatment rather than the underlying disease (Simon et al., 2015). However, compared with no treatment, treatments such as those involving TNFi (incidence rate ratios [IRR] = 1.1, 95%CI 0.8–1.6) and tocilizumab (IRR = 1.2, 95%CI 0.5–2.9) were not associated with increased melanoma risk (Mercer et al., 2017). Further, compared with csDMARD, TNFi did not increase the risk of lymphomas, solid cancers excluding non-melanoma skin cancer, non-melanoma skin cancer, and melanoma skin cancer in patients with RA (Chen et al., 2016; De Cock and Hyrich, 2018).

By targeting Syk, fostamatinib has been used to treat a wide range of diseases, such as graft-versus-host disease (Flynn et al.,

2015), follicular lymphoma (Fruchon et al., 2012), chronic lymphocytic leukemia (Quiroga et al., 2009), Waldenström macroglobulinemia (Kuiatse et al., 2015), ulcerative colitis (Can et al., 2015), and idiopathic thrombocytopenic purpura (Bajpai, 2009), in addition to treating RA. The mechanisms underlying the treatment of such diseases include targeting Syk signaling in B-cells and promoting their apoptosis for graft-versus-host disease (Flynn et al., 2015); suppressing the expression of matrix metalloproteinase 9 and angiogenesis through Syk-mTOR pathway for follicular lymphoma (Fruchon et al., 2012); inhibiting the phosphorylation of B-cell downstream signaling molecules, Syk, ERK, and AKT to reduce the production of CXCL12 and CXCL13 chemokines for chronic lymphocytic leukemia (Quiroga et al., 2009); inhibiting the activation of Syk and Bruton's tyrosine kinase and suppressing downstream signaling through MAPK kinase (MEK), p44/42 MAPK, and protein kinase B/Akt to prolong the onset of tumor growth and reduce viability of primary Waldenström macroglobulinemia cells (Kuiatse et al., 2015); inhibiting tissue myeloperoxidase activity and suppressing the molecular expressions of TNFα, CD3, Syk, and phospho-Syk in tissues (Can et al., 2015); reducing inflammation through decreased major inflammatory cytokines such as TNFα, IL-1, and IL-6 and inhibiting bone degradation for the autoimmune diseases idiopathic thrombocytopenic purpura and RA (Bajpai, 2009; Boccia et al., 2020).

Moreover, fostamatinib can reduce inflammatory cell adhesion and migration, diminish macrophage survival, and normalize upregulated monocytosis and inflammatory gene expression induced by a high-cholesterol diet (Hilgendorf et al., 2011). Thus, fostamatinib can be used to treat RA. Theoretically, fostamatinib can reduce the disease activity of RA and be used to treat lymphoma; thus, the finding of this study that longer-time use of fostamatinib could increase malignant neoplasm risk should be explained with caution. Fostamatinib is only effective in approximately 50% (680/1419) of RA patients assessed by ACR20, (Kunwar et al., 2016); therefore, uncontrolled disease activity would contribute to the risk of malignant tumors. In contrast, owing to the limited data, subgroup analyses were only performed for the nature of the neoplasm, follow-up periods, and the neoplasm-originating system, but analysis of subtypes of malignant neoplasms was not conducted. Thus, we could not focus on a specific neoplasm type, such as B-cell lymphoma, as Syk is reported to be necessary for B-cell development, proliferation, and survival.

LIMITATIONS

Several limitations of this meta-analysis and systematic review should be considered. First, only RCTs were eligible for inclusion, whereas no cohort and case-control studies were included. Although RCTs can balance the baseline measurements, having the least potential bias and less likely to be affected by possible confounders, the sample size and follow-up duration were relatively small, and short. For the small smaple size, the minimal reported sample size in the subgroup in our review was only based on 96 participants. Especially for rare cases, such as neoplasm, a small sample size may be underpowered to detect the outcomes. Moreover, the number of included studies was small and only seven RCTs were included. When performing subgroup analysis, several outcomes were reported only by one study; thus, the confidence intervals were relatively wide, affecting the reliability of the results. Therefore, well-planned observational studies with large study populations, such as cohort studies, are needed to determine whether fostamatinib is associated with increased or decreased malignant neoplasm risk in RA patients.

CONCLUSION

Our findings suggest that a longer duration of fostamatinib use in RA patients increases the risk of malignant neoplasms and a higher dose of fostamatinib may increase malignant neoplasms in the digestive system. However, owing to the small sample size and short follow-up duration, further studies such as cohort studies with large study populations and longer follow-up times are required to rule out the results.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

QX, GY, and YC conceived the study, YC and HL designed the study forms, GY and QX guided this study. YC searched the literature; YC, HL, YT, and ZL screened the studies for inclusion, extracted data, and assessed methodological quality; YC, GY, and QX organized and analyzed data. All authors drafted and revised the manuscript.

FUNDING

This study is supported by Post-doctoral Research and Development Fund of West China Hospital of Sichuan University (2019HXBH090), Clinical Research Incubation Project, West China Hospital, Sichuan University (2019HXFH038 and 2021HXFH018), and Sichuan Science and Technology Program (2021JDRC0045, 2021YFS0164, 2021YJ0472, and 2021JDRC0169).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2022.768980/full#supplementary-material

REFERENCES

Askling, J., Berglind, N., Franzen, S., Frisell, T., Garwood, C., Greenberg, J. D., et al. (2016). How Comparable Are Rates of Malignancies in Patients with Rheumatoid Arthritis across the World? A Comparison of Cancer Rates, and Means to Optimise Their Comparability, in Five RA Registries. *Ann. Rheum. Dis.* 75 (10), 1789–1796. doi:10.1136/annrheumdis-2015-208105

Baecklund, E., Iliadou, A., Askling, J., Ekbom, A., Backlin, C., Granath, F., et al. (2006). Association of Chronic Inflammation, Not its Treatment, with Increased Lymphoma Risk in Rheumatoid Arthritis. *Arthritis Rheum.* 54 (3), 692–701. doi:10.1002/art.21675

Bajpai, M., (2009). Fostamatinib, a Syk Inhibitor Prodrug for the Treatment of Inflammatory Diseases. *IDrugs* 12 (3), 174–185.

Boccia, R., Cooper, N., Ghanima, W., Boxer, M. A., Hill, Q. A., Sholzberg, M., et al. (2020). Fostamatinib Is an Effective Second-Line Therapy in Patients with Immune Thrombocytopenia. *Br. J. Haematol.* 190 (6), 933–938. doi:10.1111/bjh.16959

Buchner, M., Baer, C., Prinz, G., Dierks, C., Burger, M., Zenz, T., et al. (2010). Spleen Tyrosine Kinase Inhibition Prevents Chemokine- and Integrin-Mediated Stromal Protective Effects in Chronic Lymphocytic Leukemia. *Blood* 115 (22), 4497–4506. doi:10.1182/blood-2009-07-233692

Can, G., Ayvaz, S., Can, H., Demirtas, S., Aksit, H., Yilmaz, B., et al. (2015). The Syk Inhibitor Fostamatinib Decreases the Severity of Colonic Mucosal Damage in a Rodent Model of Colitis. *J. Crohns Colitis* 9 (10), 907–917. doi:10.1093/ecco-jcc/jjv114

Chen, Y., Liu, H., Huang, Y., Lin, S., Yin, G., and Xie, Q. (2021). The Cardiovascular Risks of Fostamatinib in Patients with Rheumatoid Arthritis: A Systematic Review and Meta-Analysis. *Front. Pharmacol.* 12, 632551. doi:10.3389/fphar.2021.632551

Chen, Y., Sun, J., Yang, Y., Huang, Y., and Liu, G. (2016). Malignancy Risk of Anti-tumor Necrosis Factor Alpha Blockers: an Overview of Systematic Reviews and Meta-Analyses. *Clin. Rheumatol.* 35 (1), 1–18. doi:10.1007/s10067-015-3115-7

De Cock, D., and Hyrich, K. (2018). Malignancy and Rheumatoid Arthritis: Epidemiology, Risk Factors and Management. *Best Pract. Res. Clin. Rheumatol.* 32 (6), 869–886. doi:10.1016/j.berh.2019.03.011

Dougados, M. (2016). Comorbidities in Rheumatoid Arthritis. *Curr. Opin. Rheumatol.* 28 (3), 282–288. doi:10.1097/BOR.0000000000000267

Efficacy and Safety Study of R935788 Tablets to Treat Rheumatoid Arthritis (Taski-2). Available at: https://www.clinicaltrials.gov/ct2/show/NCT00665925?term=NCT00665925&draw=2&rank=1.

Evaluation of Effectiveness of Two Dosing Regimens of Fostamatinib Compared to Placebo in Patients with Rheumatoid Arthritis (RA) Who Are Taking Disease Modifying Anti-rheumatic Drug (DMARD) but Not Responding. (OSKIRA - 2). Available at: https://www.clinicaltrials.gov/ct2/show/NCT01197534?term=NCT01197534&draw=1&rank=1.

Evaluation of Effectiveness of Two Dosing Regimens of Fostamatinib Compared to Placebo in Patients with Rheumatoid Arthritis (RA) Who Are Taking Methotrexate and Have Had Inadequate Response to Single TNF-Alpha Antagonist (OSKIRA - 3). Available at: https://www.clinicaltrials.gov/ct2/show/NCT01197755?term=NCT01197755&draw=2&rank=1.

Evaluation of Effectiveness of Two Dosing Regimens of Fostamatinib Compared to Placebo in Patients with Rheumatoid Arthritis (RA) Who Are Taking Methotrexate but Not Responding. (OSKIRA - 1). Available at: https://www.clinicaltrials.gov/ct2/show/NCT01197521?term=NCT01197521&draw=2&rank=1.

Evaluation of Efficacy and Safety of Fostamatinib Monotherapy Compared with Adalimumab Monotherapy in Patients with Rheumatoid Arthritis (RA) (OSKIRA -4). https://www.clinicaltrials.gov/ct2/show/NCT01264770?term=NCT01264770&draw=1&rank=1.

Evaluation of Long-Term Safety and Effectiveness of Fostamatinib in the Treatment of Rheumatoid Arthritis (RA) (OSKIRA-X). Available at: https://www.clinicaltrials.gov/ct2/show/NCT01242514?term=NCT01242514&draw=2&rank=1.

Flynn, R., Allen, J. L., Luznik, L., MacDonald, K. P., Paz, K., Alexander, K. A., et al. (2015). Targeting Syk-Activated B Cells in Murine and Human Chronic Graft-Versus-Host Disease. *Blood* 125 (26), 4085–4094. doi:10.1182/blood-2014-08-595470

Fruchon, S., Kheirallah, S., Al Saati, T., Ysebaert, L., Laurent, C., Leseux, L., et al. (2012). Involvement of the Syk-mTOR Pathway in Follicular Lymphoma Cell Invasion and Angiogenesis. *Leukemia* 26 (4), 795–805. doi:10.1038/leu.2011.248

Higgins, J. P., Thompson, S. G., Deeks, J. J., and Altman, D. G. (2003). Measuring Inconsistency in Meta-Analyses. *BMJ* 327 (7414), 557–560. doi:10.1136/bmj.327.7414.557

Hilgendorf, I., Eisele, S., Remer, I., Schmitz, J., Zeschky, K., Colberg, C., et al. (2011). The Oral Spleen Tyrosine Kinase Inhibitor Fostamatinib Attenuates Inflammation and Atherogenesis in Low-Density Lipoprotein Receptor-Deficient Mice. *Arterioscler Thromb. Vasc. Biol.* 31 (9), 1991–1999. doi:10.1161/ATVBAHA.111.230847

Hoellenriegel, J., Coffey, G. P., Sinha, U., Pandey, A., Sivina, M., Ferrajoli, A., et al. (2012). Selective, Novel Spleen Tyrosine Kinase (Syk) Inhibitors Suppress Chronic Lymphocytic Leukemia B-Cell Activation and Migration. *Leukemia* 26 (7), 1576–1583. doi:10.1038/leu.2012.24

JPT, H., and So, G. (2011). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updatedAssessing Risk of Bias in Included Studies. Www. cochrane-handbook. Org. *The Cochrane Collaboration*, Chapter 82011

Kang, Y., Jiang, X., Qin, D., Wang, L., Yang, J., Wu, A., et al. (2019). Efficacy and Safety of Multiple Dosages of Fostamatinib in Adult Patients with Rheumatoid Arthritis: A Systematic Review and Meta-Analysis. *Front. Pharmacol.* 10, 897. doi:10.3389/fphar.2019.00897

Kuiatse, I., Baladandayuthapani, V., Lin, H. Y., Thomas, S. K., Bjorklund, C. C., Weber, D. M., et al. (2015). Targeting the Spleen Tyrosine Kinase with Fostamatinib as a Strategy against Waldenström Macroglobulinemia. *Clin. Cancer Res.* 21 (11), 2538–2545. doi:10.1158/1078-0432.CCR-14-1462

Kunwar, S., Devkota, A. R., and Ghimire, D. K. (2016). Fostamatinib, an Oral Spleen Tyrosine Kinase Inhibitor, in the Treatment of Rheumatoid Arthritis: a Meta-Analysis of Randomized Controlled Trials. *Rheumatol. Int.* 36 (8), 1077–1087. doi:10.1007/s00296-016-3482-7

Mercer, L. K., Askling, J., Raaschou, P., Dixon, W. G., Dreyer, L., Hetland, M. L., et al. (2017). Risk of Invasive Melanoma in Patients with Rheumatoid Arthritis Treated with Biologics: Results from a Collaborative Project of 11 European Biologic Registers. *Ann. Rheum. Dis.* 76 (2), 386–391. doi:10.1136/annrheumdis-2016-209285

Mercer, L. K., Lunt, M., Low, A. L., Dixon, W. G., Watson, K. D., Symmons, D. P., et al. (2015). Risk of Solid Cancer in Patients Exposed to Anti-tumour Necrosis Factor Therapy: Results from the British Society for Rheumatology Biologics Register for Rheumatoid Arthritis. *Ann. Rheum. Dis.* 74 (6), 1087–1093. doi:10.1136/annrheumdis-2013-204851

Moher, D., Liberati, A., Tetzlaff, J., and Altman, D. G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA Statement. *BMJ* 339, b2535. doi:10.1136/bmj.b2535

Quiroga, M. P., Balakrishnan, K., Kurtova, A. V., Sivina, M., Keating, M. J., Wierda, W. G., et al. (2009). B-cell Antigen Receptor Signaling Enhances Chronic Lymphocytic Leukemia Cell Migration and Survival: Specific Targeting with a Novel Spleen Tyrosine Kinase Inhibitor, R406. *Blood* 114 (5), 1029–1037. doi:10.1182/blood-2009-03-212837

Salgado, E., Maneiro, J. R., Carmona, L., and Gomez-Reino, J. J. (2014). Safety Profile of Protein Kinase Inhibitors in Rheumatoid Arthritis: Systematic Review and Meta-Analysis. *Ann. Rheum. Dis.* 73 (5), 871–882. doi:10.1136/annrheumdis-2012-203116

Scott, D. L., Wolfe, F., and Huizinga, T. W. (2010). Rheumatoid Arthritis. *Lancet* 376 (9746), 1094–1108. doi:10.1016/S0140-6736(10)60826-4

Simon, T. A., Thompson, A., Gandhi, K. K., Hochberg, M. C., and Suissa, S. (2015). Incidence of Malignancy in Adult Patients with Rheumatoid Arthritis: a Meta-Analysis. *Arthritis Res. Ther.* 17 (1), 212. doi:10.1186/s13075-015-0728-9

Smitten, A. L., Simon, T. A., Hochberg, M. C., and Suissa, S. (2008). A Meta-Analysis of the Incidence of Malignancy in Adult Patients with Rheumatoid Arthritis. *Arthritis Res. Ther.* 10 (2), R45. doi:10.1186/ar2404

Smolen, J. S., Aletaha, D., and McInnes, I. B. (2016). Rheumatoid Arthritis. *Lancet* 388 (10055), 2023–2038. doi:10.1016/S0140-6736(16)30173-8

Treatment of Arthritis with Syk Kinase Inhibition (TASKI-1). Available at: https://www.clinicaltrials.gov/ct2/show/NCT00326339?term=NCT00326339&draw=1&rank=1.

WHO The Tenth Revision of the International Classification of Diseases (ICD-10). Available at: http://www.cdc.gov/nchs/icd/.

Widdifield, J., Paterson, J. M., Huang, A., and Bernatsky, S. (2018). Causes of Death in Rheumatoid Arthritis: How Do They Compare to the General Population? *Arthritis Care Res. (Hoboken)* 70 (12), 1748–1755. doi:10.1002/acr.23548

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Chen, Liu, Tian, Luo, Yin and Xie. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Public Health

ORIGINAL RESEARCH
published: 04 March 2022
doi: 10.3389/fpubh.2022.842104

Prediction of Atrial Fibrillation in Hospitalized Elderly Patients With Coronary Heart Disease and Type 2 Diabetes Mellitus Using Machine Learning: A Multicenter Retrospective Study

*Qian Xu[1,2,3], Yan Peng[4], Juntao Tan[5], Wenlong Zhao[1,2], Meijie Yang[1] and Jie Tian[2,6,7]**

[1] College of Medical Informatics, Chongqing Medical University, Chongqing, China, [2] Medical Data Science Academy, Chongqing Medical University, Chongqing, China, [3] Collection Development Department of Library, Chongqing Medical University, Chongqing, China, [4] Department of Cardiology, University-Town Hospital of Chongqing Medical University, Chongqing, China, [5] Operation Management Office, Affiliated Banan Hospital of Chongqing Medical University, Chongqing, China, [6] Department of Cardiology, Ministry of Education Key Laboratory of Child Development and Disorders, National Clinical Research Center for Child Health and Disorders, China International Science and Technology Cooperation Base of Child Development and Critical Disorders, Children's Hospital of Chongqing Medical University, Chongqing, China, [7] Chongqing Key Laboratory of Pediatrics, Chongqing, China

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Kurt Neumann,
Independent Researcher,
Kerékteleki, Hungary
Primoz Kocbek,
University of Maribor, Slovenia

***Correspondence:**
Jie Tian
jietian@cqmu.edu.cn

Specialty section:
This article was submitted to Aging and Public Health, a section of the journal Frontiers in Public Health

Received: 23 December 2021
Accepted: 09 February 2022
Published: 04 March 2022

Citation:
Xu Q, Peng Y, Tan J, Zhao W, Yang M and Tian J (2022) Prediction of Atrial Fibrillation in Hospitalized Elderly Patients With Coronary Heart Disease and Type 2 Diabetes Mellitus Using Machine Learning: A Multicenter Retrospective Study. Front. Public Health 10:842104. doi: 10.3389/fpubh.2022.842104

Background: The objective of this study was to use machine learning algorithms to construct predictive models for atrial fibrillation (AF) in elderly patients with coronary heart disease (CHD) and type 2 diabetes mellitus (T2DM).

Methods: The diagnosis and treatment data of elderly patients with CHD and T2DM, who were treated in four tertiary hospitals in Chongqing, China from 2015 to 2021, were collected. Five machine learning algorithms: logistic regression, logistic regression+least absolute shrinkage and selection operator, classified regression tree (CART), random forest (RF) and extreme gradient lifting (XGBoost) were used to construct the prediction models. The area under the receiver operating characteristic curve (AUC), sensitivity, specificity, and accuracy were used as the comparison measures between different models.

Results: A total of 3,858 elderly patients with CHD and T2DM were included. In the internal validation cohort, XGBoost had the highest AUC (0.743) and sensitivity (0.833), and RF had the highest specificity (0.753) and accuracy (0.735). In the external verification, RF had the highest AUC (0.726) and sensitivity (0.686), and CART had the highest specificity (0.925) and accuracy (0.841). Total bilirubin, triglycerides and uric acid were the three most important predictors of AF.

Conclusion: The risk prediction models of AF in elderly patients with CHD and T2DM based on machine learning algorithms had high diagnostic value. The prediction models constructed by RF and XGBoost were more effective. The results of this study can provide reference for the clinical prevention and treatment of AF.

Keywords: coronary heart disease, type 2 diabetes mellitus, atrial fibrillation, machine learning, prediction models

INTRODUCTION

Coronary heart disease (CHD) is the most common cardiovascular disease among the elderly (1). According to the data from the World Health Organization (2014), CHD is the main cause of death globally, accounting for more than 7 million deaths every year (2). As the most common type of diabetes mellitus (DM), type 2 diabetes mellitus (T2DM) has become the leading cause of morbidity and mortality worldwide. In 2019, there were 463 million individuals affected by diabetes globally (3). By 2040, the number of patients is expected to increase to 629 million, accounting for ∼90% of all cases (4). T2DM is one of the most important complications of CHD. It is an important risk factor for the development of CHD and associated with the death of patients (5).

Atrial fibrillation (AF) is the most common arrhythmia, which is characterized by atrial activation disorder, that results in the deterioration of atrial function (6, 7). The prevalence of AF increased with age in recent studies. The overall prevalence was estimated to be 5.5%, rising from 0.7% in the 55–59 age group to 17.8% in the 85 years-and-above age group (8). Owing to the aging population, the prevalence of AF is expected to double by 2050 (9). AF in hospitalized elderly patients with CHD and T2DM is a very serious cardiac adverse event. Once patients develop AF, further thromboembolism will occur, leading to a high risk of disability and death (10, 11). Therefore, early identification and the timely management of AF in elderly patients with CHD and T2DM is very important. However, early detection of AF based on traditional indicators is difficult.

Therefore, this study is aimed at developing AF prediction models for the elderly patients with CHD and T2DM by using different machine learning algorithms to provide reference for the clinical prevention and treatment of AF. In order to the interpretability of results, the machine learning algorithms included in this study are mainly composed of logistic regression (LR), logistic regression+least absolute shrinkage and selection operator (LR+LASSO), classification and regression trees (CART), random forests (RF), and extreme gradient boosting (XGBoost).

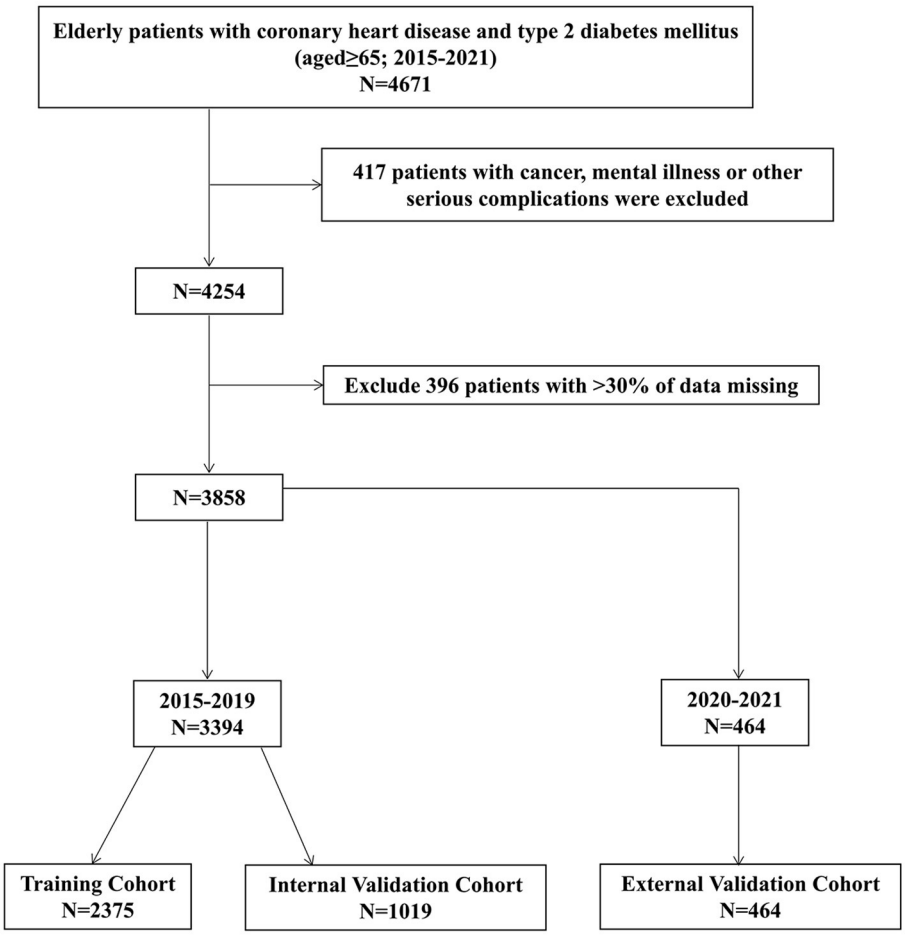

FIGURE 1 | Flow of inclusions and exclusions.

METHODS

Data Source

The study data were obtained from four tertiary hospitals in Chongqing, China, namely the Second Affiliated Hospital of Chongqing Medical University, the Third Affiliated Hospital of Chongqing Medical University, the University-Town Hospital of the Chongqing Medical University, and the Yong Chuan Hospital of the Chongqing Medical University. The transparent reporting of a multivariable predictive model for individual prognosis or diagnosis guidelines were followed for model development and validation. The clinical data were collected using electronic medical record systems. Data on patients with CHD and T2DM obtained for the period of 2015 to 2021 were used in the study.

The Ethics Committee of Chongqing Medical University approved the study. Written informed consent for participation was not required for this study owing to its retrospective design, and the study was undertaken in accordance with the national legislation and institutional requirements.

Definition

The diagnosis of AF (ICD-10, I48) was based on AHA/ACC/ESC 2006 atrial fibrillation guidelines (12). Two physicians with rich clinical experience were convened to diagnose together with echocardiography, electrocardiogram, and laboratory examinations.

Inclusion and Exclusion Criteria

Inclusion criteria comprised the following: (i) data obtained from 2015 to 2021, (ii) patients aged $\geq$ 65 years, and (iii) hospitalization (s) for CHD and T2DM.

Exclusion criteria comprised the following: (i) patients with cancer, mental illness, or other serious complications and (ii) patients with >30% of missing data. The study selection process is depicted in the following flow chart (**Figure 1**).

Data Collection

Based on previous studies, 28 possible risk factors for predicting AF were selected, namely age, gender, smoking status, drinking status, hypertension, diastolic blood pressure (DBP), γ-glutamyltransferase (GGT), white blood cell (WBC), neutrophil-to-lymphocyte ratio (NLR), lymphocyte-to-monocyte ratio (LMR), platelet-lymphocyte (PLR), blood creatinine, total bilirubin (TB), uric acid (UA), urea Nitrogen (UN), low-density lipoprotein (LDL), hemoglobin, alanine aminotransferase (ALT), albumin (ALB), triglycerides (TGs), total cholesterol (TC), high-density lipoprotein (HDL), glycated hemoglobin (HbA1c), alkaline phosphatase (ALP), aspartate aminotransferase (AST), blood potassium, blood calcium, and blood phosphorus.

TABLE 1 | Comparison of continuous variables before and after interpolation.

Variables	Before interpolation	After interpolation	*P*-values
DBP (IQR, mmHg)	77.00 (69.00, 85.00)	77.00 (69.00, 85.00)	0.894
GGT (IQR, IU/L)	26.00 (18.00, 44.00)	26.00 (18.00, 44.67)	0.760
WBC (IQR, ×109/L)	6.78 (5.52, 8.56)	6.78 (5.52, 8.56)	1.000
NLR (IQR)	3.33 (2.32, 5.51)	3.33 (2.32, 5.52)	0.969
LMR (IQR)	3.63 (2.37, 5.23)	3.63 (2.37, 5.23)	0.999
Blood creatinine	73.35 (57.70, 97.73)	73.35 (57.70, 97.78)	0.981
TB (IQR, umol/l)	10.30 (7.60, 14.00)	10.30 (7.60, 14.00)	0.944
UA (IQR, umol/L)	339.30 (273.48, 417.03)	338.90 (273.30, 417.08)	0.962
UN (IQR, mmol/L)	6.54 (5.20, 8.58)	6.54 (5.20, 8.57)	0.973
LDL (IQR, mmol/L)	2.23 (1.67, 2.91)	2.24 (1.67, 2.91)	0.871
Hemoglobin (IQR, g/L)	127.00 (115.00, 138.00)	127.00 (115.00, 138.00)	1.000
PLR (IQR)	128.43 (94.69, 178.29)	127.98 (94.40, 177.71)	0.776
ALT (IQR, IU/L)	18.16 (13.00, 27.14)	18.12 (13.00, 27.58)	0.987
ALB (IQR, g/L)	39.50 (36.50, 42.60)	39.50 (36.50, 42.60)	0.966
TGs (IQR, mmol/L)	1.41 (1.02, 2.02)	1.42 (1.02, 2.03)	0.774
TC (IQR, mmol/L)	4.10 (3.37, 4.96)	4.09 (3.37, 4.96)	0.958
HDL (IQR, mmol/L)	1.10 (0.91, 1.33)	1.09 (0.91, 1.32)	0.587
HbA1c (IQR, %)	7.41 (6.60, 9.04)	7.44 (6.60, 9.10)	0.743
ALP (IQR, IU/L)	74.00 (61.00, 92.00)	74.00 (61.00, 92.00)	0.878
AST (IQR, IU/L)	20.20 (16.00, 27.00)	20.20 (16.01, 27.00)	0.968
Blood potassium (IQR, mmol/L)	4.02 (3.72, 4.34)	4.02 (3.71, 4.33)	0.639
Blood calcium (IQR, mmol/L)	2.23 (2.13, 2.33)	2.23 (2.13, 2.33)	0.921
Blood phosphorus (IQR, mmol/L)	1.09 (0.95, 1.24)	1.09 (0.94, 1.23)	0.160

DBP, diastolic blood pressure; GGT, γ-glutamyltransferase; WBC, white blood cell; NLR, neutrophil-to-lymphocyte ratio; LMR, lymphocyte-to-monocyte ratio; TB, total bilirubin; UA, uric acid; UN, urea nitrogen; LDL, low-density lipoprotein; PLR, platelet-lymphocyte ratio; ALT, alanine aminotransferase; ALB, albumin; TGs, triglycerides; TC, total cholesterol; HDL, high-density lipoprotein; HbA1c, glycated hemoglobin; ALP, alkaline phosphatase; AST, aspartate aminotransferase; IQR, interquartile range.

Statistical Analyses

The multiple imputation method (predictive mean matching) was used to fill in the missing continuous variables. Since the visit time period of the patients included in this study spanned from 2015 to 2021, this study can further subjected to temporal validation (13, 14). Therefore, the data included in this study were divided into two parts: 2015 to 2019 and 2020 to 2021 according to date. The data from 2015 to 2019 were divided into the training cohort and internal validation cohort based on the ratio of 7:3. The data from 2020 to 2021 were used as an external validation cohort. The training cohort was used to develop predictive models using machine learning (CART, RF, and XGBoost) and logistic regression algorithms, and then the final performance of each model was verified and compared in the internal validation cohort and external validation cohort. The area under the receiver operating characteristic curve (AUC), sensitivity, specificity, and accuracy were used as the comparison measures between different models.

Continuous variables were expressed as mean ± standard deviation (SD) or median (interquartile range), according to the distribution of normality. Categorical variables were reported as counts with percentages. Continuous variables were tested by t-test or Mann-Whitney U-test whereas categorical variables were analyzed by Chi-square test (χ^2-test). All statistical analyses used two-sided tests, and $P < 0.05$ was considered statistically significant. Statistical analyses were performed using SPSS 22.0 and R (version 4.0.2, Vienna, Austria).

RESULTS

Patient Characteristics

A total of 3,858 patients with CHD and T2DM were included in this study. They were divided into a training cohort ($N = 2{,}375$), an internal validation cohort ($N = 1{,}019$), and an external validation cohort ($N = 464$). The Mann-Whitney U-test showed that there was no significant difference before and after interpolation (**Table 1**).

In the training cohort, the results showed that in comparison with patients in the non-AF group, patients who suffered from AF had a higher age ($P < 0.001$), higher DBP ($P = 0.005$), GGT (P

TABLE 2 | Univariate analyses of variables associated with AF.

Variables	AF (*N* = 270)	Non-AF (*N* = 2,105)	*P*-values
Gender (*n*, %)	119 (44.07%)	924 (43.90%)	0.999
Smoking status (*n*, %)	78 (28.89%)	574 (27.27%)	0.625
Drinking status (*n*, %)	55 (20.37%)	411 (19.52%)	0.804
Hypertension (*n*, %)	206 (76.30%)	1,670 (79.33%)	0.283
Age (IQR, years)	78.00 (73.00, 83.00)	75.00 (70.00, 81.00)	<0.001
DBP (IQR, mmHg)	79.00 (71.00, 88.00)	77.00 (69.00, 85.00)	0.005
GGT (IQR, IU/L)	36.00 (22.55, 69.00)	25.00 (17.60, 43.00)	<0.001
WBC (IQR, ×109/L)	6.82 (5.57, 8.39)	6.73 (5.53, 8.50)	0.848
NLR (IQR)	4.44 (2.74, 6.36)	3.23 (2.28, 5.29)	<0.001
LMR (IQR)	3.18 (2.02, 4.69)	3.81 (2.46, 5.44)	<0.001
Blood creatinine	80.30 (62.78, 105.18)	72.40 (56.90, 97.20)	0.001
TB (IQR, umol/l)	12.30 (9.20, 17.55)	10.00 (7.40, 13.70)	<0.001
UA (IQR, umol/L)	370.40 (296.43, 469.68)	332.10 (270.40, 410.00)	<0.001
UN (IQR, mmol/L)	7.09 (5.41, 9.46)	6.51 (5.23, 8.39)	0.001
LDL (IQR, mmol/L)	1.92 (1.47, 2.62)	2.27 (1.69, 2.95)	<0.001
Hemoglobin (IQR, g/L)	125.00 (112.25, 137.75)	127.00 (115.00, 139.00)	0.377
PLR (IQR)	129.96 (95.24, 180.14)	127.95 (94.27, 175.86)	0.356
ALT (IQR, IU/L)	19.00 (12.02, 28.00)	18.00 (13.00, 27.00)	0.722
ALB (IQR, g/L)	38.40 (35.79, 41.00)	39.60 (36.60, 42.70)	<0.001
TGs (IQR, mmol/L)	1.17 (0.83, 1.72)	1.42 (1.04, 2.03)	<0.001
TC (IQR, mmol/L)	3.64 (3.02, 4.54)	4.16 (3.41, 4.97)	<0.001
HDL (IQR, mmol/L)	1.05 (0.85, 1.27)	1.10 (0.92, 1.32)	0.007
HbA1c (IQR, %)	7.35 (6.50, 8.50)	7.40 (6.60, 9.00)	0.129
ALP (IQR, IU/L)	74.80 (58.43, 96.08)	74.00 (61.00, 91.90)	0.579
AST (IQR, IU/L)	22.17 (17.25, 31.23)	20.00 (16.00, 27.00)	0.001
Blood potassium (IQR, mmol/L)	3.97 (3.69, 4.36)	4.02 (3.72, 4.32)	0.694
Blood calcium (IQR, mmol/L)	2.21 (2.11, 2.29)	2.24 (2.14, 2.34)	0.001
Blood phosphorus (IQR, mmol/L)	1.09 (0.96, 1.25)	1.10 (0.95, 1.24)	0.942

AF, atrial fibrillation; DBP, diastolic blood pressure; GGT, γ-glutamyltransferase; WBC, white blood cell; NLR, neutrophil-to-lymphocyte ratio; LMR, lymphocyte-to-monocyte ratio; TB, total bilirubin; UA, uric acid; UN, urea nitrogen; LDL, low-density lipoprotein; PLR, platelet-lymphocyte ratio; ALT, alanine aminotransferase; ALB, albumin; TGs, triglycerides; TC, total cholesterol; HDL, high-density lipoprotein; HbA1c, glycated hemoglobin; ALP, alkaline phosphatase; AST, aspartate aminotransferase; IQR, interquartile range.

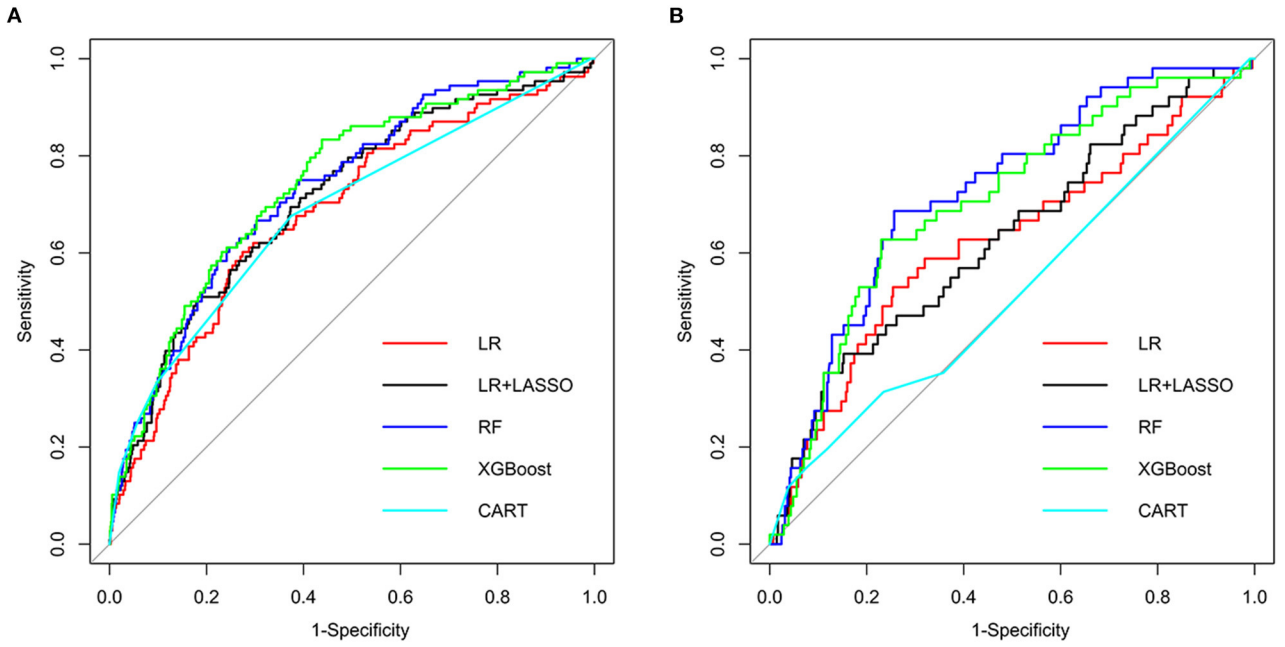

FIGURE 2 | Receiver operating characteristic (ROC) curves of five different models in internal validation cohort **(A)** and external validation cohort **(B)**.

TABLE 3 | Detailed performance metrics for the five models in internal validation.

Model	Sensitivity	Specificity	Accuracy	AUC (95%CI)
LR	0.602	0.726	0.712	0.684 (0.629–0.739)
LR+LASSO	0.694	0.627	0.633	0.712 (0.659–0.765)
CART	0.676	0.626	0.631	0.686 (0.632–0.739)
RF	0.611	0.753	0.735	0.733 (0.683–0.783)
XGBoost	0.833	0.562	0.587	0.743 (0.693–0.792)

LR, logistic regression; CART, Classification and Regression Tree; RF, random forest; AUC, area under the receiver operating characteristic curve; CI, confidence interval.

TABLE 4 | Detailed performance metrics for the five models in external validation.

Model	Sensitivity	Specificity	Accuracy	AUC (95%CI)
LR	0.529	0.746	0.722	0.621 (0.532–0.711)
LR+LASSO	0.353	0.889	0.828	0.630 (0.659–0.765)
CART	0.157	0.925	0.841	0.523 (0.444–0.601)
RF	0.686	0.743	0.733	0.726 (0.655–0.797)
XGBoost	0.627	0.770	0.754	0.705 (0.628–0.781)

LR, logistic regression; CART, Classification and Regression Tree; RF, random forest; AUC, area under the receiver operating characteristic curve; CI, confidence interval.

< 0.001), NLR ($P < 0.001$), blood creatinine ($P = 0.001$), higher TB ($P < 0.001$), UA ($P < 0.001$), UN ($P = 0.001$), and AST ($P < 0.001$). The levels of LMR ($P < 0.001$), LDL ($P < 0.001$), ALB ($P < 0.001$), TGs ($P < 0.001$), TC ($P < 0.001$), HDL ($P = 0.007$), and blood calcium ($P < 0.001$) in the AF group are lower compared with those in the non-AF group (**Table 2**).

TABLE 5 | Ranks of feature importance in RF, XGBoost, and LR+LASSO for predicting AF.

Rank	RF	XGBoost	LR+LASSO
1	TB	TGs	TC
2	GGT	UA	TGs
3	TGs	TC	Age
4	UA	GGT	TB
5	DBP	TB	UA

LR, logistic regression; RF, random forest; DBP, diastolic blood pressure; GGT, γ-glutamyltransferase; TB, total bilirubin; UA, uric acid; TGs, triglycerides; TC, total cholesterol.

Predictive Effects of Different Models

Five models were generated, including LR, LR+LASSO, CART, RF, and XGBoost, to predict the development of AF in elderly patients with CHD and T2DM after admission. **Figure 2** shows the performance of the five different models in predicting AF in the internal validation cohort and external validation cohort in terms of ROC curves. In the internal validation cohort, AUC shows that the XGBoost model has the best prediction effectiveness on AF, and AUC was 0.743 compared with other models. In the external validation cohort, the highest AUC model was RF.

Tables 3, **4**, respectively, show the detailed performance indicators of the five models in the internal verification cohort

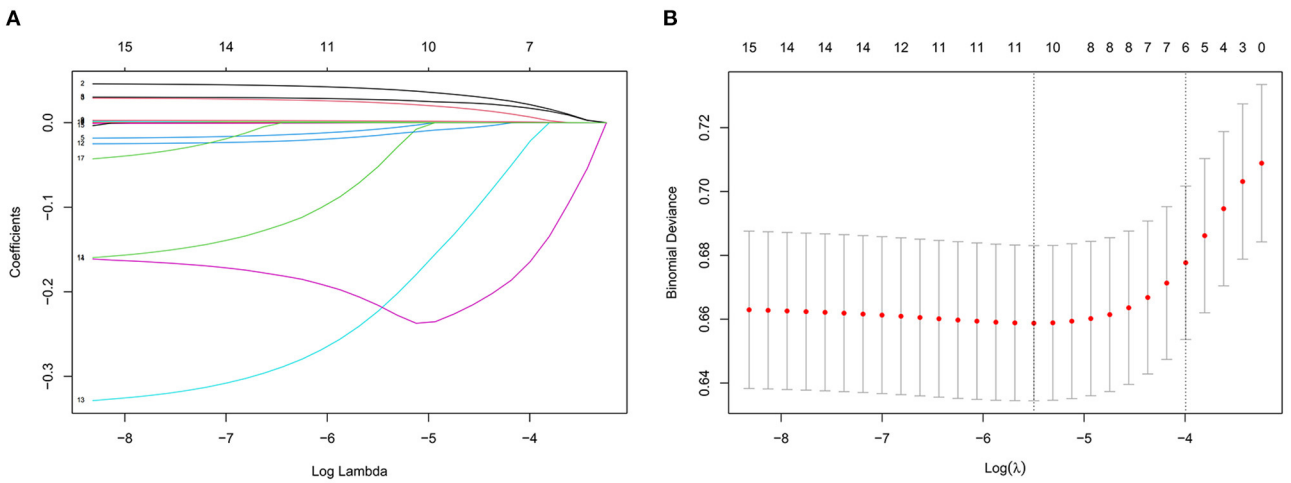

FIGURE 3 | Features selection by LASSO. **(A)** LASSO coefficients profiles (y-axis) of the 16 features. The upper x-axis is the average numbers of predictors and the lower x-axis is the log (λ). **(B)** Five-fold cross-validation for tuning parameter selection in the LASSO model.

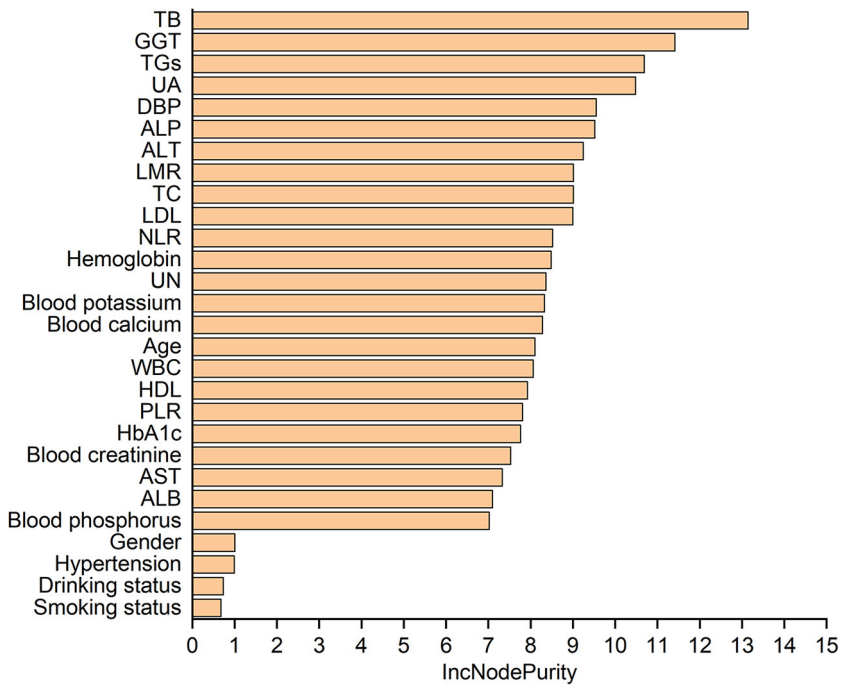

FIGURE 4 | Importance analysis of indexes in RF model.

and external verification cohort. In the internal validation cohort, XGBoost had the highest AUC (0.743) and sensitivity (0.833), and RF had the highest specificity (0.753) and accuracy (0.735). In the external verification, RF had the highest AUC (0.726) and sensitivity (0.686), and CART had the highest specificity (0.925) and accuracy (0.841).

Table 5 lists the importance levels of the top five features in the top three AUC models in both the internal validation cohort and the external validation cohort. As shown in **Table 5**, TB, TGs and UA were the top three features that promoted the development of the prediction models for AF in coronary heart disease and type 2 diabetes mellitus patients. **Figure 3** shows the feature screening process of LR+LASSO model, **Figures 4**, **5** shows the feature importance ranking of the RF model and the XGBoost model, respectively.

DISCUSSION

In this study, A total of 3,858 patients with CHD and T2DM were included, and then 28 possible risk factors were selected

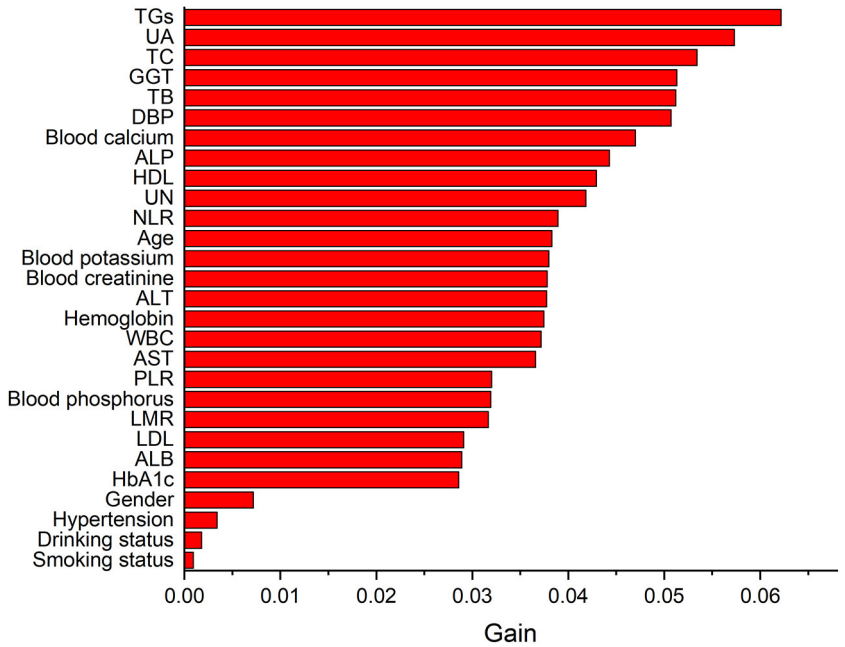

FIGURE 5 | Importance analysis of indexes in XGBoost model.

for predicting AF. Five machine learning algorithms were used to construct the prediction models, and then the final performance of each model was verified and compared in the internal validation cohort and the external validation cohort. The results showed that the prediction models constructed by RF and XGBoost were more effective (delong test showed that there was no significant difference between the two models, $P = 0.654$). TB, TGs and UA were the three most important predictors of AF.

Machine learning algorithms could deeply mine and analyze big data, and it had been widely used in disease predictions and prognostication (15–17). RF is one of the widely used algorithms in machine learning. It uses a set model of various decision trees (18). RF rely on computers to learn all of the complex non-linear interactions between variables by minimizing the error between observation results and prediction results (19), and it uses bootstrap aggregation and randomization of predictors to obtain high disease prediction accuracy (20, 21). In addition, XGBoost is also an efficient and widely used machine learning method (22). It is an integrated algorithm belonging to a class of boosting algorithms. The core of the boosting algorithm is to integrate many weak classifiers to form a strong classifier, so as to improve the accuracy of classification. XGBoost helps to reduce overfitting compared to gradient tree boosting by only a random subset of descriptors in building a tree and is known as the "regularized boosting" technique (23).

Our study shows that high TB was a risk factor for AF in elderly patients with CHD and T2DM. Recently, studies found that TB levels are related to CVD and AF (24, 25). Bilirubin is a product of hemoglobin catabolism and has long been used as a diagnostic indicator for hepatobiliary diseases and hemolytic diseases (26). AF promotes the occurrence of inflammation and the increase of inflammatory markers, while TB, as an antioxidant stress factor, increases compensatory in order to balance oxidation and antioxidation in the body (27). Therefore, for elderly patients with CHD and T2DM, when TB level is significantly increased, there is risk of AF.

TGs are the most abundant lipids in the human body. Most tissues can use TG decomposition products to supply energy. At the same time, liver, fat, and other tissues can also synthesize TGs and store them in adipose tissue (28, 29). TGs have been proven to be a risk factor for cardiovascular disease (CVD), but its role in the development of AF in the elderly is unclear (30). The study found that low TGs were a risk factor for AF in elderly patients with CHD and T2DM. A cross-sectional survey showed that Low TGs may be a marker of CVD risk in Chinese patients with long-term T2DM (31), which is consistent with the results of our study. Of note, clinicians usually pay attention to the risk caused by elevated TGs and ignore the impact of low TGs in patients, especially in elderly patients with CHD and T2DM. The results of this study suggest that clinicians should pay attention to the TG level of elderly patients with CHD and T2DM in order to prevent atrial fibrillation in advance.

This study also showed that high UA was a risk factor for AF in elderly patients with CHD and T2DM. In recent years, several cross-sectional studies and prospective studies had shown that a high UA level is a potential risk factor for AF (32–34). Another study showed that high UA is associated with an increased prevalence of AF in hospitalized patients with T2DM, independent of multiple risk factors and potential confounders (35). Yutong Ji et al. also found that there is an association between AF and high UA in elderly patients, high UA is associated with persistent or permanent AF (36). Therefore, reducing the level of UA can reduce the risk of AF in elderly patients to a certain extent.

This study is the first to investigate the predictive model for AF in hospitalized elderly patients with coronary heart disease and type 2 diabetes mellitus. However, this study also has several limitations. Firstly, selection bias may exist due to the retrospective nature of the investigation and the imbalanced dataset. However, a multicenter and relatively large training cohort were used to build models, which was further subjected to temporal validation. Secondly, data on social support, socioeconomic status, and some other important factors were not available. Further research is warranted to explore the impact of these important indicators.

In conclusion, compared with classic logistic regression models, machine learning models (XGBoost and RF) have better performance in predicting AF in elderly patients with CHD and T2DM using features that were easily available at admission. Predictive models using machine learning algorithms can help clinicians predict AF early and implement targeted treatment measures.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The Ethics Committee of Chongqing Medical University approved the study. Written informed consent for participation was not required for this study due to its retrospective design and the study was undertaken in accordance with national legislation and institutional requirements.

AUTHOR CONTRIBUTIONS

QX, YP, JTa, and MY: conception, methodology, and draft of the manuscript. WZ: suggestions for amendment. JTi: review and supervision. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by Chongqing Federation of Social Sciences (No. 2021NDQN65) and the Intelligent Medical Foundation of Chongqing Medical University (No. ZHYX202029).

REFERENCES

1. Tang M, Wang SH, Li HL, Chen H, Ma SJ. Mental health status and quality of life in elderly patients with coronary heart disease. *PeerJ.* (2021) 9:e10903. doi: 10.7717/peerj.10903
2. Nowbar AN, Howard JP, Finegold JA, Asaria P, Francis DP. 2014 Global geographic analysis of mortality from ischaemic heart disease by country, age and income: Statistics from World Health Organisation and United Nations. *Int J Cardiol.* (2014) 174:293–8. doi: 10.1016/j.ijcard.2014.04.096
3. Rong S, Birong W, Zheyun N, Hui S, Fan H. Nomogram based on risk factors for type 2 diabetes mellitus patients with coronary heart disease. *Diabetes Metab Syndr Obes.* (2020) 13:5025–36. doi: 10.2147/DMSO.S273880
4. Găman M-A, Cozma M-A, Dobrică E-C, Bacalba?a N, Bratu OG, Diaconu CC. Dyslipidemia: a trigger for coronary heart disease in romanian patients with diabetes. *Metabolites.* (2020) 10:195. doi: 10.3390/metabo10050195
5. Valgeirsson KB, Hreinsson JP, Björnsson ES. Increased incidence of autoimmune hepatitis is associated with wider use of biological drugs. *Liver Int.* (2019) 39:2341–9. doi: 10.1111/liv.14224
6. Matthew K, Tom B, Mark D, Derek C, John E, Brendan M, et al. A cohort study examination of established and emerging risk factors for atrial fibrillation: the Busselton Health Study. *Eur J Epidemiol.* (2014) 29:181–90. doi: 10.1007/s10654-013-9875-y
7. Kashou AH, Adedinsewo DA, Noseworthy PA. Subclinical atrial fibrillation: a silent threat with uncertain implications. *Ann Rev Med.* (2022) 73:355–62. doi: 10.1146/annurev-med-042420-105906
8. Humbert X, Roule V, Chequel M, Fedrizzi S, Brionne M, Lelong-Boulouard V, et al. Non-vitamin K oral anticoagulant treatment in elderly patients with atrial fibrillation and coronary heart disease. *Int J Cardiol.* (2016) 222:1079–83. doi: 10.1016/j.ijcard.2016.07.212
9. Yoko M, Barnes ME, Gersh BJ, Cha SS, Bailey KR, Abhayaratna WP, et al. Secular trends in incidence of atrial fibrillation in Olmsted County, Minnesota, 1980 to 2000 and implications on the projections for future prevalence. *Circulation.* (2006) 114:119–25. doi: 10.1161/CIRCULATIONAHA.105.595140
10. Feng L, Yi W. Coronary heart disease and atrial fibrillation: a vicious cycle. *Am J Physiol Heart Circ Physiol.* (2020) 320:H1–12. doi: 10.1152/ajpheart.00702.2020
11. Abed HS, Wittert GA, Leong DP, Shirazi MG, Bahrami B, Middeldorp ME, et al. Effect of weight reduction and cardiometabolic risk factor management on symptom burden and severity in patients with atrial fibrillation: a randomized clinical trial. *JAMA.* (2013) 310:2050–60. doi: 10.1001/jama.2013.280521
12. Fuster V, Rydén LE, Cannom DS, Crijns HJ, Curtis AB, Ellenbogen KA, et al. ACC/AHA/ESC (2006). Guidelines for the management of patients with atrial fibrillation—executive summary. *J Am Coll Cardiol.* (2006) 48:854–906. doi: 10.1016/j.jacc.2006.07.009
13. Lucia C, Massimo C, Rosie C, Deborah L. External validation of the improving partial risk adjustment in surgery (PRAIS-2) model for 30-day mortality after paediatric cardiac surgery. *BMJ Open.* (2020) 10:e039236. doi: 10.1136/bmjopen-2020-039236
14. Tao P, JianMing X, Lin L, BingJie P, XiangKe N, XiaoHui Z, et al. Can machine learning-based analysis of multiparameter MRI and clinical parameters improve the performance of clinically significant prostate cancer diagnosis? *Int J Comput Assist Radiol Surg.* (2021) 16:2235–49. doi: 10.1007/s11548-021-02507-w
15. Ke W, Jing T, Chu Z, Hong Y, Jia R, Yanling L, et al. Interpretable prediction of 3-year all-cause mortality in patients with heart failure caused by coronary heart disease based on machine learning and SHAP. *Comput Biol Med.* (2021) 137:104813. doi: 10.1016/j.compbiomed.2021.104813
16. Chen W, Yue Z, Bingyu J, Xuedong G, Bin L, Yang X, et al. Development and validation of a predictive model for coronary artery disease using machine learning&#13. *Front Cardiovasc Med.* (2021) 8:614204. doi: 10.3389/fcvm.2021.614204
17. Hung M, Lauren E, Hon E, Xu J, Negrón BR, Rosales M, et al. Using machine learning to predict 30-day hospital readmissions in patients with atrial fibrillation undergoing catheter ablation. *J Pers Med.* (2020) 10:82. doi: 10.3390/jpm10030082
18. Breiman L. Random Forests. *Mach Learn.* (2001) 45:5–32. doi: 10.1023/A:1010933404324
19. Dreiseitl S, Ohno-Machado L. Logistic regression and artificial neural network classification models: a methodology review. *J Biomed Inform.* (2002) 35:352–9. doi: 10.1016/S1532-0464 (03)00034-0
20. Li-Hong X, Pei-Ran C, Zhong-Ping G, Yong-Zhong L, Mei L, Liang-Cheng X, et al. Prostate cancer prediction using the random forest algorithm that takes into account transrectal ultrasound findings, age, and

serum levels of prostate-specific antigen. *Asian J Androl.* (2017) 19:586–90. doi: 10.4103/1008-682X.186884
21. Alessia S, Antonio C, Aldo Q. Random forest algorithm for the classification of neuroimaging data in Alzheimer's disease: a systematic review. *Front Aging Neurosci.* (2017) 9:329. doi: 10.3389/fnagi.2017.00329
22. Zheng H, Yuan J, Chen L. Short-term load forecasting using EMD-LSTM neural networks with a xgboost algorithm for feature importance evaluation. *Energies.* (2017) 10:1168. doi: 10.3390/en10081168
23. Omer S, Lior R. Approximating XGBoost with an interpretable decision tree. *Inform Sci.* (2021) 572:522–42. doi: 10.1016/j.ins.2021.05.055
24. Zheng-Yun Z, Lu-Qin B, Sung-Joo K, Cai-Cun Z, Yoon-Ho C. Inverse relation of total serum bilirubin to coronary artery calcification score detected by multidetector computed tomography in males. *Clin Cardiol.* (2012) 35:301–6. doi: 10.1002/clc.21964
25. Shao-Peng L, Pei-Yi L, Hui-Lin J, You-Ming L, Xiao-Hui C. Is serum total bilirubin useful to differentiate cardioembolic stroke from other stroke subtypes? *Neurol Res.* (2015) 37:727–31. doi: 10.1179/1743132815Y.0000000038
26. Karl-Heinz W, Marlies W, Christine M, Silvia G, Cameron BA, Claudio T, et al. Looking to the horizon: the role of bilirubin in the development and prevention of age-related chronic diseases. *Clin Sci.* (2015) 129:1–25. doi: 10.1042/CS20140566
27. Troughton JA, Woodside JV, Young IS, Arveiler D, Amouyel P, Ferrières J, et al. Bilirubin and coronary heart disease risk in the prospective epidemiological study of myocardial infarction (PRIME). *Eur J Cardiovasc Prevent Rehabil.* (2006) 14:79–84. doi: 10.1097/01.hjr.0000230097.81202.9f
28. Wiesner P, Watson KE. Triglycerides: a reappraisal. *Trends Cardiovasc Med.* (2017) 27:428–32. doi: 10.1016/j.tcm.2017.03.004
29. Sergio R, Xavier R, Belén O, Leticia F, Mendiguren JM, Vicente A, et al. Triglycerides and residual atherosclerotic risk. *J Am Coll Cardiol.* (2021) 77:3031–41. doi: 10.1016/j.jacc.2021.04.059
30. Waters DD, Bangalore S, Fayyad R, DeMicco DA, Laskey R, Melamed S, et al. Visit-to-visit variability of lipid measurements as predictors of cardiovascular events. *J Clin Lipidol.* (2018) 12:356–66. doi: 10.1016/j.jacl.2017.12.003
31. Yanfeng R, Qian R, Juming L, Xiaohui G, Xiaoxu H, Linong J, et al. Low triglyceride as a marker for increased risk of cardiovascular diseases in patients with long-term type 2 diabetes: a cross-sectional survey in China. *Diabetes Metab Res Rev.* (2018) 34:e2960. doi: 10.1002/dmrr.2960
32. Kuwabara M, Niwa K, Nishihara S, Nishi Y, Takahashi O, Kario K, et al. Hyperuricemia is an independent competing risk factor for atrial fibrillation. *Int J Cardiol.* (2016) 231:137–42. doi: 10.1016/j.ijcard.2016.11.268
33. Kwon CH, Lee SH, Lee J-Y, Ryu S, Sung K-C. Uric acid and risk of atrial fibrillation in the korean general population. *Circ J.* (2018) 82:2728–35. doi: 10.1253/circj.CJ-18-0748
34. Liu T, Zhang X, Korantzopoulos P, Wang S, Li G. Uric acid levels and atrial fibrillation in hypertensive patients. *Internal Medicine.* (2011) 50:799–803. doi: 10.2169/internalmedicine.50.4587
35. Bajaj JS, Reddy KR, Tandon P, Wong F, Kamath PS, Garcia-Tsao G, et al. The 3-month readmission rate remains unacceptably high in a large North American cohort of patients with cirrhosis. *Hepatology.* (2016) 64:200–8. doi: 10.1002/hep.28414
36. Ji Y, Feng G, Dong F, Xie Z. A8380 Investigation of serum uric acid levels in elderly patients with persistent atrial fibrillation. *J Hypertens.* (2018) 36 (Suppl. 3):e188. doi: 10.1097/01.hjh.0000548766.81870.fb

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Xu, Peng, Tan, Zhao, Yang and Tian. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 04 March 2022
doi: 10.3389/fphar.2022.858531

Associations of Fentanyl, Sufentanil, and Remifentanil With Length of Stay and Mortality Among Mechanically Ventilated Patients: A Registry-Based Cohort Study

*Wen Wang[1,2], Qiao He[1,2], Mingqi Wang[1,2], Yan Kang[3], Peng Ji[3], Shichao Zhu[4], Rui Zhang[5], Kang Zou[1,2] and Xin Sun[1,2]**

[1]Chinese Evidence-based Medicine Center, West China Hospital, Sichuan University, Chengdu, China, [2]NMPA Key Laboratory for Real World Data Research and Evaluation in Hainan, Chengdu, China, [3]Intensive Care Unit, West China Hospital, Sichuan University, Chengdu, China, [4]Department of Infection Control, West China Hospital, Sichuan University, Chengdu, China, [5]Information Center, West China Hospital, Sichuan University, Chengdu, China

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Xiaochen Shu,
Soochow University, China
Hua Jiang,
Sichuan Academy of Medical Sciences and Sichuan Provincial People's Hospital, China

***Correspondence:**
Xin Sun
sunxin@wchscu.cn

Specialty section:
This article was submitted to Drugs Outcomes Research and Policies,
a section of the journal Frontiers in Pharmacology

Received: 20 January 2022
Accepted: 07 February 2022
Published: 04 March 2022

Citation:
Wang W, He Q, Wang M, Kang Y, Ji P, Zhu S, Zhang R, Zou K and Sun X (2022) Associations of Fentanyl, Sufentanil, and Remifentanil With Length of Stay and Mortality Among Mechanically Ventilated Patients: A Registry-Based Cohort Study. Front. Pharmacol. 13:858531. doi: 10.3389/fphar.2022.858531

Background: As the first-line treatment for mechanically ventilated patients with critical illness, fentanyl and its analogs (e.g., sufentanil and remifentanil) are commonly used in the intensive care unit (ICU). However, the pharmacokinetics, metabolism, and potency of these agents differed. Their effects on clinical outcomes have not been well-understood.

Materials and Methods: Using a well-established registry, we conducted a cohort study. Patients who consistently underwent mechanical ventilation (MV) for more than 24 h were identified. We used a time-varying exposure definition, in which we coded each type of opioids as prescribed or not prescribed on each day from initiation of MV to extubation and ICU discharge. We used Fine-Gray competing risk models to compare the effects of fentanyl, sufentanil, and remifentanil on hazards for extubation, ventilator mortality, ICU discharge, and ICU mortality. All models were adjusted using a combination of fixed-time and time-varying covariates. Missing data were imputed using multiple imputation by chained equations.

Results: A total of 8,165 patients were included. There were, respectively, 4,778, 4,008, and 2,233 patients receiving at least 1 day of fentanyl, sufentanil, and remifentanil dose. Compared to fentanyl, sufentanil was associated with shorter duration to extubation (hazard ratio 1.31, 95% CI, 1.20–1.41) and ICU discharge (hazard ratio 1.63, 95% CI, 1.38–1.92), and remifentanil was associated with shorter duration to extubation (hazard ratio 1.60, 95% CI, 1.40–1.84) and ICU discharge (hazard ratio 2.02, 95% CI, 1.43–2.84). No significant differences in time to extubation (Hazard ratio 1.14, 95% CI, 0.92–1.41) and ICU discharge (Hazard ratio 1.31, 95% CI, 0.81–2.14) were found between sufentanil and remifentanil. No differences were observed between any two of the agents regarding ventilator mortality or ICU mortality. The effects were similar in patients with versus without surgery.

Conclusion: Sufentanil and remifentanil may be superior to fentanyl in shortening the time to extubation and ICU discharge. The effects on ventilator mortality and ICU mortality appeared similar across these agents, while further research is warranted.

Keywords: mechanical ventilation, intensive care unit, fentanyl, sufentanil, remifentanil

INTRODUCTION

Pain is common among patients admitted to the intensive care unit (ICU) (Jeitziner et al., 2012; Ehieli et al., 2017). Intravenous opioids are usually recommended as the first-line treatment for non-neuropathic pain among patients with critical illness, especially for those receiving mechanical ventilation (MV) (Faust et al., 2016; Devlin et al., 2018; Jungquist et al., 2019). Previous studies showed that 90% of patients received opioids during MV treatment, and fentanyl and fentanyl analogs were the most commonly prescribed opioids; about 75–85% of patients received fentanyl, sufentanil, and remifentanil treatment during MV (Payen et al., 2007).

Although widely used, these opioid agents vary in their pharmacokinetic characteristics and risk of accumulation in organ failure (Ehieli et al., 2017; Zhu et al., 2017; Armenian et al., 2018). For instance, remifentanil is metabolized by unspecific esterases independent of liver or renal function, and the duration of action (5–10 min) is shorter relative to fentanyl and sufentanil. While sufentanil is 5–10 times potent than fentanyl and remifentanil, the risk of opioid accumulation in organs may be higher than that of remifentanil (Maciejewski, 2012). The accumulation of opioids may cause respiratory depression, sedation, hypotension, and immunosuppression, which may further result in adverse outcomes (Zhu et al., 2017; Devlin et al., 2018; Gupta et al., 2018).

The pharmacokinetic differences among these agents may lead to differential patient-important clinical outcomes, such as respiratory depression and death. However, evidence regarding the effects of these agents on ICU patients receiving MV is inadequate. Several randomized controlled trials (RCTs) investigated the effects of these agents, but were small in sample sizes and had inconsistent findings.(Breen et al., 2005; Djian et al., 2006; Rozendaal et al., 2009; Spies et al., 2011; Bhavsar et al., 2016; Zhu et al., 2017). One meta-analysis involving 1905 mechanically ventilated patients compared the effectiveness of analgesic regimens with versus without remifentanil; however, this study compared remifentanil versus all other opioids, leaving the comparison among fentanyl and its analogs an unanswered question (Zhu et al., 2017). One additional important issue that existing studies did not address is how varying exposure to opioids during MV treatment in routine practice may have affected treatment outcomes (Klompas et al., 2016b).

Therefore, we conducted a cohort study to compare the effects of fentanyl, sufentanil, and remifentanil on extubation, ICU discharge, and ICU mortality. This study included a large number of patients and used statistical methods to handle daily change in the exposure to fentanyl and its analogs.

MATERIALS AND METHODS

Data Source

This cohort study was conducted using a registry of healthcare-associated infections (HAIs) at ICUs from the West China Hospital (WCH) system (Wang et al., 2019). The WCH system includes three independent healthcare organizations and contains six ICUs, and is a critical care center that covers population in West China. Every year, there were more than 200,000 inpatient visits, and 8,000 patients were admitted to ICUs.

The registry was established by integrating three databases, including the electronic medical record (EMR) system, ICU system, and ICU-HAI system. All patients admitted to any of six ICUs since April 1, 2015 were included in the registry. Till December, 2018, a total of 29,480 ICU admissions were included. Detailed information about the registry has been published elsewhere (Wang et al., 2019). The registry has a high level of quality and comprehensiveness of information and has been utilized for addressing various clinical research questions(Zhu et al., 2019; Wang et al., 2021a; Wang et al., 2021b; He et al., 2021; Zhu et al., 2021). Quality assessment showed that the accuracy of data extraction and linkage was 100%, and the completeness of important laboratory tests was more than 98%. The International Classification of Diseases 10th Revision (ICD-10) codes at WCH have been validated, and the completeness and accuracy were 99 and 88%, respectively (Wang et al., 2019; Wang et al., 2021b).

This study was approved by the West China Hospital Institutional Review Board (WCH 2018–409), who granted a waiver of patient consent.

Study Population

The patients were eligible in the cohort if they consistently received MV for more than 24 h between April 1, 2015 and December 31, 2018. The patients were excluded if they met any of the following criteria: patients with incomplete information including date of birth, sex, and discharged diagnosis; age < 18 years or admitted to the pediatric ICU; without receiving any opioids during ICU stay. The clinical characteristics differed among patients with and without extremely long ICU stay (>90 days). To minimize indication bias, we also excluded patients with extremely long ICU stay. If patients experienced multiple episodes of MV, we only measured the first episode of MV treatment for more than 24 h.

Drug Exposure and Outcomes

The exposures of interest included fentanyl, sufentanil, and remifentanil. We used a time-varying exposure definition. Information regarding daily exposures to fentanyl, sufentanil, and remifentanil was extracted from the prescription data of the EMR system. We coded each type of opioids (fentanyl, sufentanil, and remifentanil) as prescribed or not prescribed on each day from initiation of MV to extubation and each day from initiation of MV to ICU discharge, respectively.

The outcomes included duration of MV, ventilator mortality, length of ICU stay, and ICU mortality, which were measured as time-to-event variables. Patients who died within 1 day after extubation and 1 day after ICU discharge were defined as ventilator mortality and ICU mortality, respectively.

Confounding Factors

We identified potential confounding factors based on clinical plausibility. We also consulted experts in critical care medicine, anesthesia, and respiratory therapy to identify potential

confounding factors. The potential confounding factors included demographic characteristics (age, sex, and ICU type), acute conditions at ICU admission (acute respiratory distress syndrome (ARDS), shock, gastrointestinal bleeding, and pneumonia), chronicity comorbidities (diabetes, malignant tumor, chronic lung disease, cardiovascular disease, heart failure, liver and renal failure, and hypertension), APACHE II score, fluid balance, surgery before the event of interest (thoracic surgery, cranial surgery, cardiac surgery, and abdominal surgery), fiberoptic bronchoscopy examination, tracheotomy, concomitant medications (non-opioid analgesics and other types of opioids (e.g., morphine, pethidine), sedatives, neuroleptic agents, thromboembolism agents, acid inhibitors, neuromuscular blockers, intestinal probiotics, expectorants, antibiotics, immunosuppressive agents, and vasopressors), and other treatment (blood transfusion, gastrointestinal decompression, prone position ventilation, enteral nutrition, mandatory ventilation, esophagogastroduodenoscopy, and head-of-bed elevation). We also included days from ICU admission to initiation of MV, and opioids prescribed before the initiation of MV as potential confounding factors.

We identified concomitant medications from prescription data and comorbidity from discharge diagnoses using ICD-10 codes.

Statistical Analysis

We assessed the impact of daily fentanyl, sufentanil, and remifentanil exposures on time to extubation, ventilator mortality, time to ICU discharge, and ICU mortality using Fine-Gray competing risk models. Fine-Gray competing risks models are proportional subdistribution hazard models which can disentangle effects of competing risks and outcomes of interest by generating separate hazard ratios for each competing event (Mohammad et al., 2017). The reason for extubation and ICU discharge varies considerably depending on the patient's health condition: clinical improvement versus death. Therefore, we used Fine-Gray competing risk models to measure the competing risks for extubation alive versus ventilator mortality and ICU discharge alive versus ICU mortality.(Klompas et al., 2016a; Klompas et al., 2016b). We calculated hazard ratios (HRs) for sufentanil versus fentanyl, remifentanil versus fentanyl, and sufentanil versus remifentanil.

All models were adjusted using a combination of fixed-time and time-varying covariates. We defined demographic characteristics, acute conditions, chronicity comorbidities, APACHE II score, surgery, fiberoptic bronchoscopy examination, tracheotomy, days from ICU admission to initiation of MV, and opioids prescribed before the initiation of MV as fixed-time variables. We defined exposures to concomitant medications and other treatment as time-varying variables. We measured time-varying variables as daily exposure on each day of MV treatment for the model of extubation alive versus ventilator mortality and each day from initiation of MV to ICU discharge for the model of ICU discharge versus ICU death. We additionally adjusted the duration of mechanical ventilation and measured mandatory ventilation and prone position ventilation as fixed-time variables for the analysis of ICU discharge versus ICU death.

The percentages of missing data for APACHE II score and fluid balance during ICU stay were 9.18 and 8.68%, respectively. We handled missing data of APACHE II score and fluid balance using multiple imputation by chained equations (MICE), and the iterations were set to five times. By creating a number of datasets to handle uncertainty in missing value imputation, multiple imputation is thought to be superior to single imputation and complete-case analysis (Zhang, 2016).

Subgroup Analyses and Sensitivity Analyses

We calculated HRs for patients with versus without receiving surgery separately, given that the clinical characteristics and treatment approach differed variously between these two populations.

We performed several sensitivity analyses to test the robustness of the primary findings. First, we used the traditional Cox model with time-varying covariates to calculate the effect estimates. Second, we used complete case analysis without imputation to handle missing data. Third, to address the potential confounding factors by indication, we restricted to patients who did not receive non-opioid analgesics. All analyses were performed using R version 3·6·1.

RESULTS

The study included 8,165 patients who consistently received MV for at least 24 h and treated with opioids during ICU stays from April 1, 2015 to December 31, 2018. Of 8,165 included patients, 4,778 received at least 1 day of fentanyl, 4,008 received at least 1 day of sufentanil, and 2,233 received at least 1 day of remifentanil (see **Figure 1**).

Among all included patients, the median age was 57 (interquartile range (IQR) years, 46–68); 3,147 (38.5%) patients were females; the median APACHE II score was 19 (IQR, 14–24). There were 2,457 (30.1%) patients admitted to the general ICU and 1705 (20.9%) admitted to the surgical ICU. The most common comorbidities were hypertension (21.2%), heart failure (12.5%), and chronic lung disease (9.8%), and the most common acute conditions were pneumonia (7.7%) and shock (5.4%). There were 1,444 (17.7%) patients who underwent cardiac surgery, 809 (9.9%) who underwent thoracic surgery, 1,465 (17.9%) who underwent cranial surgery, and 1,573 (19.3%) who underwent abdominal surgery. The clinical characteristics differed among patients receiving different opioids. For instance, the proportions of cardiac surgery were 4.7% among patients receiving fentanyl and 24.7% among patients receiving remifentanil (see **Table 1**).

Among all included patients, the median day of MV was 5 (IQR, 3–11), the median ICU stay was 10 (IQR, 5–19) days, and crude ICU mortality was 11.9%. Among patients receiving fentanyl, sufentanil, and remifentanil, the median ICU stay was 13 (IQR, 7–22), 11 (IQR, 6–21) and 9 (IQR, 5–19) days; the crude mortality was 15.5, 11.1, and 7.2% (see **Table 1**).

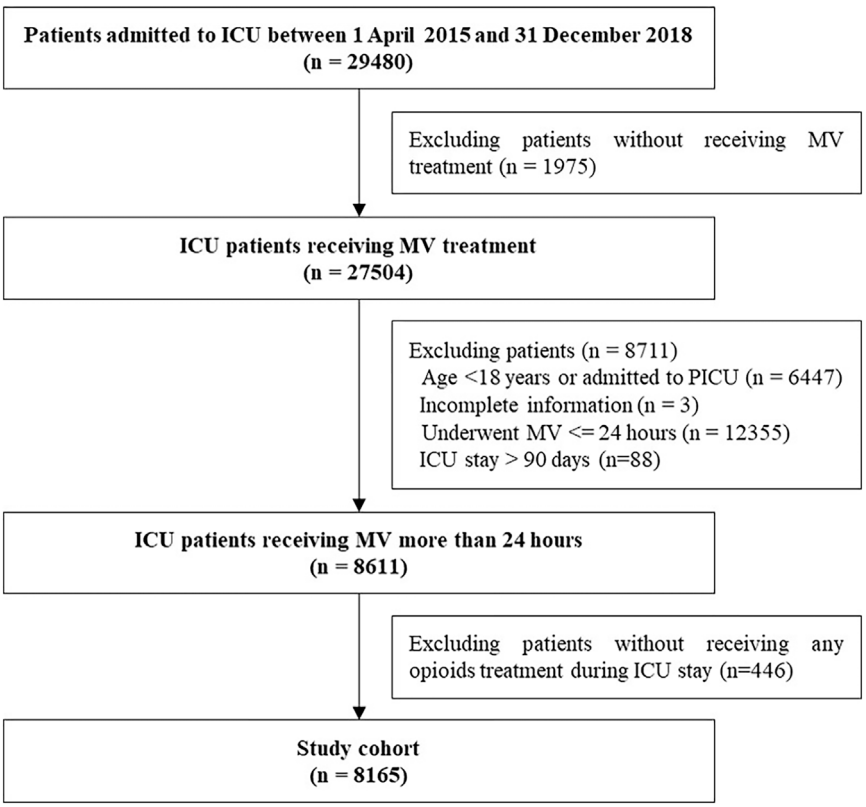

FIGURE 1 | Study flow chart.

Opioid Treatment Among Included Patients

Opioids were often used concurrently in patients receiving MV. Among patients receiving fentanyl, sufentanil, and remifentanil, the proportion of patients treated with other opioids during ICU stay was 60.9, 83.8, and 97.2%, respectively (see **Figure 2**). Moreover, combined treatment of opioids with sedatives was common; the proportion among patients receiving fentanyl, sufentanil, and remifentanil was 74.9, 72.6, and 70.4%, respectively.

The median cumulative dose fentanyl, sufentanil and remifentanil administrated during ICU stays was 4.5 mg (IQR, 1.5–10), 0.4 mg (IQR, 0.14–1.1) and 1.0 mg (IQR, 1.0–2.0), respectively. The median days of use of fentanyl [4 days (IQR, 2–9)] was longer than sufentanil [1 day (IQR, 1–5)] and remifentanil [1 day (IQR, 1–1)] (See **Supplementary Table S1**).

Associations Between Opioid Exposures and Outcomes

The adjusted estimates of opioids on the predefined clinical outcomes are shown in **Figures 3**, **4**. Compared to fentanyl, sufentanil was associated with an increased hazard for extubation (HR 1.31, 95% CI, 1.20–1.41) and ICU discharge (HR 1.63, 95% CI, 1.38–1.92), which suggested that sufentanil was associated with less time to extubation and ICU discharge. There were no significant differences between fentanyl and sufentanil regarding hazards for ventilator mortality (HR 0.95, 95% CI, 0.77–1.17) and ICU mortality (HR 1.03, 95% CI, 0.84–1.26).

Remifentanil was associated with higher hazards for extubation (HR 1.60, 95% CI, 1.40–1.84) and ICU discharge (HR 2.02, 95% CI, 1.43–2.84), relative to fentanyl. However, no significant differences between fentanyl and sufentanil were found for ventilator mortality (HR 0.87, 95% CI, 0.50–1.49) and ICU mortality (HR 1.25, 95% CI, 0.73–2.16). No significant differences were observed between sufentanil and remifentanil regarding hazards for extubation (HR 1.14, 95% CI, 0.92–1.41), ventilator mortality (HR 0.91, 95% CI, 0.47–1.76), ICU discharge (HR 1.31, 95% CI, 0.81–2.14), and ICU mortality (HR 0.66, 95% CI, 0.34–1.28).

Subgroup Analyses and Sensitivity Analyses

Figures 3, **4** showed the subgroup analyses by surgery. The hazard ratios were similar across the subgroups (test of interaction, $p > 0.05$). Compared to fentanyl, sufentanil was associated with increased hazards for extubation and ICU discharge both in patients with and without surgery. Similarly, compared to patients receiving fentanyl, patients treated with remifentanil had higher hazard ratios for

TABLE 1 | Clinical characteristics of included patients.

Characteristics	Overall patients (*n* = 8,165)	Fentanyl (*n* = 4,778)	Sufentanil (*n* = 4,008)	Remifentanil (*n* = 2,233)
Age, median [IQR]	57 [46, 68]	58.00 [45, 70]	56 [45, 67]	53.00 [44, 65]
Sex, female (%)	3,147 (38.5)	1,696 (35.5)	1,524 (38.0)	915 (41.0)
APACHE II, median [IQR]	19 [14, 24]	20 [15, 25]	19 [14, 23]	18 [13, 22]
ICU type (%)				
General ICU	2,457 (30.1)	1712 (35.8)	1,557 (38.8)	472 (21.1)
Neurological ICU	1,338 (16.4)	1,080 (22.6)	423 (10.6)	412 (18.5)
Respiratory ICU	1,139 (13.9)	1,074 (22.5)	384 (9.6)	92 (4.1)
Surgical ICU	1705 (20.9)	675 (14.1)	957 (23.9)	692 (31.0)
Thoracic ICU	1,526 (18.7)	237 (5.0)	687 (17.1)	565 (25.3)
Acute conditions (%)				
ARDS	76 (0.9)	60 (1.3)	46 (1.1)	14 (0.6)
Shock	441 (5.4)	290 (6.1)	302 (7.5)	150 (6.7)
Gastrointestinal bleeding	112 (1.4)	81 (1.7)	66 (1.6)	25 (1.1)
Pneumonia	632 (7.7)	433 (9.1)	325 (8.1)	121 (5.4)
Comorbidities (%)				
Diabetes	646 (7.9)	399 (8.4)	293 (7.3)	147 (6.6)
Cardiovascular disease	148 (1.8)	63 (1.3)	49 (1.2)	30 (1.3)
Heart failure	1,020 (12.5)	250 (5.2)	467 (11.7)	341 (15.3)
Chronic lung disease	798 (9.8)	640 (13.4)	334 (8.3)	94 (4.2)
Malignant tumor	676 (8.3)	408 (8.5)	366 (9.1)	140 (6.3)
Liver failure	158 (1.9)	112 (2.3)	81 (2.0)	25 (1.1)
Hypertension	1734 (21.2)	961 (20.1)	798 (19.9)	516 (23.1)
Kidney failure	485 (5.9)	350 (7.3)	216 (5.4)	79 (3.5)
Surgery (%)				
Cardiac surgery	1,444 (17.7)	226 (4.7)	672 (16.8)	552 (24.7)
Thoracic surgery	809 (9.9)	560 (11.7)	432 (10.8)	168 (7.5)
Cranial surgery	1,465 (17.9)	927 (19.4)	754 (18.8)	656 (29.4)
Abdominal surgery	1,573 (19.3)	871 (18.2)	1,117 (27.9)	621 (27.8)
Outcomes				
Days of MV, median (IQR)	5 [3, 11]	7 [4, 13]	5 [3, 11]	5 [2, 9]
ICU stays, median (IQR)	10 [5, 19]	13 [7, 22]	11 [6, 21]	9 [5, 19]
ICU mortality (%)	971 (11.9)	742 (15.5)	446 (11.1)	161 (7.2)

Abbreviations: IQR, interquartile range; ICU, intensive care unit; ARDS, acute respiratory distress syndrome; MV, mechanical ventilation.

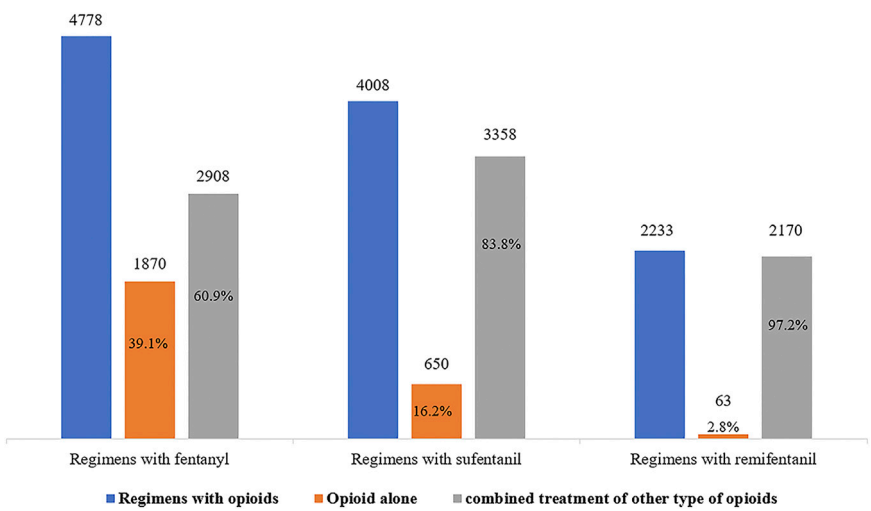

FIGURE 2 | Treatment of opioids during ICU stay.

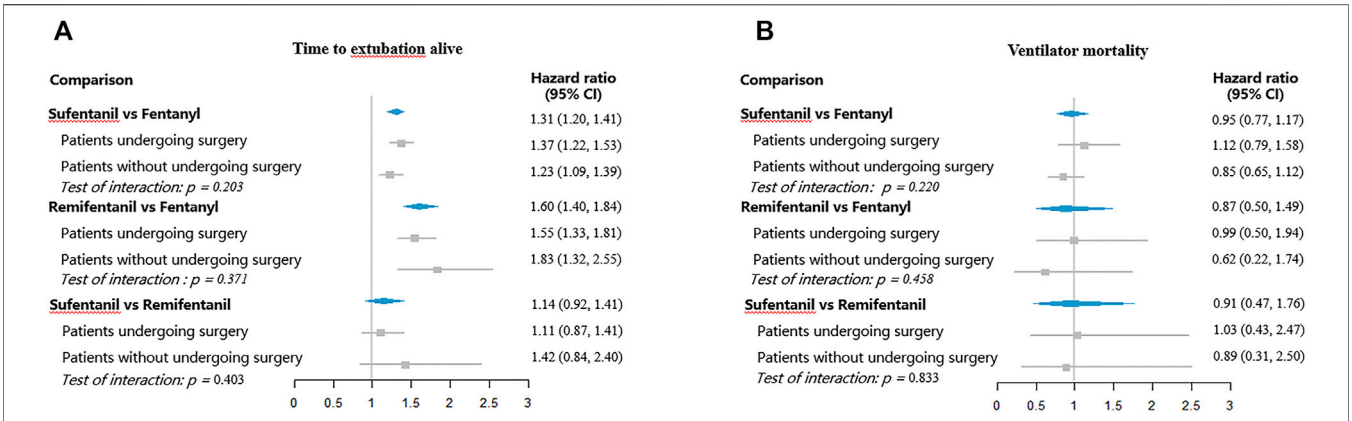

FIGURE 3 | Hazard ratios for extubation and ventilator mortality with different opioids. **(A)** Time to extubation alive; **(B)** Ventilator mortality.

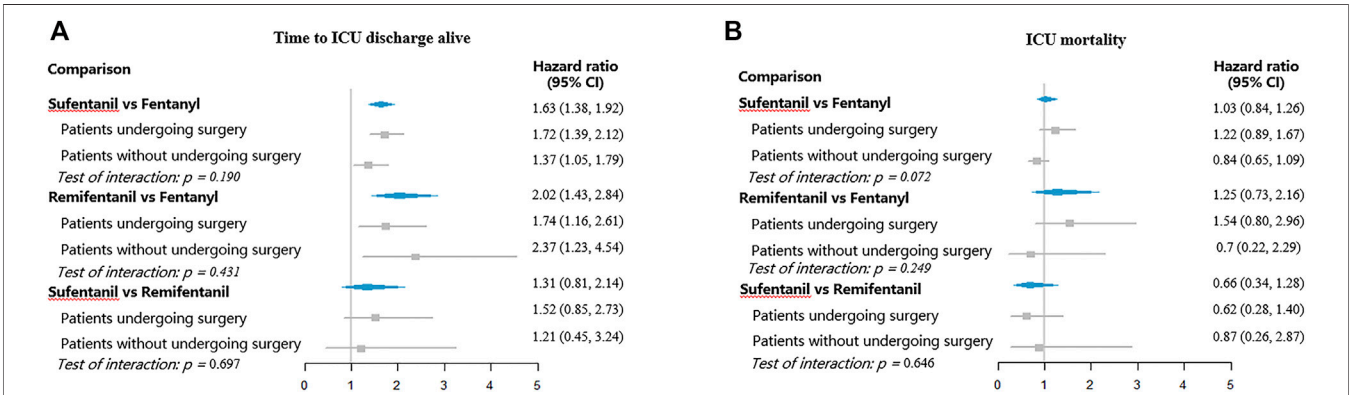

FIGURE 4 | Hazard ratios for ICU discharge and ICU mortality with different opioids. **(A)** Time to ICU discharge alive; **(B)** ICU mortality.

extubation and ICU discharge both among patients with and without surgery.

Sensitivity analyses using the traditional Cox model with time-varying covariates and excluding patient treated with non-opioid analgesics did not show change in interpretation. In the comparisons of sufentanil versus remifentanil, however, the sensitivity analyses using complete cases without imputation generated hazard ratios that excluded the null for extubation (HR 1.36, 95% CI, 1.03–1.80), ICU discharge (HR 2.05, 95% CI, 1.07–3.93), and ICU mortality (HR 0.44, 95% CI, 0.22–0.88) (See **Supplementary Table S2**.).

DISCUSSION

In this large observational study, we assessed the association of fentanyl, sufentanil, and remifentanil with important clinical outcomes under the real-world conditions. Our findings showed that treatment of opioids varied during MV treatment in routine practice. Compared to fentanyl, use of sufentanil or remifentanil was associated with shorter time to extubation and ICU discharge; no significant differences in duration of MV and length of ICU stay were found between sufentanil and remifentanil. No significant differences were observed between any of these agents in ventilator mortality and ICU mortality. Subgroup analyses suggested that the effects were similar among patients with and without receiving surgery.

The findings may be due to pharmacokinetic and pharmacodynamic differences of these agents. As important safety concerns, the accumulation of opioids may cause respiratory depression, sedation, immunosuppression, and ileus (Smith et al., 2014; Lee et al., 2015; Gomes et al., 2017). Studies suggested that opioids with short half-life may enable greater titration accuracy with limited drug accumulation (Futier et al., 2012). Compared to fentanyl, the half-life of sufentanil and remifentanil was shorter, which may have potential advantages in reducing accumulation in organs and further result in shortening duration of ventilation and length of ICU stay (Glass et al., 1999; Breen et al., 2004; Rozendaal et al., 2009). However, no differences were observed between sufentanil and remifentanil. This may be because sufentanil is approximately 5–10 times more potent than remifentanil. Due to acute tolerance and opioid-induced hyperalgesia, more pain was reported in the remifentanil group (Welzing et al., 2013; Fletcher and Martinez, 2014;

Bhavsar et al., 2016). The adverse effect caused by pain may counteract the effect of limited accumulation in organs.

In our study, the study setting was more complex than RCTs. Nevertheless, the findings under real-world conditions were consistent with those from previous RCTs—our study found that remifentanil may shorten duration of MV and length of ICU stay when compared to fentanyl (Breen et al., 2005; Rozendaal et al., 2009). Our study also suggested that sufentanil and remifentanil may be similar in their clinical effects, consistent with a previous trial (Bhavsar et al., 2016).

Our study also extended the knowledge about these agents in their use for MV patients, which was not identified from previous RCTs. In particular, there were just few studies comparing sufentanil and fentanyl in mechanically ventilated patients. We found that sufentanil was associated with less time to extubation and ICU discharge than fentanyl. A small double-blind RCT, including 20 newborns receiving MV for more than 24 h, suggested that compared to fentanyl, sufentanil did not reduce the weaning time (Schmidt et al., 2010). However, this study exclusively included newborns and had a very small sample size.

Strengths and Limitations

In this study, we utilized a large database with high level of quality and comprehensiveness of information. The sample size of this study was relatively large. We considered day-to-day changes in opioid exposures and used statistical methods to handle time-dependent variates. We also measured the competing risks and adjusted for an extensive array of potential confounding factors.

However, our study is tempered by important limitations. First, owing to the observational nature, residual confounding from unknown and unmeasured variables is possible, especially for confounding by indication. Prescription of opioid was not randomly performed but determined by clinicians. Although we consider and adjusted for an extensive array of potential confounding, confounding by indication may not be excluded. Second, relatively few patients were exposed to remifentanil, and the inference on the effect of remifentanil is weakened. Finally, this study was conducted exclusively based on a database from a homogenous healthcare system, which may limit the generalization.

In conclusion, our findings suggested that remifentanil was superior to fentanyl in shortening the duration of MV and length of ICU stay, which was consistent with previous RCTs. The effects on ventilator mortality and ICU mortality appeared comparable across these agents. Our finding also showed that sufentanil was associated with less time to extubation and ICU stay, which extends the prior RCT. Further studies are needed to confirm these results.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

XS conceived and designed the research and assisted in revision of the manuscript; WW designed the research, wrote the initial manuscript, and interpreted research results; QH and MW cleared, analyzed, and interpreted research data. YK and PJ collected the research data and assisted in interpretation of research results; SZ and RZ assisted in acquisition and clean data; KZ assisted in interpretation of research results. All authors read and approved the final manuscript.

FUNDING

WW received support from the National Natural Science Foundation of China (Grant No. 72104155) and the National Key Research and Development Program (Grant No. 2020YFC2009003). XS received support from Sichuan Youth Science and Technology Innovation Research Team (Grant No. 2020JDTD0015), and the 1·3·5 Project for Disciplines of Excellence, West China Hospital, Sichuan University (Grant No. ZYYC08003).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2022.858531/full#supplementary-material

REFERENCES

Armenian, P., Vo, K. T., Barr-Walker, J., and Lynch, K. L. (2018). Fentanyl, Fentanyl Analogs and Novel Synthetic Opioids: A Comprehensive Review. *Neuropharmacology* 134 (Pt A), 121–132. doi:10.1016/j.neuropharm.2017.10.016

Bhavsar, R., Ryhammer, P. K., Greisen, J., Rasmussen, L. A., and Jakobsen, C. J. (2016). Remifentanil Compared with Sufentanil Does Not Enhance Fast-Track Possibilities in Cardiac Surgery-A Randomized Study. *J. Cardiothorac. Vasc. Anesth.* 30 (5), 1212–1220. doi:10.1053/j.jvca.2015.12.021

Breen, D., Karabinis, A., Malbrain, M., Morais, R., Albrecht, S., Jarnvig, I. L., et al. (2005). Decreased Duration of Mechanical Ventilation when Comparing Analgesia-Based Sedation Using Remifentanil with Standard Hypnotic-Based Sedation for up to 10 Days in Intensive Care Unit Patients: A Randomised Trial [ISRCTN47583497]. *Crit. Care* 9 (3), R200–R210. doi:10.1186/cc3495

Breen, D., Wilmer, A., Bodenham, A., Bach, V., Bonde, J., Kessler, P., et al. (2004). Offset of Pharmacodynamic Effects and Safety of Remifentanil in Intensive Care Unit Patients with Various Degrees of Renal Impairment. *Crit. Care* 8 (1), R21–R30. doi:10.1186/cc2399

Devlin, J. W., Skrobik, Y., Gélinas, C., Needham, D. M., Slooter, A. J. C., Pandharipande, P. P., et al. (2018). Clinical Practice Guidelines for the Prevention and Management of Pain, Agitation/Sedation, Delirium, Immobility, and Sleep Disruption in Adult Patients in the ICU. *Crit. Care Med.* 46 (9), e825–e73. doi:10.1097/CCM.0000000000003299

Djian, M. C., Blanchet, B., Pesce, F., Sermet, A., Disdet, M., Vazquez, V., et al. (2006). Comparison of the Time to Extubation after Use of Remifentanil or Sufentanil in Combination with Propofol as Anesthesia in Adults Undergoing Nonemergency Intracranial Surgery: a Prospective,

Randomized, Double-Blind Trial. *Clin. Ther.* 28 (4), 560–568. doi:10.1016/j.clinthera.2006.04.001

Ehieli, E., Yalamuri, S., Brudney, C. S., and Pyati, S. (2017). Analgesia in the Surgical Intensive Care Unit. *Postgrad. Med. J.* 93 (1095), 38–45. doi:10.1136/postgradmedj-2016-134047

Faust, A. C., Rajan, P., Sheperd, L. A., Alvarez, C. A., McCorstin, P., and Doebele, R. L. (2016). Impact of an Analgesia-Based Sedation Protocol on Mechanically Ventilated Patients in a Medical Intensive Care Unit. *Anesth. Analg* 123 (4), 903–909. doi:10.1213/ANE.0000000000001393

Fletcher, D., and Martinez, V. (2014). Opioid-induced Hyperalgesia in Patients after Surgery: A Systematic Review and a Meta-Analysis. *Br. J. Anaesth.* 112 (6), 991–1004. doi:10.1093/bja/aeu137

Futier, E., Chanques, G., Cayot Constantin, S., Vernis, L., Barres, A., Guerin, R., et al. (2012). Influence of Opioid Choice on Mechanical Ventilation Duration and ICU Length of Stay. *Minerva Anestesiol* 78 (1), 46–53.

Glass, P. S., Gan, T. J., and Howell, S. (1999). A Review of the Pharmacokinetics and Pharmacodynamics of Remifentanil. *Anesth. Analg* 89 (4 Suppl. l), S7–S14. doi:10.1097/00000539-199910001-00003

Gomes, T., Juurlink, D. N., Antoniou, T., Mamdani, M. M., Paterson, J. M., and van den Brink, W. (2017). Gabapentin, Opioids, and the Risk of Opioid-Related Death: A Population-Based Nested Case-Control Study. *Plos Med.* 14 (10), e1002396. doi:10.1371/journal.pmed.1002396

Gupta, K., Prasad, A., Nagappa, M., Wong, J., Abrahamyan, L., and Chung, F. F. (2018). Risk Factors for Opioid-Induced Respiratory Depression and Failure to rescue: A Review. *Curr. Opin. Anaesthesiol* 31 (1), 110–119. doi:10.1097/ACO.0000000000000541

He, Q., Wang, W., Zhu, S., Wang, M., Kang, Y., Zhang, R., et al. (2021). The Epidemiology and Clinical Outcomes of Ventilator-Associated Events Among 20,769 Mechanically Ventilated Patients at Intensive Care Units: an Observational Study. *Crit. Care* 25 (1), 44. doi:10.1186/s13054-021-03484-x

Jeitziner, M. M., Schwendimann, R., Hamers, J. P., Rohrer, O., Hantikainen, V., and Jakob, S. M. (2012). Assessment of Pain in Sedated and Mechanically Ventilated Patients: An Observational Study. *Acta Anaesthesiol Scand.* 56 (5), 645–654. doi:10.1111/j.1399-6576.2012.02660.x

Jungquist, C. R., Quinlan-Colwell, A., Vallerand, A., Carlisle, H. L., Cooney, M., Dempsey, S. J., et al. (2020). American Society for Pain Management Nursing Guidelines on Monitoring for Opioid-Induced Advancing Sedation and Respiratory Depression: Revisions. *Pain Manage. Nurs.* 21, 7–25. doi:10.1016/j.pmn.2019.06.007

Klompas, M., Li, L., Kleinman, K., Szumita, P. M., and Massaro, A. F. (2016a). Associations between Ventilator Bundle Components and Outcomes. *JAMA Intern. Med.* 176 (9), 1277–1283. doi:10.1001/jamainternmed.2016.2427

Klompas, M., Li, L., Szumita, P., Kleinman, K., and Murphy, M. V. (2016b). Associations between Different Sedatives and Ventilator-Associated Events, Length of Stay, and Mortality in Patients Who Were Mechanically Ventilated. *Chest* 149 (6), 1373–1379. doi:10.1378/chest.15-1389

Lee, L. A., Caplan, R. A., Stephens, L. S., Posner, K. L., Terman, G. W., Voepel-Lewis, T., et al. (2015). Postoperative Opioid-Induced Respiratory Depression: A Closed Claims Analysis. *Anesthesiology* 122 (3), 659–665. doi:10.1097/aln.0000000000000564

Maciejewski, D. (2012). Sufentanil in Anaesthesiology and Intensive Therapy. *Anaesthesiol Intensive Ther.* 44 (1), 35–41.

Mohammad, K. A., Fatima-Tuz-Zahura, M., and Bari, W. (2017). Fine and Gray Competing Risk Regression Model to Study the Cause-specific Under-five Child Mortality in Bangladesh. *BMC Int. Health Hum. Rights* 17 (1), 3. doi:10.1186/s12914-017-0112-8

Payen, J. F., Chanques, G., Mantz, J., Hercule, C., Auriant, I., Leguillou, J. L., et al. (2007). Current Practices in Sedation and Analgesia for Mechanically Ventilated Critically Ill Patients: A Prospective Multicenter Patient-Based Study. *Anesthesiology* 106 (4), 687–692. doi:10.1097/01.anes.0000264747.09017.da

Rozendaal, F. W., Spronk, P. E., Snellen, F. F., Schoen, A., van Zanten, A. R., Foudraine, N. A., et al. (2009). Remifentanil-propofol Analgo-Sedation Shortens Duration of Ventilation and Length of ICU Stay Compared to a Conventional Regimen: a centre Randomised, Cross-Over, Open-Label Study in the Netherlands. *Intensive Care Med.* 35 (2), 291–298. doi:10.1007/s00134-008-1328-9

Schmidt, B., Adelmann, C., Stützer, H., Welzing, L., Hünseler, C., Kribs, A., et al. (2010). Comparison of Sufentanil versus Fentanyl in Ventilated Term Neonates. *Klin Padiatr* 222 (2), 62–66. doi:10.1055/s-0029-1225348

Smith, M. A., Hibino, M., Falcione, B. A., Eichinger, K. M., Patel, R., and Empey, K. M. (2014). Immunosuppressive Aspects of Analgesics and Sedatives Used in Mechanically Ventilated Patients: an Underappreciated Risk Factor for the Development of Ventilator-Associated Pneumonia in Critically Ill Patients. *Ann. Pharmacother.* 48 (1), 77–85. doi:10.1177/1060028013510698

Spies, C., Macguill, M., Heymann, A., Ganea, C., Krahne, D., Assman, A., et al. (2011). A Prospective, Randomized, Double-Blind, Multicenter Study Comparing Remifentanil with Fentanyl in Mechanically Ventilated Patients. *Intensive Care Med.* 37 (3), 469–476. doi:10.1007/s00134-010-2100-5

Wang, W., He, Q., Zhu, S., Wang, M., Kang, Y., Zhang, R., et al. (2021a). Association between Blood Transfusion and Ventilator-Associated Events: a Nested Case-Control Study. *Infect. Control. Hosp. Epidemiol.* Online ahead of print, 1–6. doi:10.1017/ice.2021.178

Wang, W., Zhu, S., He, Q., Zhang, R., Kang, Y., Wang, M., et al. (2019). Developing a Registry of Healthcare-Associated Infections at Intensive Care Units in West China: Study Rationale and Patient Characteristics. *Clin. Epidemiol.* 11, 1035–1045. doi:10.2147/CLEP.S226935

Wang, W., Zhu, S., He, Q., Wang, M., Kang, Y., Zhang, R., et al. (2021b). Fluid Balance and Ventilator-Associated Events Among Patients Admitted to ICUs in China. *Crit. Care Med.* 50, 307–316. doi:10.1097/ccm.0000000000005227

Welzing, L., Link, F., Junghaenel, S., Oberthuer, A., Harnischmacher, U., Stuetzer, H., et al. (2013). Remifentanil-induced Tolerance, Withdrawal or Hyperalgesia in Infants: a Randomized Controlled Trial. RAPIP Trial: Remifentanil-Based Analgesia and Sedation of Paediatric Intensive Care Patients. *Neonatology* 104 (1), 34–41. doi:10.1159/000348790

Zhang, Z. (2016). Multiple Imputation with Multivariate Imputation by Chained Equation (MICE) Package. *Ann. Transl Med.* 4 (2), 30. doi:10.3978/j.issn.2305-5839.2015.12.63

Zhu, S., Kang, Y., Wang, W., Cai, L., Sun, X., and Zong, Z. (2019). The Clinical Impacts and Risk Factors for Non-central Line-Associated Bloodstream Infection in 5046 Intensive Care Unit Patients: an Observational Study Based on Electronic Medical Records. *Crit. Care* 23 (1), 52. doi:10.1186/s13054-019-2353-5

Zhu, S., Wang, W., Kang, Y., He, Q., Zhang, H., Deng, Y., et al. (2021). Clinical Outcomes and Risk Factors for Mortality from Ventilator-Associated Events: A Registry-Based Cohort Study Among 30,830 Intensive Care Unit Patients. *Infect. Control. Hosp. Epidemiol.* 43, 48–55. doi:10.1017/ice.2021.64

Zhu, Y., Wang, Y., Du, B., and Xi, X. (2017). Could Remifentanil Reduce Duration of Mechanical Ventilation in Comparison with Other Opioids for Mechanically Ventilated Patients? A Systematic Review and Meta-Analysis. *Crit. Care* 21 (1), 206. doi:10.1186/s13054-017-1789-8

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Wang, He, Wang, Kang, Ji, Zhu, Zhang, Zou and Sun. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Public Health

BRIEF RESEARCH REPORT
published: 04 March 2022
doi: 10.3389/fpubh.2022.779910

One-Bed-One-Team—Does an Integrated General Hospital Inpatient Model Improve Care Outcomes and Productivity: An Observational Study

Yi Feng Lai[1,2,3,4], Shi Qi Lee[5], Yi-Roe Tan[1], Zheng Yi Lau[5], Jason Phua[6], See Meng Khoo[6], Satya Pavan Kumar Gollamudi[6], Cher Wee Lim[1] and Yee Wei Lim[6,7]*

[1] MOH Office for Healthcare Transformation, Singapore, Singapore, [2] Department of Pharmacy, Alexandra Hospital, Singapore, Singapore, [3] Department of Pharmacy, National University of Singapore, Singapore, Singapore, [4] School of Public Health, University of Illinois at Chicago, Chicago, IL, United States, [5] Division of Policy Research and Evaluation, Ministry of Health, Singapore, Singapore, [6] Alexandra Hospital, National University Health System, Singapore, Singapore, [7] Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

OPEN ACCESS

Edited by:
Bin Jiang,
Peking University, China

Reviewed by:
Yunhwan Lee,
Ajou University, South Korea
Jean-Sébastien Rachoin,
Cooper University Hospital,
United States

****Correspondence:***
Yi Feng Lai
yifeng.lai@moht.com.sg

Specialty section:
This article was submitted to Aging and Public Health, a section of the journal Frontiers in Public Health

Received: *20 September 2021*
Accepted: *25 January 2022*
Published: *04 March 2022*

Citation:
Lai YF, Lee SQ, Tan Y-R, Lau ZY, Phua J, Khoo SM, Gollamudi SPK, Lim CW and Lim YW (2022) One-Bed-One-Team—Does an Integrated General Hospital Inpatient Model Improve Care Outcomes and Productivity: An Observational Study. Front. Public Health 10:779910. doi: 10.3389/fpubh.2022.779910

Introduction: With the increasing complexity of healthcare problems worldwide, the demand for better-coordinated care delivery is on the rise. However, current hospital-based practices remain largely disease-centric and specialist-driven, resulting in fragmented care. This study aimed to evaluate the effectiveness and feasibility of an integrated general hospital (IGH) inpatient care model.

Methods: Retrospective analysis of medical records between June 2018 and August 2019 compared patients admitted under the IGH model and patients receiving usual care in public hospitals. The IGH model managed patients from one location with a multidisciplinary team, performing needs-based care transition utilizing acuity tagging to match the intensity of care to illness acuity.

Results: 5,000 episodes of IGH care entered analysis. In the absence of care transition in intervention and control, IGH average length of stay (ALOS) was 0.7 days shorter than control. In the group with care transition in intervention but not in control, IGH acute ALOS was 2 days shorter, whereas subacute ALOS was 4.8 days longer. In the presence of care transition in intervention and control, IGH acute ALOS was 6.4 and 10.2 days shorter and subacute ALOS was 15.8 and 26.9 days shorter compared with patients under usual care at acute hospitals with and without co-located community hospitals, respectively. The 30- and 60-days readmission rates of IGH patients were marginally higher than usual care, though not clinically significant.

Discussions: The IGH care model maybe associated with shorter ALOS of inpatients and optimize resource allocation and service utilization. Patients with dynamic acuity transition benefited from a seamless care transition process.

Keywords: integrated care, acuity transition, inpatient, care coordination, generalist

INTRODUCTION

Globally, the rapidly aging population with a growing chronic disease prevalence has placed an increasing burden on healthcare systems (1, 2). Healthcare systems are facing the challenge of caring for patients with multimorbidity. The prevalence of multimorbidity increases with age, affecting over 60% of people aged 65 years or older (3–6). With the increasing complexity of multimorbidity, the demand for a comprehensive and coordinated delivery of healthcare services is expected to rise, and healthcare expenditure is expected to soar with inpatient load as the main cost driver (7–9). However, current hospital-based practices remain largely disease-centric with a specialist focus. This is partly driven by the population's preference to seek care from specialists with the belief that they provide better enhanced quality of care (10, 11). Consequently, individuals have multiple specialists managing different health problems during their inpatient stay, resulting in fragmented care and an inefficient hospital system.

To tackle this changing healthcare landscape, integrated care models have been implemented in many countries (12). Integrated care is generally defined as care which involves greater coordination of health and social services, leading to seamless and holistic management (13). Beyond process reorganization, integrated care places the needs of an individual, a family and a community at its center, shifting the approach from disease-centric to person-centric. For the context of this paper, we examined the inpatient aspect of integrated care, in which acute and subacute care are integrated; hence, patients are managed at one setting by one multidisciplinary team, usually led by a generalist (14, 15). Integrated care programmes have been shown to improve the quality of care, patient and caregiver satisfaction, access and cost (12, 16).

In Singapore, the government has implemented various initiatives over the years to provide highly efficient and enhanced coordinated care. These initiatives include reorganization of healthcare services into three integrated regional health systems, increasing a primary care system's capacity and establishing national registries (17, 18). The regional health systems were established to help reduce care fragmentation by including institutions which span the entire health continuum, from acute hospitals (AH) to community hospitals (CH), in each cluster. The government envisioned a seamless transition across different care settings via shared clinical pathways (19, 20). Nevertheless, having acute and subacute care in separate institutions means that patients still need to be transferred between different settings when they transit between episodic acute care and subacute rehabilitative care. Care across multiple sites not only leads to transfer delays and avoidable bed-days, but gaps in knowledge transfer and information sharing also exist between the sites and care teams (17).

The current ecosystem in Singapore still focuses largely on episodic care (21). To meet the evolving complex needs of a population with a finite pool of manpower resources and infrastructure capacities, new care models which offer consolidated care in a highly efficient manner should be explored. Although evidence indicates varying extents of success of integrated care models, most studies were based in the United States and Europe, with limited studies in Asia (12). The current study evaluated a programme in which patients received integrated care in a public hospital in Singapore. Findings from this study could prove useful for other hospitals attempting to implement such a care model in a similar context.

METHODS

Settings

"FAST" Programme

Since June 2018, "FAST" wards in Alexandra Hospital, National University Health System, have operated on the basis of the integrated general hospital (IGH) model which encourages patients to move away from seeking care from multiple specialists. The one-bed-one-team approach is adopted, and patients are managed from one location by a multidisciplinary team across the entire care continuum. This approach is made possible by integrating acute and subacute inpatient care within the same ward. A generalist-led team assesses patient acuity at regular intervals, and the amount of clinical resources devoted is adjusted accordingly in a dynamic manner. The acuity level of a patient is loosely defined as L3 for acute care, L2 for subacute care and L1 for rehab care, with L3 being equivalent to AH care and L2/L1 being equivalent to CH care under the usual care model. Lower acuity levels receive less contact time with physicians (L1<L2<L3) but still receive daily management by a nursing team with a lower intensity of monitoring. The intention is to ensure that resources are allocated in an efficient and cost-effective manner and patient care is uncompromised.

Usual Care

By contrast, usual care separates acute and subacute care, with some hospitals having CHs co-located within the same premise as an AH, whereas most hospitals do not. For AHs that do not have co-located CHs, transfers of care between the institutions are usually more cumbersome. In addition, patients with multimorbidity are usually managed by separate specialists instead of one multidisciplinary team. This has inadvertently resulted in the fragmentation of care, in which patients must be transferred from one team to another or from one facility to another as they progress through their illness. Whilst IGH optimizes resources by having nurse-led teams manage patients with lower acuity, nurses in usual care play a more passive role in patient management.

Study Design

Electronic medical records of patients, including demographic variables, clinical indicators and financial factors, were extracted from Ministry of Health's national casemix datasets. Further detailed financing and costs data were extracted from National University Health System, one of the three regional health systems in Singapore, which included data from IGH wards in Alexandra Hospital, usual care general medicine wards in National University Hospital (NUH, without co-located CH) and usual care general medicine wards in Ng Teng Fong General

Hospital (NTFGH, with co-located Jurong Community Hospital, JCH). All data were de-identified prior to analysis.

Comparisons were planned between IGH patients and two separate groups of controls- Usual care with co-located CHs and usual care without co-located CHs.

For phase one of analysis for primary outcomes, Ministry of Health's national casemix datasets from June 2018 to March 2019 were used. Patients admitted under the IGH wards were compared with patients receiving usual care with data included in the de-identified national cohort from the datasets. Propensity score matching was performed using variables that included age, gender, race, residence type, primary diagnosis, Charlson Comorbidity Index and procedures undergone during hospitalization.

For phase two of analysis for secondary outcomes, financing and costs data from National University Health System between July 2018 and August 2019 were used. Patients admitted under the IGH wards in Alexandra Hospital were compared with patient cohorts from general medicine wards in NUH and NTFGH. Propensity score matching was also performed using variables that included age, gender, race, residence type, primary diagnosis, Charlson Comorbidity Index and procedures undergone during hospitalization.

Prior to propensity score matching, there were a total of 7,087 cases and more than 500,000 controls from national casemix databases where both control groups were selected from. After propensity score matching, a total of 5,000 episodes of IGH care, 9,078 controls without co-located CHs and 7,919 controls with co-located CHs entered analysis. In our propensity score matching, unmatched IGH patients and controls were discarded from the analysis. The quality of matching was evaluated by analyzing the standardized percentage bias across the covariates, which was satisfactorily and significantly reduced post-matching.

Data Analysis

Patient demographics and baseline characteristics were summarized through descriptive statistics. Primary outcome measures included the average length of stay (ALOS) and 30- and 60-days readmission rates, which were compared between the matched intervention group and two groups of usual care controls: (1) patients who received care in hospitals with co-located CH and (2) patients who received care in hospitals without co-located CH. Differences in ALOS between intervention and control were calculated in number of days and compared through *T*-test for each of the following categories: (1) absence of care transition (patients with only L3 acuity level in intervention vs.patients with only AH stay in control); (2) heterogeneous care transition (patients with L3 to L2/L1 acuity transition in intervention vs. patients eligible for acuity transition but not transferred to CH in control) and (3) presence of care transition (patients with L3 to L2/L1 acuity transition in intervention vs. patients transferred to CH in control). Differences in 30- and 60-days readmission rates were compared using one-way ANOVA.

Secondary outcome measures included productivity and cost measures, specifically discharge rates to step-down care services/acuity transition rates, hospitalization costs and service utilization. These were performed among a subset of selected cases and controls were further matched based on their care trajectories- (1) patients without acuity transition, (2) patients with acuity transition in IGH but no CH transfer in controls, and (3) patients with acuity transition in IGH and CH transfer in controls. Differences in care transition rates were reported in proportions, hospitalization costs were analyzed to identify cost drivers in intervention and control, and inpatient service utilization was reported in ratios. Financial data showed that the main cost drivers of hospitalization were room charges, investigations and daily treatment fee. We used the hospitals' daily treatment fees instead of hospitalization bill sizes as the surrogate measure to account for manpower and equipment resourcing because other fixed costs including room charges accounted for in gross hospitalization bills were insignificantly different across institutions. The differences in investigation costs in the current analysis were also attributable to economies of scale and logistics irrelevant for care model comparisons. All statistical analyses were performed with Stata version 16 with $P < 0.05$ regarded as significant.

Ethics Approval

This study was approved by the National Healthcare Group Domain Specific Review Board (Ref: 2020/00023).

RESULTS

Patient Characteristics

The characteristics of patients under IGH care and the two subgroups of usual care controls are presented in **Table 1**.

TABLE 1 | Demographics of patients in different care models.

Patient characteristics	Usual care (without CH co-location)	Usual care (with CH co-location)	IGH patients
No. of patients	9,078	7,919	5,000
Age, mean	69.0	68.1	70.4
Gender = Male	4,176 (46%)	3,722 (47%)	2,350 (47%)
Ethnicity			
Chinese	6,264 (69%)	5,306 (67%)	3,400 (68%)
Malay	999 (11%)	1,029 (13%)	600 (12%)
Indian	999 (11%)	871 (11%)	550 (11%)
Others	817 (9%)	713 (9%)	450 (9%)
Housing			
1–2 rooms flat	1,362 (15%)	1,109 (14%)	800 (16%)
3–4 rooms flat	5,356 (59%)	4,672 (59%)	2,950 (59%)
5-rooms and executive flat	1,543 (17%)	1,584 (20%)	850 (17%)
Private housing	817 (9%)	554 (7%)	400 (8%)
Average CCI score	3.2	3.2	3.2
Among those with SGO data			
Live alone	545 (6%)	317 (4%)	350 (7%)
Little/No interaction with family (less than once a month)	272 (3%)	158 (2%)	200 (4%)
Mobility-impaired	1,634 (18%)	1,267 (16%)	800 (16%)
Need assistance to doctor	2,633 (29%)	2,138 (27%)	1,450 (29%)

IGH, Integrated General Hospital; CH, Community Hospital; CCI, Charlson Comorbidity Index; SGO, Silver Generation Office.

Primary Outcomes

As shown in **Table 2**, though the 30- and 60-days AH readmission rates of IGH patients were marginally higher than those of patients under usual care, the values were not statistically or clinically significant.

Figure 1 shows the ALOS comparison between IGH patients and controls. Among patients without acuity transition, the IGH acute care episodes (AH ALOS) was 0.7 days shorter than that of patients under usual care in hospitals without CH co-location. Among patients with acuity transition in IGH but no CH transfer in controls, the IGH AH ALOS was 2 days shorter, but subacute care episodes (CH ALOS) was 4.8 days longer, indicating that for every AH day saved, there was a corresponding additional two lower resourced CH days expended (with an ALOS trade-off ratio of 1:2). Among patients with acuity transition in IGH and CH transfers in controls, the IGH AH ALOS was 10.2 and 6.4 days shorter and the CH ALOS was 26.9 and 15.8 days shorter compared with those of patients under usual care in hospitals without CH co-location and with CH co-location, respectively.

TABLE 2 | Readmission rates comparison between intervention and usual care.

Clinical outcomes	Variables	Usual care (without CH co-location)	Usual care (with CH co-location)
AH readmission rates within 30 days	IGH	16.5%	16.2%
	Comparator	15.9%	14.6%
	Difference	+0.6%	+1.6%
AH readmission rates within 60 days	IGH	22.8%	22.6%
	Comparator	23.2%	20.6%
	Difference	−0.4%	+2.0%

IGH, Integrated General Hospital; AH, Acute Hospital; CH, Community Hospital.

Secondary Outcomes

The majority of IGH patients had medical diagnoses that were commensurate with a significant proportion of regular general medicine patients at other AHs. However, significant differences in acuity transition rates existed. For the same diagnosis code, 41% of the patients in IGH underwent acuity transition from L3 to L2/L1, whereas only 1% of the patients under usual care were discharged from AHs to CHs.

Table 3 shows the cost and productivity analyses categorized by acuity transition. Cost categories included medication, investigation, treatment services, consumables and consultation fees accounting for approximately 30% of total costs. Amongst all episodes, IGH and NUH's costs were comparable (1%), whereas NTFGH's cost was significantly lower (22%). Among patients without acuity transition, the IGH cost per episode was 8% lower than that for usual care at NTFGH and 34% lower than that at NUH. This difference was ascribed to the lower IGH ALOS (by 3 and 13%, respectively) and lower unit cost for IGH service items, especially for daily treatment fee.

Among patients with acuity transition in IGH but no CH transfer in controls, the IGH overall cost per episode was 38% higher than that for usual care at NTFGH and 27% higher than that at NUH. When analyzed by variable costs and care model productivity by using daily treatment fee as surrogate, the IGH cost was then 17.7 and 33.7% lower than those at NTFGH and NUH, respectively, primarily because of the lower L3 daily cost (**Table 3**). The cost trade-off ratio was ∼1:2, with one L3 day under usual care equivalent to two L1/L2 days under the IGH model. This cost trade-off ratio was similar to the ALOS trade-off ratio of 1:2 indicated above. Thus, limited savings were achieved from a lower IGH daily treatment fee. The relatively modest productivity gain was also negated by larger cost drivers, such as room charges, resulting in an overall higher hospitalization cost for IGH. Among patients with acuity transition in IGH and CH transfer in controls, the IGH overall cost per episode was 53% lower than that for usual care at NTFGH and 162% lower than

(1) No acuity transition in both groups

		Usual care (without CH co-location)	Usual care (with CH co-location)
AH ALOS	IGH	3.6	3.5
	Comparator	4.3	3.4
	Difference (%)	-0.7 (15%)***	+0.1 (2%)

(2) Acuity transition in IGH, no CH transfer in usual care

		Usual care (without CH co-location)	Usual care (with CH co-location)
AH ALOS	IGH	3.2	3.2
	Comparator	5.4	5.2
	Difference (%)	-2.2 (41%)***	-2.0 (39%)***
CH ALOS	IGH	4.8	4.8
	Comparator	-	-
	Difference (%)	+4.8***	+4.8***
Total ALOS	IGH	8.0	8.0
	Comparator	5.4	5.2
	Difference (%)	+2.6 (48%)***	+2.8 (54%)***

(3) Acuity transition in IGH, CH transfer in usual care

		Usual care (without CH co-location)	Usual care (with CH co-location)
AH ALOS	IGH	3.3	3.3
	Comparator	13.5	9.7
	Difference (%)	-10.2 (76%)***	-6.4 (66%)***
CH ALOS	IGH	5.0	5.1
	Comparator	31.9	20.9
	Difference (%)	-26.9 (84%)***	-15.8 (76%)***
Total ALOS	IGH	8.3	8.4
	Comparator	45.4	30.6
	Difference (%)	-37.1 (82%)***	-22.2 (73%)***

FIGURE 1 | Average length of stay (days) comparison between intervention and usual care. IGH, Integrated General Hospital; AH, Acute Hospital; CH, Community Hospital; ALOS, Average Length of Stay. ***P<0.001.

TABLE 3 | Cost and productivity analyses between intervention and usual care.

	IGH (Intervention)	NUH (Usual care without co-located CH	NTFGH (Usual care with co-located CH)
Overall cost comparisons			
No. of matched episodes	1,182		
Cost per episode (% difference)	$5,249	$5,315 (+1%)	$4,075 (−22%)
Among patients without acuity transition			
No. of matched episodes	619		
Cost per episode (% difference)	$3,247	$4,340 (+34%)	$3,515 (+8%)
Among patients with acuity transition in IGH but no CH transfer in controls			
No. of matched episodes	595		
Cost per episode (% difference)	$7,307	$5,370 (−27%)	$4,509 (−38%)
Among patients with acuity transition in IGH and CH transfer in controls			
No. of matched episodes	68 (vs. NUH) 112 (vs. NTFGH)	68	112
Cost per episode (% difference)	$9,907 (vs. NUH) $12,891 (vs. NTFGH)	$25,958 (+162%)	$19,785 (+53%)
Productivity comparisons			
	IGH	**NUH**	**NTFGH**
Daily treatment fee			
L1	$111	$280	$256
L2	$135		
L3	$169		
ALOS			
Acute hospital/L3	3.2	5.4	5.2
Community hospital/L1.L2	4.8	-	-
Overall cost			
L3 cost	$540.80 (3.2days x $169)	$1,512 (5.4days x $280)	$1,331.20 (5.2days x $256)
L1/L2 cost	$590.40 (4.8days x $123)	-	-
Total cost (% difference)	$1,131.20	$1,512 (+33.7%)	$1,331.20 (+17.7%)

IGH, Integrated General Hospital; NUH, National University Hospital; NTFGH, Ng Teng Fong General Hospital; CH, Community Hospital; ALOS, Average Length of Stay.

that at NUH, primarily driven by the much higher CH ALOS for NUH and NTFGH patients.

DISCUSSION

Multi-Element Integrated Care Model

Many integrated care models have been implemented and evaluated internationally (22). Studies which examined the length of stay as an outcome have shown inconsistent evidence, with some reporting a reduction, whereas others found an insignificant effect (12). In our study, we found that the IGH inpatient model resulted in an overall shorter ALOS. This finding could be due to several differentiating elements of the IGH model.

Firstly, having one care team to overlook a patient's entire care continuum in the hospital instead of having multiple specialists was proven more efficient without the need for case handovers and duplicative investigations. Secondly, the IGH model reduced administrative hassle and transfer delays by managing patients from one location without the need to move patients physically across institutions, resulting in timely and frictionless care transition. Lastly, a seamless acuity transition process allowed regular adjustment of patient acuity when a patient's condition improved or deteriorated during the care episode, contributing to a smooth continuity of care. The interplay of these elements contributed to decreased ALOS in IGH patients and reduced cost per episode. This finding has been similarly demonstrated in other studies that incorporated multiple elements into their integrated care models (23).

Resource Optimisation Through Smooth Acuity Transition

Our study found that for the same diagnosis code, much lesser patients in usual care underwent care transition compared with those under IGH care. This group of patients reflects the potential pool of patients that an IGH model could have an effect on because the only current mechanism to influence acuity transition in usual care systematically is through an AH–CH transfer. However, this mechanism is seldom meaningfully practiced for vast majority of patients due to barriers in transfer, complex transfer application and documentation and

operational and financial frameworks at CHs (24, 25). Hence, the implementation of an IGH model could result in process optimisation and resource savings by targeting patients in a high-risk group who have a high likelihood of undergoing acuity transition. As demonstrated in the IGH model and in the literature, this group could include patients who have multimorbidity, low socioeconomic status and mobility impairment and are older and living alone (26).

For patients requiring acuity transition, the current ALOS trade-off ratio is similar to the cost trade-off ratio at 1:2, negating the effects of any productivity gains and causing the overall hospitalization cost to be high for IGH. The ALOS and cost trade-off ratios should be maximized to make IGH financially favorable compared with usual care. This objective can be achieved through workforce substitution and increased efficiency in care transition. A multidisciplinary team can efficiently match patients' care plans to their illness trajectories by familiarizing with the acuity tagging system. This strategy allows service utilization to be tagged to the right acuity levels and clinical resources to be allocated in an efficient manner (14). Communication within the care team and administrative protocols can also be improved to achieve enhanced efficiency in care transition across acuity levels (27). To yield further cost savings, an IGH model may be explored in CH-like settings and resources instead of acute care settings.

Lessons from this study suggested that the areas where holistic integrated care models such as IGH model could provide the greatest benefits are in whole-of-person care. Such an approach reduced fragmentation and improved discharge planning especially for the increasing proportion of general medicine patients not requiring specialist care as a result of aging population and rising chronic diseases. These patients could be looked after by a generalist care team to allow for cost-effective single-team holistic management.

Strengths and Limitations

The strengths of this study are that the data were extracted from the national cohort and a comprehensive analysis was conducted on a nationwide level including all public hospitals. This study is also one of the few integrated care programmes in the region which looked at the use of acuity labeling and its effects on care outcomes and productivity. The limitations are that patients' disease severity was not an outright variable in the propensity score matching, considering that no standardized measurement is available for disease severity across institutions. Some of the large differences in ALOS observed, especially in group 3 comparison, could be magnified due to this poor matching of disease severity. We tried to account for this limitation by including variables which are equivalent predictors, such as the Charlson Comorbidity Index, primary diagnosis and procedures performed during the length of stay. In addition, while other analyses such as cost-effectiveness could be useful to be carried out, this retrospective study was limited by the amount and type of extractable data available for review, with the aim of providing early signals to the care model validation. The team would be pursuing a follow-up prospective cohort study to further validate the care model outcomes during hospitalization, as well as potential benefits in post-discharge healthcare utilization.

Implications and Outlook

Findings from this study contribute to the global pool of evidence validating this new model of care, especially in the Asian context where evidence is scarce. Given that we build insights into clinical efficiency, manpower sustainability and healthcare financing, policy owners can make a highly informed decision on hospital care redesign and whether such an integrated care model is feasible and sustainable to be developed as mainstream care in our ecosystem. Moving forward, the perspectives and acceptability of patients, caregivers and various stakeholders of this IGH model implementation should be explored. Beyond inpatient care and hospital-based measures, the short- and medium-term effects of the IGH care model could not be evaluated in this retrospective analysis. Future studies, such as a prospective cohort study, could provide additional insights into the downstream effects of IGH and the degree of care integration beyond hospital walls. Such areas as discharge planning, post-discharge care experience, care continuity and post-discharge utilization of healthcare resources are potential factors that will contribute to a comprehensive IGH model.

CONCLUSION

The IGH care model has shown promising results in shortening the ALOS of inpatients by providing a holistic and coordinated management of their illnesses through a one-bed-one-team approach. The IGH model has the potential to allow for highly efficient care transition and resource optimisation without compromising on the quality of patient care by accurately and dynamically tagging a patient's trajectory of illness to appropriate acuity levels. Thus, general medicine patients who require acuity transition, which is the majority of inpatient load, should be the patient segment that the IGH model focuses on to become a mainstream care model that tackles the evolving complex healthcare needs of patients.

DATA AVAILABILITY STATEMENT

The datasets analyzed during the current study are available from the corresponding author on reasonable request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by National Healthcare Group Domain Specific Review Board. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

YL, SL, JP, SK, SG, CL, and YWL were involved in the study conception and study design. JP and YL were involved in the extraction of data. SL and ZL were involved in data analysis. Y-RT, YL, JP, SK, SG, CL, and YL were involved in the interpretation of

findings. Y-RT was involved in the writing of the manuscript. All authors participated in the review of the manuscript, provided critical revisions, and gave the final approval of the version to be published.

FUNDING

This research was internally funded by the Ministry of Health Office for Healthcare Transformation (MOHT) to support the research for Integrated General Hospital Care Model Evaluation in the National University Health System in Singapore.

ACKNOWLEDGMENTS

The study team would like to acknowledge the efforts put in by the team from Alexandra Hospital in implementing the FAST/IGH programme.

REFERENCES

1. GBD 2015 Mortality and Causes of Death Collaborators. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the global burden of disease study 2015. *Lancet.* (2016) 388:1459–544. doi: 10.1016/S0140-6736(16)31012-1
2. Wen WK. Futures of ageing in Singapore. *J Futures Stud.* (2013) 17:81–102. Available online at: https://jfsdigital.org/wp-content/uploads/2013/10/173-A05.pdf (accessed February 7, 2022).
3. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet.* (2012) 380:37–43. doi: 10.1016/S0140-6736(12)60240-2
4. Melis R, Marengoni A, Angleman S, Fratiglioni L. Incidence and predictors of multimorbidity in the elderly: a population-based longitudinal study. *PLoS ONE.* (2014) 9:e103120. doi: 10.1371/journal.pone.0103120
5. Marengoni A, Winblad B, Karp A, Fratiglioni L. Prevalence of chronic diseases and multimorbidity among the elderly population in Sweden. *Am J Public Health.* (2008) 98:1198–200. doi: 10.2105/AJPH.2007.121137
6. Tinetti ME, Fried TR, Boyd CM. Designing health care for the most common chronic condition—multimorbidity. *JAMA.* (2012) 307:2493–4. doi: 10.1001/jama.2012.5265
7. Low LL, Liu N, Wang S, Thumboo J, Ong MEH, Lee KH. Predicting frequent hospital admission risk in Singapore: a retrospective cohort study to investigate the impact of comorbidities, acute illness burden and social determinants of health. *BMJ Open.* (2016) 6:e012705. doi: 10.1136/bmjopen-2016-012705
8. Choo F. *Proportion of Older Adults with Multiple Chronic Diseases Surges.* The Straits Times (2019) pB6. Available online at: https://www.nus.edu.sg/newshub/news/2019/2019-05/2019-05-08/CHRONIC-st-8may-pB6.pdf (accessed February 7, 2022).
9. Vogeli C, Shields AE, Lee TA, Gibson TB, Marder WD, Weiss KB, et al. Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. *J Gen Intern Med.* (2007) 22(Suppl. 3):391–5. doi: 10.1007/s11606-007-0322-1
10. van Weel C, Kassai R. Expanding primary care in South and East Asia. *BMJ.* (2017) 356:j634. doi: 10.1136/bmj.j634
11. Montegut AJ, Cartwright CA, Schirmer JM, Cummings S. An international consultation: the development of family medicine in Vietnam. *Fam Med.* (2004) 36:352–6.
12. Baxter S, Johnson M, Chambers D, Sutton A, Goyder E, Booth A. The effects of integrated care: a systematic review of UK and international evidence. *BMC Health Serv Res.* (2018) 18:350. doi: 10.1186/s12913-018-3161-3
13. World Health Organization. WHO global strategy on people-centred and integrated health services: interim report. Geneva: World Health Organization (2015).
14. Onder G, Palmer K, Navickas R, Jurevičiene E, Mammarella F, Strandzheva M, et al. Time to face the challenge of multimorbidity. A European perspective from the joint action on chronic diseases and promoting healthy ageing across the life cycle (JA-CHRODIS). *Eur J Intern Med.* (2015) 26:157–9. doi: 10.1016/j.ejim.2015.02.020
15. Hébert R, Durand PJ, Dubuc N, Tourigny A. PRISMA a new model of integrated service delivery for the frail older people in Canada. *Int J Integr Care.* (2003) 3:e08. doi: 10.5334/ijic.73
16. Ow Yong LM, Cameron A. Learning from elsewhere: Integrated care development in Singapore. *Health Policy.* (2019) 123:393–402. doi: 10.1016/j.healthpol.2018.12.004
17. Sumner J, Lim YW. Back to the hospital for chronic care: a hybrid generalist model. *J Gen Intern Med.* (2020) 36:782–5. doi: 10.1007/s11606-020-06271-6
18. Ministry of Health. *Reorganisation of Healthcare System into Three Integrated Clusters to Better Meet Future Healthcare Needs* (2018).
19. Thomas JS, Ong SE, Chia KS, Lee HP. A brief history of public health in Singapore. *Singapore's Healthc Syst.* (2016) 33–56. doi: 10.1142/9789814696067_0002
20. Nurjono M, Yoong J, Yap P, Wee SL, Vrijhoef HJM. Implementation of integrated care in Singapore: a complex adaptive system perspective. *Int J Integr Care.* (2018) 18:4. doi: 10.5334/ijic.4174
21. Saxena N, You AX, Zhu Z, Sun Y, George PP, Teow KL, et al. Singapore's regional health systems—a data-driven perspective on frequent admitters and cross utilization of healthcare services in three systems. *Int J Health Plann Manage.* (2017) 32:36–49. doi: 10.1002/hpm.2300
22. Kodner DL, Spreeuwenberg C. Integrated care: meaning, logic, applications, and implications–a discussion paper. *Int J Integr Care.* (2002) 2:e12. doi: 10.5334/ijic.67
23. Hughes G, Shaw SE, Greenhalgh T. Rethinking integrated care: a systematic hermeneutic review of the literature on integrated care strategies and concepts. *Milbank Q.* (2020) 98:446–92. doi: 10.1111/1468-0009.12459
24. Kadir FS. *Role of the Multi-Disciplinary Team in Integrated Care*. Singapore: The Singapore Family Physician (2011).
25. Community Hospital Care - Handbook for Healthcare Professionals. *Ministry of Health Singapore* (2018). Available online at: https://www.moh.gov.sg/docs/librariesprovider5/default-document-library/handbook-on-ch-care-for-healthcare-professionals-2nd-edition-200918-for.pdf (accessed February 7, 2022).
26. Puvanendran R. *Discharge Planning in Integrated Care*. Singapore: The Singapore Family Physician. (2011).
27. Sykes S, Baillie L, Thomas B, Scotter J, Martin F. Enhancing care transitions for older people through interprofessional simulation: a mixed method evaluation. *Int J Integr Care.* (2017) 17:3. doi: 10.5334/ijic.3055

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Lai, Lee, Tan, Lau, Phua, Khoo, Gollamudi, Lim and Lim. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 04 March 2022
doi: 10.3389/fphar.2022.815882

Statin Therapy for Hyperlipidemic Patients With Chronic Kidney Disease and End-Stage Renal Disease: A Retrospective Cohort Study Based on 925,418 Adults in Taiwan

Fung-Chang Sung[1,2,3†], *Ying-Chin Jong*[4†], *Chih-Hsin Muo*[2,5], *Chih-Cheng Hsu*[1,6,7], *Wen-Chen Tsai*[1] *and Yueh-Han Hsu*[4,8,9*]

[1]Department of Health Services Administration, China Medical University, Taichung, Taiwan, [2]Management Office for Health Data, China Medical University Hospital, Taichung, Taiwan, [3]Department of Food Nutrition and Health Biotechnology, Asia University, Taichung, Taiwan, [4]Division of Nephrology, Department of Internal Medicine, Ditmanson Medical Foundation Chia-Yi Christian Hospital, Chiayi, Taiwan, [5]Graduate Institute of Clinical Medical Science, College of Medicine, China Medical University, Taichung, Taiwan, [6]Institute of Population Health Sciences, National Health Research Institutes, Zhunan, Taiwan, [7]Department of Family Medicine, Min-Sheng General Hospital, Taoyuan, Taiwan, [8]Department of Medical Research, China Medical University Hospital, Taichung, Taiwan, [9]Department of Nursing, Min-Hwei Junior College of Health Care Management, Tainan, Taiwan

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Pilly Chillo,
Muhimbili University of Health and Allied Sciences, Tanzania
Tauqeer Hussain Mallhi,
Al Jouf University, Saudi Arabia

***Correspondence:**
Yueh-Han Hsu
cych07023@gmail.com

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Drugs Outcomes Research and Policies,
a section of the journal Frontiers in Pharmacology

Received: 16 November 2021
Accepted: 24 January 2022
Published: 04 March 2022

Citation:
Sung F-C, Jong Y-C, Muo C-H, Hsu C-C, Tsai W-C and Hsu Y-H (2022) Statin Therapy for Hyperlipidemic Patients With Chronic Kidney Disease and End-Stage Renal Disease: A Retrospective Cohort Study Based on 925,418 Adults in Taiwan. Front. Pharmacol. 13:815882. doi: 10.3389/fphar.2022.815882

Background: For non-dialysis patients with hyperlipidemia, statins may provide clinical benefits in reducing mortality risk; however, the optimal treatment for dialysis patients with hyperlipidemia remains debatable. We evaluated the mortality risks for hyperlipidemic patients with renal disorders associated with statin therapy (ST), using the insurance claims data of Taiwan.

Methods: From hyperlipidemic patients diagnosed in 2000–2011, we identified 555,153 patients receiving statin treatment for at least 90 days continuously and 1,141,901 non-statin users, and then randomly selected, from both groups, the propensity score-matched subcohorts of statin users and nonusers in a 1:1 pair by renal function: 415,453 pairs with normal renal function , 43,632 pairs with chronic kidney disease (CKD), and 3,624 pairs with end-stage renal disease (ESRD). We compared the mortalities, by the end of 2016, from all causes, cancer, heart disease, and septicemia between statin users and non-users and between hydrophilic-statin users and lipophilic-statin users. The Cox method estimated ST users to non-user hazard ratios. The time-dependent model was also conducted as sensitivity analysis.

Results: The mean ages were 58.7 ± 10.7, 64.2 ± 10.7, and 62.2 ± 10.8 years in normal renal function, CKD, and ESRD groups, respectively. Compared with non-users, statin users had reduced mortality risks from all causes for 32%–38%, from cancer for 37%–46%, from heart disease for 6%–24%, and from septicemia for 17%–21% in all three renal groups. The hydrophilic statin therapy was superior than the lipophilic statin therapy, particularly for reducing deaths from all-causes and cancer. The results under the time-dependent model were similar.

Conclusion: Statin therapy is associated with reduced all-causes and non-cardiovascular mortality in ESRD patients.

Keywords: cancer, chronic kidney disease, end-stage renal disease, heart disease, mortality, propensity-score, septicemia, statins

INTRODUCTION

The mortality is higher in dialysis patients than in the general population, approximately 16-fold higher for the US patients and 5-fold higher for Japanese patients (De Jager et al., 2009; Wakasugi et al., 2013). The main causes of deaths for patients with end-stage renal disease (ESRD) included cardiovascular disease (CVD), infection, cancer, and dialysis withdrawal (Den Hoedt et al., 2013; Vogelzang et al., 2015; Steenkamp et al., 2018). CVD has been the leading cause of deaths. However, the UK Renal Registry Annual Report indicated that the deaths from CVD decreased from 34% to 24% in 2000–2015 in adult dialysis patients (Steenkamp et al., 2018). The proportion of non-cardiovascular deaths for ESRD patients was increasing, especially in the elderly group (Den Hoedt et al., 2013; Vogelzang et al., 2015). The study from the ERA-EDTA registry showed an 82-fold increase in infection-related mortality and a near 3-fold increase in cancer-related mortality for dialysis patients compared to the general population (Vogelzang et al., 2015).

Statins are the first-line lipid-lowering medications for the patients of chronic kidney disease (CKD) with hyperlipidemia. Statin therapy (ST) could reduce deaths and CVD for 20% in CKD patients not requiring dialysis (Palmer et al., 2014). However, the optimal treatment for dialysis patients with hyperlipidemia remains inconclusive. For hemodialysis (HD) patients, Manson et al. reported that ST could reduce the mortality risks from all causes for 31%, from cardiac disease for 23%, and from noncardiac diseases for 44% (Fellstrom et al., 2009). In two randomized control trials, the AURORA study and the 4D (Die Deutsche Diabetes Dialyse) study, focusing on dialysis cohorts, found that rosuvastatin and atorvastatin could effectively reduce the serum levels of low-density lipoprotein cholesterol. However, both studies failed to observe the clinical benefits in reducing CVD mortality, all-cause mortality (ACM), and cardiovascular events (Wanner et al., 2005; Fellstrom et al., 2009). A meta-analysis has commented that ST contributed little benefits on mortality in dialysis adults (Palmer et al., 2013). In the *post-hoc* analyses of these two trials, the AURORA investigators remodeled the study design and found a 32% risk reduction from cardiac events in those with rosuvastatin treatment for patients with diabetes alone (Holdaas et al., 2011). For diabetic dialysis patients with pretreatment LDL-cholesterol >145 mg/dl in the 4D trial, März et al. reported that atorvastatin treatment could reduce 28%–52% of the risk for composite cardiovascular outcomes, cardiac deaths, and all-cause deaths (März et al., 2011). Recent studies using real-world data showed that ST could reduce the ACM for dialysis patients, particularly for HD patients (De Rango et al., 2016; Chung et al., 2017; Kim et al., 2018; Ota et al., 2019). A recent Korean population-based study found that ST was associated with reduced ACM for over 41% in HD patients (Jung et al., 2020).

Previous studies rarely used population-based data to examine simultaneously the effectiveness of ST associated with deaths from different causes for hyperlipidemia patients with renal disorders. We, therefore, used insurance claims data to investigate the effectiveness of ST in reducing the death risks not only from all causes but also from cancer, heart disease, and septicemia, for hyperlipidemic patients with kidney disorders, including patients with CKD and ESRD, compared with hyperlipidemic patients with normal renal function (NRF). Furthermore, the effectiveness between using hydrophilic statins (HSs) and lipophilic statins (LSs) was compared.

MATERIALS AND METHODS

Data Source

We obtained the data from the Ministry of Health and Welfare Health and Welfare Data Science Center database, consisting of the National Health Insurance Research database (NHIRD) and the death registry of all insured people in Taiwan. NHIRD was composed of the information on the medical records of outpatient and inpatient claims for all residents in Taiwan. The death registry was composed of the information on the demographic status and cause of death of each deceased. These data sets were linked by recoded identifications to protect the privacy of the insured population. Drug and disease classifications conformed to the Anatomical Therapeutic Chemical (ATC) Classification System and the International Statistical Classification of Diseases and Related Health Problems [9th Revision (ICD-9) before 2016 and 10th Revision (ICD-10) since 2016].

Study Population Selection

ST cohort. We identified from the NHIRD 4,638,044 the patients with hyperlipidemia who were newly diagnosed in 2000–2011 in Taiwan. Among these patients, 812,024 patients who had received ST for at least 90 days continuously were selected for the potential cohorts with ST (**Figure 1**). The 91st day with ST was defined as the index date to address immortal time bias. We excluded patients aged <40 or >80 (n = 71,482), patients with the history of ESRD or CKD for less than 90 days (183,486), patients with HIV (n = 436), patients with kidney transplant history (n = 728), and ESRD patients without dialysis information (n = 739). The remaining 555,153 patients eligible for the ST cohort were stratified into three subcohorts by the renal function: the subcohorts of patients with NRF, with CKD, and with ESRD. Patients without the diagnosis of CKD and ESRD were selected in the NRF group.

Non-ST cohort. From 2,449,148 patients with hyperlipidemia, but without ST, we identified 1,141,901 patients eligible for the non-ST cohort, applying the similar exclusion criteria to the corresponding ST cohort. Likewise, we categorized them into three groups by the renal function: NRF, CKD, and ESRD groups (**Figure 1**). From each of these three groups, we further randomly selected a non-ST subcohort, as a comparison, corresponding to the specific ST subcohort with a 1:1 ratio, matched by the index date and propensity score (PS). The PS was calculated by multivariable logistic regression, including the variables of age, gender, monthly income, living areas, comorbidities, year hyperlipidemia diagnosed, and index date for NRF group. In the CKD group, we added the CKD diagnosis-year for adjustment. In the ESRD group, we further added ESRD diagnosis-year and HD for adjustments.

Ethic Statement

All personal identifications in the database had been scrambled and replaced with surrogate numbers before data were released to

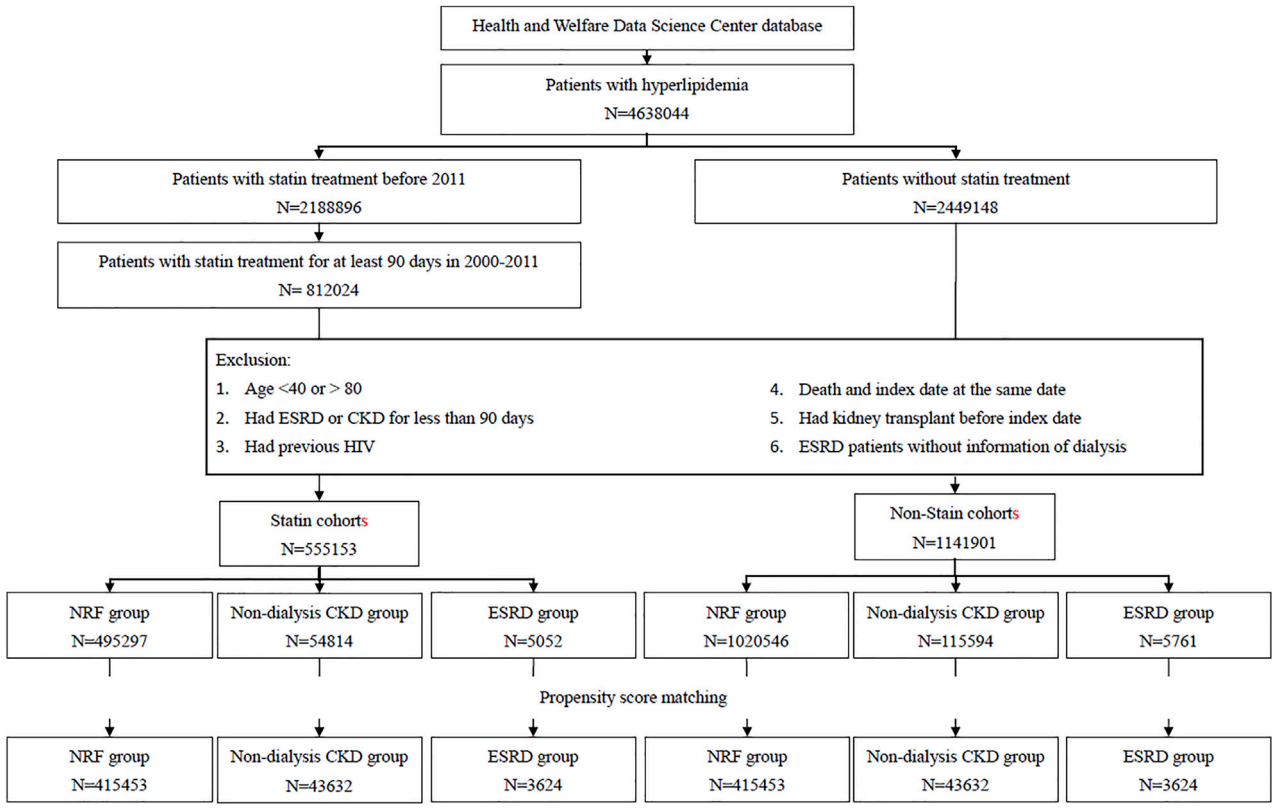

FIGURE. 1 | Flow chart for establishing study cohorts.

researchers to protect privacy. This study was approved by the Research Ethics Committee at China Medical University and Hospital (CRREC-107-021) and Ditmanson Medical Foundation Chiayi Christian Hospital (CYCH-IRB-2019063). Because personal privacy was protected from the claims data, the requirement for patient consent was waived. This study was conducted in accordance with the ethics code of the Declaration of Helsinki.

Outcome and Comorbidity

All study subjects were followed from the index date until death, withdrawal from the insurance, or the end of 2016. The primary outcomes were ACM and three other cause-specific deaths from cancer, heart disease, and septicemia, retrieved from the death registry. Mortalities were compared between patients with and without ST for the three renal groups. Comorbidities including diabetes, hypertension, peripheral arterial occlusive disease, ischemic heart disease, hepatitis B infection, hepatitis C infection, and stroke were considered as covariates. All comorbidities were defined by clinical diagnosis within 2 years before the index date.

Statistical Analysis

The baseline distributions of the demographic status and comorbidities between each paired subcohort with and without ST in each NRF, CKD, and ESRD group were compared. We calculated the standardized difference of each variable between each paired subcohort. The Kaplan–Meier method calculated and plotted proportional survivals and compared (A) between statin users and nonusers (B) among HS (pravastatin and rosuvastatin) users, LS (including simvastatin, lovastatin, fluvastatin, atorvastatin, cerivastatin, and pitavastatin) users, and nonusers. The ACM rate and each cause specific mortality rate were calculated for each subcohort: the sum of deaths divided by the sum of follow-up years (person-years). We used Cox proportional hazards regression analysis to estimate the ST subcohort-to-non-ST subcohort adjusted hazard ratio (aHR) and a 95% confidence interval (CI) of mortality after controlling for covariates. We further compared the age-specific mortality rates between the ST users and non-users by cause and the renal function group. In order to estimate the short-term impact of ST, we also used a time-dependent Cox regression model to estimate the age-specific ST users to non-user aHRs. All statistical tests were two sided, and the statistical significance was defined as p-value <0.05. We used SAS, version 9.4 (SAS Institute, Cary, NC, United States) to conduct data analyses.

RESULTS

We established three pairs of ST subcohort and non-ST subcohort by the renal function: 415,453 pairs in the NRF group, 43,632 pairs in the CKD group, and 3,624 pairs in the ESRD group (**Figure 1**). The CKD groups were older with more

TABLE 1 | Demographic and clinical characteristics of statin cohort and propensity-score matched non-statin cohort in patients with NRF, CKD, and ESRD.

Variable	Normal renal function					Non-dialysis CKD					ESRD				Standardized difference
	Non-statin N = 415,453 n %		Statin N = 415,453 n %		Standardized difference	Non-statin N = 43,632 n %		Statin N = 43,632 n %		Standardized difference	Non-statin N = 3,624 n %		Statin N = 3,624 n %		
Age, mean (SD)	58.7	(10.7)	59.3	(9.94)	0.059	64.2	(10.7)	64.1	(9.99)	0.010	62.2	(10.8)	61.8	(9.89)	0.035
Men, n (%)	199,414	48.0	194,026	46.7	0.026	23,968	54.9	22,680	52.0	0.059	1,690	46.6	1,605	44.3	0.047
Income, NTD															
≤19,200	108,710	26.2	111,341	26.8	0.014	12,795	29.3	12,702	29.1	0.005	1,128	31.1	1,112	30.7	0.010
19,201–21,000	127,763	30.8	127,228	30.6	0.003	14,600	33.5	15,055	34.5	0.022	1,188	32.8	1,220	33.7	0.019
21,001–34,800	83,391	20.1	82,112	19.8	0.008	7,861	18.0	7,648	17.5	0.013	734	20.3	710	19.6	0.017
>34,800	95,589	23.0	94,772	22.8	0.005	8,376	19.2	8,227	18.9	0.009	574	15.8	582	16.1	0.006
Living area															
Northern	18,8952	45.5	192,093	46.2	0.015	16,887	38.7	17,072	39.1	0.009	1,437	39.7	1,440	39.7	0.002
Central	78,779	19.0	77,074	18.6	0.011	9,401	21.6	9,425	21.6	0.001	743	20.5	728	20.1	0.010
Southern	121,452	29.2	118,696	28.6	0.015	14,833	34.0	14,438	33.1	0.019	1,232	34.0	1,231	34.0	0.001
Eastern and offshore islands	26,270	6.32	27,590	6.64	0.013	2,511	5.75	2,697	6.18	0.018	212	5.85	225	6.21	0.015
Comorbidity, n (%)															
DM	125,511	30.2	132,964	32.0	0.039	15,163	34.8	16,721	38.3	0.074	1,887	52.1	1,876	51.8	0.006
Hypertension	181,994	43.8	182,661	44.0	0.003	25,396	58.2	25,146	57.6	0.012	2,495	68.9	2,473	68.2	0.013
PAOD	8,899	2.14	9,254	2.23	0.006	1,516	3.47	1,550	3.55	0.004	314	8.66	326	9.00	0.012
IHD	57,994	14.0	64,611	15.6	0.045	8,688	19.9	9,367	21.5	0.038	1,242	34.3	1,245	34.4	0.002
HBV	7,516	1.81	7,041	1.69	0.009	835	1.91	662	1.52	0.031	119	3.28	118	3.26	0.002
HCV	3,219	0.77	2,744	0.66	0.014	572	1.31	459	1.05	0.024	130	3.59	132	3.64	0.003
Stroke	17,002	4.09	22,993	5.53	0.067	3,050	6.99	3,584	8.21	0.046	291	8.03	299	8.25	0.008
Cancer	15,432	3.71	13,985	3.37	0.019	2,595	5.95	2,133	4.89	0.047	299	8.25	281	7.75	0.018
Hemodialysis											3,243	89.5	3,177	87.7	0.057

Abbreviations: CKD, chronic kidney disease; ESRD, end-stage renal disease; DM, diabetes mellitus; PAOD, peripheral arterial occlusive disease; IHD, ischemic heart disease; HBV, hepatitis B virus infection; HCV, hepatitis C virus infection

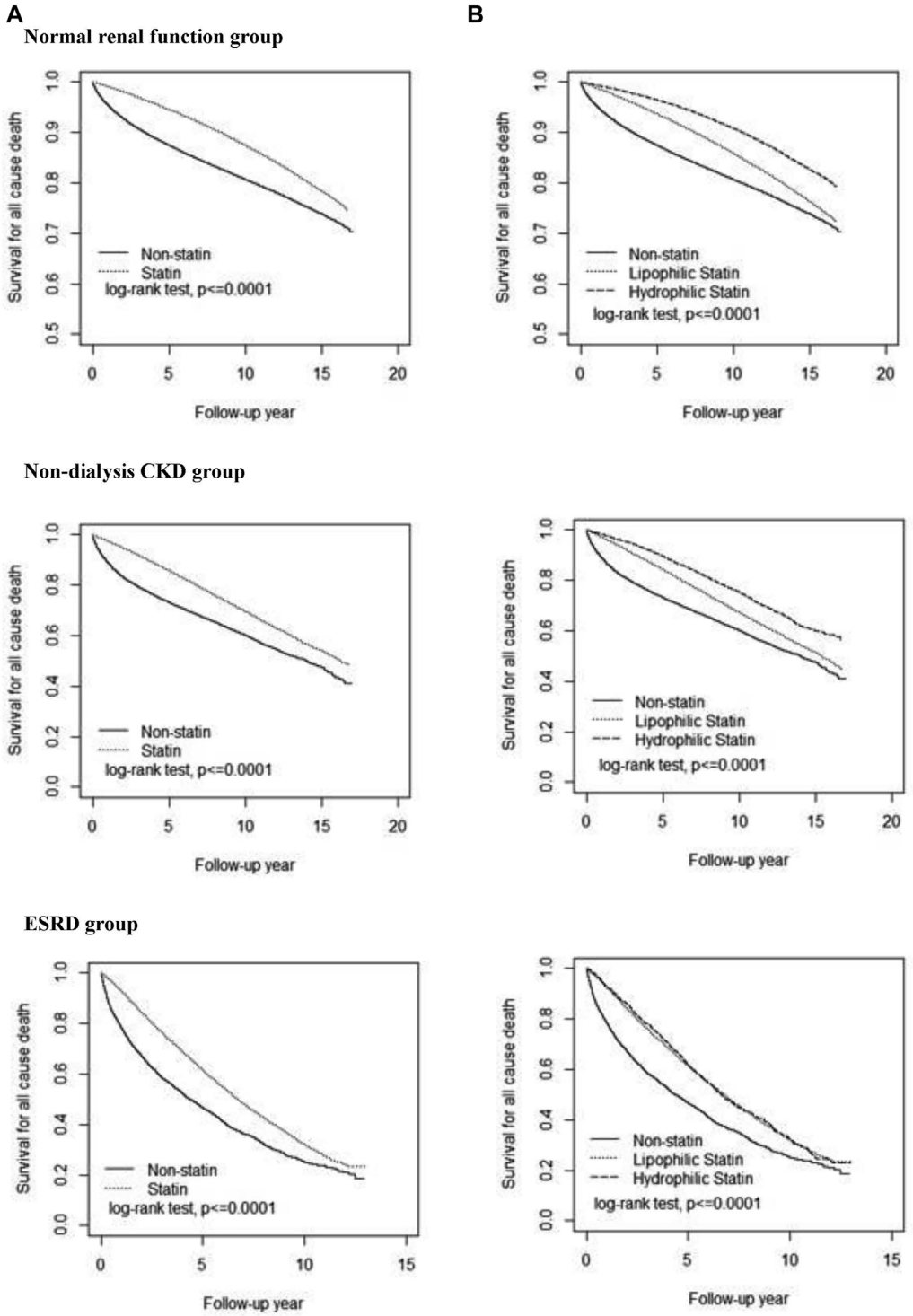

FIGURE. 2 | Cumulative survival probability among three renal groups. Comparisons of all-cause death in patients with statin therapy versus non-statin therapy **(A)** and with hydrophilic statin versus lipophilic statin **(B)**. CKD, chronic kidney disease; ESRD, end-stage renal disease.

men than the groups of NRF and ESRD (**Table 1**). The distributions of income, living area, and comorbidities were similar between each pair of subcohorts. However, the CKD groups and ESRD groups were more likely from lower-income households in central and southern Taiwan and were more prevalent with comorbidities.

Kaplan–Meier plots showed that the survival probability was the highest in the NRF group, followed by the CKD group, and

TABLE 2 | Mortality compared between statin users and non-users by cause and renal group.

Outcome	Non-statin			Statin			aHR (95%CI)
	Event no.	PYs	Rate, per 1,000 PYs	Event no.	PYs	Rate, per 1,000 PYs	
All-cause death							
Normal renal function	73,560	3,249,626	22.6	54,517	3,916,082	13.9	0.62 (0.60–0.63)[b]
Non-dialysis CKD	15,377	272,621	56.4	11,642	324,339	35.9	0.64 (0.62–0.65)[b]
ESRD	2,278	14,734	154.6	2,186	20,904	104.6	0.68 (0.64–0.72)[b]
p for trend			<0.0001			<0.0001	
Cancer death							
Normal renal function	23,821	3,249,626	7.33	15,868	3,916,082	4.05	0.56 (0.55–0.57)[b]
Non-dialysis CKD	3,226	272,621	11.8	2,078	324,339	6.41	0.54 (0.51–0.57)[b]
ESRD	231	14,734	15.7	204	20,904	9.76	0.63 (0.52–0.75)[b]
p for trend			<0.0001			<0.0001	
Heart disease death							
Normal renal function	7,214	3,249,626	2.22	8,196	3,916,082	2.09	0.94 (0.91–0.97)[b]
Non-dialysis CKD	1,780	272,621	6.53	1,825	324,339	5.63	0.86 (0.80–0.91)[b]
ESRD	228	14,734	15.5	291	20,904	13.9	0.89 (0.74–1.06)
p for trend			<0.0001			<0.0001	
Septicemia death							
Normal renal function	1,100	3,249,626	0.37	1,033	3,916,082	0.26	0.75 (0.69–0.82)[b]
Non-dialysis CKD	307	272,621	1.13	308	324,339	0.95	0.83 (0.71–0.97)[a]
ESRD	73	14,734	4.95	71	20,904	3.40	0.69 (0.50–0.95)[a]
p for trend			<0.0001			<0.0001	

CKD, chronic kidney disease; ESRD, end-stage renal disease; PYs, person-year; aHR, adjusted hazard ratio and CI, confidence interval, estimated controlling for age, gender, and all comorbidities.

[a]$p < 0.05$

[b]$p < 0.001$

the lowest in the ESRD group, higher in the ST subcohort than in the non-ST subcohort in all three renal groups (**Figure 2**). The survival probability was higher for HS users than for LS users in NRF and CKD groups, but similar to LS users in the ESRD group.

Table 2 shows that ACM and the other three cause-specific mortality rates were substantially greater for patients with ESRD than for those with CKD and NRF in both statin and non-statin subcohorts. The ACM mortality in ESRD patients was 154.6 per 1,000 person-years in the non-ST subcohort and 104.6 per 1,000 person-years in the ST subcohort, with an aHR of 0.68 (95% CI = 0.64–0.72) for the ST users. The corresponding ACM rates in CKD patients were 56.4 versus 35.9 per 1,000 person-years, with an aHR of 0.64 (95% CI = 0.62–0.65) for the ST users. The corresponding ACM rates in the NRF group were 22.6 versus 13.9 per 1,000 person-years, with an aHR of 0.62 (95% CI = 0.60–0.63) for the ST users. The mortality rates of hyperlipidemic patients with cancer, heart disease, and septicemia were all lower for ST users than for non-ST users in the three renal groups. However, the ST users to non-users aHR of death from heart disease were not significant for ESRD patients.

The results from the time-dependent Cox regression analysis showed that the short-term impact of ST was also associated with the reduced aHRs of deaths, which is also significant for ACM and cancer deaths (**Supplementary Table S2**). The aHR of ACM for ST users with an ESRD of 40–59 years old was 0.33 (95% CI = 0.28–0.40), which was reduced slightly to 0.31 (95% CI = 0.26–0.37) for 70–80 years old. However, the short-term effect of ST on deaths from heart disease and septicemia in 40–59 years old with NRF was not significant. The treatment effects of ST were also not significant on deaths from heart disease for older patients with ESRD.

Figure 3 shows, in general, that the ST treatment effectiveness was superior in HS users than in LS users, with lower mortality rates from most types of death compared with nonusers. All aHRs were significant for deaths from all causes and from heart disease. For ACM in the ESRD group, the mortality rate for HS users was 51 per 1,000 person-years lower than that for non-statin patients, with an aHR of 0.68 (95% CI = 0.62–0.74). For patients with NRF or CKD, the hazards of deaths from heart disease were significantly reduced in HS users but not in LS users. The ST treatment effectiveness was superior in HS users than in LS users for reducing deaths from septicemia in the NRF and CKD groups.

DISCUSSION

Findings

Our data showed that hyperlipidemic dialysis patients receiving ST were significantly associated with reduced mortality hazards for 32% from ACM, for 37% from cancer, and for 31% from septicemia-related events, whereas the benefit of reducing deaths from heart disease was less obvious (**Table 2**). The dialysis patients receiving HS or LS were also associated with reduced risks of deaths from ACM and cancer. The beneficial effects on deaths from ACM were similar for using HS and LS, whereas the reduction effect on deaths from cancer was slightly greater for using HS than using LS (**Figure 2** and **Figure 3**). Our age-specific data showed that the ACM rate increased with age (**Supplementary Table S1**). However, statin users' to non-

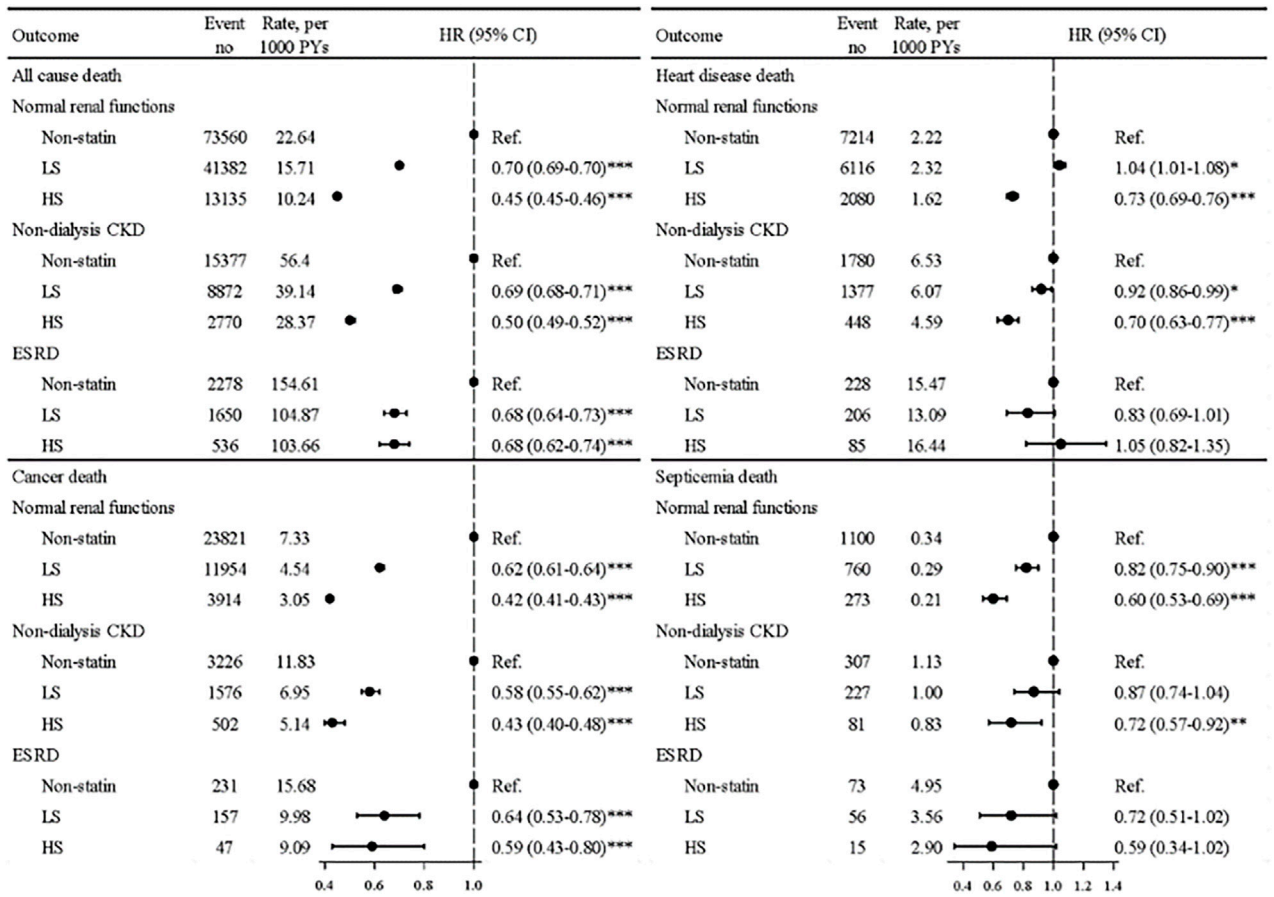

FIGURE. 3 | Mortality rates from all causes, cancer, heart disease, and septicemia and aHRs in patients using hydrophilic statins, lipophilic statins, and nonusers among three renal groups.

users' aHR of deaths from ACM and from other causes tended to decrease with age. These data suggested that the treatment effectiveness of statins might exert a greater impact in the death risk reduction for the elderly than for youngsters.

A simulated 4D trial of ST on diabetic ESRD patients also reported significant cardiovascular benefits for those with the therapy (Chan et al., 2010). This simulated observational trial may help in providing evidence to develop potential studies on treatment effectiveness. Several recent studies using real-world data have also reported the ST benefit in the associations with reduced ACM for patients undergoing HD (De Rango et al., 2016; Chung et al., 2017; Kim et al., 2018; Jung et al., 2020). An Italian study comparing the ACM risk between HD patients with ST (n = 127) and without ST (n = 232) found that ST reduced the risk for nearly 50%, 5 years after a vascular access operation (De Rango et al., 2016). A population-based study in Taiwan comparing 790 pairs of HD patients with and without ST after acute MI episodes found the ACM risk reduced for 24%–30% in ST users (Chung et al., 2017). A Korean population-based study found that ST was associated with 46% reduced risk of composite cardiovascular events, including ACM (Kim et al., 2018). Another Korean population-based observational study revealed that in adult HD patients, statin and ezetimibe combined treatment was associated with a halved risk of ACM (Jung et al., 2020). These reports support our research findings relating to ACM in dialysis patients.

The increasing noncardiovascular deaths in ESRD patients in recent years, especially in the elderly, have attracted attention (Den Hoedt et al., 2013; Vogelzang et al., 2015; Steenkamp et al., 2018). In Taiwan, infection and cancer are the major non-cardiovascular causes for the cost of admission healthcare for dialysis patients (Hsu, 2020). Studies have reported that statins possess anti-inflammatory capacities and ST was associated with improved survival in patients with Gram-positive or Gram-negative bacterial infections and sepsis (Ou et al., 2014; Mehl et al., 2015; Ouellette et al., 2015; Shrestha et al., 2016; Caffrey et al., 2017). A propensity-score matched data analysis found that the high-potency statin was associated with a lower risk of sepsis-related mortality than low-potency statin uses (Ou et al., 2014). A meta-analysis based on six cohort studies with 7,553 patients of bacterial infection without renal dysfunction showed that ST could reduce the mortality risk for nearly 30% (Shrestha et al., 2016), although, in a systemic study, Collins et al. have a different perspective and commented the absence of evidence that ST

provided clinical benefits in non-cardiovascular health outcomes such as cancer or infection (Collins et al., 2016).

However, recent studies provided more evidence of the clinical benefits of ST in noncardiovascular conditions. For cancer patients, evidence has revealed that adherence to ST is essential to achieve the beneficial effects (Lee et al., 2020; Feng and Qin, 2021; Okada et al., 2021). A Korean population-based retrospective cohort study found that the treatment effectiveness of ST relied on good statin adherence (Lee et al., 2020); poor adherence is associated with an increased risk of deaths from cancer or CVD, with the HR of nearly 1.3 compared to patients with good adherence. An Australian Cancer database study reported that good adherence to lipid-lowering medications was associated with reduced mortality from breast cancer, colorectal cancer, or melanoma (Feng and Qin, 2021). A Japanese study with 2,536 patients of type 2 diabetes reported that ST was associated with a 40% reduction in cancer mortality (Okada et al., 2021).

Statins are medicines effective in inhibiting the enzyme HMG-CoA reductase and the following mevalonate/isoprenoid pathways, exerting pleiotropic effects but widely used as lipid-lowering medications (Ahmadi et al., 2017). With anti-inflammatory effects, statins can inhibit the production of TNFa and IL-6 by mast cells, and activate the endothelial cells, exerting antioxidant effects (Nseir et al., 2010). With pleiotropic effects, the anti-inflammatory and antiproliferative properties of statin might partially explain the clinical benefits (Barbalata et al., 2020). Besides, prolific basic lab researches have provided strong evidences supporting the anti-cancer effects of statins, which may block diverse carcinogenic pathways (Bonsu et al., 2016). These might be the proposed mechanism of protection of statins for the reduction of all-cause mortality and non-CVD mortality.

A previous research comparing the mortality associated with HS and LS was conducted mainly in non-CKD patients with CVDs, with inconsistent findings. One meta-analysis reported that the LS medication is associated with reduced ACM and cardiovascular mortality in patients with congestive heart failure (Bonsu et al., 2016). A meta-analysis and a randomized trial reported the similar treatment effectiveness of the medications of HS and LS on major adverse cardiac events (Izawa et al., 2015; Bytyçi et al., 2017). Another PS-matched retrospective cohort study using the insurance data of Taiwan found a greater ACM risk reduction in patients receiving HS than in LS users for 35% (Chung et al., 2018), which is similar to our finding in the NRF group.

It is also noteworthy in our study that the treatment effectiveness of ST in reducing deaths from heart disease was not significant for the ESRD group, whereas in NRF and CKD patients, the ST effectiveness was significant by providing nearly a 30% risk reduction. These findings in non-ESRD group are agreeable to current literature (Palmer et al., 2014). In patients with ESRD, there is a complicated cardiovascular pathogenesis process, involving vascular calcification in both intimal and medial layers, sympathetic hyperactivity, and myocardial fibrosis, leading to a complicated cardiovascular scenario and poor response to a certain therapy that is helpful to non-dialysis groups (Santoro and Mandreoli, 2014; Laucyte-Cibulskiene et al., 2018).

Limitations

There are several limitations in our research. First, this is an observational study retrieving the information of patients from a secondary database. Due to the characteristics of the database, certain key information such as lab data and a personal lifestyle of smoking, dietary habits, and exercise was unavailable for adjustment in the data analysis. In order to reduce potential bias and the confounding effects of these factors, we addressed immortal time bias and conducted a time-variable model analysis. We used PS matching to establish study cohorts to reduce the personal variations and effects of comorbidities, even though residual confounding effects may be present since we did not adjust for lab data. Furthermore, the information on the exact patient medication compliances was unavailable for adjustment as well. Second, our research was conducted in a Taiwanese population-based setting; the generalizability to other ethnical cohorts or healthcare systems might be limited. Lastly, we used a death registry to retrieve the data of heart disease death as a proxy for cardiovascular death. Although the death registry provides highly accurate information about the causes of death, heart disease deaths included some deaths other than the scenario of major cardiovascular death, which might contribute a bias in the risk measurement.

Strength

There are strengths in our study. The sample sizes of the study cohorts were very large, allowing us to conduct a comprehensive stratification analysis and adjustment without jeopardizing the power of analysis. Our study investigated the effectiveness of ST on the association with mortality risks in three different renal groups. The mortality risk reductions in patients with NRF and CKD were consistent with current evidences (Palmer et al., 2014; Lin et al., 2020). The sample sizes of most previous studies were likely not large enough to include patients with ESRD for analysis. To the best of our knowledge, no previous study has ever evaluated simultaneously the risks of death not only from all causes but also from cancer, heart disease, and septicemia. Septicemia deaths are relatively rare compared to other causes. Our sample is large and enables us to observe the effectiveness of ST in reducing the mortality from septicemia, even for patients with ESRD. Previous studies rarely compared the treatment effectiveness between HS and LS for all these causes. Our study demonstrated that the effectiveness of HS might be superior than LS. However, our reports shed light on further studies to investigate the mechanisms leading to the variety of effectiveness of ST in patients with and without renal dysfunction.

The findings from the present study could derive clinical implications. First, the absence of a significant clinical benefit from ST in reducing the mortality from heart disease in the ESRD group indicates the complicated cardiovascular scenario in ESRD and potential harmful effects associated with ST. Both factors call for the need for novel and more promising lipid-lowering medications, such as proprotein convertase subtilisin/kexin Type 9 (PCSK9) inhibitors. We confirmed the efficacies and

indications of ST as an effective lipid-lowering medication for hyperlipidemic dialysis patients, preferably in the older population. Clinicians might need to have a thorough understanding on statins, especially in caring for dialysis patients with effective statin products. Second, the high need for an effective treatment for ESRD calls for more studies. However, the research with an RCT is under financial constraints. Observational studies help to provide clues for further evidence. Our findings may pave the way for investigating the evidence with which the relationship between statins and ESRD patients might be delineated.

CONCLUSION

Patients with ESRD receiving ST are at reduced risks for deaths from all causes, cancer, heart disease, and septicemia but not significant for deaths from the heart disease. The medication with HS is likely associated with better death risk reduction than the medication of LS. The age-specific analysis results suggest that the ST is appropriate for the elderly patients. Further large-scale prospective trials are needed to confirm our preliminary findings.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

This study was approved by the Research Ethics Committee at China Medical University and Hospital (CRREC-107-021) and Ditmanson Medical Foundation Chiayi Christian Hospital (CYCH-IRB-2019063). Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

F-CS and Y-CJ are authors contributed equally to this work. F-CS, C-HM and Y-HH had access to the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis. F-CS, Y-CJ, C-HM, and Y-HH initiated study concept and design. C-HM and Y-HH conducted statistical analysis. F-CS, Y-CJ, C-HM, and Y-HH acquired, analyzed, or interpreted data. F-CS, Y-CJ, C-HM, and Y-HH drafted manuscript. F-CS, C-CH, W-CT, and Y-HH critically revised manuscript. Y-HH obtained funding for Management Office for Health Data. F-CS, C-HM, C-CH, W-CT, and Y-HH provided administrative, technical, or material support. C-CH, W-CT, and Y-HH supervised the conduction. All authors read and approved the final manuscript.

FUNDING

This study is supported in part by Taiwan Ministry of Health and Welfare Clinical Trial Center (MOHW109-TDU-B-212-114004), MOST Clinical Trial Consortium for Stroke (MOST 109-2321-B-039-002), China Medical University Hospital (DMR-108-154), Tseng-Lien Lin Foundation of Taichung, Taiwan, and Ditmanson Medical Foundation Chiayi Christian Hospital (R-109-17).

ACKNOWLEDGMENTS

We are grateful to Health Data Science Center, China Medical University Hospital, and Ditmanson Medical Foundation Chiayi Christian Hospital for providing administrative, technical and funding support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2022.815882/full#supplementary-material

REFERENCES

Ahmadi, Y., Ghorbanihaghjo, A., and Argani, H. (2017). The Balance between Induction and Inhibition of Mevalonate Pathway Regulates Cancer Suppression by Statins: A Review of Molecular Mechanisms. *Chem. Biol. Interact* 273, 273–285. doi:10.1016/j.cbi.2017.06.026

Barbalata, C. I., Tefas, L. R., Achim, M., Tomuta, I., and Porfire, A. S. (2020). Statins in Risk-Reduction and Treatment of Cancer. *World J. Clin. Oncol.* 11, 573–588. doi:10.5306/wjco.v11.i8.573

Bonsu, K. O., Reidpath, D. D., and Kadirvelu, A. (2016). Lipophilic Statin versus Rosuvastatin (Hydrophilic) Treatment for Heart Failure: a Meta-Analysis and Adjusted Indirect Comparison of Randomised Trials. *Cardiovasc. Drugs Ther.* 30, 177–188. doi:10.1007/s10557-015-6636-z

Bytyçi, I., Bajraktari, G., Bhatt, D. L., Morgan, C. J., Ahmed, A., Aronow, W. S., and Banach, M. (2017). Lipid, and Blood Pressure Meta-Analysis CollaborationHydrophilic vs Lipophilic Statins in Coronary Artery Disease: A Meta-Analysis of Randomized Controlled Trials. *J. Clin. Lipidol.* 11, 624–637. doi:10.1016/j.jacl.2017.03.003

Caffrey, A. R., Timbrook, T. T., Noh, E., Sakoulas, G., Opal, S. M., Nizet, V., et al. (2017). Evidence to Support Continuation of Statin Therapy in Patients with *Staphylococcus aureus* Bacteremia. *Antimicrob. Agents Chemother.* 61. doi:10.1128/AAC.02228-16

Chan, K. E., Thadhani, R., Lazarus, J. M., and Hakim, R. M. (2010). Modeling the 4D Study: Statins and Cardiovascular Outcomes in Long-Term Hemodialysis Patients with Diabetes. *Clin. J. Am. Soc. Nephrol.* 5, 856–866. doi:10.2215/CJN.07161009

Chung, C. M., Lin, M. S., Chang, C. H., Cheng, H. W., Chang, S. T., Wang, P. C., et al. (2017). Moderate to High Intensity Statin in Dialysis Patients after Acute Myocardial Infarction: A National Cohort Study in Asia. *Atherosclerosis* 267, 158–166. doi:10.1016/j.atherosclerosis.2017.09.018

Chung, C. M., Lin, M. S., Liu, C. H., Lee, T. H., Chang, S. T., Yang, T. Y., et al. (2018). Discontinuing or Continuing Statin Following Intracerebral Hemorrhage from the View of a National Cohort Study. *Atherosclerosis* 278, 15–22. doi:10.1016/j.atherosclerosis.2018.08.049

Collins, R., Reith, C., Emberson, J., Armitage, J., Baigent, C., Blackwell, L., et al. (2016). Interpretation of the Evidence for the Efficacy and Safety of Statin Therapy. *Lancet* 388, 2532–2561. doi:10.1016/S0140-6736(16)31357-5

De Jager, D. J., Grootendorst, D. C., Jager, K. J., Van Dijk, P. C., Tomas, L. M., Ansell, D., et al. (2009). Cardiovascular and Noncardiovascular Mortality Among Patients Starting Dialysis. *JAMA* 302, 1782–1789. doi:10.1001/jama.2009.1488

De Rango, P., Parente, B., Farchioni, L., Cieri, E., Fiorucci, B., Pelliccia, S., et al. (2016). Effect of Statins on Survival in Patients Undergoing Dialysis Access for End-Stage Renal Disease. *Semin. Vasc. Surg.* 29, 198–205. doi:10.1053/j.semvascsurg.2017.03.001

Den Hoedt, C. H., Bots, M. L., Grooteman, M. P., Mazairac, A. H., Penne, E. L., Van Der Weerd, N. C., et al. (2013). Should We Still Focus that Much on Cardiovascular Mortality in End Stage Renal Disease Patients? the CONvective TRAnsport STudy. *PLoS One* 8, e61155. doi:10.1371/journal.pone.0061155

Fellström, B. C., Jardine, A. G., Schmieder, R. E., Holdaas, H., Bannister, K., Beutler, J., et al. (2009). Rosuvastatin and Cardiovascular Events in Patients Undergoing Hemodialysis. *N. Engl. J. Med.* 360, 1395–1407. doi:10.1056/NEJMoa0810177

Feng, J. L., and Qin, X. (2021). Does Adherence to Lipid-Lowering Medications Improve Cancer Survival? A Nationwide Study of Breast and Colorectal Cancer, and Melanoma. *Br. J. Clin. Pharmacol.* 87, 1847–1858. doi:10.1111/bcp.14573

Holdaas, H., Holme, I., Schmieder, R. E., Jardine, A. G., Zannad, F., Norby, G. E., et al. (2011). Rosuvastatin in Diabetic Hemodialysis Patients. *J. Am. Soc. Nephrol.* 22, 1335–1341. doi:10.1681/ASN.2010090987

Hsu, C. C. (2020). 2019 Annual Report on Kidney Disease in Taiwan. Available at: https://www.tsn.org.tw/UI/L/L002.aspx: Taiwan Society of Nephrology.

Izawa, A., Kashima, Y., Miura, T., Ebisawa, S., Kitabayashi, H., Yamamoto, H., et al. (2015). Assessment of Lipophilic vs. Hydrophilic Statin Therapy in Acute Myocardial Infarction – ALPS-AMI Study. *Circ. J.* 79, 161–168. doi:10.1253/circj.CJ-14-0877

Jung, J., Bae, G. H., Kang, M., Kim, S. W., and Lee, D. H. (2020). Statins and All-Cause Mortality in Patients Undergoing Hemodialysis. *J. Am. Heart Assoc.* 9, e014840. doi:10.1161/JAHA.119.014840

Kim, S. H., Jeong, H. Y., Yang, D. H., Kim, J., and Lee, S. Y. (2018). Beneficial Effect of Statins in Patients Receiving Chronic Hemodialysis Following Percutaneous Coronary Intervention: A Nationwide Retrospective Cohort Study. *Sci. Rep.* 8, 9692. doi:10.1038/s41598-018-27941-w

Laucyte-Cibulskiene, A., Petraviciute, M., Gudynaite, M., Gumbys, L., Valanciene, D., Galiauskiene, K., et al. (2018). Mismatch between Stiffness in Elastic and Muscular Arteries as a Predictor of Vascular Calcification in Dialysis Patients. *Aging Clin. Exp. Res.* 30, 375–382. doi:10.1007/s40520-017-0787-7

Lee, Y. R., Oh, S. S., Jang, S. I., and Park, E. C. (2020). Statin Adherence and Risk of All-Cause, Cancer, and Cardiovascular Mortality Among Dyslipidemia Patients: A Time-dependent Analysis. *Nutr. Metab. Cardiovasc. Dis.* 30, 2207–2214. doi:10.1016/j.numecd.2020.07.024

Lin, Y. C., Lai, T. S., Wu, H. Y., Chou, Y. H., Chiang, W. C., Lin, S. L., et al. (2020). Effects and Safety of Statin and Ezetimibe Combination Therapy in Patients with Chronic Kidney Disease: A Systematic Review and Meta-Analysis. *Clin. Pharmacol. Ther.* 108, 833–843. doi:10.1002/cpt.1859

März, W., Genser, B., Drechsler, C., Krane, V., Grammer, T. B., Ritz, E., et al. (2011). Atorvastatin and Low-Density Lipoprotein Cholesterol in Type 2 Diabetes Mellitus Patients on Hemodialysis. *Clin. J. Am. Soc. Nephrol.* 6, 1316–1325. doi:10.2215/CJN.09121010

Mehl, A., Harthug, S., Lydersen, S., Paulsen, J., Åsvold, B. O., Solligård, E., et al. (2015). Prior Statin Use and 90-day Mortality in Gram-Negative and Gram-Positive Bloodstream Infection: a Prospective Observational Study. *Eur. J. Clin. Microbiol. Infect. Dis.* 34, 609–617. doi:10.1007/s10096-014-2269-6

Nseir, W., Khateeb, J., Tatour, I., Haiek, S., Samara, M., and Assy, N. (2010). Long-term Statin Therapy Affects the Severity of Chronic Gastritis. *Helicobacter* 15, 510–515. doi:10.1111/j.1523-5378.2010.00803.x

Okada, S., Morimoto, T., Ogawa, H., Soejima, H., Matsumoto, C., Sakuma, M., et al. (2021). Association between Statins and Cancer Incidence in Diabetes: a Cohort Study of Japanese Patients with Type 2 Diabetes. *J. Gen. Intern. Med.* 36, 632–639. doi:10.1007/s11606-020-06167-5

Ota, Y., Kitamura, M., Muta, K., Yamashita, H., Uramatsu, T., Obata, Y., et al. (2019). Effect of Statin on Life Prognosis in Japanese Patients Undergoing Hemodialysis. *PLoS One* 14, e0224111. doi:10.1371/journal.pone.0224111

Ou, S. Y., Chu, H., Chao, P. W., Ou, S. M., Lee, Y. J., Kuo, S. C., et al. (2014). Effect of the Use of Low and High Potency Statins and Sepsis Outcomes. *Intensive Care Med.* 40, 1509–1517. doi:10.1007/s00134-014-3418-1

Ouellette, D. R., Moscoso, E. E., Corrales, J. P., and Peters, M. (2015). Sepsis Outcomes in Patients Receiving Statins Prior to Hospitalization for Sepsis: Comparison of In-Hospital Mortality Rates between Patients Who Received Atorvastatin and Those Who Received Simvastatin. *Ann. Intensive Care* 5, 9. doi:10.1186/s13613-015-0049-9

Palmer, S. C., Navaneethan, S. D., Craig, J. C., Johnson, D. W., Perkovic, V., Hegbrant, J., et al. (2014). HMG CoA Reductase Inhibitors (Statins) for People with Chronic Kidney Disease Not Requiring Dialysis. *Cochrane Database Syst Rev* 31 (5), CD007784. doi:10.1002/14651858.CD007784.pub2

Palmer, S. C., Navaneethan, S. D., Craig, J. C., Johnson, D. W., Perkovic, V., Nigwekar, S. U., et al. (2013). HMG CoA Reductase Inhibitors (Statins) for Dialysis Patients. *Cochrane Database Syst Rev* 11 (9), CD004289. doi:10.1002/14651858.CD004289.pub5

Santoro, A., and Mandreoli, M. (2014). Chronic Renal Disease and Risk of Cardiovascular Morbidity-Mortality. *Kidney Blood Press. Res.* 39, 142–146. doi:10.1159/000355789

Shrestha, P., Poudel, D. R., Pathak, R., Ghimire, S., Dhital, R., Aryal, M. R., et al. (2016). Effect of Statins on the Mortality of Bacteremic Patients: A Systematic Review and Meta-Analysis of Clinical Trials. *N. Am. J. Med. Sci.* 8, 250–251. doi:10.4103/1947-2714.185034

Steenkamp, R., Pyart, R., and Fraser, S. (2018). Chapter 5 Survival and Cause of Death in UK Adult Patients on Renal Replacement Therapy in 2016. *Nephron* 139 (Suppl. 1), 117–150. doi:10.1159/000490963

Vogelzang, J. L., Van Stralen, K. J., Noordzij, M., Diez, J. A., Carrero, J. J., Couchoud, C., et al. (2015). Mortality from Infections and Malignancies in Patients Treated with Renal Replacement Therapy: Data from the ERA-EDTA Registry. *Nephrol. Dial. Transpl.* 30, 1028–1037. doi:10.1093/ndt/gfv007

Wakasugi, M., Kazama, J. J., Yamamoto, S., Kawamura, K., and Narita, I. (2013). Cause-specific Excess Mortality Among Dialysis Patients: Comparison with the General Population in Japan. *Ther. Apher. Dial.* 17, 298–304. doi:10.1111/j.1744-9987.2012.01144.x

Wanner, C., März, W., Marz, J. F., Olschewski, E., Mann, J. F., Ruf, G., et al. (2005). Atorvastatin in Patients with Type 2 Diabetes Mellitus Undergoing Hemodialysis. *N. Engl. J. Med.* 353, 238–248. doi:10.1056/NEJMoa043545

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Sung, Jong, Muo, Hsu, Tsai and Hsu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Pharmacology

SYSTEMATIC REVIEW
published: 16 March 2022
doi: 10.3389/fphar.2022.862640

The Relationship Between Short-Term Surrogate Endpoint Indicators and mPFS and mOS in Clinical Trials of Malignant Tumors: A Case Study of Approved Molecular Targeted Drugs for Non-Small-Cell Lung Cancer in China

*Mingjun Rui[1,2], Zijing Wang[1,2], Zhengyang Fei[1,2], Yao Wu[1,2], Yingcheng Wang[1,2], Lei Sun[1], Ye Shang[1,2] and Hongchao Li[1,2]**

[1]School of International Pharmaceutical Business, China Pharmaceutical University, Nanjing, China, [2]Center for Pharmacoeconomics and Outcomes Research, China Pharmaceutical University, Nanjing, China

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Ni Yuan,
Dalian Medical University, China
Yujia Wei,
Mayo Clinic, United States

***Correspondence:**
Hongchao Li
lihongchao@cpu.edu.cn

Specialty section:
This article was submitted to Drugs Outcomes Research and Policies, a section of the journal Frontiers in Pharmacology

Received: 26 January 2022
Accepted: 28 February 2022
Published: 16 March 2022

Citation:
Rui M, Wang Z, Fei Z, Wu Y, Wang Y, Sun L, Shang Y and Li H (2022) The Relationship Between Short-Term Surrogate Endpoint Indicators and mPFS and mOS in Clinical Trials of Malignant Tumors: A Case Study of Approved Molecular Targeted Drugs for Non-Small-Cell Lung Cancer in China. Front. Pharmacol. 13:862640. doi: 10.3389/fphar.2022.862640

Objective: Due to the initiation of the priority review program in China, many antitumor drugs have been approved for marketing based on phase II clinical trials and short-term surrogate endpoint indicators. This study used approved targeted drugs for the treatment of non-small-cell lung cancer (NSCLC) in China as an example to evaluate the association between short-term surrogate endpoints [objective response rate (ORR) and disease control rate (DCR)] and median progression-free survival (mPFS) and median overall survival (mOS).

Methods: Five databases, i.e., MEDLINE, Embase, Cochrane Library, China National Knowledge Infrastructure (CNKI), and Wanfang Data were searched, for phase II or phase III clinical trials of all molecular targeted drugs that have been marketed in China for the treatment of NSCLC. After screening the literature and extracting information, both univariate and multivariate linear regression were performed on the short-term surrogate indicators and mPFS and mOS to explore the relationship.

Results: A total of 63 studies were included (25 studies with only ORR, DCR, and mPFS and 39 studies with ORR, DCR, mPFS, and mOS). In terms of the targeted drugs for the treatment of NSCLC, in addition to the good but not excellent linear relationship between DCR and mOS ($0.4 < R^2_{adj} = 0.5653 < 0.6$), all other short-term surrogate endpoint indicators had excellent linear relationships with mPFS and mOS ($R^2_{adj} \geq 0.6$), while mPFS and mOS had the most excellent linear relationships ($R^2_{adj} = 0.8036$).

Conclusion: For targeted drugs for the treatment of NSCLC, short-term surrogate endpoint indicators such as ORR and DCR may be reliable surrogate indicators for mPFS and mOS. However, whether short-term surrogate endpoint indicators can be used to predict final endpoints remains to be verified.

Keywords: malignant tumor, short-term surrogate endpoint, median progression-free survival, median overall survival, non-small-cell lung cancer

INTRODUCTION

Malignant tumors are a high-risk factor for death and severely hinder increases in the average life expectancy of the population (1). They are the leading cause of death in the urban population. In 2019, approximately 25.73% of urban population deaths in China were caused by malignant tumors, with a mortality rate of approximately 161.56/100,000 people (2). In 2018, there were 3.804 million new cases of malignant tumors in China, accounting for more than 20% of the global cases. The incidence of malignant tumors was 278.07 per 100,000 people, and the mortality rate was 167.89 per 100,000 people (Ma and Yu, 2020). Malignant tumors seriously threaten the lives and health of people. From the perspective of disease burden, malignant tumors have caused a substantial loss of disability-adjusted life years (DALYs). Studies have shown that (3) the proportion of DALYs caused by trachea, bronchus, and lung cancers was 4.1% of the total DALYs, ranking fourth only after stroke (11.9%), ischemic heart disease (8.1%), and chronic obstructive pulmonary disease (5.5%). From the economic burden perspective, the average medical costs for malignant tumor patients are increasing year by year. In 2005, the average cost of a single hospitalization for discharged patients in China was 10,777 yuan (RMB), increasing to 13,322 yuan in 2011, 15,672 yuan in 2013, and 17,567 yuan in 2016 (Wei-jing and xiao-lu, 2019).

To increase patients' accessibility to new drugs and to improve the quality of life, the National Medical Products Administration (NMPA) in China launched a priority review program to allow more innovative drugs to be approved as soon as possible to bring patients with malignant tumors benefits. The NMPA priority review processes mainly include three policies: one review process for breakthrough therapeutic drugs, one review process for the conditional approval of drugs for marketing, and one priority review process for drug marketing authorization (Adminstration, 2020). The priority review program greatly shortens the time to market for some new anticancer drugs which often focus on the rare targets, and many of them do not have abundant clinical data based on Chinese patients. Many of these clinical studies are often single-arm with a small sample size and short follow-ups, and even primary endpoint indicators such as progression-free survival (PFS) and overall survival (OS) were not reported.

Table 1 summarized the reported status of clinical trial indicators for anticancer drugs approved in China from 2017 to November 2021. An increasing number of drugs were approved using only short-term surrogate endpoint indicators. Among them, only 16 new drugs reported both PFS and OS data. However, the lack of primary endpoint indicators causes challenges in reliably determining the safety and efficacy of anticancer drugs and, likewise, poses a significant challenge for economic evaluations. In the economic evaluation of anticancer drugs, the partitioned survival model (PSM) and the Markov model are most popular model types (Rui et al., 2021). The construction of both the PSM and the Markov model requires the support of mature PFS and OS data (6). Therefore, when only short-term surrogate endpoint indicators available, it is worth investigating whether there is a significant relationship between such indicators and primary endpoint indicators.

To explore this issue, this study will use clinical trials related to targeted therapy for the treatment of NSCLC approved in China as an example to explore the correlation between short-term surrogate endpoint indicators and primary endpoint indicators.

MATERIALS AND METHODS

Literature Search Strategy

Chinese and English databases and platforms were searched for Phase II or phase III clinical trials of molecular targeted drugs for the treatment of NSCLC. The Chinese search included China National Knowledge Infrastructure (CNKI) and Wanfang Data. The English search included MEDLINE through the PubMed search platform, Embase, and the Cochrane Library. The date ranges for the searches were from the establishment of the databases to 20 March 2021. The searches were based on a combination of subject headings and free-text. Chinese search terms and English search terms included non-small-cell lung cancer and clinical trial, among other terms. The English search strategy was shown in the supplementary materials.

Inclusion and Exclusion Criteria

The inclusion criteria for this study were as follows: (Bray et al., 2021) phase II or phase III clinical trials, including single-arm clinical trials and placebo-controlled clinical trials (Ma and Yu, 2020); patients diagnosed with NSCLC by laboratory tests, imaging examinations and clinical signs and symptoms (Wei-jing and xiao-lu, 2019); intervention measures that included molecular targeted drugs for the treatment of NSCLC approved for marketing in China as of March 2021, including gefitinib, erlotinib, icotinib, crizotinib, dacomitinib, afatinib, osimertinib, almonertinib, alectinib, ceritinib, brigatinib, lorlatinib, selpercatinib, entrectinib, dabrafenib + trametinib, erlotinib + linsitinib, erlotinib + pazotinib, erlotinib + sorafenib, and anlotinib (Adminstration, 20202020); short-term surrogate endpoint indicators included ORR or DCR; and (Rui et al., 2021) primary endpoint indicators included median progression-free survival (mPFS) and median overall survival (mOS). The exclusion criteria for this study were as follows: (Bray et al., 2021): duplicate literature (Ma and Yu, 2020); non-Chinese or non-English literature (Wei-jing and xiao-lu, 2019); conference abstract (Adminstration, 20202020); trials other than phase II or phase III clinical trials (Rui et al., 2021); no simultaneous reporting of DCR, ORR, and mPFS; and (Coyle and Coyle, 2014) intervention measures that included molecular targeted drugs combined with other types of therapeutic measures.

Literature Screening and Data Extraction

Two researchers independently screened the literature, extracted the data, and cross-checked the data. Disagreements were resolved through consultation with a third party. Data extraction mainly included ① basic characteristics of the included studies (title, authors, year, etc.); ② sample size of each group; ③ treatment measures and their usage and dosage; ④ key elements of bias risk assessments; and ⑤ endpoint indicators (ORR, DCR, mPFS and mOS).

TABLE 1 | Summary of clinical endpoints of new anti-cancer drugs approved from January 2017 to November 2021.

Drug	Approved year	Disease	mPFS (month)	mOS (month)	ORR	mDOR (month)	DCR
PFS and OS							
Vemurafenib	2017	Melanoma	8.3	13.5	52%	—	46%
Regorafenib	2017	mCRC GIST	1.9	6.4	1%	—	41%
Bevacizumab	2017	mCRC NSCLC	4.2	9.3	41.20%	8.1	—
Anlotinib	2018	NSCLC	5.37	9.46	9.18%	—	80.95%
Pembrolizumab	2018	Melanoma	2.8	12.1	16.70%	8.4	38.20%
Lenvatinib	2018	HCC	7.4	13.6	24.10%	—	—
Fruquintinib	2018	mCRC	3.7	9.3	4.70%	5.6	62.20%
Bendamustine	2018	Lymphoma	18.6	74%	—	16.5	—
Eribulin	2019	Breast Cancer	2.8	13.4	30.70%	—	—
FTD/TPI	2019	mCRC	2	7.8	—	—	—
Pralatrexate	2020	T cell Lymphoma	3.6	14.5	39%	10.1	—
Atezolizumab	2020	SCLC	5.2	12.3	60.20%	4.2	—
Savolitinib	2021	NSCLC	6.8	12.5	42.90%	8.3	82.90%
Utidelone	2021	Breast Cancer	8.44	16.13	40.40%	7.59	53.90%
Donafenib	2021	HCC	12.1	3.7	4.60%	—	30.8%
Carfilzomib	2021	Myeloma	5.6	16.6	35.80%	—	—
PFS and ORR/DCR/DOR							
Afatinib	2017	NSCLC	11.01	—	67.80%	9.72	92.60%
Osimertinib	2017	NSCLC	9.7	—	62.70%	9.9	88.60%
Ibrutinib	2017	Lymphoma	13.9	—	67.60%	14.9	—
Pazopanib	2017	RCC	11.1	—	30%	—	—
Erlotinib	2017	NSCLC	10.4	—	65.10%	—	—
Ixazomib	2018	Myeloma	19.6	—	78.30%	—	—
Pyrotinib	2018	Breast Cancer	18.1	—	78.50%	16.7	—
Alectinib	2018	NSCLC	34.8	—	83%	—	—
Olaparib	2018	Ovarian Cancer/PPC	19.1	—	19.40%	—	81.00%
Toripalimab	2018	Melanoma	3.6	—	17.30%	—	57.50%
Sintilimab	2018	HL	15.4	—	84%	—	100%
Dacomitinib	2019	NSCLC	14.7	—	75%	14.8	—
Denosumab	2019	GCT	—	—	25%	—	—
Camrelizumab	2019	HL	11.3	—	80.30%	—	98.50%
Almonertinib	2020	NSCLC	12.3	—	68.90%	12.4	93.40%
Surufatinib	2020	NEN	9.2	—	10.30%	5.6	86.50%
Inetetamab	2020	Breast Cancer	9.12	—	46.70%	—	79.72%
Ensartinib	2020	NSCLC	11.2	—	52.60%	—	87.80%
Ripretinib	2021	GIST	6.3	—	11.80%	—	—
Furmonertinib	2021	NSCLC	7.6	—	73.60%	—	—
Donafenib	2021	HCC	12.1	3.7	4.60%	—	30.8%
Carfilzomib	2021	Myeloma	5.6	16.6	35.80%	—	—
ORR/mDOR/DCR							
Denosumab	2019	GCT	—	—	25%	—	—
Tislelizumab	2019	HL	—	—	76.90%	—	90.80%
Rituximab	2019	Lymphoma	—	—	94.10%	—	—
Zanubrutinib	2020	T cell Lymphoma	—	—	83.70%	19.5	—
Orelabrutinib	2020	Leukemia/SLL	—	—	73.80%	—	—
Neratinib	2020	Breast Cancer	—	—	32.80%	—	—
Fluzoparib	2020	PC	—	—	64.10%	—	—
Venetoclax	2020	AML	—	—	—	5.5	—
Pamiparib	2021	Ovarian Cancer/PFTC/PPC	—	—	68.30%	13.8	—
Avapritinib	2021	GIST	—	—	62.50%	—	—
Pralsetinib	2021	NSCLC	—	—	65%	—	93%
Daratumumab	2021	Myeloma	—	—	83%	—	34.3%
Penpulimab	2021	HL	—	—	84.70%	—	94.10%
Zimberelimab	2021	HL	—	—	90.48%	—	96.43%

PFS, progression-free survival; OS, overall survival; DOR, duration of response; mPFS, median progression-free survival; mOS, median overall survival; ORR, overall response rate; mDOR, median duration of response; DCR, disease control rate;/: not erported; mCRC, metastatic colorectal cancer; GIST, gastrointestinal stromal tumor; NSCLC, non-small cell lung cancer; HCC, hepatocellular carcinoma; SCLC, small cell lung cancer; RCC, renal cell carcinoma; PPC, primary peritoneal carcinoma; HL, hodgkin lymphoma; GCT, giant cell tumor of bone; NEN, neuroendocrine neoplasm; SLL, small lymphocytic lymphoma; PC, peritoneal carcinomatosis; AML, acute myeloid leukemia; PFTC, primary fallopian tube carcinoma; FTD/TPI, trifluridine/tipiracil.

Quality of the Included Studies

Two investigators independently conducted quality evaluations of the included studies and cross-checked the results. For randomized controlled clinical trials (RCTs), the quality of the included studies was evaluated using the risk of bias assessment tool for RCTs recommended by the Cochrane Manual (Higgins et al., 2011). The Newcastle–Ottawa scale (NOS), recommended by the Cochrane Non-Randomized Studies Methods Group (NRSMG), was used to evaluate the quality of single-arm clinical trials (Margulis et al., 2014).

Data Processing

This study used STATA 15.1 to perform both univariate and multivariate linear regression analysis of the relationship between DCR and ORR and mPFS, as well as the relationship between DCR and ORR and mOS. In the case of a poor linear relationship between the shot-term surrogate endpoint indicators and the primary endpoint indicators, ln transformation was performed on the short-term surrogate endpoint indicators to explore the linear relationship between ln (short-term surrogate endpoint indicators) and the primary endpoint indicators. For the different dosage, medication or duration included in the analysis, the treatments were categorized for inclusion in multivariate regression analysis. In addition, some studies showed that the OS is largely affected by the number of previous treatment lines, which means that patients received more lines of treatments often have a worse prognosis (Gisselbrecht et al., 2010; Rule et al., 2017). Therefore, the subgroup analyses were performed for first-line treatment and second-line or post-second-line treatment based on the number of treatment lines in the univariate linear regression analysis to separate patients with different treatment lines to reduce heterogeneity. The scatter plots for DCR and ORR vs. mPFS and mOS were plotted using Microsoft Excel. Adjusted goodness-of-fit R^2_{adj} was used to evaluate the degree of fit of the model. According to Lassere et al. (Lassere et al., 2012), $R^2_{adj}\geq0.6$ indicates excellent goodness-of-fit, $R^2_{adj}\geq0.4$ indicates good goodness-of-fit, $R^2_{adj}\geq0.2$ indicates fair goodness-of-fit, and $R^2_{adj}<0.2$ indicates poor goodness-of-fit.

RESULTS

Literature Screening Results

A total of 5,058 articles were obtained in the preliminary searches, and a total of 4,547 articles were included in the preliminary screening after excluding duplicates. After reading the titles and abstracts, 4,019 papers were excluded, and 528 papers were included in the full-text rescreening. After reading the full text of the 528 papers, 63 articles were included in the final sample for the quantitative analysis of DCR, ORR, mPFS and mOS. The literature screening process is shown in **Figure 1**.

Results of the Data Extraction From the Included Studies

Among the 63 included studies, 25 (Inoue et al., 2006; Wen et al., 2007; Kim et al., 2012; Yamada et al., 2013; Hu et al., 2015; Yoshimura et al., 2015; Park et al., 2016; Yang et al., 2017a; Leighl et al., 2017; Peters et al., 2017; Han et al., 2018a; Akamatsu et al., 2018; Wu et al., 2018a; Nie et al., 2018; Yang et al., 2018; Zhou et al., 2018; Deng et al., 2019; Landi et al., 2019; Michels et al., 2019; Ohe et al., 2019; Saito et al., 2019; Yang et al., 2019; Yokoyama et al., 2019; Yang et al., 2020; de Marinis et al., 2021) only reported ORR, DCR, and mPFS, and 39 (Guan et al., 2005; Zhou et al., 2009; Gaafar et al., 2011; Lee et al., 2011; Spigel et al., 2011; Ciuleanu et al., 2012; Deng et al., 2012; Miller et al., 2012; Pallis et al., 2012; Pérol et al., 2012; Ramalingam et al., 2012; Shi et al., 2013; Wu et al., 2013; Ramalingam et al., 2014; Choi et al., 2015; Juan et al., 2015; Wu et al., 2015; An et al., 2016; Ding-guo and Xiu-li, 2016; Neal et al., 2016; Yang et al., 2017b; Han et al., 2017; Ikezawa et al., 2017; Miyawaki et al., 2017; Han et al., 2018b; Wu et al., 2018b; Garon et al., 2018; Kiura et al., 2018; Spigel et al., 2018; Uchibori et al., 2018; Arrieta et al., 2019; Landi et al., 2019; Xu et al., 2019; Ahn et al., 2020; Eide et al., 2020; Gadgeel et al., 2020; Goldman et al., 2020; Huber et al., 2020; Scagliotti et al., 2020) reported ORR, DCR, mPFS, and mOS concurrently. One paper (19) grouped the patients for whom ORR, DCR, and mPFS were reported and the patients for whom ORR, DCR, mPFS and mOS were reported concurrently. Basic information and the ORR, DCR, mPFS and mOS values reported in the included studies are provided in **Tables 2**, **3**.

The publication dates for all the included articles were concentrated from 2005 to 2021, and the target population was patients with advanced NSCLC. Among the studies, 21 (33.3%) were single-arm clinical trials, 42 (66.7%) were double-arm or multi-arm clinical trials, 34 (54.0%) enrolled patients in the first-line treatment stage, and 29 (46.0%) enrolled patients in the late-line or multi-line treatment stage. The sample sizes ranged from 11 to 479. The targeted therapies included 15 targeted drugs (avapritinib, afatinib, icotinib, alectinib. Anlotinib, osimertinib, brigatinib, dacomitinib, erlotinib, ensatinib, gefitinib, crizotinib, linsitinib, ceritinib, and sorafenib). DCRs ranged from 14.9% to 100.0%; ORRs values ranged from 5.3% to 87.0%; mPFS ranged from 1.5 to 20.0 months; and mOS ranged from 3.2 to 34.0 months.

Evaluation of the Quality of the Included Studies

Risk of bias in RCTs: The results of the risk of bias analyses for 42 two-arm or multi-arm RCTs were provided in **Figure 2**. "Selective reporting," "incomplete outcome data" and "random sequence generation" had a low risk of bias, and "blinding of outcome assessment" and "blinding of participants and personnel" had a high risk of bias. The risk of bias results for "allocation concealment" and "other bias" were not clear.

Quality evaluation results for single-arm trials: The NOS scores for the 21 included single-arm clinical trials were shown in **Figure 3**. The NOS scores for all studies ranged from 4 to 6, with an average score of 5.6, indicating that the overall quality of the studies was high. Among them, the NOS score for one paper was four points, the NOS score for six papers was five points, and the NOS score for the remaining 14 papers was six points.

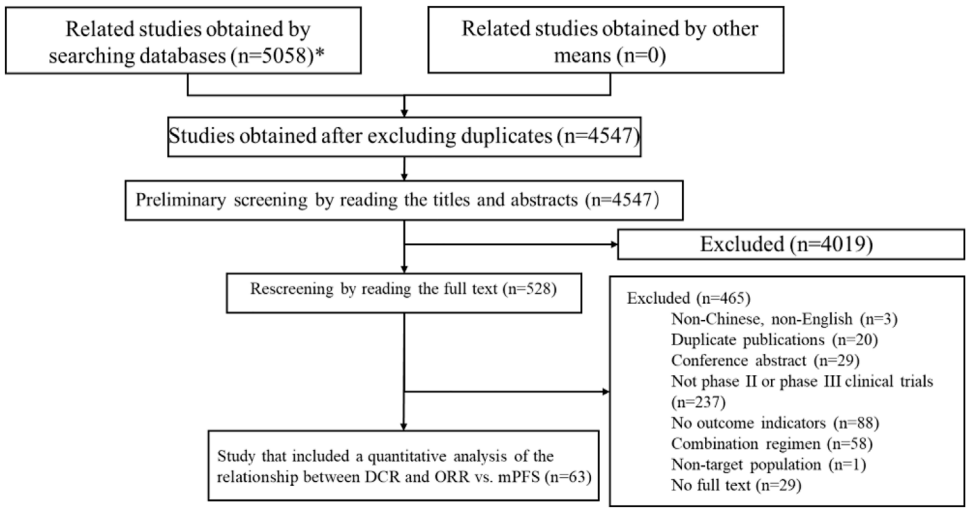

FIGURE 1 | Flowchart of literature screening.

RELATIONSHIP BETWEEN SHORT-TERM SURROGATE ENDPOINT INDICATORS AND PRIMARY ENDPOINT INDICATORS

Analysis of ORR and mPFS

Taking the natural logarithm of mPFS, the adjusted goodness-of-fit of the univariate regression between ORR and ln (mPFS) was excellent (R^2_{adj} = 0.7356 > 0.6), which was shown in **Figure 4** and **Supplementary Table S4**. After controlling the treatment factors, the adjusted goodness-of-fit of the multivariate regression between ORR and ln (mPFS) was excellent (R^2_{adj} = 0.7772 > 0.6), which was shown in the **Supplementary Table S5**.

Analysis of DCR and mPFS

Taking the natural logarithm of mPFS, the adjusted goodness-of-fit of the univariate regression between DCR and ln (mPFS) was excellent (R^2_{adj} = 0.7642 > 0.6), which was shown in **Figure 5** and **Supplementary Table S6**. After controlling the treatment factors, the adjusted goodness-of-fit of the multivariate regression between DCR and ln (mPFS) was excellent (R^2_{adj} = 0.7806 > 0.6), which was shown in the **Supplementary Table S7**.

Analysis of ORR and mOS

The adjusted goodness-of-fit of the univariate regression between ORR and mOS was excellent (R^2_{adj} = 0.7633 > 0.6), which was shown in **Figure 6** and **Supplementary Table S8**. After controlling the treatment factors, the adjusted goodness-of-fit of the multivariate regression between ORR and mOS was excellent (R^2_{adj} = 0.7813 > 0.6), which was shown in the **Supplementary Table S9**.

Analysis of DCR and mOS

Taking the natural logarithm of mOS, the adjusted goodness-of-fit of the univariate regression between DCR and ln (mOS) was good (R^2_{adj} = 0.5653 > 0.4), which was shown in **Figure 7** and **Supplementary Table S10**. After controlling the treatment factors, the adjusted goodness-of-fit of the multivariate regression between DCR and ln (mOS) was excellent (R^2_{adj} = 0.6331 > 0.6), which was shown in the **Supplementary Table S11**.

Analysis Results of mPFS and mOS

The adjusted goodness-of-fit of the univariate regression between mPFS and mOS was excellent (R^2_{adj} = 0.7616 > 0.6), which was shown in **Figure 8** and **Supplementary Table S12**. After controlling the treatment factors, the adjusted goodness-of-fit of the multivariate regression between mPFS and mOS was excellent (R^2_{adj} = 0.8036 > 0.6), which was shown in the **Supplementary Table S13**.

SUBGROUP ANALYSIS

Results for First-Line Treatment Only

Taking the natural logarithm of mPFS, the adjusted goodness-of-fit of the univariate regression between ORR and ln (mPFS) was excellent (R^2_{adj} = 0.6188 > 0.6), which was shown in **Figure 9**.

Taking the natural logarithm of mPFS, the adjusted goodness-of-fit of the univariate regression between DCR and ln (mPFS) was excellent (R^2_{adj} = 0.7128 > 0.6), which was shown in **Figure 10**.

The adjusted goodness-of-fit of the univariate regression between ORR and mOS was excellent (R^2_{adj} = 0.7074 > 0.6), which was shown in **Figure 11**.

The adjusted goodness-of-fit of the univariate regression between mPFS and mOS was excellent (R^2_{adj} = 0.7764 > 0.6), which was shown in **Figure 12**.

Results of Second-Line or Post-Second-Line Treatment Only

Taking the natural logarithm of mPFS, the adjusted goodness-of-fit of the univariate regression between ORR and ln (mPFS) was excellent (R^2_{adj} = 0.6926 > 0.6), which was shown in **Figure 13**.

TABLE 2 | Overview of general characteristics of studies reporting ORR, DCR, and mPFS.

References	Population	Brain metastasis	Line	Arm	Sample size	Treatment	Dosage	DCR (%)	ORR (%)	mPFS (month)
de Marinis et al. (2021)	Locally advanced or metastatic EGFR mutation-positive NSCLC	No	First	Single	479	AfatiTnib	40 mg/qd	85.80	49.20	13.40
Yang et al. (2019)	Locally advanced or metastatic NSCLC	No	1st	Double	17	Osimertinib	80 mg/qd	100.00	80.00	19.30
Ohe et al (2019)	Advanced or metastatic NSCLC	No	First	Double	65	Osimertinib	80 mg/qd	96.90	75.40	19.10
Ohe et al (2019)	Advanced or metastatic NSCLC	No	First	Double	55	Gefitinib	250 mg/qd	96.40	76.40	13.80
Deng et al. (2019)	Advanced or metastatic NSCLC	Yes	First + Second	Single	47	Crizotinib	-	93.60	61.70	19.00
Yokoyama et al. (2019)	Advanced NSCLC	No	First	Single	46	Afatinib	20mg/qd	93.20	81.80	15.20
Michels et al. (2019)	Metastatic ROS1-rearranged NSCLC	No	First	Single	30	Crizotinib	250 md/bid	83.30	73.00	20.00
Nie et al. (2018)	EGFR T790M mutated NSCLC	Yes	Third	Double	73	Osimertinib	80 mg/qd	87.70	61.60	10.20
Akamatsu et al. (2018)	EGFR T790M mutated NSCLC	Yes	First	Double	41	Osimertinib	80 mg/qd	95.10	70.70	10.10
Wu et al. (2018a)	Advanced NSCLC	Yes	First	Double	75	Osimertinib	80 mg/qd	40.00	87.00	5.50
Wu et al. (2018b)	Advanced NSCLC	Yes	First	Double	30	Osimertinib	80 mg/qd	93.00	70.00	5.60
Saito et al. (2019)	EGFR-positive advanced NSCLC	Yes	First	Double	112	Erlotinib	150 mg/qd	96.00	67.00	13.30
Han et al. (2018a)	Advanced NSCLC	No	3rd	Double	60	Anlotinib	12 mg/qd	83.30	10.00	4.80
Peters et al. (2017)	ALK-positive NSCLC	Yes	First	Double	151	Ceritinib	500 mg/qd	91.00	75.00	11.10
Landi et al. (2019)	ROS1-rearranged NSCLC	No	Second	Single	33	Crizotinib	250 mg/bid	85.00	65.00	22.80
Yang et al. (2018)	Advanced NSCLC	No	First	Double	35	Gefitinib	250 mg/qd	80.00	57.10	8.40
Zhou et al. (2018)	ALK-positive advanced NSCLC	Yes	First	Double	31	Crizotinib	-	100.00	78.10	16.10
Kim et al. (2012)	Advanced NSCLC	No	Second	Double	48	Erlotinib	150 mg/qd	66.70	39.60	3.10
Kim et al. (2012)	Advanced NSCLC	No	Second	Double	48	Gefitinib	250 mg/qd	72.90	47.90	4.90
Inoue et al. (2006)	Advanced NSCLC	No	First	Single	16	Gefitinib	250 mg/qd	88.00	75.00	9.70
Leighl et al. (2017)	Advanced NSCLC	No	First	Double	44	Erlotinib	150 mg/qd	77.30	47.70	8.40
Leighl et al. (2017)	Advanced NSCLC	No	First	Double	44	Linsitinib + Erlotinib	Linsitinib: 150 mg/bid Erlotinib: 150 mg/qd	95.50	75.00	12.40
Wen et al. (2007)	Advanced NSCLC	No	Second	Double	20	Gefitinib	250 mg/qd	85.00	75.00	9.50
Wen et al. (2007)	Advanced NSCLC	No	Second	Double	30	Gefitinib	250 mg/qd	43.00	20.00	3.70
Yamada et al. (2013)	EGFR T790M positive advanced NSCLC	No	Second	Single	14	Erlotinib	150 mg/qd	80.80	53.80	9.30
Yang et al. (2020)	Crizotinib-resistant, ALK-positive NSCLC	Yes	Second	Single	160	Ensartinib	225 mg/qd	93.00	52.00	9.60
Yang et al. (2017a)	T790M-Positive Advanced NSCLC	Yes	Second + Third	Single	201	Osimertinib	80 mg/dq	90.00	62.00	12.30
Park et al. (2016)	EGFR mutation-positive NSCLC	Yes	First	Double	160	Afatinib	40 mg/d	91.00	70.00	11.00
Hu et al. (2015)	Advanced NSCLC	No	Second	Single	124	Icotinib	125 mg/tid	67.70	25.80	5.00
Yoshimura et al. (2015)	EGFR mutated, advanced non-squamous NSCLC	No	First	Single	26	Pemetrexed + Gefitinib	250 mg/q3w	96.20	84.60	18.00

DCR, disease control rate; ORR, overall response rate; mPFS, median progression-free survival; NSCLC, non-small cell lung cancer; EGFR-TKI(s), EGFR tyrosine kinase inhibitors; qd, once a day; bid, twice a day; tid, three times a day; d, day; q3w, once every 3 weeks;—, not mentioned.

TABLE 3 | Overview of general characteristics of studies reporting ORR, DCR, mPFS, and OS.

References	Population	Brain metastasis	Line	Arm	Sample size	Treatment	Dosage	DCR (%)	ORR (%)	mPFS (month)	mOS (month)
Yang et al. (2017b)	Locally advanced or metastatic ALK-positive NSCLC	No	First	Single	225	Alectinib	600 md/bid	78.80	51.30	8.30	26.00
Wu et al. (2018a)	Advanced NSCLC	No	First	Double	242	Afatinib	50 mg/qd	92.20	66.80	11.00	31.60
Uchibori et al. (2018)	Advanced NSCLC	No	First	Single	33	Gefitinib	250mg/qd	81.81	22.90	6.70	24.30
Ikezawa et al. (2019)	EGFR mutated advanced NSCLC	Yes	Third + Fourth	Double	19	Erlotinib	150 mg/qd	42.10	15.80	1.60	8.00
Garon et al. (2018)	Advanced NSCLC	No	First	Double	33	Erlotinib	150 mg/qd	79.10	12.10	3.50	9.50
Kiura et al. (2018)	ALK-rearranged advanced NSCLC	Yes	First	Double	11	Ceritinib	750 mg/qd	90.90	54.50	9.80	23.90
Cheng et al. (2021)	Advanced NSCLC	Yes	Third	Double	27	Anlotinib	12 mg/qd	82.89	9.65	4.80	10.70
Cheng et al. (2021)	Advanced squamous NSCLC	Yes	Third	Double	36	Anlotinib	12 mg/qd	71.70	7.55	5.50	9.60
Miyawaki et al. (2017)	EGFR mutated advanced NSCLC	Yes	Second + Third	Single	13	Erlotinib	150 mg/qd	69.20	53.80	7.80	25.10
Miyawaki et al. (2017)	wild-type EGFR advanced NSCLC	Yes	Second + Third	Single	22	Erlotinib	150 mg/qd	31.80	9.10	2.10	14.9
Landi et al. (2019)	MET-deregulated or ROS1-rearranged NSCLC	No	Second	Single	37	Crizotinib	250 mg/bid	69.00	27.00	4.40	5.4
Pérol et al. (2012)	Advanced NSCLC	No	Second	Multi	155	Erlotinib	150 mg/qd	35.64	10.89	2.90	11.40
Spigel et al. (2011)	Advanced NSCLC	No	Second	Double	111	Erlotinib + Sorafenib	Sorafenib: 400 mg/ bidErlotinib: 150 mg/qd	54.00	8.10	3.38	7.62
Spigel et al. (2011)	Advanced NSCLC	No	Second	Double	55	Erlotinib	150 mg/qd	38.20	10.90	1.94	7.23
Deng et al. (2012)	Advanced NSCLC	No	Second	Single	40	Gefitinib	250 mg/qd	92.50	62.50	13.00	20.00
Lee et al. (2011)	Advanced or metastatic NSCLC	Yes	First	Single	24	Erlotinib	100 mg/qd	25.00	21.00	1.50	3.20
Spigel et al. (2018)	Advanced NSCLC	No	Second	Double	127	Pazopanib + Erlotinib	Pazopanib: 600 mg/ qdErlotinib: 150 mg/qd	44.00	10.00	2.60	6.90
Spigel et al. (2018)	Advanced NSCLC	No	Second	Double	65	Erlotinib	150 mg/qd	34.00	5.00	1.80	7.00
Ciuleanu et al. (2012)	Advanced or metastatic NSCLC	Yes	Second	Double	203	Erlotinib	150 mg/qd	34.50	7.90	1.47	5.30
Arrieta et al. (2019)	Advanced NSCLC	No	First	Double	70	EGFR-TKIs	-	91.40	54.30	9.90	17.50
Han et al. (2018a)	Advanced NSCLC	No	Third	Double	296	Anlotinib	12 mg/qd	81.00	9.20	5.40	9.60
Xu et al. (2019)	Advanced NSCLC	No	First	Double	89	Icotinib	125 mg/tid	79.80	64.00	10.00	34.00
Han et al. (2017)	Advanced NSCLC	No	First	Multi	41	Gefitinib	250 mg/qd	97.60	65.90	11.90	25.80
Zhou et al. (2009)	Advanced NSCLC	—	Second	Single	112	Erlotinib	150 mg/qd	76.80	35.70	6.30	12.30
Miller et al. (2012)	Advanced NSCLC	No	Second + Third	Double	390	Afatinib	50 mg/qd	58.00	7.00	3.30	10.80
Pallis et al. (2012)	Advanced or metastatic NSCLC	Yes	First	Single	49	Erlotinib	150 mg/qd	69.40	24.50	6.70	11.50
Gaafar et al. (2011)	Advanced or metastatic NSCLC	Yes	Second	Double	86	Gefitinib	250 mg/qd	79.10	11.60	4.10	10.90
Guan et al. (2005)	Advanced or metastatic NSCLC	Yes	Second + Third	Single	153	Gefitinib	250 mg/qd	54.10	27.00	3.23	10.00
Ramalingam et al. (2012)	Advanced NSCLC	No	First	Double	94	Dacomitinib	45 mg/qd	29.80	17.00	2.86	9.53
Ramalingam et al. (2012)	Advanced NSCLC	No	First	Double	94	Erlotinib	150 mg/qd	14.90	5.30	1.91	7.44

(Continued on following page)

TABLE 3 | (*Continued*) Overview of general characteristics of studies reporting ORR, DCR, mPFS, and OS.

References	Population	Brain metastasis	Line	Arm	Sample size	Treatment	Dosage	DCR (%)	ORR (%)	mPFS (month)	mOS (month)
Ahn et al. (2020)	EGFR T790M positive advanced NSCLC	Yes	Second	Single	62	Osimertinib	80 mg/qd	95.00	74.00	10.90	29.20
Eide et al. (2020)	EGFR mutated advanced NSCLC	No	Second	Single	199	Osimertinib	80 mg/qd	83.00	48.00	8.90	17.90
Goldman et al. (2020)	Advanced NSCLC	No	Second	Double	183	Erlotinib	150 mg/qd	31.70	2.70	1.90	7.80
Huber et al. (2020)	Crizotinib-refractory ALK positive NSCLC	Yes	Second	Double	112	Brigatinib	90mg/qd	78.00	51.00	9.20	29.50
Ramalingam et al. (2014)	Advanced NSCLC	No	Second	Double	439	Dacomitinib	45 mg/qd	48.97	11.39	2.60	8.10
Ramalingam et al. (2014)	Advanced NSCLC	No	Second	Double	439	Erlotinib	150 mg/qd	49.66	8.20	2.60	8.50
Scagliotti et al. (2020)	Advanced NSCLC	No	First	Double	70	Erlotinib	150 mg/qd	98.60	65.70	9.50	25.40
Shi et al. (2013)	Advanced NSCLC	No	Second	Double	199	Icotinib	125 mg/tid	75.40	27.60	4.60	13.30
Shi et al. (2013)	Advanced NSCLC	No	Second	Double	196	Gefitinib	250 mg/qd	74.90	27.20	3.40	13.90
An et al. (2016)	EGFR mutated non-squamous NSCLC	No	First	Double	45	Gefitinib	250 mg/q3w	86.67	73.33	14.00	32.00
Neal et al. (2016)	EGFR wild-type advanced NSCLC	Yes	Second + Third	Triple	42	Erlotinib	150 mg	18.40	3.00	1.80	5.10
Ding-guo and Xiu-li (2016)	Advanced NSCLC	—	Second	Double	50	Gefitinib	250 mg/d	64.00	24.00	5.20	7.90
Choi et al. (2015)	Chemotherapy-naïve NSCLC	No	First	Double	43	Paclitaxel + Carboplatin + Gefitinib	250 mg/q3w	74.40	41.90	4.10	9.30
Wu et al. (2015)	Advanced EGFR mutation-positive NSCLC	No	First	Double	110	Erlotinib	150 mg/qd	89.10	62.70	11.00	26.30
Juan et al. (2015)	Advanced NSCLC	—	Second	Double	33	Docetaxel + Erlotinib	150 mg/qd	52.00	3.00	3.00	7.50
Wu et al. (2015)	Advanced NSCLC	No	First	Double	226	Chemo + Erlotinib	150 mg/d	80.50	42.90	10.00	18.30

DCR, disease control rate; ORR, overall response rate; mPFS, median progression-free survival; mOS, median overall survival; NSCLC, non-small cell lung cancer; EGFR-TKI(s), EGFR tyrosine kinase inhibitors; qd, once a day; bid, twice a day; tid, three times a day; d, day;—, not reported; Chemo, chemotherapy.

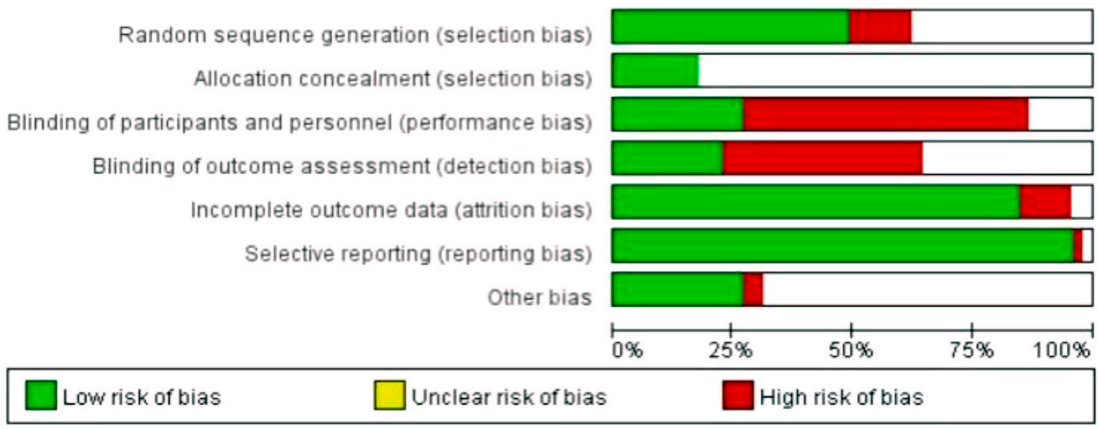

FIGURE 2 | Risk of bias in the included RCTs.

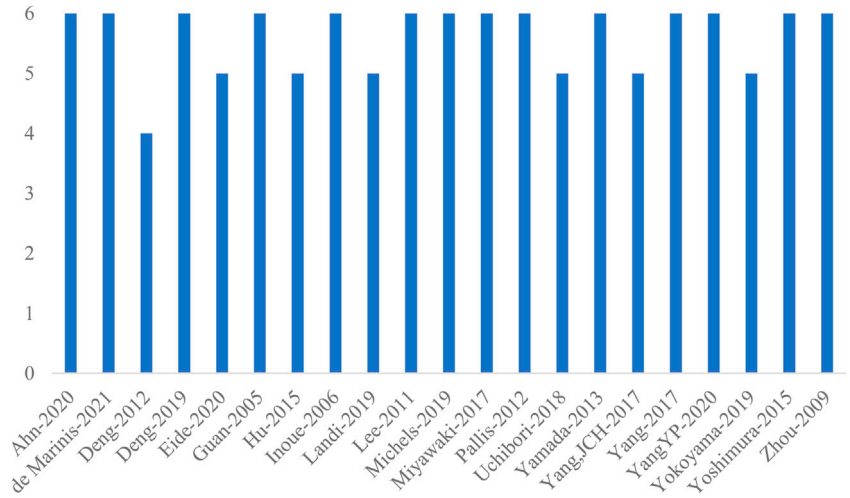

FIGURE 3 | NOS quality scores for the single-arm trials.

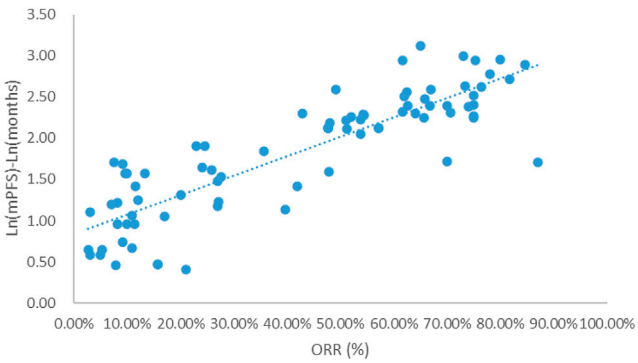

FIGURE 4 | Linear fitting results for ORR and ln (mPFS).

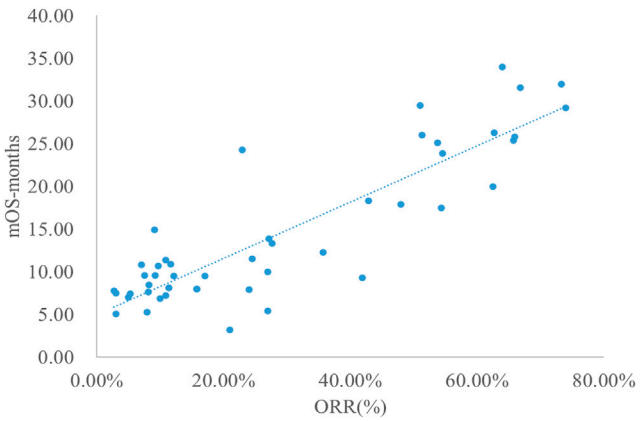

FIGURE 6 | Linear fitting results for ORR and mOS.

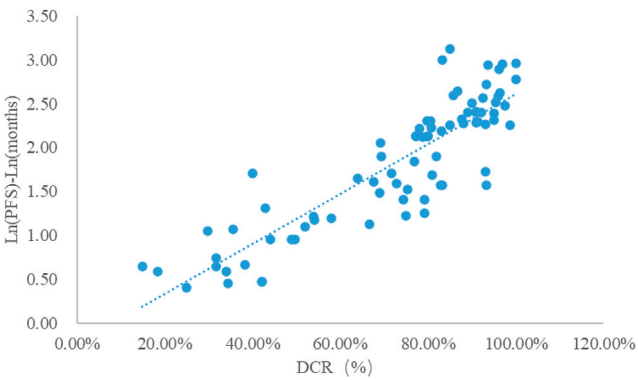

FIGURE 5 | Linear fitting results for DCR and ln (mPFS).

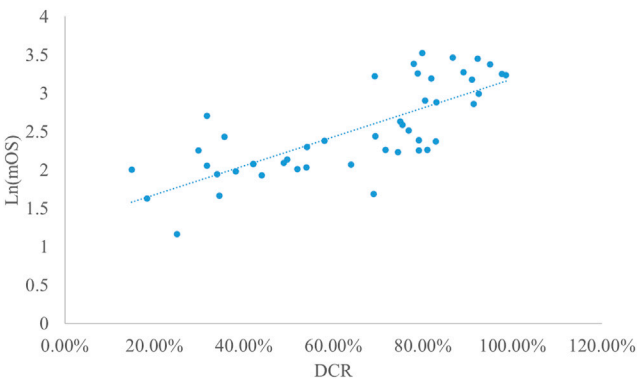

FIGURE 7 | Linear fitting results for DCR and ln (mOS).

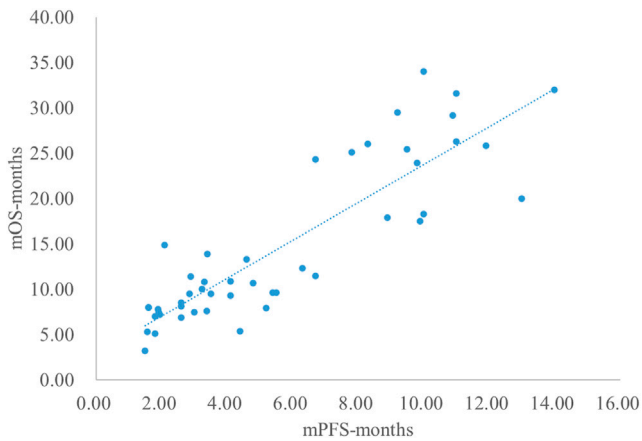

FIGURE 8 | Linear fitting results for mPFS and mOS.

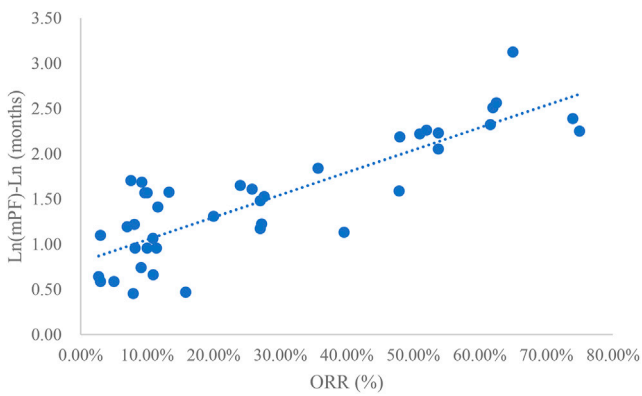

FIGURE 9 | Linear fitting results for ORR and ln (mPFS).

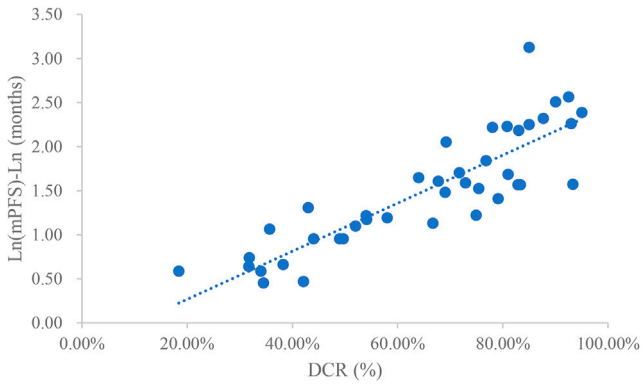

FIGURE 10 | Linear fitting results for DCR and ln (mPFS).

Taking the natural logarithm of mPFS, the adjusted goodness-of-fit of the univariate regression between mPFS and mOS was excellent (R^2_{adj} = 0.7497 > 0.6), which was shown in **Figure 14**.

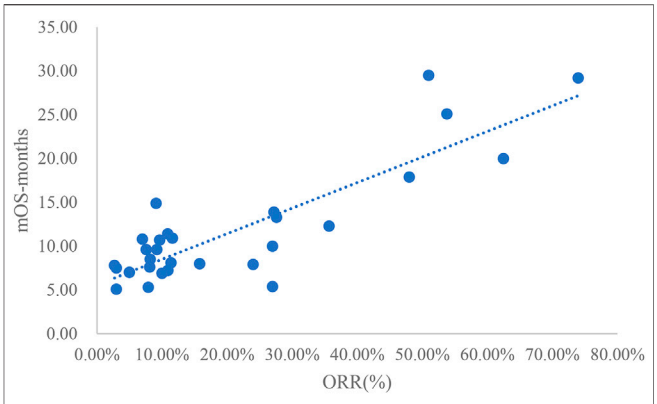

FIGURE 11 | Linear fitting results for ORR and mOS.

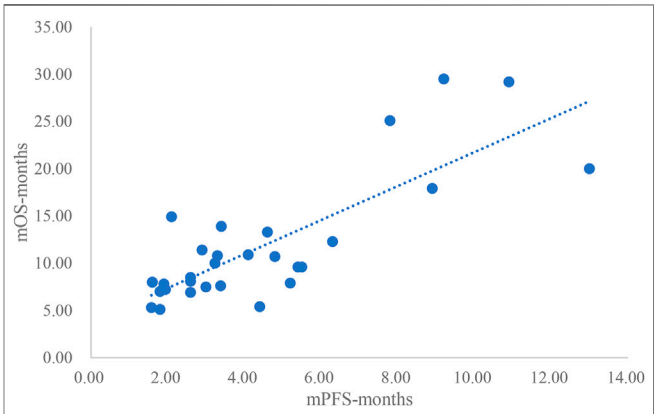

FIGURE 12 | Linear fitting results for mPFS and mOS.

The adjusted goodness-of-fit of the univariate regression between ORR and mOS was excellent (R^2_{adj} = 0.7324 > 0.6), which was shown in **Figure 15**.

DISCUSSION

This study summarized all clinical trials of molecular targeted drugs for the treatment of NSCLC approved for marketing in China as of March 2021. Studies that included DCR, ORR, mPFS concurrently and studies that included DCR, ORR, mPFS, and mOS concurrently were extracted for univariate linear regression analysis. This study included a total of 25 articles that reported DCR, ORR, and mPFS concurrently and 39 articles that reported DCR, ORR, mPFS, and mOS concurrently. In the relationship between DCR and ORR and mPFS, DCR and ORR had an excellent linear relationship with ln (mPFS), and the adjusted goodness-of-fit R^2_{adj} was >0.6. However, the linear relationships between DCR and ORR and mPFS were slightly weaker. For the relationships between DCR and ORR vs. mPFS and mOS, the linear relationship between DCR and mOS or ln (mOS) was good but not excellent (0.4 < R^2_{adj}<0.6). mPFS and mOS had most excellent linear relationships (R^2_{adj} = 0.8036).

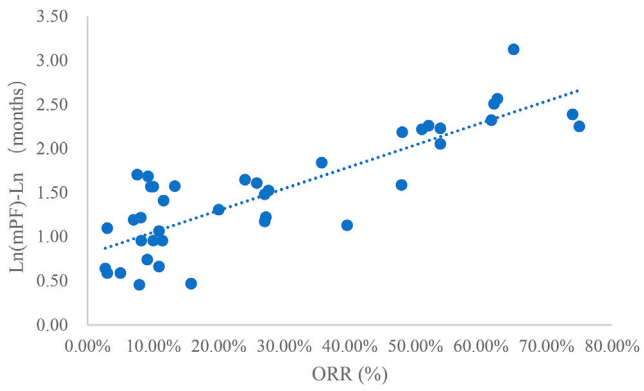

FIGURE 13 | Linear fitting results for ORR and ln (mPFS).

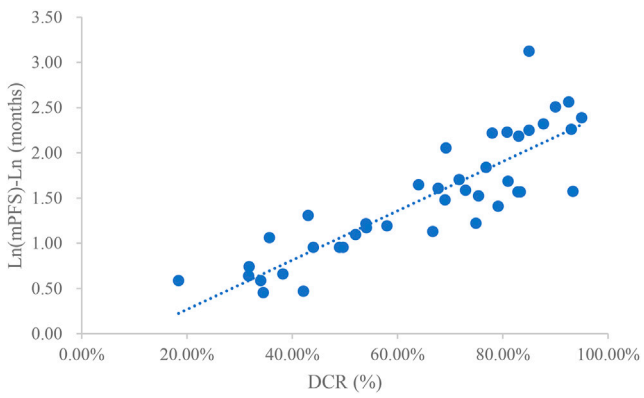

FIGURE 14 | Linear fitting results for ORR and ln (mPFS).

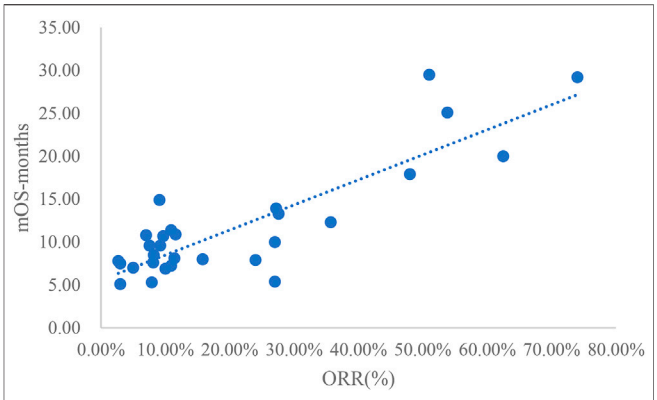

FIGURE 15 | Linear fitting results for ORR and mOS.

Cooper et al. (2020) (75) conducted a systematic review of the literature that explored the relationship between short-term surrogate endpoint indicators and primary endpoint indicators in malignancy clinical trials. A systematic search of relevant literature in five databases (from the establishment of the database to March 2019) was conducted, resulting in the inclusion of 64 articles. The results of the study suggested that short-term surrogate endpoint indicators such as ORR and CR cannot replace primary endpoint indicators such as PFS and OS and that the correlation between the two is weak and unstable. Our result was inconsistent with those reported by Cooper et al. (Cooper et al., 2020) Although Cooper et al. reported that there was no significant correlation between short-term surrogate endpoint indicators and primary endpoint indicators, the conclusion was likely due to the wide selection of disease types and treatment regimens included in the study. Moreover, Cooper et al. did not address whether there was a correlation between short-term surrogate endpoint indicators and primary endpoint indicators in the clinical trials of specific types of anticancer drugs, which might be one of the main reasons for the difference between the results of this study and the study by Cooper et al.

The results of this study revealed that short-term surrogate endpoint indicators (ORR and DCR) might have a linear relationship with mPFS and mOS, potentially providing the option to use short-term surrogate endpoint indicators to predict mPFS and mOS. In the pharmacoeconomic evaluation of tumors, PFS and OS are the most important evaluation indicators to verify drug efficacy and determine the success of the construction of pharmacoeconomic models. In the traditional pharmacoeconomic models for advanced cancer, 3-state models are often used to construct Markov models or PSMs for pharmacoeconomic evaluations (5). Markov models indicate the transition probability between health states using PFS and OS curves, and PSMs use PFS and OS curves to divide the area under the survival curve into three regions to calculate the area under the curve. If short-term surrogate endpoint indicators are used to predict mPFS and mOS, only two median values can be obtained, posing a challenge for pharmacoeconomic evaluations. We recommend that when only mPFS and mOS are available, assuming the PFS and OS curves obey an exponential distribution, mPFS and mOS should be used to construct an exponential distribution survival curve, thereby allowing the construction of a Markov model and an economic evaluation (Latimer, 2011). Although this method has strong assumptions, it can also provide a certain reference value in the absence of data.

This study has some limitations. Firstly, the molecular targeted drugs included in this study were limited to targeted drugs for the treatment of NSCLC that were approved for marketing in China as of March 2021; many targeted drugs approved for marketing in the other countries were not included in this study. Therefore, the extrapolation of the results is limited. Secondly, although only molecular targeted drugs approved in China were included, the race distribution of the included patients were not considered in the analyses. For many of these drugs, especially the recently approved drugs, were approved based on the published clinical data of the published international population plus the unpublished clinical data of a small sample of the Chinese population. Thirdly, this study did not use a large amount of real-world data for prediction and validation for the focus was to establish the statistical relationship between short-term surrogate endpoint indicators and primary endpoint indicators. Finally, for mOS, in addition to short-term surrogate endpoint indicators, other factors, such as the choice of subsequent treatment, will have a significant impact on mOS;

however, the univariate linear regression used in this study did not include enough influencing factors other than treatments.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

Conception of the study: MR, ZW, and HL; literature search: MR, ZF and ZW; data extraction: MR, ZW, ZF, and LS; statistical analysis: MR, YW, and YCW; drafting the manuscript: MR, YCW, and YW; revising and completion of final work: LS, YS, ZF, and HL; all authors reviewed and approved the final manuscript.

FUNDING

The publishing fees were funded by the corresponding author.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2022.862640/full#supplementary-material

REFERENCES

Adminstration, N. M. P. (20202020). National Medical Products Adminstration on Issuing Three Documents Including "Breakthrough Therapy Drug Review Work Procedures (Trial). Available at: https://www.nmpa.gov.cn/xxgk/ggtg/qtggtg/20200708151701834.html.

Ahn, M. J., Han, J. Y., Kim, D. W., Cho, B. C., Kang, J. H., Kim, S. W., et al. (2020). Osimertinib in Patients with T790M-Positive Advanced Non-small Cell Lung Cancer: Korean Subgroup Analysis from Phase II Studies. *Cancer Res. Treat.* 52 (1), 284–291. doi:10.4143/crt.2019.200

Akamatsu, H., Katakami, N., Okamoto, I., Kato, T., Kim, Y. H., Imamura, F., et al. (2018). Osimertinib in Japanese Patients with EGFR T790M Mutation-Positive Advanced Non-small-cell Lung Cancer: AURA3 Trial. *Cancer Sci.* 109 (6), 1930–1938. doi:10.1111/cas.13623

An, C., Zhang, J., Chu, H., Gu, C., Xiao, F., Zhu, F., et al. (2016). Study of Gefitinib and Pemetrexed as First-Line Treatment in Patients with Advanced Non-small Cell Lung Cancer Harboring EGFR Mutation. *Pathol. Oncol. Res.* 22 (4), 763–768. doi:10.1007/s12253-016-0067-4

Arrieta, O., Barrón, F., Padilla, M. S., Avilés-Salas, A., Ramírez-Tirado, L. A., Arguelles Jiménez, M. J., et al. (2019). Effect of Metformin Plus Tyrosine Kinase Inhibitors Compared with Tyrosine Kinase Inhibitors Alone in Patients with Epidermal Growth Factor Receptor-Mutated Lung Adenocarcinoma: A Phase 2 Randomized Clinical Trial. *JAMA Oncol.* 5 (11), e192553. doi:10.1001/jamaoncol.2019.2553

Bray, F., Laversanne, M., Weiderpass, E., and Soerjomataram, I. (2021). The Ever-Increasing Importance of Cancer as a Leading Cause of Premature Death Worldwide. *Cancer* 127 (16), 3029–3030. doi:10.1002/cncr.33587

Cheng, Y., Wang, Q., Li, K., Shi, J., Liu, Y., Wu, L., et al. (2021). Anlotinib vs. Placebo as Third- or Further-Line Treatment for Patients With Small Cell Lung Cancer: A Randomised, Double-Blind, Placebo-Controlled Phase 2 Study. *Br. J. Cancer* 125, 366–371. doi:10.1038/s41416-021-01356-3

Choi, Y. J., Lee, D. H., Choi, C. M., Lee, J. S., Lee, S. J., Ahn, J. H., et al. (2015). Randomized Phase II Study of Paclitaxel/carboplatin Intercalated with Gefitinib Compared to Paclitaxel/carboplatin Alone for Chemotherapy-Naïve Non-small Cell Lung Cancer in a Clinically Selected Population Excluding Patients with Non-smoking Adenocarcinoma or Mutated EGFR. *BMC Cancer* 15, 763. doi:10.1186/s12885-015-1714-y

Ciuleanu, T., Stelmakh, L., Cicenas, S., Miliauskas, S., Grigorescu, A. C., Hillenbach, C., et al. (2012). Efficacy and Safety of Erlotinib versus Chemotherapy in Second-Line Treatment of Patients with Advanced, Non-small-cell Lung Cancer with Poor Prognosis (TITAN): a Randomised Multicentre, Open-Label, Phase 3 Study. *Lancet Oncol.* 13 (3), 300–308. doi:10.1016/S1470-2045(11)70385-0

Cooper, K., Tappenden, P., Cantrell, A., and Ennis, K. (2020). A Systematic Review of Meta-Analyses Assessing the Validity of Tumour Response Endpoints as Surrogates for Progression-free or Overall Survival in Cancer. *Br. J. Cancer* 123 (11), 1686–1696. doi:10.1038/s41416-020-01050-w

Coyle, D., and Coyle, K. (2014). The Inherent Bias from Using Partitioned Survival Models in Economic Evaluation. *Value in Health* 17 (3), A194. doi:10.1016/j.jval.2014.03.1130

de Marinis, F., Laktionov, K. K., Poltoratskiy, A., Egorova, I., Hochmair, M., Passaro, A., et al. (2021). Afatinib in EGFR TKI-Naïve Patients with Locally Advanced or Metastatic EGFR Mutation-Positive Non-small Cell Lung Cancer: Interim Analysis of a Phase 3b Study. *Lung Cancer* 152, 127–134. doi:10.1016/j.lungcan.2020.12.011

Deng, H., Li, B., Li, L., Peng, J., Lv, T., Liu, Y., et al. (2019). Clinical Observation of Crizotinib in the Treatment of ALK-Positive Advanced Non-small Cell Lung Cancer. *Pathol. Res. Pract.* 215 (12), 152695. doi:10.1016/j.prp.2019.152695

Deng, J., Fang, W. J., Zhang, X. C., Wu, D. P., Fang, H. M., Chen, J., et al. (2012). Phase II Trial of Gefitinib in Pretreated Chinese Women with Advanced Non-small-cell Lung Cancer. *Med. Oncol.* 29 (2), 595–599. doi:10.1007/s12032-011-9891-2

Ding-guo, Z., and Xiu-li, R. (2016). A Randomized Clinical Study of Gefitinib and Pemetrexed as Second Line Therapy for Advanced Non-small Cell Lung Cancer. *Oncol. Prog.* 14 (01), 78–80+6.

Eide, I. J. Z., Helland, Å., Ekman, S., Mellemgaard, A., Hansen, K. H., Cicenas, S., et al. (2020). Osimertinib in T790M-Positive and -negative Patients with EGFR-Mutated Advanced Non-small Cell Lung Cancer (The TREM-Study). *Lung Cancer* 143, 27–35. doi:10.1016/j.lungcan.2020.03.009

Gaafar, R. M., Surmont, V. F., Scagliotti, G. V., Van Klaveren, R. J., Papamichael, D., Welch, J. J., et al. (2011). A Double-Blind, Randomised, Placebo-Controlled Phase III Intergroup Study of Gefitinib in Patients with Advanced NSCLC, Non-progressing after First Line Platinum-Based Chemotherapy (EORTC 08021/ILCP 01/03). *Eur. J. Cancer* 47 (15), 2331–2340. doi:10.1016/j.ejca.2011.06.045

Gadgeel, S., Rodríguez-Abreu, D., Speranza, G., Esteban, E., Felip, E., Dómine, M., et al. (2020). Updated Analysis from KEYNOTE-189: Pembrolizumab or Placebo Plus Pemetrexed and Platinum for Previously Untreated Metastatic Nonsquamous Non-small-cell Lung Cancer. *J. Clin. Oncol.* 38 (14), 1505–1517. doi:10.1200/jco.19.03136

Garon, E. B., Siegfried, J. M., Stabile, L. P., Young, P. A., Marquez-Garban, D. C., Park, D. J., et al. (2018). Randomized Phase II Study of Fulvestrant and Erlotinib Compared with Erlotinib Alone in Patients with Advanced or Metastatic Non-small Cell Lung Cancer. *Lung Cancer* 123, 91–98. doi:10.1016/j.lungcan.2018.06.013

Gisselbrecht, C., Glass, B., Mounier, N., Singh Gill, D., Linch, D. C., Trneny, M., et al. (2010). Salvage Regimens with Autologous Transplantation for Relapsed Large B-Cell Lymphoma in the Rituximab Era. *J. Clin. Oncol.* 28 (27), 4184–4190. doi:10.1200/JCO.2010.28.1618

Goldman, J. W., Mazieres, J., Barlesi, F., Dragnev, K. H., Koczywas, M., Göskel, T., et al. (2020). A Randomized Phase III Study of Abemaciclib versus Erlotinib in Patients with Stage IV Non-small Cell Lung Cancer with a Detectable KRAS Mutation Who Failed Prior Platinum-Based Therapy: JUNIPER. *Front. Oncol.* 10, 578756. doi:10.3389/fonc.2020.578756

Guan, Z. Z., Zhang, L., Li, L. Y., Jiang, G. L., Liu, X. Y., Chu, D. T., et al. (2005). Efficacy of Gefitinib on Chinese Patients with Locally Advanced or Metastatic Non-small Cell Lung Cancer: a Clinical Trial. *Ai Zheng* 24 (8), 980–984.

Han, B., Jin, B., Chu, T., Niu, Y., Dong, Y., Xu, J., et al. (2017). Combination of Chemotherapy and Gefitinib as First-Line Treatment for Patients with Advanced Lung Adenocarcinoma and Sensitive EGFR Mutations: A Randomized Controlled Trial. *Int. J. Cancer* 141 (6), 1249–1256. doi:10.1002/ijc.30806

Han, B., Li, K., Wang, Q., Zhang, L., Shi, J., Wang, Z., et al. (2018). Effect of Anlotinib as a Third-Line or Further Treatment on Overall Survival of Patients with Advanced Non-small Cell Lung Cancer: The ALTER 0303 Phase 3 Randomized Clinical Trial. *JAMA Oncol.* 4 (11), 1569–1575. doi:10.1001/jamaoncol.2018.3039

Han, B., Li, K., Zhao, Y., Li, B., Cheng, Y., Zhou, J., et al. (2018). Anlotinib as a Third-Line Therapy in Patients with Refractory Advanced Non-small-cell Lung Cancer: a Multicentre, Randomised Phase II Trial (ALTER0302). *Br. J. Cancer* 118 (5), 654–661. doi:10.1038/bjc.2017.478

Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., et al. (2011). The Cochrane Collaboration's Tool for Assessing Risk of Bias in Randomised Trials. *BMJ* 343, d5928. doi:10.1136/bmj.d5928

Hu, X., Zhang, L., Shi, Y., Zhou, C., Liu, X., Wang, D., et al. (2015). The Efficacy and Safety of Icotinib in Patients with Advanced Non-small Cell Lung Cancer Previously Treated with Chemotherapy: A Single-Arm, Multi-Center, Prospective Study. *PLoS One* 10 (11), e0142500. doi:10.1371/journal.pone.0142500

Huber, R. M., Hansen, K. H., Paz-Ares Rodríguez, L., West, H. L., Reckamp, K. L., Leighl, N. B., et al. (2020). Brigatinib in Crizotinib-Refractory ALK+ NSCLC: 2-Year Follow-Up on Systemic and Intracranial Outcomes in the Phase 2 ALTA Trial. *J. Thorac. Oncol.* 15 (3), 404–415. doi:10.1016/j.jtho.2019.11.004

Ikezawa, Y., Asahina, H., Oizumi, S., Watanabe, M., Takamura, K., Kawai, Y., et al. (2017). A Randomized Phase II Trial of Erlotinib vs. S-1 as a Third- or Fourth-Line Therapy for Patients with Wild-type EGFR Non-small Cell Lung Cancer (HOT1002). *Cancer Chemother. Pharmacol.* 80 (5), 955–963. doi:10.1007/s00280-017-3432-4

Inoue, A., Suzuki, T., Fukuhara, T., Maemondo, M., Kimura, Y., Morikawa, N., et al. (2006). Prospective Phase II Study of Gefitinib for Chemotherapy-Naive Patients with Advanced Non-small-cell Lung Cancer with Epidermal Growth Factor Receptor Gene Mutations. *J. Clin. Oncol.* 24 (21), 3340–3346. doi:10.1200/JCO.2005.05.4692

Juan, Ó., Aparisi, F., Sánchez-Hernández, A., Muñoz-Langa, J., Esquerdo, G., García-Sánchez, J., et al. (2015). Intercalated Dosing Schedule of Erlotinib and Docetaxel as a Therapeutic Strategy to Avoid Antagonism and Optimize its Benefits in Advanced Non-small-cell Lung Cancer. A Randomized Phase II Clinical Trial. *Clin. Lung Cancer* 16 (3), 193–199. doi:10.1016/j.cllc.2014.11.006

Kim, S. T., Uhm, J. E., Lee, J., Sun, J. M., Sohn, I., Kim, S. W., et al. (2012). Randomized Phase II Study of Gefitinib versus Erlotinib in Patients with Advanced Non-small Cell Lung Cancer Who Failed Previous Chemotherapy. *Lung Cancer* 75 (1), 82–88. doi:10.1016/j.lungcan.2011.05.022

Kiura, K., Imamura, F., Kagamu, H., Matsumoto, S., Hida, T., Nakagawa, K., et al. (2018). Phase 3 Study of Ceritinib vs Chemotherapy in ALK-Rearranged NSCLC Patients Previously Treated with Chemotherapy and Crizotinib (ASCEND-5): Japanese Subset. *Jpn. J. Clin. Oncol.* 48 (4), 367–375. doi:10.1093/jjco/hyy016

Landi, L., Chiari, R., Tiseo, M., D'Incà, F., Dazzi, C., Chella, A., et al. (2019). Crizotinib in MET-Deregulated or ROS1-Rearranged Pretreated Non-small Cell Lung Cancer (METROS): A Phase II, Prospective, Multicenter, Two-Arms Trial. *Clin. Cancer Res.* 25 (24), 7312–7319. doi:10.1158/1078-0432.CCR-19-0994

Lassere, M. N., Johnson, K. R., Schiff, M., and Rees, D. (2012). Is Blood Pressure Reduction a Valid Surrogate Endpoint for Stroke Prevention? an Analysis Incorporating a Systematic Review of Randomised Controlled Trials, a By-Trial Weighted Errors-In-Variables Regression, the Surrogate Threshold Effect (STE) and the Biomarker-Surrogacy (BioSurrogate) Evaluation Schema (BSES). *BMC Med. Res. Methodol.* 12 (1), 27. doi:10.1186/1471-2288-12-27

Latimer, N. (2011). *NICE DSU Technical Support Document 14: Undertaking Survival Analysis for Economic Evaluations Alongside Clinical Trials - Extrapolation with Patient-Level Data*. London, United Kingdom: The National Institute for Health and Care Excellence.

Lee, D. H., Kim, S. W., Suh, C., Han, Y. H., and Lee, J. S. (2011). Phase II Study of Erlotinib for Chemotherapy-Naïve Patients with Advanced or Metastatic Non-small Cell Lung Cancer Who Are Ineligible for Platinum Doublets. *Cancer Chemother. Pharmacol.* 67 (1), 35–39. doi:10.1007/s00280-010-1280-6

Leighl, N. B., Rizvi, N. A., de Lima, L. G., Jr., Arpornwirat, W., Rudin, C. M., Chiappori, A. A., et al. (2017). Phase 2 Study of Erlotinib in Combination with Linsitinib (OSI-906) or Placebo in Chemotherapy-Naive Patients with Non-small-cell Lung Cancer and Activating Epidermal Growth Factor Receptor Mutations. *Clin. Lung Cancer* 18 (1), 34–e2. doi:10.1016/j.cllc.2016.07.007

Ma, W., and Yu, J.Commission NH (2020). *China Health Statistical Yearbook 2020*. Beijing: Peking Union Medical College Press.

Margulis, A. V., Pladevall, M., Riera-Guardia, N., Varas-Lorenzo, C., Hazell, L., Berkman, N. D., et al. (2014). Quality Assessment of Observational Studies in a Drug-Safety Systematic Review, Comparison of Two Tools: the Newcastle-Ottawa Scale and the RTI Item Bank. *Clin. Epidemiol.* 6, 359–368. doi:10.2147/CLEP.S66677

Michels, S., Massutí, B., Schildhaus, H. U., Franklin, J., Sebastian, M., Felip, E., et al. (2019). Safety and Efficacy of Crizotinib in Patients with Advanced or Metastatic ROS1-Rearranged Lung Cancer (EUCROSS): A European Phase II Clinical Trial. *J. Thorac. Oncol.* 14 (7), 1266–1276. doi:10.1016/j.jtho.2019.03.020

Miller, V. A., Hirsh, V., Cadranel, J., Chen, Y. M., Park, K., Kim, S. W., et al. (2012). Afatinib versus Placebo for Patients with Advanced, Metastatic Non-small-cell Lung Cancer after Failure of Erlotinib, Gefitinib, or Both, and One or Two Lines of Chemotherapy (LUX-Lung 1): a Phase 2b/3 Randomised Trial. *Lancet Oncol.* 13 (5), 528–538. doi:10.1016/S1470-2045(12)70087-6

Miyawaki, M., Naoki, K., Yoda, S., Nakayama, S., Satomi, R., Sato, T., et al. (2017). Erlotinib as Second- or Third-Line Treatment in Elderly Patients with Advanced Non-small Cell Lung Cancer: Keio Lung Oncology Group Study 001 (KLOG001). *Mol. Clin. Oncol.* 6 (3), 409–414. doi:10.3892/mco.2017.1154

Neal, J. W., Dahlberg, S. E., Wakelee, H. A., Aisner, S. C., Bowden, M., Huang, Y., et al. (2016). Erlotinib, Cabozantinib, or Erlotinib Plus Cabozantinib as Second-Line or Third-Line Treatment of Patients with EGFR Wild-type Advanced Non-small-cell Lung Cancer (ECOG-ACRIN 1512): a Randomised, Controlled, Open-Label, Multicentre, Phase 2 Trial. *Lancet Oncol.* 17 (12), 1661–1671. doi:10.1016/S1470-2045(16)30561-7

Nie, K., Zhang, Z., Zhang, C., Geng, C., Zhang, L., Xu, X., et al. (2018). Osimertinib Compared Docetaxel-Bevacizumab as Third-Line Treatment in EGFR T790M Mutated Non-small-cell Lung Cancer. *Lung Cancer* 121, 5–11. doi:10.1016/j.lungcan.2018.04.012

Ohe, Y., Imamura, F., Nogami, N., Okamoto, I., Kurata, T., Kato, T., et al. (2019). Osimertinib versus Standard-Of-Care EGFR-TKI as First-Line Treatment for EGFRm Advanced NSCLC: FLAURA Japanese Subset. *Jpn. J. Clin. Oncol.* 49 (1), 29–36. doi:10.1093/jjco/hyy179

Pallis, A. G., Voutsina, A., Kentepozidis, N., Giassas, S., Papakotoulas, P., Agelaki, S., et al. (2012). A Phase II Trial of Erlotinib as Front-Line Treatment in Clinically Selected Patients with Non-small-cell Lung Cancer. *Clin. Lung Cancer* 13 (2), 129–135. doi:10.1016/j.cllc.2011.08.004

Park, K., Tan, E. H., O'Byrne, K., Zhang, L., Boyer, M., Mok, T., et al. (2016). Afatinib versus Gefitinib as First-Line Treatment of Patients with EGFR Mutation-Positive Non-small-cell Lung Cancer (LUX-Lung 7): a Phase 2B, Open-Label, Randomised Controlled Trial. *Lancet Oncol.* 17 (5), 577–589. doi:10.1016/S1470-2045(16)30033-X

Pérol, M., Chouaid, C., Pérol, D., Barlési, F., Gervais, R., Westeel, V., et al. (2012). Randomized, Phase III Study of Gemcitabine or Erlotinib Maintenance Therapy versus Observation, with Predefined Second-Line Treatment, after Cisplatin-Gemcitabine Induction Chemotherapy in Advanced Non-small-cell Lung Cancer. *J. Clin. Oncol.* 30 (28), 3516–3524.

Peters, S., Camidge, D. R., Shaw, A. T., Gadgeel, S., Ahn, J. S., Kim, D. W., et al. (2017). Alectinib versus Crizotinib in Untreated ALK-Positive Non-small-cell Lung Cancer. *N. Engl. J. Med.* 377 (9), 829–838. doi:10.1056/NEJMoa1704795

Ramalingam, S. S., Blackhall, F., Krzakowski, M., Barrios, C. H., Park, K., Bover, I., et al. (2012). Randomized Phase II Study of Dacomitinib (PF-00299804), an Irreversible Pan-Human Epidermal Growth Factor Receptor Inhibitor, versus Erlotinib in Patients with Advanced Non-small-cell Lung Cancer. *J. Clin. Oncol.* 30 (27), 3337–3344. doi:10.1200/JCO.2011.40.9433

Ramalingam, S. S., Jänne, P. A., Mok, T., O'Byrne, K., Boyer, M. J., Von Pawel, J., et al. (2014). Dacomitinib versus Erlotinib in Patients with Advanced-Stage, Previously Treated Non-small-cell Lung Cancer (ARCHER 1009): a Randomised, Double-

Blind, Phase 3 Trial. *Lancet Oncol.* 15 (12), 1369–1378. doi:10.1016/S1470-2045(14)70452-8

Rui, M., Wang, Y., Fei, Z., Zhang, X., Shang, Y., and Li, H. (2021). Will the Markov Model and Partitioned Survival Model lead to Different Results? A Review of Recent Economic Evidence of Cancer Treatments. *Expert Rev. Pharmacoeconomics Outcomes Res.* 21 (3), 373–380. doi:10.1080/14737167.2021.1893167

Rule, S., Dreyling, M., Goy, A., Hess, G., Auer, R., Kahl, B., et al. (2017). Outcomes in 370 Patients with Mantle Cell Lymphoma Treated with Ibrutinib: a Pooled Analysis from Three Open-Label Studies. *Br. J. Haematol.* 179 (3), 430–438. doi:10.1111/bjh.14870

Saito, H., Fukuhara, T., Furuya, N., Watanabe, K., Sugawara, S., Iwasawa, S., et al. (2019). Erlotinib Plus Bevacizumab versus Erlotinib Alone in Patients with EGFR-Positive Advanced Non-squamous Non-small-cell Lung Cancer (NEJ026): Interim Analysis of an Open-Label, Randomised, Multicentre, Phase 3 Trial. *Lancet Oncol.* 20 (5), 625–635. doi:10.1016/S1470-2045(19)30035-X

Scagliotti, G., Moro-Sibilot, D., Kollmeier, J., Favaretto, A., Cho, E. K., Grosch, H., et al. (2020). A Randomized-Controlled Phase 2 Study of the MET Antibody Emibetuzumab in Combination with Erlotinib as First-Line Treatment for EGFR Mutation-Positive NSCLC Patients. *J. Thorac. Oncol.* 15 (1), 80–90. doi:10.1016/j.jtho.2019.10.003

Shi, Y., Zhang, L., Liu, X., Zhou, C., Zhang, L., Zhang, S., et al. (2013). Icotinib versus Gefitinib in Previously Treated Advanced Non-small-cell Lung Cancer (ICOGEN): a Randomised, Double-Blind Phase 3 Non-inferiority Trial. *Lancet Oncol.* 14 (10), 953–961. doi:10.1016/S1470-2045(13)70355-3

Spigel, D. R., Burris, H. A., Greco, F. A., Shih, K. C., Gian, V. G., Lipman, A. J., et al. (2018). Erlotinib Plus Either Pazopanib or Placebo in Patients with Previously Treated Advanced Non-small Cell Lung Cancer: A Randomized, Placebo-Controlled Phase 2 Trial with Correlated Serum Proteomic Signatures. *Cancer* 124 (11), 2355–2364. doi:10.1002/cncr.31290

Spigel, D. R., Burris, H. A., Greco, F. A., Shipley, D. L., Friedman, E. K., Waterhouse, D. M., et al. (2011). Randomized, Double-Blind, Placebo-Controlled, Phase II Trial of Sorafenib and Erlotinib or Erlotinib Alone in Previously Treated Advanced Non-small-cell Lung Cancer. *J. Clin. Oncol.* 29 (18), 2582–2589. doi:10.1200/JCO.2010.30.7678

Uchibori, K., Satouchi, M., Sueoka-Aragane, N., Urata, Y., Sato, A., Imamura, F., et al. (2018). Phase II Trial of Gefitinib Plus Pemetrexed after Relapse Using First-Line Gefitinib in Patients with Non-small Cell Lung Cancer Harboring EGFR Gene Mutations. *Lung Cancer* 124, 65–70. doi:10.1016/j.lungcan.2018.07.031

Wei-jing, W., and xiao-lu, Z. (2019). Current Status and Disease burden of Cancer in China. *China Econ. Rep.* 2019 (04), 63–73.

Wen, Z. S., Chen, X. Q., Wu, H. Y., Wei, W. D., and Rong, T. H. (2007). Efficacy of Gefitinib on Advanced Non-small Cell Lung Cancer of Bilateral Diffuse or Unilateral Giant Mass Type. *Ai Zheng* 26 (4), 415–417.

Wu, Y. L., Ahn, M. J., Garassino, M. C., Han, J. Y., Katakami, N., Kim, H. R., et al. (2018). CNS Efficacy of Osimertinib in Patients with T790M-Positive Advanced Non-small-cell Lung Cancer: Data from a Randomized Phase III Trial (AURA3). *J. Clin. Oncol.* 36 (26), 2702–2709. doi:10.1200/JCO.2018.77.9363

Wu, Y. L., Lee, J. S., Thongprasert, S., Yu, C. J., Zhang, L., Ladrera, G., et al. (2013). Intercalated Combination of Chemotherapy and Erlotinib for Patients with Advanced Stage Non-small-cell Lung Cancer (FASTACT-2): a Randomised, Double-Blind Trial. *Lancet Oncol.* 14 (8), 777–786. doi:10.1016/S1470-2045(13)70254-7

Wu, Y. L., Xu, C. R., Hu, C. P., Feng, J., Lu, S., Huang, Y., et al. (2018). Afatinib versus Gemcitabine/cisplatin for First-Line Treatment of Chinese Patients with Advanced Non-small-cell Lung Cancer Harboring EGFR Mutations: Subgroup Analysis of the LUX-Lung 6 Trial. *Onco Targets Ther.* 11, 8575–8587. doi:10.2147/OTT.S160358

Wu, Y. L., Zhou, C., Liam, C. K., Wu, G., Liu, X., Zhong, Z., et al. (2015). First-line Erlotinib versus Gemcitabine/cisplatin in Patients with Advanced EGFR Mutation-Positive Non-small-cell Lung Cancer: Analyses from the Phase III, Randomized, Open-Label, ENSURE Study. *Ann. Oncol.* 26 (9), 1883–1889. doi:10.1093/annonc/mdv270

Xu, L., Qi, Q., Zhang, Y., Cui, J., Liu, R., and Li, Y. (2019). Combination of Icotinib and Chemotherapy as First-Line Treatment for Advanced Lung Adenocarcinoma in Patients with Sensitive EGFR Mutations: A Randomized Controlled Study. *Lung Cancer* 133, 23–31. doi:10.1016/j.lungcan.2019.05.008

Yamada, K., Takayama, K., Kawakami, S., Saruwatari, K., Morinaga, R., Harada, T., et al. (2013). Phase II Trial of Erlotinib for Japanese Patients with Previously Treated Non-small-cell Lung Cancer Harboring EGFR Mutations: Results of Lung Oncology Group in Kyushu (LOGiK0803). *Jpn. J. Clin. Oncol.* 43 (6), 629–635. doi:10.1093/jjco/hyt056

Yang, J. C., Ahn, M. J., Kim, D. W., Ramalingam, S. S., Sequist, L. V., Su, W. C., et al. (2017). Osimertinib in Pretreated T790M-Positive Advanced Non-small-cell Lung Cancer: AURA Study Phase II Extension Component. *J. Clin. Oncol.* 35 (12), 1288–1296. doi:10.1200/JCO.2016.70.3223

Yang, J. C., Ou, S. I., De Petris, L., Gadgeel, S., Gandhi, L., Kim, D. W., et al. (2017). Pooled Systemic Efficacy and Safety Data from the Pivotal Phase II Studies (NP28673 and NP28761) of Alectinib in ALK-Positive Non-small Cell Lung Cancer. *J. Thorac. Oncol.* 12 (10), 1552–1560. doi:10.1016/j.jtho.2017.06.070

Yang, J. C., Shepherd, F. A., Kim, D. W., Lee, G. W., Lee, J. S., Chang, G. C., et al. (2019). Osimertinib Plus Durvalumab versus Osimertinib Monotherapy in EGFR T790M-Positive NSCLC Following Previous EGFR TKI Therapy: CAURAL Brief Report. *J. Thorac. Oncol.* 14 (5), 933–939. doi:10.1016/j.jtho.2019.02.001

Yang, X. B., Chai, X. S., Wu, W. Y., Long, S. Q., Deng, H., Pan, Z. Q., et al. (2018). Gefitinib Plus Fuzheng Kang'ai Formula () in Patients with Advanced Non-small Cell Lung Cancer with Epidermal Growth Factor Receptor Mutation: A Randomized Controlled Trial. *Chin. J. Integr. Med.* 24 (10), 734–740. doi:10.1007/s11655-017-2819-8

Yang, Y., Zhou, J., Zhou, J., Feng, J., Zhuang, W., Chen, J., et al. (2020). Efficacy, Safety, and Biomarker Analysis of Ensartinib in Crizotinib-Resistant, ALK-Positive Non-small-cell Lung Cancer: a Multicentre, Phase 2 Trial. *Lancet Respir. Med.* 8 (1), 45–53. doi:10.1016/S2213-2600(19)30252-8

Yokoyama, T., Yoshioka, H., Fujimoto, D., Demura, Y., Hirano, K., Kawai, T., et al. (2019). A Phase II Study of Low Starting Dose of Afatinib as First-Line Treatment in Patients with EGFR Mutation-Positive Non-small-cell Lung Cancer (KTORG1402). *Lung Cancer* 135, 175–180. doi:10.1016/j.lungcan.2019.03.030

Yoshimura, N., Kudoh, S., Mitsuoka, S., Yoshimoto, N., Oka, T., Nakai, T., et al. (2015). Phase II Study of a Combination Regimen of Gefitinib and Pemetrexed as First-Line Treatment in Patients with Advanced Non-small Cell Lung Cancer Harboring a Sensitive EGFR Mutation. *Lung Cancer* 90 (1), 65–70. doi:10.1016/j.lungcan.2015.06.002

Zhou, J., Zheng, J., Zhang, X., Zhao, J., Zhu, Y., Shen, Q., et al. (2018). Crizotinib in Patients with Anaplastic Lymphoma Kinase-Positive Advanced Non-small Cell Lung Cancer versus Chemotherapy as a First-Line Treatment. *BMC Cancer* 18 (1), 10. doi:10.1186/s12885-017-3720-8

Zhou, S., Ren, S., Yan, L., Zhang, L., Tang, L., Zhang, J., et al. (2009). Clinical Efficacy of Erlotinib in Patients Previously Treated for Advanced Non-small Cell Lung Cancer. *Respirology* 14 (5), 709–715. doi:10.1111/j.1440-1843.2009.01564.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Rui, Wang, Fei, Wu, Wang, Sun, Shang and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Pharmacology

ORIGINAL RESEARCH
published: 21 March 2022
doi: 10.3389/fphar.2022.859495

Effects of Motherwort Injection Versus Intramuscular Oxytocin for Preventing Postpartum Hemorrhage Among Women Who Underwent Cesarean Section

Ming-xi Li[1,2,3], Chun-rong Liu[1,2,3], Meng Chen[4], Hong-cai Shang[5], Wen Wang[1,2,3], Xiao-chao Luo[1,2,3], Ling Li[1,2,3], Ya-na Qi[1,2,3], Yi-quan Xiong[1,2,3], Shi-yao Huang[1,6], Jing Wang[1,2,3], Kang Zou[1,2,3], Xing-hui Liu[4], Jing Tan[1,2,3] and Xin Sun[1,2,3]**

[1]Chinese Evidence-based Medicine Center, West China Hospital, Sichuan University, Chengdu, China, [2]NMPA Key Laboratory for Real World Data Research and Evaluation in Hainan, Chengdu, China, [3]Sichuan Center of Technology Innovation for Real World Data, Chengdu, China, [4]Department of Obstetrics and Gynecology, and Key Laboratory of Birth Defects and Related Diseases of Women and Children (Sichuan University), Ministry of Education, West China Second University Hospital, Sichuan University, Chengdu, China, [5]Beijing University of Chinese Medicine and Pharmacology, Key Laboratory of Chinese Internal Medicine of MOE and Beijing, Beijing, China, [6]Sichuan Evidence-Based Medicine Center of Traditional Chinese Medicine, Hospital of Chengdu University of Traditional Chinese Medicine, Chengdu, China

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Jiarui Wu,
Beijing University of Chinese Medicine,
China
Yutong Fei,
Beijing University of Chinese Medicine,
China
Kurt Neumann,
Independent Rsearcher, Kerékteleki,
Hungary
Yingzi Cui,
The Affiliated Hospital of Changchun
University of Traditional Chinese
Medicine, China

***Correspondence:**
Jing Tan
tanjing84@outlook.com
Xin Sun
sunxin@wchscu.cn

Specialty section:
This article was submitted to
Drugs Outcomes Research and
Policies,
a section of the journal
Frontiers in Pharmacology

Received: 21 January 2022
Accepted: 07 February 2022
Published: 21 March 2022

Citation:
Li M, Liu C, Chen M, Shang H, Wang W, Luo X, Li L, Qi Y, Xiong Y, Huang S, Wang J, Zou K, Liu X, Tan J and Sun X (2022) Effects of Motherwort Injection Versus Intramuscular Oxytocin for Preventing Postpartum Hemorrhage Among Women Who Underwent Cesarean Section.
Front. Pharmacol. 13:859495.
doi: 10.3389/fphar.2022.859495

Objectives: Subject to ethical constraints, real-world data are an important resource for evaluating treatment effects of medication use during pregnancy and the postpartum period. This study investigated whether motherwort injection, a traditional Chinese medicine preparation, was more effective than intramuscular (IM) oxytocin for preventing postpartum hemorrhage (PPH) in a real-world setting when intravenous (IV) oxytocin is administered.

Methods: We conducted an active-controlled, propensity-score matched cohort study using an established pregnancy registry database. Women who underwent cesarean section and received IV oxytocin at the third stage of labor were included. We used an active-comparator design to minimize indication bias, in which we compared IM motherwort injection in the uterus versus IM oxytocin, both on top of IV oxytocin use. We applied 1:1 propensity-score matching (PSM) to balance patient baseline characteristics and used a logistic regression model to estimate treatment effect (i.e., risk difference (RD) and odds ratio (OR)) by using the counterfactual framework. The outcomes of interest were blood loss over 500 ml within 2 h after delivery (PPH, primary) and blood loss over 1,000 ml (severe PPH, secondary). We conducted four sensitivity analyses to examine the robustness of the results.

Results: A total of 22,519 pregnant women underwent cesarean sections, among which 4,081 (18.12%) PPH and 480 (2.13%) severe PPH occurred. Among included women, 586 (2.60%) were administrated with IM motherwort injection, and 21,933 (97.40%) used IM oxytocin. After PSM, patient baseline characteristics were well balanced. Compared with IM oxytocin, the use of IM motherwort injection was associated with significantly lower risk of PPH (RD −25.26%, 95% CI −30.04% to −20.47%, $p < 0.001$; OR 0.25, 95% CI 0.18

to 0.32, $p < 0.001$) and severe PPH (RD –3.58%, 95% CI –5.87% to –1.30%, $p < 0.001$; OR 0.39, 95% CI 0.20 to 0.71, $p < 0.002$). Sensitivity analyses showed that the results were similar.

Conclusion: With the use of data from a real-world setting, the findings consistently showed that among women undergoing cesarean section who had received IV oxytocin, the additional use of IM motherwort injection could achieve a lower risk of PPH as compared to the additional use of IM oxytocin. Our study suggested a paradigm for investigating the treatment effect of Chinese herbal medicine in the real-world practice setting.

Keywords: motherwort injection, postpartum hemorrhage, Chinese herbal medicine preparation, pregnancy registry database, real-world study

INTRODUCTION

Postpartum hemorrhage (PPH), defined as blood loss above 500 ml within the first 24 h after delivery, is the major cause of maternal mortality worldwide (World Health Organization, 2012). It accounts for nearly 13% of maternal deaths in developed countries and 28% in developing countries (Khan et al., 2006; World Health Organization, 2012). Uterine atony is recognized as the primary cause of PPH. Intravenous or intramuscular (IV/IM) oxytocin is recommended by World Health Organization (WHO) as the first-line treatment preventing PPH (World Health Organization, 2012). However, due to a short half-life (4–15 min), oxytocin has a short-term muscular contraction effect and may have serious adverse effects when reaching a high dose, such as tachycardia, hypotension, and water retention with hyponatremia (Pantoja et al., 2016; Pacheco et al., 2019). Thus, additional uterotonics may often be used in combination to lower the risk of PPH (Tan et al., 2016; Gallos et al., 2018).

Motherwort injection is a modern purified Chinese herbal medicine preparation derived from *Leonurus japonicus* Houtt and has a long history and a wide range of clinical use for lowering postpartum bleeding. Its main ingredients are multiple alkaloids, including leonurine and stachydrine (Ruan et al., 2003; Chen et al., 2010; Liu et al., 2017), which can stimulate contractions of uterine smooth muscle both *in vitro* and *in vivo* (Jin et al., 2004; Xie et al., 2015; Chen et al., 2019).

Motherwort injection is often IM administered to the uterus in addition to IV oxytocin (Zhao et al., 2014). In clinical practice, this Chinese herbal injection and IM oxytocin are often used in addition to IV oxytocin. These additions are added for maintaining the muscular contraction effects to lower the risk of PPH. Although earlier trials showed that the addition of motherwort injection to IV oxytocin had a better effect than IV oxytocin alone (Chen et al., 2018; Yu et al., 2019), these trials applied strict exclusion criteria, as a result of which the study population was highly selected. In addition, these trials primarily administered motherwort injection between 12 and 72 h after cesarean section, while its targeting use and effect at the point of cesarean section were not tested.

More importantly, there is no clear evidence regarding the comparative effectiveness of IM motherwort injection versus IM oxytocin to the uterus for preventing PPH when either of them is added to IV oxytocin. The lack of this evidence has a limited appropriate choice of treatment options, and guideline recommendations for this specific question are opinion based. Therefore, by using an established large pregnancy registry database, we conducted a cohort study to investigate the comparative effectiveness of IM motherwort injection versus IM oxytocin for preventing PPH during cesarean section.

PATIENTS AND METHODS

Study Setting and Data Source

We performed a cohort study using a pregnancy registry database at West China Second University Hospital (i.e., West China Women and Children Hospital) of Sichuan University. This hospital is the leading medical center for obstetrics care and referral center in China and has an average of more than 8,000 deliveries every year (West China Second Univers). This research was conducted in accordance with the reporting of studies conducted using observational routinely collected health data statements for pharmacoepidemiology (RECORD-PE) (Langan et al., 2018).

A multidisciplinary research team composed of epidemiologists, statisticians, obstetricians, and informatics experts was organized to establish the pregnancy registry database. We included pregnant women who were registered in the first trimester and had records for any of the pregnancy outcomes, including miscarriage, induced abortion, and delivery, between January 1, 2015, and November 30, 2019. We excluded patients if they were not registered at the first trimester, lost to follow-up, or transferred to other hospitals after registration.

We used the information from electronic medical record (EMR) data to establish the database. Data extracted from the EMR system were saved separately into different files, and data linkage was performed by using a unique identifier or a combination of multiple identifiers. We used the unique patient ID number and the last menstrual period (LMP) and/or expected date of childbirth (EDC) to identify a pregnancy, thus

forming a cohort from the first antenatal visit to delivery that involved data about antenatal care (e.g., pre-pregnancy conditions and prenatal visits) and hospitalization (e.g., admission examination, admission diagnosis, treatments, discharge diagnoses, etc.). We cleaned the data by using transparent and pre-specified rules, including the development of variable dictionaries, standardization of semi-structured or structured variables, identification, and handling of outliers. Finally, a preliminary multidimensional database of pregnancies registry was developed.

To validate the accuracy of the database, we reviewed all information extracted from 600 randomly selected medical records. The consistency of linkage between the pregnancy registry and hospitalization for pregnancies reached 99%. The gestational comorbidities, pregnancy complications, and maternal and fetal outcomes were recorded by the *International Classification of Diseases*, 10th edition (ICD-10). This study was approved by the Ethics Committee of West China Second University Hospital of Sichuan University (2016 No.145-03) and registered at the clinicaltrials.gov (NCT04607499).

Study Population and Comparison

We included pregnant women who underwent cesarean section at or over 24 weeks of gestation and routinely received oxytocin by IV drip during the third stage of labor. We excluded pregnant women with incomplete information regarding the timing of cesarean section.

In order to reliably evaluate the treatment effect of motherwort injection in a real-world setting, we aimed to choose an active comparator to minimize indication bias. Thus, we compared the use of IM motherwort injection versus IM oxytocin, both of which were administered to the uterus (i.e., IM motherwort injection as the exposure group and IM oxytocin as the control group). Across all the patients, IV oxytocin was administered as a baseline treatment. Information regarding data of prescriptions and route of administration were uniformly extracted from EMR.

Outcomes

The primary outcome of interest was the incidence of PPH, which was defined as blood loss of no less than 500 ml within the first 2 h after delivery. The amount of blood loss was uniformly measured by clinicians and nurses according to the Guideline for Management of PPH in routine clinical practice (Liu, 2014) and was recorded into the surgery records in EMR. We also investigated the severe PPH as the secondary outcome, which was diagnosed as blood loss of more than 1,000 ml within the first 2 h.

Covariates

We selected candidate covariates based on clinical experiences and medical literature. The potentially relevant covariates were demographic and gestational characteristics, including maternal age, nation, years of education, non-permanent residents, residence of location, and pre-pregnancy body mass index (BMI), multiple gestations, parity, use of *in vitro* fertilization (IVF), and scarred uterus; gestational comorbidities, including anemia, other hematological diseases (thrombocytosis, thrombocytopenia, and antiphospholipid antibody syndrome), hepatitis virus B (HBV) infection, cardiac diseases, respiratory diseases, subclinical hypothyroidism, hyperthyroidism, pelvic diseases, and chronic hypertension during pregnancy; gestational complications, including gestational diabetes mellitus (GDM), preeclampsia, eclampsia, gestational hypertension, placenta previa, placenta implantation, intrahepatic cholestasis of pregnancy (ICP), premature rupture of membranes, polyhydramnios, oligohydramnios and fetal macrosomia; and medications during the day of cesarean section, including IV magnesium sulfate, IV dexamethasone, IM insulin, IM ademetionine, IV cephalosporins, IV lincosamides, β-lactams and β-lactamase inhibitors, IM ergometrine maleate, IV tranexamic acid, IM carboprost tromethamine injection, and IV and IV push carbetocin.

Of those, demographic characteristics were recorded in the antenatal care, and diagnosis information (i.e., comorbidities, complications, and other characteristics) was identified using the ICD-10 code, as well as surgical records, preanesthetic visits, and preoperative notes. Medications were identified from the prescription records.

Data Analysis

We described the distribution of baseline characteristics. Given the differences in baseline characteristics and indication in medications between treatment and comparator, we used the propensity-score matching (PSM) method to achieve prognostic balances, which was conducted with a 1:1 matching ratio by using the nearest neighbor method algorithm with a caliper of 0.2 SD (Caliendo and Kopeinig, 2008; Abadie and Imbens, 2016). The propensity score (i.e., the probability of receiving IM motherwort injection) was calculated by logistic regression models of covariates mentioned above (Abadie and Imbens, 2016). We used the Matching package in the R software to estimate the risk difference (RD) and odds ratio (OR) in average treatment effect.

We evaluated the balance of baseline characteristics within the cohorts before and after propensity scoring with standardized mean difference (SMD) of characteristics distribution. If SMD between two groups was less than 0.2, a good balance between groups was achieved (Abadie and Imbens, 2016).

To validate the robustness of the results, we conducted sensitivity analyses, including the use of 1:4 matching and inverse-probability weighting (Survey package and Reshape2 package) and applying traditional multivariable logistic regression in the original full cohort. We also used blood loss above 1,000 ml in the first 24 h after delivery to verify the robustness of the outcome measure, which was recorded as a diagnosis of PPH in the summary of medical records in our hospital. All analyses were two-tailed, and a p-value < 0.05 was considered statistically significant. All statistical analyses were performed using R (R version R 3.4.1).

RESULTS

The pregnancy registry database between January 1, 2015, and December 30, 2019, included a total of 51,925 pregnant women who underwent deliveries. Of those, 29,961 pregnant women

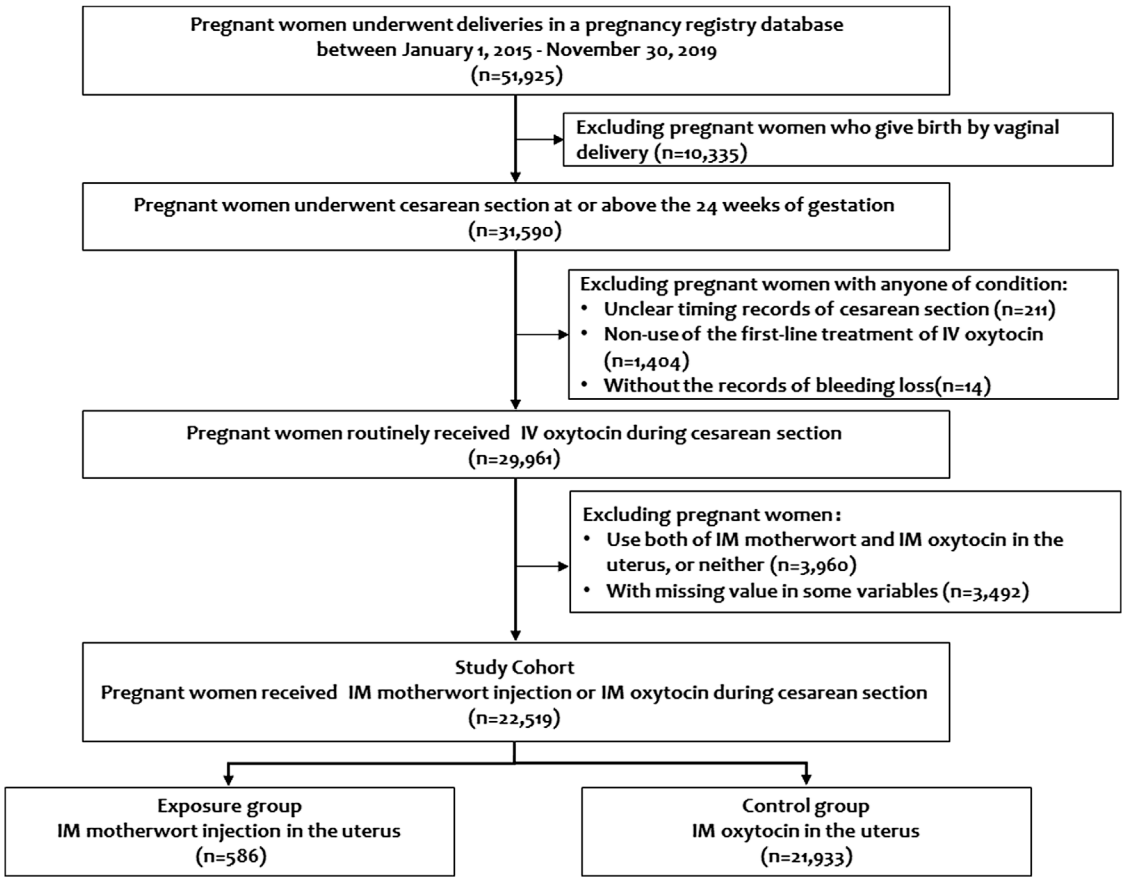

FIGURE 1 | Study flowchart.

routinely received IV oxytocin during cesarean section. After exclusion of pregnant women using both IM motherwort and IM oxytocin, or neither, our cohort finally included 22,519 pregnant women, in whom 586 pregnant women received IM motherwort injection, and 21,933 received IM oxytocin. No missing values occurred among treatments and outcomes of interest (**Figure 1**).

Patient Characteristics at Baseline

In the full cohort, there were significant differences in the distribution among baseline characteristics between IM motherwort injection and IM oxytocin. Before matching, pregnant women who received IM motherwort injection had a higher proportion of pelvic diseases, placenta previa, at baseline. They were also more likely to be older, to be urban residents and permanent residents, to be pregnant *via* IVF, to receive higher education, and to have more multiple pregnancies (**Table 1**). In addition, there were significant differences in the use of mediations during the day of cesarean section, including higher rates of IV dexamethasone, IM insulin, IV/IM cephalosporins, IM ergometrine maleate, IM carboprost tromethamine injection, and IV carbetocin among women receiving IM motherwort injection (**Table 1**).

After PSM, 586 (100%) pregnant women in the IM motherwort injection group were matched to those in the IM oxytocin group. In the PS-matched cohort (586: 586), the differences in baseline characteristics between women who received IM motherwort injection and IM oxytocin were minimized (SMD ≤ 0.2 for all covariates) (**Tables 1** and **2**).

Treatment Effects by Propensity-Score Matching Analysis

The primary and secondary outcomes after matching are listed in **Table 2**. Among 1,172 pregnant women in the PS-matched cohort, IM motherwort injection was associated with a lower risk of blood loss during cesarean section compared with the use of IM oxytocin (for blood loss of 500 ml or more, 76/586, 13.0% versus 224/586, 38.2%; absolute RD −25.26%, 95% CI −30.04% to −20.47%; OR 0.25, 95% CI 0.18–0.32; and for blood loss ≥1,000 ml, 14/586, 2.4% versus 35/586, 6.0%; RD −3.58%, 95% CI −5.87% to −1.30%; OR 0.25, 95% CI 0.20 to 0.71).

Four sensitivity analyses showed similar results, all showing that IM motherwort injection was associated with a lower risk of blood loss during cesarean section compared with the use of IM oxytocin (**Table 3**).

TABLE 1 | Baseline characteristics of pregnant women before and after PS matching.

	Before matching, n (%)			After matching, n (%)		
	IM motherwort injection (n = 586)	IM oxytocin (n = 21,933)	SMD	IM motherwort injection (n = 586)	IM oxytocin (n = 586)	SMD
Demographics and gestational characteristics						
Maternal age, mean (SD)	33.02 (4.26)	32.20 (4.25)	0.192	33.02 (4.26)	32.30 (3.93)	0.175
Years of education ≤12	435 (74.2)	14,876 (67.8)	0.142	435 (74.2)	428 (73.0)	0.027
Ethnic minorities	570 (97.3)	21,528 (98.2)	0.059	570 (97.3)	574 (98.0)	0.045
Urban residents	531 (90.6)	19,108 (87.1)	0.111	531 (90.6)	517 (88.2)	0.078
Non-permanent residents	54 (9.2)	3,490 (15.9)	0.203	54 (9.2)	69 (11.8)	0.084
Pre-pregnancy BMI, mean (SD)	21.12 (2.68)	21.29 (2.73)	0.065	21.12 (2.68)	21.24 (2.70)	0.045
Multiple gestations	61 (10.4)	1,471 (6.7)	0.133	61 (10.4)	90 (15.4)	0.148
Multipara	279 (47.6)	9,163 (41.8)	0.118	279 (47.6)	242 (41.3)	0.127
Use of IVF	49 (8.4)	552 (2.5)	0.260	49 (8.4)	34 (5.8)	0.100
Scarred uterus	236 (40.3)	8,112 (37.0)	0.068	236 (40.3)	195 (33.3)	0.145
Gestational complications						
Anemia	60 (10.2)	2,061 (9.4)	0.028	60 (10.2)	75 (12.8)	0.080
Other hematological diseases	14 (2.4)	1,237 (5.6)	0.166	14 (2.4)	23 (3.9)	0.088
HBV infection	31 (5.3)	1,266 (5.8)	0.021	31 (5.3)	32 (5.5)	0.008
Cardiac diseases	20 (3.4)	1,611 (7.3)	0.175	20 (3.4)	29 (4.9)	0.077
Respiratory diseases	7 (1.2)	154 (0.7)	0.051	7 (1.2)	6 (1.0)	0.016
Subclinical hypothyroidism	37 (6.3)	1,012 (4.6)	0.075	37 (6.3)	29 (4.9)	0.059
Hyperthyroidism	39 (6.7)	1,810 (8.3)	0.061	39 (6.7)	45 (7.7)	0.040
Pelvic diseases	29 (4.9)	387 (1.8)	0.177	29 (4.9)	19 (3.2)	0.086
Chronic hypertension during pregnancy	9 (1.5)	634 (2.9)	0.092	9 (1.5)	16 (2.7)	0.083
Gestational complications						
Gestational diabetes mellitus	136 (23.2)	4,619 (21.1)	0.052	136 (23.2)	112 (19.1)	0.100
Preeclampsia	16 (2.7)	614 (2.8)	0.004	16 (2.7)	24 (4.1)	0.075
Gestational hypertension	12 (2.0)	648 (3.0)	0.058	12 (2.0)	19 (3.2)	0.074
Placenta previa	76 (13.0)	1,717 (7.8)	0.169	76 (13.0)	85 (14.5)	0.045
Placenta implantation	16 (2.7)	745 (3.4)	0.039	16 (2.7)	21 (3.6)	0.049
Intrahepatic cholestasis of pregnancy	52 (8.9)	1,712 (7.8)	0.039	52 (8.9)	53 (9.0)	0.006
Premature rupture of membranes	41 (7.0)	4,077 (18.6)	0.352	41 (7.0)	74 (12.6)	0.190
Polyhydramnios	33 (5.6)	1,075 (4.9)	0.033	33 (5.6)	40 (6.8)	0.049
Oligohydramnios	13 (2.2)	831 (3.8)	0.092	13 (2.2)	10 (1.7)	0.037
Fetal macrosomia	31 (5.3)	1,738 (7.9)	0.106	31 (5.3)	39 (6.7)	0.058
Concomitant medications						
IV magnesium sulfate	26 (4.4)	1,053 (4.8)	0.017	26 (4.4)	41 (7.0)	0.110
IV dexamethasone	153 (26.1)	3,760 (17.1)	0.219	153 (26.1)	163 (27.8)	0.038
IM insulin	146 (24.9)	4,097 (18.7)	0.151	146 (24.9)	109 (18.6)	0.153
IM ademetionine	18 (3.1)	780 (3.6)	0.027	18 (3.1)	18 (3.1)	<0.001
Cephalosporins	575 (98.1)	21,027 (95.9)	0.132	575 (98.1)	569 (97.1)	0.067
Lincosamides	17 (2.9)	1,146 (5.2)	0.118	17 (2.9)	25 (4.3)	0.073
β-Lactams and β-lactamase inhibitors	9 (1.5)	217 (1.0)	0.049	9 (1.5)	10 (1.7)	0.014
IM ergometrine maleate	522 (89.1)	3,078 (14.0)	2.273	522 (89.1)	518 (88.4)	0.022
IV tranexamic acid	154 (26.3)	5,219 (23.8)	0.057	154 (26.3)	200 (34.1)	0.172
IM carboprost tromethamine injection	118 (20.1)	6,226 (28.4)	0.193	118 (20.1)	138 (23.5)	0.083
Carbetocin	575 (98.1)	3,146 (14.3)	3.152	575 (98.1)	568 (96.9)	0.077

Note. IM, intramuscular; IV, intravenous; SMD, standardized mean difference; BMI, body mass index; IVF, in vitro *fertilization; HBV, hepatitis B virus; PS, propensity score.*

TABLE 2 | Primary and secondary outcomes after matching.

Outcome	IM motherwort injection	IM oxytocin	Absolute risk difference (95%CI)	*p*-Value	Odds ratio (95% CI)	*p*-Value
	n (%)	n (%)				
Primary outcome						
PPH	76 (13.0)	224 (38.2)	−25.26% (−30.04%, −20.47%)	<0.001	0.25 (0.18, 0.32)	<0.001
Secondary outcome						
Severe PPH	14 (2.4)	35 (6.0)	−3.58% (−5.87%, −1.30%)	0.007	0.39 (0.20, 0.71)	0.002

Note. Propensity score (PS) matching variables: all variables in the baseline.
PPH, postpartum hemorrhage; IM, intramuscular.

TABLE 3 | Sensitivity analyses.

Sensitivity analyses	Outcome	Absolute risk difference (95%CI)	*p* value	Odds ratio (95%CI)	*p* value
1:4 matching	PPH	−14.55%, 95%CI (−16.81%, −12.29%)	<0.001	0.39 (0.34, 0.46)	<0.001
	Severe PPH	−2.52%, 95%CI (−3.59%, −1.45%)	<0.001	0.47 (0.34, 0.65)	<0.001
Inverse-probability weighting	PPH	−22.19%, 95%CI (−27.72%, −16.66%)	0.005	0.29 (0.22, 0.39)	<0.001
	Severe PPH	−3.00%, 95%CI (−5.66%, −0.35%)	<0.001	0.47 (0.25, 0.86)	0.015
Traditional logistic regression	PPH	−12.94%, 95%CI (−14.29%, −11.59%)	<0.001	0.18 (0.13, 0.24)	<0.001
	Severe PPH	−1.09%, 95%CI (−1.64%, −0.54%)	0.001	0.36 (0.17, 0.69)	0.004
Blood loss ≥1,000 ml within 24 hours	—	−4.78%, 95%CI (−7.37%, −2.19%)	<0.001	0.37 (0.21, 0.64)	0.001

Note. Adjusted for all variables listed in the baseline.
PPH, postpartum hemorrhage.

DISCUSSION

Interpretation of Findings and Implications

Using an established large database in the integrative medicine context, our study showed that, among women undergoing cesarean section who had received IV oxytocin as a baseline treatment, the additional use of IM motherwort injection to the uterus could achieve a lower risk of PPH compared to the additional use of IM oxytocin. Our findings were also robust to a series of sensitivity analyses, suggesting higher confidence about the findings.

Subject to ethical constraints, real-world data were generally recognized as an important data source for evaluating treatments effects for medication use during pregnancy and postpartum periods, especially for identifying the optimal treatment pattern or more benefit to the population. To the best of our knowledge, this was the first study showing that IM motherwort injection during cesarean section was more effective than IM oxytocin in the lowering of PPH risk in a real-world setting. In addition, this study provided clear evidence to clinicians for the choice of medications after IV oxytocin has been used. Our findings also supported that motherwort injection, as a traditional Chinese medicine (TCM) intervention, may have an important role in obstetric care. In particular, our study suggested that if all women had received IV oxytocin, IM motherwort injection may be a better choice than IM oxytocin, as excessive administration of oxytocin was observed to reduce oxytocin receptor availability (Gabriel et al., 2015).

Chinese herbal injections were often used in combination with western medicine as a complementary treatment. This model of practice is often complex. We have clearly demonstrated that, in the real-world setting, combination therapy was often used for preventing PPH among women undergoing cesarean section. For example, IM motherwort injection may be used on the uterus during a cesarean section on the basis of IV oxytocin. These complexities not only represent the realities of clinical practice but also require more sophisticated methods for reliably parceling out the effects of TCM interventions.

In order to achieve this purpose, we have applied an active-control design to minimize indication bias, which is usually a common and major threat to the validity of findings. We additionally applied PSM and other alternative statistical approaches in the sensitivity analyses to further control confounding. As it turned out, the consistency in findings by using different methods has strengthened the strength of inference about the effect of motherwort injection. Additionally, we particularly measured the blood loss within the first 2 h after delivery, rather than overall blood loss in earlier studies, which also has helped identify the effects of motherwort injection at the point of cesarean section.

Overall, our study not only reliably demonstrated the comparative effectiveness of motherwort injection for preventing PPH during cesarean section, but also the strategies we used in our study suggested a paradigm for investigating the treatment effect of Chinese herbal medicine in the real-world setting. The methodology used in our study represents a step forward in controlling potential confounding and indication bias, which are the two primary concerns in the assessment of treatment effects of TCM interventions.

Comparison With Previous Studies

Previous clinical evidence including randomized controlled trials (RCTs) and their systematic reviews investigated the efficacy of motherwort injection for preventing PPH, by using different controls, such as motherwort injection plus oxytocin versus oxytocin alone, or motherwort injection versus oxytocin. Nevertheless, only a few trials compared motherwort injection versus IM oxytocin, and these trials did not find significant differences in PPH incidence. More importantly, serious methodological limitations were presented in those RCTs, such as unclear randomization and small sample sizes, thus substantially lowering the credibility of their findings (Chen et al., 2018).

In addition, our study included a broader and more heterogeneous population than previous RCTs. In comparison, previous trials excluded women if they had placenta previa, placental abruption, anemia or abnormal coagulation function, cardiopulmonary disease, previous placental adhesion, and other severe risk factors. The excluded pregnant women were typically high-risk populations; arbitrary exclusion would have hampered the generalizability of findings. Our study covered a broader population, and the findings may be applicable to clinical practice. Overall, our findings may be considered together with the existing trials to form a more inclusive body of evidence about motherwort injection.

Mechanism of Action

Pharmacologically, motherwort injection is primarily composed of leonurine and stachydrine (Xie et al., 2015), and earlier studies have shown that its main materials include the water extraction part and *n*-butanol part (Li et al., 2014). Leonurine hydrochloride

can significantly reduce the amount of vaginal bleeding and shorten bleeding time in rats by promoting frequency and tension of uterine contraction and increasing the level of serum estradiol (Xia et al., 2009). In addition, motherwort was found to cause increasing intrauterine pressure in women (Chan et al., 1983). Our findings together with earlier trials have confirmed that motherwort injection is effective for preventing PPH.

Strengths and Limitations

Our study has a few strengths. Firstly, our study used a cohort with a large sample size and high-quality data to reliably assess the treatment effect of motherwort injection in an integrative medicine setting. To the best of our knowledge, this is the first real-world study involving more than 20,000 pregnant women to investigate the effectiveness of add-on use of motherwort injection with the first-line treatment for preventing PPH. The findings fill a gap to evaluate the effectiveness of motherwort injection in real-world practice, and the findings had reasonably good generalizability. Secondly, given the potential confounding and indication bias, we applied a novel active-controlled study design to minimize indication bias, which was common in a real-world practice setting; we also applied PS matching to construct a counterfactual framework to ensure comparability between groups, and we further conducted sensitivity analyses to consolidate the results. We believe that the use of rigorous methods and the robustness of findings through various sensitivity analyses have substantially improved the strength of findings.

Our study also has some limitations. Although we have adjusted for a variety of confounders in PSM, residual confounding from unmeasured factors may not be ruled out. Nevertheless, based on the clinical experiences and medical literature, as well as the large database, we included a large number of covariates for adjustment compared with similar studies. These factors covered a wide range of important domains that may affect the effect estimation, including demographic and gestational characteristics, gestational comorbidities, gestational complications, and concomitant medications. Residual confounding is a challenging issue in the real-world setting, and we are hopeful that other studies could validate our results. Secondly, the data of this study were derived from West China Second University Hospital, a major referral center and teaching hospital that has more than 20,000 deliveries in our cohort. The findings may not be completely generalizable to all healthcare settings.

CONCLUSION

In summary, our study showed that IM motherwort injection during cesarean section was more effective than IM oxytocin in the lowering of PPH risk in a real-world setting. The findings supported that motherwort injection, as a TCM intervention, may be the better choice of medications after IV oxytocin has been used. In response to the complexities of combination therapy in an integrative medicine setting, our study suggested a paradigm for investigating the treatment effect of Chinese herbal medicine using real-world data, particularly the methodology used to control potential confounding and indication bias.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because of privacy restrictions. The data that support the findings of the study are available from the corresponding authors upon reasonable request. Requests to access the datasets should be directed to sunxin@wchscu.cn.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of West China Second University Hospital of Sichuan University. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements. Written informed consent was not obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

M-XL contributed to the drafting of the manuscript, data collection, and data analysis. JT and XS contributed to the study design, data analysis and management, drafting of the manuscript, interpretation of results, and critical revision of the manuscript. C-RL, MC, H-CS, WW, X-CL, LL, Y-NQ, JW, S-YH, and KZ participated in data collection and data analysis. Y-QX and X-HL participated in the interpretation of results and critically revised the manuscript. All authors have read and approved the final manuscript. All authors agree to be accountable for all aspects of work, ensuring integrity and accuracy.

FUNDING

The study was supported by funding from the National Natural Science Foundation of China (72174132, 71974138, 72004148), National Key Research and Development Program of China (2021YFC2701503), Sichuan Youth Science and Technology Innovation Research Team (2020JDTD0015), China Medical Board (CMB19-324), and 1·3·5 Project for Disciplines of Excellence, West China Hospital, Sichuan University (ZYYC08003).

ACKNOWLEDGMENTS

This study is based on data obtained from the West China Second University Hospital of Sichuan University.

REFERENCES

Abadie, A., and Imbens, G. W. (2016). Matching on the Estimated Propensity Score. *Econometrica* 84 (2), 781–807. doi:10.3982/ECTA11293

Caliendo, M., and Kopeinig, S. (2008). Some Practical Guidance for the Implementation of Propensity Score Matching. *J. Econ. Surv.* 22 (1), 31–72. doi:10.1111/j.1467-6419.2007.00527.x

Chan, W. C., Wong, Y. C., Kong, Y. C., Chun, Y. T., Chang, H. T., and Chan, W. F. (1983). Clinical Observation on the Uterotonic Effect of I-Mu Ts'ao (Leonurus Artemisia). *Am. J. Chin. Med.* 11 (01n04), 77–83. doi:10.1142/S0192415X83000136

Chen, S., Xie, B., Tian, H., Ding, S., and Lu, C. (2019). Traditional Chinese Medicine Yimucao Injection Combined with Western Medicine for Preventing Postpartum Hemorrhage after Cesarean Section: A Systematic Review and Meta-Analysis. *Evidence-Based Complement. Altern. Med.* 2019, 1–13. doi:10.1155/2019/7475151

Chen, W., Yu, J., Tao, H., Cai, Y., Li, Y., and Sun, X. (2018). Motherwort Injection for Preventing Postpartum Hemorrhage in Pregnant Women with Cesarean Section: A Systematic Review and Meta-Analysis. *J. Evid. Based Med.* 11 (4), 252–260. doi:10.1111/jebm.12300

Chen, Z., Wu, J. B., Liao, X. J., Yang, W., and Song, K. (2010). Development and Validation of an UPLC-DAD-MS Method for the Determination of Leonurine in Chinese Motherwort (Leonurus Japonicus). *J. Chromatogr. Sci.* 48 (10), 802–806. doi:10.1093/chromsci/48.10.802

Gabriel, M. A. M., Fernández, I. O., Martínez, A. M. M., Armengod, C. G., Costarelli, V., Santos, I. M., et al. (2015). Intrapartum Synthetic Oxytocin Reduce the Expression of Primitive Reflexes Associated with Breastfeeding. *Breastfeed. Med.* 10 (4), 209–213. doi:10.1089/bfm.2014.0156

Gallos, I. D., Papadopoulou, A., Man, R., Athanasopoulos, N., Tobias, A., Price, M. J., et al. (2018). Uterotonic Agents for Preventing Postpartum Haemorrhage: A Network Meta-Analysis. *Cochrane Database Syst. Rev.* 12, CD011689. doi:10.1002/14651858.CD011689.pub3

Jin, R-m., Chen, Z-s., Chen, C-x., and Zhang, H-g. (2004). Effects of Motherwort on Dymenorrhea. *Chin. J. Mod. Appl. Pharm.* 21 (2), 90–93.

Khan, K. S., Wojdyla, D., Say, L., Gülmezoglu, A. M., and Van Look, P. F. (2006). WHO Analysis of Causes of Maternal Death: A Systematic Review. *Lancet* 367 (9516), 1066–1074. doi:10.1016/S0140-6736(06)68397-9

Langan, S. M., Schmidt, S. A., Wing, K., Ehrenstein, V., Nicholls, S. G., Filion, K. B., et al. (2018). The Reporting of Studies Conducted Using Observational Routinely Collected Health Data Statement for Pharmacoepidemiology (RECORD-PE). *BMJ* 363, k3532. doi:10.1136/bmj.k3532

Li, D., Xie, X., Peng, C., Xiong, L., and Zhao, X. (2014). Effects of Different Extracts from Leonurus Japonicus Houtt. On Isolated Smooth Muscle of Uterus. *Pharm. Clin. Chin. Materia Med.* 5 (3), 24–27. doi:10.1016/j.phytochem.2017.11.003

Liu, S., Tang, Z., Cui, C., Liu, H., LIiang, Y., Zhang, Y., et al. (2017). Advances in Study on Chemical Constituents of Herba Leonuri. *Henan Traditional Aditional Chin. Med.* 37 (02), 358–362.

Liu, X-h. (2014). Guidelines for the Prevention and Management of Postpartum Haemorrhage. *Chin. J. Obstet. Gynecol.* 2014 (009), 641–646.

Pacheco, L. D., Saade, G. R., and Hankins, G. D. V. (2019). Medical Management of Postpartum Hemorrhage: An Update. *Semin. Perinatol* 43 (1), 22–26. doi:10.1053/j.semperi.2018.11.005

Pantoja, T., Abalos, E., Chapman, E., Vera, C., and Serrano, V. P. (2016). Oxytocin for Preventing Postpartum Haemorrhage (PPH) in Non-facility Birth Settings. *Cochrane Database Syst. Rev.* 4 (4), CD011491. doi:10.1002/14651858.CD011491.pub2

Ruan, J., Du, J., Zeng, Q., and Qian, Z. (2003). Advance in Studies on Chemical Components,pharmacological Effect and Clinical Application of *Leonurus Japonicus*. *Chin. Traditional Herbal Drugs* 34 (11), U015–U9.

Tan, J., Cao, Q., He, Gl., Cai, Yh., Yu, Jj., Sun, X., et al. (2016). Misoprostol versus Ergometrine-Oxytocin for Preventing Postpartum Haemorrhage: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *J. Evid. Based Med.* 9 (4), 194–204. doi:10.1111/jebm.12201

West China Second University Hospital of Sichuan University (). Available from: https://www.motherchildren.com/intro_introduce/ (Accessed December 22, 2021).

World Health Organization (2012). *WHO Recommendations for the Prevention and Treatment of Postpartum Haemorrhage*. Geneva, Switzerland: World Health Organization.

Xia, L., Feihu, C., Fenglai, Y., Weiguo, L., Fan-rong, W., Yan-na, Z., et al. (2009). Effects of Leonurine on Uterus Induced by Drug-Abortion in Rats. *Chin. J. Clin. Pharmacol. Ther.* 14 (5), 481–486.

Xie, X-f., Xiong, L., Li, D., Li, Y., Li, M., and Peng, C. (2015). Effect of the Active Site of Motherwort Injection on the Activity of the Isolated Rat Uterus. *Chin. Traditional Patent Med.* 37 (5), 1103–1106.

Yu, J., Cai, Y., Su, G., and Li, Y. (2019). Motherwort Injection for Preventing Postpartum Hemorrhage in Women with Vaginal Delivery: A Systematic Review and Meta-Analysis of Randomized Evidence. *Evidence-Based Complement. Altern. Med.* 2019, 1–9. doi:10.1155/2019/1803876

Zhao, X., Peng, C., and Xie, X. (2014). Advances in Studies on Clinical and Pharmacological Effects of Motherwort on Prevention of Postpartum Hemorrhage. *Pharm. Clin. Chin. Materia Med.* 5 (4), 41–45.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewers JW and YF declared a shared affiliation, with no collaboration, with one of the authors HS to the handling editor at the time of the review.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Li, Liu, Chen, Shang, Wang, Luo, Li, Qi, Xiong, Huang, Wang, Zou, Liu, Tan and Sun. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Public Health

BRIEF RESEARCH REPORT
published: 24 March 2022
doi: 10.3389/fpubh.2022.783153

The Incidence and Characteristics of Perinatal Stroke in Beijing: A Multicenter Study

Qianqian Xia[1†], Zhao Yang[1†], Yao Xie[1†], Ying Zhu[1], Zixin Yang[2], Mingyan Hei[2], Yingxue Ding[3], Weijing Kong[3], Limin Kang[4], Shengnan Yang[4], Yabo Mei[5], Zhichun Feng[5], Li Zhang[6], Yanzhe Lei[6], Ping Wang[7], Jingyu Dong[8], Li Yang[8], Jun Ju[9], Hesheng Chang[9], Shina Zhan[10], Jinqian Yu[10], Peng Zhang[11], Ran Wang[11], Hui Guo[12], Xinli Liu[12], Huaqing Tan[13], Yakun Liu[13], Zhenzong Zhang[14], Lixia Zhen[14], Jinting Yan[15], Zhan Liu[15], Chunxia Yang[16], Qingqing Wang[16], Jinfeng Wang[17], Lifang Sun[17], Huihui Zeng[18], Jing Li[18], Wenwen Qu[19], Xuemei Wang[19,20], Guiying Liu[21], Xi Yang[21], Xinxin Liu[22], Xuefeng Zhang[22], Xiaojing Xu[23], Yanan Gu[23], Hui Long[24], Li Zhang[24], Lili Liu[1], Zezhong Tang[1]* and Xinlin Hou[1]**

[1] Pediatric Department, Peking University First Hospital, Beijing, China, [2] Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing, China, [3] Beijing Friendship Hospital, Capital Medical University, Beijing, China, [4] Department of Neonatology, Children's Hospital, Capital Institute of Pediatrics, Beijing, China, [5] Neonatal Intensive Care Unit of Faculty of Pediatrics, Chinese People's Liberation Army (PLA) General Hospital, Beijing, China, [6] Pediatric Department, Haidian Maternal and Child Health Hospital, Beijing, China, [7] Department of Neonatology, Beijing Obstetrics and Gynecology Hospital, Capital Medical University, Beijing, China, [8] Pediatric Department, Tongzhou Maternal and Child Health Hospital, Beijing, China, [9] Pediatric Department, Beijing Chao-Yang Hospital, Capital Medical University, Beijing, China, [10] Pediatric Department, Shunyi Maternal and Children's Hospital of Beijing Children's Hospital, Beijing, China, [11] Pediatric Department, Beijing Miyun Maternal and Child Health Hospital, Beijing, China, [12] Pediatric Department, Fengtai Maternal and Child Health Hospital, Beijing, China, [13] Pediatric Department, Beijing Daxing Maternal and Child Care Hospital, Beijing, China, [14] Pediatric Department, The First Hospital of Fangshan, Beijing, China, [15] Pediatric Department, Beijing Mentougou Hospital, Beijing, China, [16] Pediatric Department, Changping Women and Children Health Care Hospital, Beijing, China, [17] Beijing Pinggu Maternal and Child Health Hospital, Beijing, China, [18] Pediatric Department, Huairou Maternal and Child Health Care Hospital, Beijing, China, [19] Department of Pediatrics, Peking University Shougang Hospital and Capital Medical University, Beijing, China, [20] General Respiratory Department of Beijing Jingdu Children's Hospital, Beijing, China, [21] Department of Pediatrics, Beijing Anzhen Hospital, Capital Medical University, Beijing, China, [22] Pediatric Department, Fifth Medical Center of Chinese People's Liberation Army (PLA) General Hospital, Beijing, China, [23] Pediatric Department, The First Hospital of Tsinghua University, Beijing, China, [24] Department of Pediatrics, Chinese People's Liberation Army (PLA) General Hospital, Beijing, China

OPEN ACCESS

Edited by:
Bin Jiang,
Peking University, China

Reviewed by:
Jeroen Dudink,
University Medical Center Utrecht, Netherlands
Dejan Nikolic,
University of Belgrade, Serbia

***Correspondence:**
Lili Liu
liulili0205@163.com
Zezhong Tang
tangzz308@163.com
Xinlin Hou
xinlin.hou@pkufh.com

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Public Health Policy, a section of the journal Frontiers in Public Health

Received: *25 September 2021*
Accepted: *31 January 2022*
Published: *24 March 2022*

Citation:
Xia Q, Yang Z, Xie Y, Zhu Y, Yang Z, Hei M, Ding Y, Kong W, Kang L, Yang S, Mei Y, Feng Z, Zhang L, Lei Y, Wang P, Dong J, Yang L, Ju J, Chang H, Zhan S, Yu J, Zhang P, Wang R, Guo H, Liu X, Tan H, Liu Y, Zhang Z, Zhen L, Yan J, Liu Z, Yang C, Wang Q, Wang J, Sun L, Zeng H, Li J, Qu W, Wang X, Liu G, Yang X, Liu X, Zhang X, Xu X, Gu Y, Long H, Zhang L, Liu L, Tang Z and Hou X (2022) The Incidence and Characteristics of Perinatal Stroke in Beijing: A Multicenter Study. Front. Public Health 10:783153. doi: 10.3389/fpubh.2022.783153

Objective: To assess the incidence, risk factors, and clinical characteristics of perinatal stroke in Beijing.

Methods: This multicenter prospective study included all the live births from 17 representative maternal delivery hospitals in Beijing from March 1, 2019 to February 29, 2020. Neonates with a stroke were assigned to the study group. Clinical data, including general information, clinical manifestations, and risk factors, were collected. Up until 18 months after birth, neonates were routinely assessed according to the Ages and Stages Questionnaire (ASQ) and/or the Bayley scale. Statistical analysis was done using the chi-squared, *t*-tests, and logistic regression analysis using SPSS version 26.0.

Outcomes: In total, 27 cases were identified and the incidence of perinatal stroke in Beijing was 1/2,660 live births, including 1/5,985 for ischemic stroke and 1/4,788 for hemorrhagic stroke. Seventeen cases (62.96%) of acute symptomatic stroke and convulsions within 72 h (10 cases, 37.04%) were the most common presentations. Ten patients showed no neurological symptoms and were found to have had a stroke through routine cranial ultrasonography after being hospitalized for non-neurological diseases.

The risk factors include primiparity, placental or uterine abruption/acute chorioamnionitis, intrauterine distress, asphyxia, and severe infection. In the study group, 11.1% (3/27) of patients had adverse neurodevelopmental outcomes. The patients in the study group had lower scores for the ASQ than those in the control group in the communication, gross, and fine motor dimensions.

Conclusion: The incidence of perinatal stroke in Beijing was consistent with that in other countries. Routine neuroimaging of infants with risk factors may enable identification of asymptomatic strokes in more patients. Patients who have suffered from a stroke may have neurological sequelae; therefore, early detection, treatment, and regular follow-ups are beneficial for improving their recovery outcomes.

Keywords: neonatal stroke, incidence, diagnosis, clinical characteristics, prognosis

INTRODUCTION

Perinatal stroke is a cerebrovascular disease that occurs in newborns from 20 weeks of gestation to 28 days after birth. The incidence of this type of stroke is one in every 1,600–3,000 live births (1–4). This can involve either the arteries and/or veins, being categorized as hemorrhagic stroke and ischemic stroke (1, 5). Specifically, perinatal strokes can be divided into neonatal arterial ischemic strokes, neonatal cerebral venous sinus thrombosis, and neonatal hemorrhagic strokes. Although the incidence of perinatal strokes is low, it seriously increases the likelihood of an added prognosis. Of those children with perinatal strokes, 40% develop epilepsy after the neonatal period (6), 68% of neonates diagnosed with ischemic strokes develop dyskinesia (7), 43% of these children suffer from cognitive dysfunctions (8, 9), and neurological sequelae occurs in 47% of neonates with hemorrhagic strokes (10). Some patients suffer from an acute onset of symptoms, including seizures, changes in tone, irritability, feeding difficulties, and dystonia, which are classified as an acute symptomatic perinatal stroke.

The leading unit of this project, the Peking University First Hospital, has a fully equipped neonatal neurointensive care unit. This unit is one of the main referral centers for neonatal brain injury in Beijing, actively taking part in the referral, treatment, and long-term follow-ups of neonates with these conditions. In the author's experiences, although the incidence of perinatal strokes is lower than that of hypoxic-ischemic encephalopathy, which is the major cause of neurologic disabilities in term neonates, the former is still rather common. When these children are diagnosed early, particularly through imaging screenings for asymptomatic high-risk infants and receive physical therapy later in their childhood, their prognosis could be improved (11).

The main purpose of this study was to assess the incidence, risk factors, and clinical characteristics of perinatal strokes in Beijing. Specifically, it aimed to analyze its high-risk factors, disease onset time, symptoms, and prognosis. The hope was to supply a basis for the diagnosis and treatment of strokes that occur during the perinatal period among children and to improve their future prognosis. The Peking University First Hospital led the study, but it also includes 16 representative maternity hospitals and three neonatal referral hospitals, thus covering the main administrative regions of Beijing.

METHODS

This was a multicenter prospective study, with 17 of the recruiting hospitals being maternity hospitals and three neonatal referral hospitals for children with severe illnesses in Beijing. The participating maternal delivery hospitals included secondary and tertiary hospitals that covered the central and suburban areas of the city, all of which are the main delivery hospitals in the administrative region of Beijing. The 17 maternity hospitals then conducted follow-ups on all the newborns, who were alive at birth until the diagnosis of perinatal stroke was either excluded or confirmed. Some infants with neurologic presentations were assessed with MRI and then for further evaluation. Neurological symptoms or signs included, but was not limited to the following manifestations: convulsions, frequent myoclonus, altered consciousness (drowsiness, responsiveness, and irritability), abnormal respiratory rhythm, weakened muscle strength, dystonia, postural abnormalities, feeding difficulties (inappetence and emesis), and the abnormal Neonatal Behavioral Neurological Assessment (NBNA) score. Those who did not display neurological symptoms, but who had risk factors that could lead to neurologic disabilities in an infant, maternal and peripartum characteristics, first underwent a craniocerebral ultrasound. If the ultrasound, with an abnormal echo in parenchyma or sinus, showed symptoms or signs of a suspected perinatal stroke, the neonate would be further assessed with an MRI (**Figure 1**). If the participating maternity hospital teams believed it too difficult to diagnose, treat, or follow-up with any of the children, they were referred to the Pediatrics Department of Peking University First Hospital or one of the three neonatal referral hospitals for further diagnosis and treatment. The cases of the control group were matched based on the perinatal stroke cases.

Clinical data collection and regular follow-ups were included in the standardized diagnosis and treatment process of the enrolled newborns. The results of the follow-ups and related data were analyzed. This study's enrollment period was from

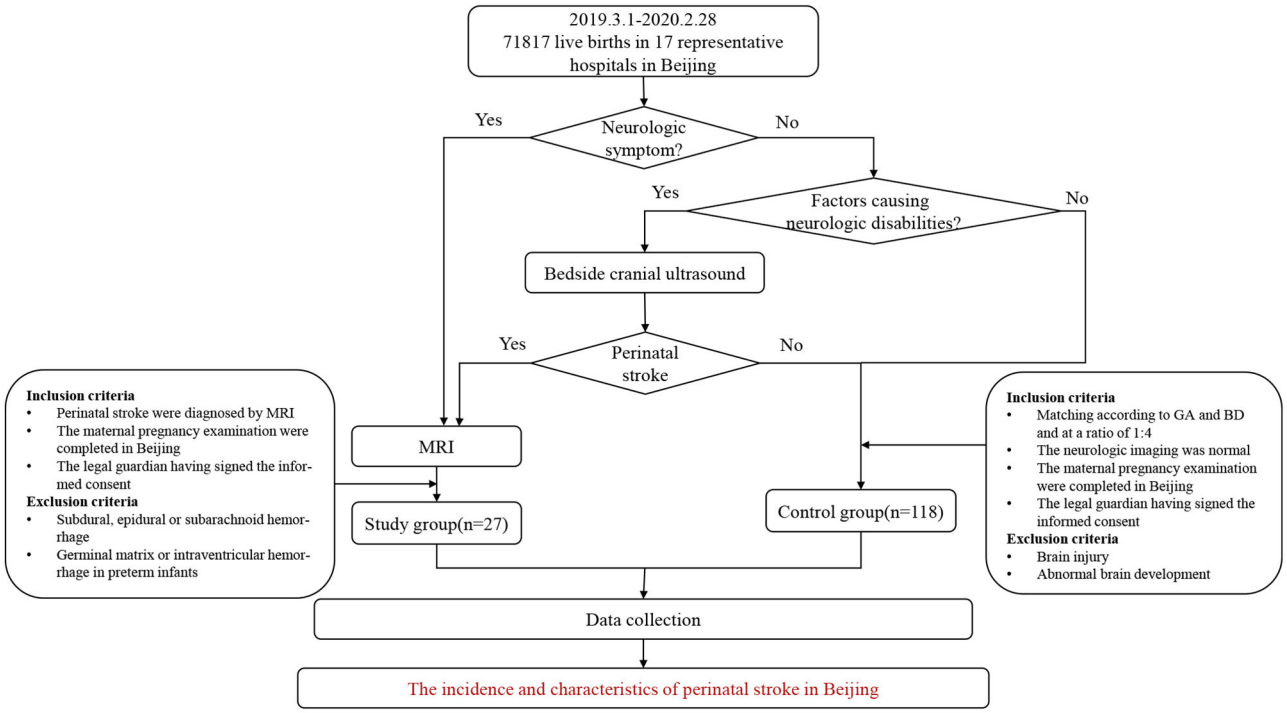

FIGURE 1 | Case ascertainment methods. GA, gestational age; BD, birth date.

March 1, 2019 to February 29, 2020. This study was approved by the Ethics Committee of Peking University First Hospital and a written informed consent was obtained from the legal guardians of the newborns.

PATIENTS

Study Group

All the newborns delivered in the participating units from March 1, 2019 to February 29, 2020 were followed-up by the delivery unit team until 42 days after their birth. If any neurological symptoms or signs appeared during this period, cranial ultrasounds and/or MRIs were performed in a timely manner. For children hospitalized owing to other symptoms (e.g., jaundice and infection), cranial ultrasound and/or MRIs were performed as a routine procedure. All the children in the study group met the inclusion criteria described below and their diagnoses were confirmed by MRIs.

The inclusion criteria were: (1) MRI performed during the neonatal period and the diagnosis of a stroke; in this study, hemorrhagic strokes refer to cerebral parenchymal hemorrhages in full-term infants and near-term infants over 35 weeks of gestational age (12), while ischemic strokes include arterial ischemic strokes and neonatal intracranial venous sinus thrombosis (13); (2) the maternal pregnancy examinations were completed in Beijing; (3) the date of birth of the neonate was between March 1, 2019 and February 29, 2020; (4) the place of birth was a unit involved in this study; and (5) the legal guardian signed the informed consent form.

The exclusion criteria were: (1) subdural, epidural, or subarachnoid hemorrhages; (2) germinal matrix or intraventricular hemorrhages in premature infants; and (3) refusal to sign the informed consent form.

Control Group

The researchers matched the control group to the study group according to gestational age (with a difference in gestational age of 3 days) and date of birth (with a difference of 2 weeks from the study group, at a ratio of 1:4).

The control group also attended follow-ups and met the inclusion criteria below.

The inclusion criteria were: (1) all the maternity examinations during pregnancy were completed in Beijing; (2) the date of birth of the neonate was between March 1, 2019 and February 29, 2020; (3) the place of birth was a unit involved in this study; (4) the imaging examination (cranial ultrasounds and/or cranial MRIs) was normal; and (5) the legal guardian signed the informed consent form.

The exclusion criteria were: (1) brain injury and abnormal brain development caused by other etiologies (e.g., hypoxic-ischemic encephalopathy, bilirubin encephalopathy, hypoglycemic encephalopathy, and brain structural malformations); (2) refusal to undergo brain imaging examination; and (3) refusal to sign the informed consent form.

NEUROIMAGING

Owing to economic restrictions and human resource limitations, this study did not perform routine imaging screening on all

of the participants in the cohort. Only the participants with neurological symptoms and/or pathological conditions were screened. However, all the patients in the cohort were uniformly followed-up to reduce the missed detection rate.

DATA COLLECTION

Specific data were collected using standardized information forms. First, population information, including the number of live births delivered in each unit per month, was collected. Second, collection included general information, such as gestational age, age, and sex. Third, clinical data, including age at presentation, symptoms, signs, treatment history, and medication, were collected. The fourth set of collected information included the potential risk factors, including maternal factors (age at time of delivery, pregnancy complications, primiparity, and parity), perinatal factors (delivery method, intrauterine distress, and acute intrapartum events, including placental or uterine abruption, postnatal resuscitation, meconium-stained amniotic fluid, and umbilical cord around the neck), and neonatal factors (birth weight and infection status). Fifth, imaging data from all the patients diagnosed with or with difficult to diagnose conditions, comprising a 3.0-T MRI [T1, T2, diffusion-weighted imaging, susceptibility-weighted imaging, and T2 fluid-attenuated inversion recovery (FLAIR) sequence] were collected.

Data Preservation

The maternity hospitals involved in this study provided the Peking University First Hospital, the leading unit, monthly with data on the delivery volume and detailed information on confirmed and suspected cases. The specific follow-up period was included in the database only after an expert group's diagnosis was verified and sent back to the corresponding unit. The control group was selected according to the final confirmed cases and all the participating units provided the leading unit with detailed information on the control group every month. After the unit obtained information on a case and on a specific follow-up destination, the leading unit contacted the follow-up unit to check the case information and incorporate it into the research process. The researchers archived the data and checked it every 3 months. Any logical errors or missing information were reported to the source unit. All the follow-up information was uploaded to a network database.

FOLLOW-UP AFTER BIRTH

The initial plan of this study was to perform outpatient follow-ups at the clinic. However, due to coronavirus disease 2019 (COVID-19) pandemic, some cases required WeChat video calls in addition to the outpatient follow-ups. If the partner hospital did not have the required conditions to perform the follow-up procedures, the leading hospital and/or one of the three neonatal referral hospitals would do so. A chief neonatal physician appointed by each unit and who had more than 10 years of work experience conducted the follow-ups. All the enrolled children were contacted by the pediatrics specialist of the Peking University First Hospital, who then conducted the follow-ups at 1, 3, 6, 9, 12, and 18 months postpartum at the designated chief physician's outpatient clinic.

Follow-ups included clinical manifestation, symptoms, and physical and auxiliary examinations of the children. Neurodevelopmental evaluations were performed using the Ages and Stages Questionnaire (ASQ) for all the enrolled children at 36 months postpartum (14). It covers five areas of child development, each with six items: communication, gross motor, fine motor, problem solving, and personal–social domains. Each item represents a developmental milestone for that specific age group. Items responses are based on a three-item scale that showed the capacity of the child to perform that item, with "yes" being assigned 10 points, "sometimes" being assigned five points, and "not yet" being assigned zero points. The sum of the six items for one area forms the area score. This score then aids with figuring out whether the child is developing normally, at risk of suspected developmental delays (borderline), and a diagnosis of developmental delays in an area. Development delays in communication, gross motor, fine motor, problem solving, and personal–social domains were defined as a score below 13.06, 37.38, 34.32, 25.74, and 27.19. Meanwhile, suspected developmental delays in the above five items were defined as a score below 29, 46, 44, 36, and 38, respectively.

The Bayley Scales of Infant and Toddler Development, third edition (Bayley-III), was used for the diagnostic evaluation, if the ASQ results were abnormal (15). Neurodevelopment delays were defined as the Bayley-III cognitive score of <85 or the Bayley-III motor composite score of <85.

QUALITY CONTROL

Diagnosis

The leading unit provided all the cooperating units with diagnostic training by videos and meetings and explained the standards. The cranial MRIs of all the newborns whose diagnosis was difficult were performed by two radiologists of the leading units. Both have over 5 years of experience in reviewing neonatal neurological diseases. Imaging specialists conducted a blind review. The diagnosis was confirmed or excluded only when the neonatal imaging team (which consists of two neuroimaging specialists with more than 10 years of experience from the imaging department of Peking University) agreed. Upon non-agreement, the team discussed the case further until reaching a consensus.

Data Collection

Each cooperating unit arranged fixed data collectors. When the project started, all the cooperating units received training on the process to be followed in terms of the case data collection. The monthly delivery volume reported by the 16 hospitals was verified and recorded and the contents of the various cases' data on the strokes reported by all the cooperating units were verified and feedback every 3 months.

Follow-Up

The leading hospital designed a standardized follow-up manual and distributed it to all the units expected to conduct them. All

the professional physicians involved in the follow-up and who worked in units that had the ability to conduct the ASQ and the Bayley scale were trained on the follow-up procedures before the project began.

Project Progress

The project leader of each cooperating unit monitored the progress longitudinally. They communicated with the leading unit through an online meeting every 3 months to clarify the project progress and conducted offline or online meetings with personnel from all the centers every 6 months. The leading unit reported on project progress and supplied feedback on the problems occurring in the research process of the partner hospitals during the meeting that took place every 3 months.

STATISTICAL ANALYSIS

Numerical data were expressed as frequencies (percentages) and analyzed using the chi-squared tests. Measurement data were expressed as median (minimum-maximum) or mean ± SD and were analyzed by *t*-tests. The logistic regression analysis was applied to the risk factors. Statistical significance was set at $p < 0.05$. The researchers analyzed the statistical data with SPSS software version 26.0.

RESULTS

Demographics and Incidence

During the study period, 71,817 live births were delivered by all the cooperating units, including 42,919 live births in the central area and 28,898 live births in the suburbs. The number of live births in the spring, summer, autumn, and winter seasons were 18,802, 18,186, 18,954, and 17,875, respectively.

A total of 27 cases (**Table 1**) of perinatal strokes were diagnosed in 14 male (52%) and 13 female (48%) neonates. Among the confirmed cases, 15 had hemorrhagic and 12 had ischemic strokes (seven cases of arterial ischemic stroke and five cases of venous sinus thrombosis). Eighteen cases were from the central areas of the city and nine cases were from the suburbs. Spring and summer each saw eight diagnosed cases of perinatal strokes, while autumn saw six diagnosed cases of perinatal strokes and winter saw only five diagnoses cases of perinatal strokes.

Based on this calculation, the incidence of perinatal stroke in Beijing was one in 2,660 live births. In urban areas, this incidence was one in every 2,384 live births and one in 3,210 live births in suburban areas. The incidence rates in spring, summer, autumn, and winter were 1/2,350, 1/2,273, 1/3,159, and 1/3,175, respectively. For perinatal ischemic strokes, the incidence was 1/5,985 and 1/4,788 for hemorrhagic strokes.

Clinical Features

Onset Time of Symptoms

Among neonates with acute symptomatic perinatal strokes (17 neonates), some who had convulsions started to experience the symptoms within 12 h after birth (two neonates). Most symptoms occurred between 12 and 72 h after birth (eight neonates).

There were eight cases of convulsions among neonates with ischemic stroke, with six of them developing their symptoms within 12–72 h after birth. Of these eight patients, six patients had arterial ischemic strokes: five patients involving the left middle cerebral artery and one patient involving the right middle cerebral artery. Additionally, two cases had venous sinus thrombosis involving the bilateral transverse sinuses and sigmoid sinuses, respectively. Of the two cases who developed symptoms within 12 h after birth, one case had an arterial ischemic stroke involving the left middle cerebral artery and the other had a venous sinus thrombosis involving the bilateral sigmoid sinuses.

Among the patients with hemorrhagic stroke, two cases showed convulsions, both of which had symptom onset within 12–72 h after birth. The one case involved the right frontal lobe and the other case involved the left temporal lobe simultaneously.

Onset Symptoms

Of the 27 perinatal stroke cases, 17 (62.96%) patients had acute symptomatic strokes. The onset symptoms included convulsions (10, 37.04%), muscular hypertonia (3, 11.1%), frequent apnea (2, 7.40%), crying sharply (1, 3.70%), and decreased responsiveness (1, 3.70%). Routine cranial ultrasound examinations were performed for other children who were diagnosed with other conditions. These included perinatal hyperbilirubinemia, hypoxic asphyxia, and central nervous system infection and the diagnosis was further confirmed by cranial MRIs (**Table 2**).

Among the children with ischemic strokes, 75% (9/12) patients had an acute symptomatic stroke, of which 66.67% (8/12) patients suffered from convulsions. Five patients had a stroke involving the left middle cerebral artery. In one case, the right middle cerebral artery, bilateral transverse sinuses, and bilateral sigmoid sinuses were also involved. Only one (out of 12; 8.33%) patient showed muscular hypertonia and hemorrhage of the left thalamus. The other three cases were diagnosed based on the results of a conventional cranial ultrasound and their strokes involved the left middle cerebral artery, superior sagittal sinus, and bilateral transverse sinuses.

Among children with hemorrhagic strokes, 53.5% (8/15) presented with an acute symptomatic stroke, of which two cases had convulsions. Of these two, one involved the right frontal lobe and the other involved the right temporal lobe. Two patients had hypertonia, with one case involving multiple parts of the temporo-occipital lobe and cerebellum and the other case involving the left ventricle and left frontal lobe (at the same time). Two patients had frequent apnea, with one case involving the left frontal lobe and the other case involving the two hemorrhage sites in multiple parts of the lateral frontal lobe, left basal ganglia, and bilateral cerebellum. In one case, the patient displayed symptoms of crying sharply, vomiting, and the stroke involved the left frontal lobe. Another patient showed irritability and the stroke involved the bilateral cerebellum.

The remaining seven cases of hemorrhagic strokes with no neurological symptoms were diagnosed based on the results of a routine cranial ultrasound. Additionally, they had a history of perinatal hypoxia and asphyxia, were children older than gestational age, younger than gestational age, born from diabetic

TABLE 1 | Patients' characteristics.

Stroke	Case	Sex	Stroke territory	Co-morbidities	Outcome
Neonatal arterial ischemic stroke	1	F	Anterior LMCA	Asphyxia	N
	2	M	Posterior LMCA	Dehydration; Cholestasis; IDM	N
	3	M	Anterior LMCA	Dehydration	N
	4	M	Main RMCA	Asphyxia	Epilepsy
	5	F	Posterior LMCA	Sepsis; Twins	N
	6	M	Middle LMCA	Hyperbilirubinemia	N
	7	F	Main LMCA	Neonatal infection	N
Cranial venous sinus thrombosis	8	F	Thalamus	IVH	N
	9	M	Superior sagittal sinus	Bacterial meningitis; Sepsis; Pneumorrhagia	N
	10	M	Bilateral transverse sinus	Sepsis (GBS); ARDS; AKI; Thrombocytopenia	N
	11	M	Bilateral sigmoid sinus	Sepsis; Asphyxia; AKI; Coagulation disorders; LGA	N
	12	M	Bilateral transverse sinus	Neonatal pneumonia (Klebsiella pneumoniae); Subarachnoid hemorrhage; Coagulation disorders	N
Neonatal hemorrhagic stroke	13	F	Cerebellum	IDM; LGA; Subarachnoid hemorrhage; Subdural hemorrhage	N
	14	F	Cerebellum	Hyperbilirubinemia; Craniotabes	N
	15	M	Cerebellum	Neonatal pneumonia	N
	16	M	Cerebellum	Hyperbilirubinemia; Subdural hemorrhage	N
	17	M	Frontal lobe	Dehydration; Hypernatremia; Metabolic acidosis; Subdural hemorrhage; IVH; Subarachnoid hemorrhage	Motor delay
	18	F	Cerebellum; Temporal lobe; Occipital lobe	Asphyxia; Metabolic acidosis; Cephalohematoma; IDM	N
	19	F	Temporal lobe; Cerebellum	Hypernatremia	N
	20	F	Frontal lobe	Subdural hemorrhage; IDM	N

(Continued)

TABLE 1 | Continued

Stroke	Case	Sex	Stroke territory	Co-morbidities	Outcome
	21	M	Temporal lobe	Asphyxia; Skull fracture; Subdural hemorrhage; Epidural hemorrhage; Subarachnoid hemorrhage; LGA	Motor and communication delay
	22	M	Temporal lobe	IDM	N
	23	F	Parietal lobe	VSD; PDA; IDM; Metabolic acidosis	N
	24	M	Temporal lobe	Neonatal pneumonia	N
	25	M	Frontal lobe; Cerebellum; Basal ganglia	SGA	N
	26	F	Temporal lobe; Occipital lobe	Craniotabes	N
	27	M	Frontal lobe	ARDS; PPHN	N

LMCA, left middle cerebral artery; RMCA, right middle cerebral artery; F, female; M, male; IDM, infant of diabetic mother; IVH, intraventricular hemorrhage; GBS, group B streptococcus agalactiae; ARDS, acute respiratory distress syndrome; PPHN, persistent pulmonary hypertension of the newborn; AKI, acute kidney injury; LGA, large for gestational age; SGA, small for gestational age; VSD, ventricular septal defect; PDA, patent ductus arteriosus; N, normal.

TABLE 2 | MRI features of neonatal perinatal strokes.

Stroke territory	NIS *n* (%)	NHS *n* (%)
BGT	2 (8)	1 (5)
PLIC	1 (4)	-
Central sulcus	1 (4)	-
Superior sagittal sinus	1 (4)	-
Transverse sinus	2 (8)	-
Sigmoid sinus	1 (4)	-
Cerebellum	-	7 (33)
Frontal lobe	4 (15)	4 (19)
Temporal lobe	5 (19)	6 (29)
Parietal lobe	4 (15)	1 (5)
Occipital lobe	5 (19)	2 (10)

BGT, basal ganglia and thalami; PLIC, posterior limb of the internal capsule; NIS, neonatal ischemic stroke; NHS, neonatal hemorrhagic stroke.

mothers, and had hyperbilirubinemia. The left cerebellum ($n = 2$), bilateral cerebellar left temporal lobe ($n = 1$), bilateral cerebellum ($n = 1$), left temporal ($n = 1$), left parietal ($n = 1$), and occipital lobes ($n = 1$) were involved in these seven cases.

Risk Factors

The risk factors in the study and control groups are shown in **Table 3**. Results showed that the following factors were related to perinatal stroke: the maternal factor of primiparity; the perinatal factors of placental abruption/acute chorioamnionitis, premature rupture of membranes $\geq$ 18 h, intrauterine distress, and asphyxia/resuscitation; and the neonatal factor of severe infection (i.e., sepsis, bacterial meningitis). Of the 27 neonates in the study group, 23 (85%) had one or more of these factors, the most common was a history of perinatal hypoxia (e.g., having presented intrauterine distress or postnatal asphyxia; 16 cases, 59%). Among the 108 neonates in the control group, 35 (32%) had at least one of these factors ($p < 0.05$).

Based on the results of this univariate logistic regression analyses, all the aforementioned-related factors were subjected to the binary logistic regression analysis. The results demonstrated that primiparity, premature rupture of membranes $\geq$ 18 h, intrauterine distress, asphyxia/resuscitation, and severe infection were significantly related to the occurrence of perinatal stroke (details in **Table 4**). Mothers of infants with perinatal strokes were more likely to be primipara [odds ratio (OR) 3.484; 95% CI 1.032–11.759]. In 22% of patients, premature rupture of membranes $\geq$ 18 h was diagnosed compared with only 4% in the comparison group (OR 11.359; 95% CI 1.522–84.759). Cases were more likely to have intrauterine distress (OR 5.479; 95% CI 1.362–22.041), asphyxia/resuscitation (OR 7.001; 95% CI 1.657–29.589), and suffer from severe infection (OR 12.065; 95% CI 1.123–129.611).

Prognosis

All the patients were followed-up regularly until 18 months of age. Twenty-six patients completed the ASQ-3 assessment and one completed the Bayley scale. In the study group, 11.1% (3/27)

TABLE 3 | Analysis of the risk factors of perinatal strokes by maternal, perinatal, and neonatal factors and groups.

	Study group *n* (%)	Control group *n* (%)	X^2	*P*
Maternal factors				
Mother's age ≥ 35 years	4 (15)	33 (31)	2.690	0.077
Diabetes	6 (22)	22 (20)	0.045	0.508
Hypertension	3 (11)	14 (13)	0.067	0.546
Autoimmune disease	1 (4)	7 (6)	0.299	0.697
Multiple pregnancy	1 (4)	4 (4)	0.000	0.739
Primiparity	15 (56)	36 (33)	4.538	0.045
Gestational age < 35 weeks	1 (4)	4 (4)	0.000	0.739
Perinatal factors				
Cesarean delivery	11 (41)	46 (43)	0.030	0.52
MSAF	6 (22)	16 (15)	0.869	0.254
Umbilical cord around neck	3 (11)	8 (7)	0.396	0.383
PROM ≥ 18 h	6 (22)	4 (4)	10.800	0.01
Acute chorioamnionitis	2 (7)	1 (1)	7.793	0.025
Intrauterine distress	16 (59)	29 (27)	10.208	0.002
Asphyxia / resuscitation	12 (44)	9 (8)	21.443	0.000
Neonatal factors				
Male	13 (48)	53 (49)	0.007	0.552
SGA / LGA	4 (15)	19 (18)	0.118	0.493
Severe infection	4 (15)	2 (2)	8.547	0.015

PROM, premature rupture of membranes; MSAF, meconium-stained amniotic fluid; SGA, small for gestational age; LGA, large for gestational age.

TABLE 4 | The logistic regression analysis of risk factors for perinatal strokes.

Factors	*P*	OR	95% CI
Primiparity	0.044	3.484	1.032∼11.759
Intrauterine distress	0.017	5.479	1.362∼22.041
PROM ≥ 18 h	0.018	11.359	1.522∼84.795
Asphyxia / resuscitation	0.008	7.001	1.657∼29.589
Severe infection	0.000	12.065	1.123∼129.611

PROM, premature rupture of membranes; OR, odds ratio; CI, confidence interval.

of the patients had adverse neurodevelopmental outcomes. One patient had an ischemic stroke accompanied by convulsions and seizures within 12 h after birth, which affected the right middle cerebral artery. This patient was treated with oral antiepileptic drugs and had no clinical seizures. However, his electroencephalogram (EEG) was still abnormal at 18 months of age.

The other two patients had hemorrhagic strokes. One patient, whose parents withdraw the treatment shortly after birth, had a right frontal hemorrhage (4.9 × 5.9 cm) and the ASQ scores showed that gross and fine movements were below the threshold. The other patient suffered from a hemorrhage in the right temporal lobe (0.9 × 0.7 cm) and their ASD scores showed that the gross and fine movements and communication dimensions were below the threshold.

The ASQ scores of children in both the groups showed that patients who had convulsions had lower scores than those in the control group for total scores, communication, gross motor, and fine motor dimensions (**Table 5**) and the difference was statistically significant ($p < 0.05$).

DISCUSSION

This multicenter prospective study described the incidence and clinical characteristics of perinatal strokes in Beijing, with the aim to provide a basis for diagnosis and prognosis improvement. The incidence was 1 in 2,660 live births, which is consistent with earlier reports (1/1,600–1/3,000) (2). There was no statistically significant difference in incidence between regions, nor among seasons, which might be related to the limited number of cases in this study.

Previous studies have reported that arterial ischemic strokes are the most common type of stroke in neonates, but the evidence in this study showed that the incidence for hemorrhagic strokes was higher (2). This may be due to the routine neuroimaging procedures for patients with no neurological symptoms in this study, as was the case for approximately half of the neonatal hemorrhagic stroke cases in this study. Instead, these patients, who were diagnosed with perinatal strokes without neurological symptoms, had been hospitalized for neonatal hyperbilirubinemia and perinatal hypoxia and asphyxia. Through the routine use of bedside cranial ultrasounds, the strokes were found coincidently; this urged the unit team to run a brain MRI, which then confirmed the diagnosis. In addition, three asymptomatic patients (25% of patients with ischemic strokes in the study sample) were diagnosed through screening. This indicates that neonates with risk factors such

TABLE 5 | The Ages and Stages Questionnaire (ASQ) scores according to stroke type and symptoms.

ASQ dimension	$\bar{X}$	σ	*T*	*P*
Total scores				
Ischemic	249.58	15.87	−0.663	0.514
Hemorrhagic	256.33	32.21		
Convulsion	239.00	9.45	−2.253	**0.033**
Non-convulsion	263.82	6.38		
Symptom onset time <12 h	246.25	31.48	−0.671	0.516
Symptom onset time within 12–72 h	255.00	20.91		
Communication				
Ischemic	47.92	6.20	−0.535	0.598
Hemorrhagic	49.67	10.60		
Convulsion	43.50	8.51	−2.305	**0.030**
Non-convulsion	50.88	7.75		
Symptom onset time < 12 h	45.83	10.62	−1.144	0.277
Symptom onset time within 12–72 h	51.00	7.41		
Gross motor				
Ischemic	51.67	2.46	0.000	1
Hemorrhagic	51.67	7.94		
Convulsion	48.00	7.14	−2.353	**0.036**
Non-convulsion	53.82	4.15		
Symptom onset time <12 h	50.00	7.69	−0.882	0.407
Symptom onset time within 12–72 h	52.00	4.47		
Fine motor				
Ischemic	51.67	2.46	0.000	1
Hemorrhagic	51.67	7.94		
Convulsion	45.00	5.27	−3.722	**0.002**
Non-convulsion	52.35	4.37		
Symptom onset time < 12 h	47.92	6.89	0.000	1
Symptom onset time within 12–72 h	48.00	4.47		
Problem solving				
Ischemic	50.00	6.03	−0.459	0.651
Hemorrhagic	51.00	5.07		
Convulsion	50.50	6.43	−0.808	0.433
Non-convulsion	52.35	4.37		
Symptom onset time < 12 h	50.83	6.33	−1.111	0.347
Symptom onset time within 12–72 h	52.00	4.47		
Personal-social				
Ischemic	49.58	6.20	−1.556	0.135
Hemorrhagic	53.00	4.92		
Convulsion	49.50	5.99	−1.372	0.188
Non-convulsion	52.64	5.34		
Symptom onset time < 12 h	51.67	6.15	−0.107	0.917
Symptom onset time within 12–72 h	52.00	5.70		

Bold values represents $P < 0.05$.

as severe infection or perinatal hypoxia should be routinely examined by a cranial ultrasound, as this may ensure that they are diagnosed early and receive timely interventions. These results also show that there may be cases in which hemorrhagic strokes were missed in neonates because of the lack of neurological manifestations. Therefore, subsequent studies should further expand the scope of screening to obtain more accurate data.

This study showed that the most common location involved in the stroke was the left middle cerebral artery, consistent with the literature results (16). This may be related to the characteristics of fetal circulation. Specifically, at first, either a placental thrombosis can enter the umbilical vein or an embolus can form locally in the heart. Then, either one can pass through the foramen ovale and eventually enter the left

cerebral artery or directly enter the left cerebral artery through the foramen ovale, as described by prior research (17). Despite these suspicions, research shows that the etiology of neonatal hemorrhagic strokes is unclear and that the most common lesion of such strokes is in the temporal lobe (18). In this study, 40% of patients with hemorrhagic strokes had temporal lobe hemorrhages. At the same time, a relatively high proportion of patients (46.7%) had cerebellar hemorrhages and most patients were asymptomatic and diagnosed by routine cranial ultrasound screenings. Although small-volume cerebellar hemorrhages have a relatively good prognosis, those that experienced them may develop dysplasia, including motor and language impairment at about 2 years of age (19). This fact once again suggests the need for actively screening high-risk patients while paying attention to the presence of cerebellar hemorrhages, which can be imaged through the additional mastoid fontanelle windows of ultrasounds.

It has been reported in the literature that primiparity, intrauterine distress, asphyxia, and postnatal resuscitation are closely related to perinatal strokes (18, 20–25). Intrauterine distress and asphyxia often lead to the hypoxia-ischemia procedure and temporary hypoxia can lead to vasomotor dysfunction, damage to the vascular endothelium, and hypercoagulability. These complications are the cause of hemorrhages or ischemia in the arteries or veins after hypoxia (20–23). This study showed that 59% of patients in the study group had a history of intrauterine distress, asphyxia, or postnatal resuscitation and this rate was significantly higher than that in the control group (27%). Primiparity may contribute to prolonged labor, which increases the risk of perinatal hypoxia and asphyxia (20). Furthermore, infection is a common risk factor for perinatal stroke (25, 26). In this study, two patients in the study group had acute chorioamnionitis owing to placental pathology. Additionally, the proportion of patients with severe infections was significantly higher in the study group than in the control group. Acute chorioamnionitis is likely to release cytokines, causing diffuse arterial injury, thrombosis, and intracranial arteritis in the fetus (23, 27, 28). Severe infections can also induce vascular endothelial injury and hypercoagulability (1). Further, the results showed that the risk of a perinatal stroke was 12.065 times higher in the case group than in the control group. Regarding other complications, 22% of patients in the study group had a history of premature rupture of membranes and this proportion was significantly higher than that in the control group (4%). Among children with premature rupture of membranes in both the groups, the risk of perinatal strokes was 11.359 times higher in the case group than that in the control group. Indeed, earlier investigations have found that inflammatory markers in amniotic fluid are related to perinatal strokes (29). Thus, this study speculates that the premature rupture of membranes causes inflammation and results in a perinatal stroke. However, this pathophysiology requires further laboratory tests and clinical evidence.

In general, patients with perinatal strokes have a poor prognosis, as 9–40% of them show adverse outcomes (6–8). Noticeably, some patients with stroke show developmental abnormalities after 2 years of age (2, 16, 30, 31). In this study, 11% of the neonates showed adverse neurodevelopmental outcomes, denoting that the rate was generally lower than that reported in the literature. This may be related to the early detection and active follow-up enabled by the procedures taken in this study and that were proposed despite the children going through an asymptomatic period. It was early detection and active follow-ups that facilitated the early detection of neurological disabilities and timely rehabilitation. However, this study had a follow-up period of 18 months after birth, thus could not assess these abnormalities. Longer follow-up periods are required in future studies.

STUDY LIMITATIONS

In this study, not all the children had a routine neuroimaging follow-up and, therefore, the actual incidence of perinatal strokes may have been higher. Besides, the relatively small number of children with perinatal strokes in this study may limit the assessment of risk factors and outcomes. Furthermore, the relatively short duration of the follow-up limits long-term outcome assessments. In addition, the researchers compared perinatal strokes in two categories ischemic stroke and hemorrhagic stroke, which make comparisons with other studies difficult.

CONCLUSION

In this study, the incidence of perinatal strokes in Beijing was found to be one in 2,660 live births and the first symptom in most newborns within 12–72 h after birth was convulsions. The results highlight how cranial neuroimaging may be used to detect asymptomatic strokes in neonates with associated risk factors, especially hemorrhagic strokes. Further, primiparity, intrauterine distress, asphyxia, resuscitation, premature rupture of membranes over 18 h, and severe infections are all high-risk factors for perinatal strokes and the most common risk factor is perinatal hypoxia. Patients diagnosed with a perinatal stroke are at high risk for developing neurological sequelae. Therefore, early screening, treatment after diagnosis, regular follow-ups, and rehabilitation are important to improve their prognosis.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Peking University First Hospital Ethics Committee. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin. Written informed consent was obtained from the individual(s) and minor(s)' legal guardian/next of kin, for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

QX, ZhY, and YX collected and analyzed the data, wrote the manuscript, and conducted statistical analysis. YZ was responsible for the interpretation of the MRI results. ZiY, MH, YD, WK, LK, SY, YM, ZF, LZhe, YLe, PW, JD, LY, JJ, HC, SZ, JY, PZ, RW, HG, XinlL, HT, YLi, ZZ, LZha, JY, ZL, CY, QW, JW, LS, HZ, JL, WQ, XW, GL, and XY provided information on the number of births and/or data of patients in their respective hospitals. XinxL, XZ, XX, YG, HL, and LZha participated in discussions surrounding the project design and implementation and provided suggestions. LL, ZT, and XH designed the research and supervised the implementation, guided the data analysis, and revised the article. All the authors have contributed to the manuscript and approved the submitted version of the manuscript.

FUNDING

The Beijing Municipal Science and Technology Commission (No. Z191100006619049 to XH) and the Beijing Municipal Science and Technology Commission (No. Z211100002921050 to LL) supported this study.

REFERENCES

1. Nelson KB, Lynch JK. Stroke in newborn infants. *Lancet Neurol.* (2004) 3:150–8. doi: 10.1016/S1474-4422(04)00679-9
2. Dunbar M, Kirton A. Perinatal stroke: mechanisms, management, and outcomes of early cerebrovascular brain injury. *Lancet Child Adolesc Health.* (2018) 2:666–76. doi: 10.1016/S2352-4642(18)30173-1
3. Laugesaar R, Kolk A, Tomberg T, Metsvaht T, Lintrop M, Varendi H, et al. Acutely and retrospectively diagnosed perinatal stroke: a population-based study. *Stroke.* (2007) 38:2234–40. doi: 10.1161/STROKEAHA.107.483743
4. Chabrier S, Husson B, Dinomais M, Landrieu P, Nguyen S. New insights (and new interrogations) in perinatal arterial ischemic stroke. *Thromb Res.* (2011) 127:13–22. doi: 10.1016/j.thromres.2010.10.003
5. Lynch JK. Epidemiology and classification of perinatal stroke. *Semin Fetal Neonatal Med.* (2009) 14:245–9. doi: 10.1016/j.siny.2009.07.001
6. Laugesaar R, Vaher U, Lõo S, Kolk A, Männamaa M, Talvik I, et al. Epilepsy after perinatal stroke with different vascular subtypes. *Epilepsia Open.* (2018) 3:193–202. doi: 10.1002/epi4.12104
7. Golomb MR, Garg BP, Saha C, Azzouz F, Williams LS. Cerebral palsy after perinatal arterial ischemic stroke. *J Child Neurol.* (2008) 23:279–86. doi: 10.1177/0883073807309246
8. Lee CC, Lin JJ, Lin KL, Lim WH, Hsu KH, Hsu JF, et al. Clinical manifestations, outcomes, and etiologies of perinatal stroke in Taiwan: comparisons between ischemic, and hemorrhagic stroke based on 10-year experience in a single institute. *Pediatr Neonatol.* (2017) 58:270–7. doi: 10.1016/j.pedneo.2016.07.005
9. Hajek CA, Yeates KO, Anderson V, Mackay M, Greenham M, Gomes A, et al. Cognitive outcomes following arterial ischemic stroke in infants and children. *J Child Neurol.* (2014) 29:887–94. doi: 10.1177/0883073813491828
10. Bruno CJ, Beslow LA, Witmer CM, Vossough A, Jordan LC, Zelonis S, et al. Haemorrhagic stroke in term and late preterm neonates. *Arch Dis Child Fetal Neonatal Ed.* (2014) 99:F48–53. doi: 10.1136/archdischild-2013-304068
11. Douglas-Escobar M, Weiss MD. Hypoxic-ischemic encephalopathy: a review for the clinician. *J Am Med Assoc Pediatr.* (2015) 169:397–403. doi: 10.1001/jamapediatrics.2014.3269
12. Cole L, Dewey D, Letourneau N, Kaplan BJ, Chaput K, Gallagher C, et al. Clinical characteristics, risk factors, and outcomes associated with neonatal hemorrhagic stroke: A population-based case-control study. *J Am Med Assoc Pediatr.* (2017) 171:230–8. doi: 10.1001/jamapediatrics.2016.4151
13. Dudink J, Mercuri E, Al-Nakib L, Govaert P, Counsell SJ, Rutherford MA, et al. Evolution of unilateral perinatal arterial ischemic stroke on conventional and diffusion-weighted MR imaging. *Am J Neuroradiol.* (2009) 30:998–1004. doi: 10.3174/ajnr.A1480
14. Singh A, Yeh CJ, Blanchard SB. Ages and stages questionnaire: a global screening scale. *Bol Med Hosp Infant Mex.* (2017) 74:5–12. doi: 10.1016/j.bmhimx.2016.07.008
15. Albers CA, Grieve AJ. Test review: Bayley, N. (2006). Bayley scales of infant and toddler development– third edition. San Antonio, TX: Harcourt Assessment. *J Psychoeduc Assess.* (2007) 25:180–90. doi: 10.1177/0734282906297199
16. Wagenaar N, Martinez-Biarge M, van der Aa NE, van Haastert IC, Groenendaal F, Benders MJNL, et al. Neurodevelopment after perinatal arterial ischemic stroke. *Pediatrics.* (2018) 142:e20174164. doi: 10.1542/peds.2017-4164
17. Arnaez J, Garcia-Alix A. Extracerebral thrombosis in symptomatic neonatal arterial ischemic stroke. *Eur J Paediatr Neurol.* (2017) 21:687–8. doi: 10.1016/j.ejpn.2017.05.004
18. Fluss J, Dinomais M, Chabrier S. Perinatal stroke syndromes: similarities and diversities in aetiology, outcome and management. *Eur J Paediatr Neurol.* (2019) 23:368–83. doi: 10.1016/j.ejpn.2019.02.013
19. Williamson WD, Percy AK, Fishman MA, Cheek WR, Desmond MM, Lafevers N, et al. Cerebellar hemorrhage in the term neonate: developmental and neurologic outcome. *Pediatr Neurol.* (1985) 1:356–60. doi: 10.1016/0887-8994(85)90072-4
20. Li C, Miao JK, Xu Y, Hua YY, Ma Q, Zhou LL, et al. Prenatal, perinatal and neonatal risk factors for perinatal arterial ischaemic stroke: a systematic review and meta-analysis. *Eur J Neurol.* (2017) 24:1006–15. doi: 10.1111/ene.13337
21. Lee J, Croen LA, Backstrand KH, Yoshida CK, Henning LH, Lindan C, et al. Maternal and infant characteristics associated with perinatal arterial stroke in the infant. *J Am Med Assoc.* (2005) 293:723–9. doi: 10.1001/jama.293.6.723
22. Armstrong-Wells J, Johnston SC, Wu YW, Sidney S, Fullerton HJ. Prevalence and predictors of perinatal hemorrhagic stroke: results from the Kaiser Pediatric Stroke Study. *Pediatrics.* (2009) 123:823–28. doi: 10.1542/peds.2008-0874
23. Sorg AL, Von Kries R, Klemme M, Gerstl L, Beyerlein A, Lack N, et al. Incidence and risk factors of cerebral sinovenous thrombosis in infants. *Dev Med Child Neurol.* (2021) 63:697–704. doi: 10.1111/dmcn.14816
24. Fitzgerald KC, Williams LS, Garg BP, Carvalho KS, Golomb MR. Cerebral sinovenous thrombosis in the neonate. *Arch Neurol.* (2006) 63:405–9. doi: 10.1001/archneur.63.3.405
25. Wu YW, Miller SP, Chin K, Collins AE, Lomeli SC, Chuang NA, et al. Multiple risk factors in neonatal sinovenous thrombosis. *Neurology.* (2002) 59:438–40. doi: 10.1212/WNL.59.3.438
26. Nwosu ME, Williams LS, Edwards-Brown M, Eckert GJ, Golomb MR. Neonatal sinovenous thrombosis: presentation and association with imaging. *Pediatr Neurol.* (2008) 39:155–61. doi: 10.1016/j.pediatrneurol.2008.06.001
27. Giraud A, Guiraut C, Chevin M, Chabrier S, Sébire G. Role of perinatal inflammation in neonatal arterial ischemic stroke. *Front Neurol.* (2017) 8:612. doi: 10.3389/fneur.2017.00612
28. Guiraut C, Cauchon N, Lepage M, Sébire G. Perinatal arterial ischemic stroke is associated to materno-fetal immune activation and intracranial arteritis. *Int J Mol Sci.* (2016) 17:1980. doi: 10.3390/ijms17121980
29. Bernson-Leung ME, Boyd TK, Meserve EE, Danehy AR, Kapur K, Trenor CC, et al. Placental pathology in perinatal stroke: a retrospective case-control study. *J Pediatr.* (2018) 195:39–47. doi: 10.1016/j.jpeds.2017.11.061
30. Grunt S, Mazenauer L, Buerki SE, Boltshauser E, Mori AC, Datta AN, et al. Incidence and outcomes of symptomatic neonatal arterial ischemic stroke. *Pediatrics.* (2015) 135:e1220–8. doi: 10.1542/peds.2014-1520

31. Billinghurst LL, Beslow LA, Abend NS, Uohara M, Jastrzab L, Licht DJ, et al. Incidence and predictors of epilepsy after pediatric arterial ischemic stroke. *Neurology.* (2017) 88:630–7. doi: 10.1212/WNL.0000000000003603

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor declared a shared affiliation with several of the authors WQ and XW at time of review.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Xia, Yang, Xie, Zhu, Yang, Hei, Ding, Kong, Kang, Yang, Mei, Feng, Zhang, Lei, Wang, Dong, Yang, Ju, Chang, Zhan, Yu, Zhang, Wang, Guo, Liu, Tan, Liu, Zhang, Zhen, Yan, Liu, Yang, Wang, Wang, Sun, Zeng, Li, Qu, Wang, Liu, Yang, Liu, Zhang, Xu, Gu, Long, Zhang, Liu, Tang and Hou. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Pharmacology

SYSTEMATIC REVIEW
published: 31 March 2022
doi: 10.3389/fphar.2022.781192

The Use of Oral Anticoagulation Is Not Associated With a Reduced Risk of Mortality in Patients With COVID-19: A Systematic Review and Meta-Analysis of Cohort Studies

Meng-Fei Dai[1,2†], Si-Tong Guo[3†], Yi-Jun Ke[4], Bao-Yan Wang[1,5], Feng Yu[2], Hang Xu[1,5], Zhi-Chun Gu[6*] and Wei-Hong Ge[1,5*]*

[1]Department of Pharmacy, China Pharmaceutical University Nanjing Drum Tower Hospital, Nanjing, China, [2]School of Basic Medicine and Clinical Pharmacy, China Pharmaceutical University, Nanjing, China, [3]Department of Pharmacy, The People's Hospital of Guangxi Zhuang Autonomous Region, Nanning, China, [4]Department of Pharmacy, The Anqing Hospital Affiliated to Anhui Medical University, Anqing, China, [5]Department of Pharmacy, Nanjing Drum Tower Hospital, The Affiliated Hospital of Nanjing University Medical School, Nanjing, China, [6]Department of Pharmacy, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

Background: Hypercoagulability and thromboembolic events are associated with poor prognosis in coronavirus disease 2019 (COVID-19) patients. Whether chronic oral anticoagulation (OAC) improve the prognosis is yet controversial. The present study aimed to investigate the association between the chronic OAC and clinical outcomes in COVID-19 patients.

Methods: PubMed, Embase, Web of Science, and the Cochrane Library were comprehensively searched to identify studies that evaluated OAC for COVID-19 until 24 July 2021. Random-effects model meta-analyses were performed to pool the relative risk (RR) and 95% confidence interval (CI) of all-cause mortality and intensive care unit (ICU) admission as primary and secondary outcomes, respectively. According to the type of oral anticoagulants [direct oral anticoagulants (DOACs) or vitamin K antagonists (VKAs)], subgroup and interaction analyses were performed to compare DOACs and VKAs. Meta-regression was performed to explore the potential confounders on all-cause mortality.

Results: A total of 12 studies involving 30,646 patients met the inclusion criteria. The results confirmed that chronic OAC did not reduce the risk of all-cause mortality (RR: 0.92; 95% CI 0.82–1.03; $p = 0.165$) or ICU admission (RR: 0.65; 95% CI 0.40–1.04; $p = 0.073$) in patients with COVID-19 compared to those without OAC. The chronic use of DOACs did not reduce the risk of all-cause mortality compared to VKAs ($P_{interaction} = 0.497$) in subgroup and interaction analyses. The meta-regression failed to detect any potential confounding on all-cause mortality.

Conclusion: COVID-19 patients with chronic OAC were not associated with a lower risk of all-cause mortality and ICU admission compared to those without OAC, and the results were consistent across DOACs and VKA subgroups.

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Klara Komici,
University of Molise, Italy
Amelia Filippelli,
University of Salerno, Italy

***Correspondence:**
Hang Xu
njglyyxh@126.com
Zhi-Chun Gu
guzhichun213@163.com
Wei-Hong Ge
6221230@sina.com

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Drugs Outcomes Research and Policies, a section of the journal Frontiers in Pharmacology

Received: 22 September 2021
Accepted: 07 March 2022
Published: 31 March 2022

Citation:
Dai M-F, Guo S-T, Ke Y-J, Wang B-Y, Yu F, Xu H, Gu Z-C and Ge W-H (2022) The Use of Oral Anticoagulation Is Not Associated With a Reduced Risk of Mortality in Patients With COVID-19: A Systematic Review and Meta-Analysis of Cohort Studies. Front. Pharmacol. 13:781192. doi: 10.3389/fphar.2022.781192

Systematic Review Registration: clinicaltrials.gov, identifier CRD42021269764.

Keywords: COVID-19, oral anticoagulant, direct oral anticoagulants, vitamin K antagonists, mortality, meta-analysis

1 INTRODUCTION

Coronavirus disease 2019 (COVID-19) is a respiratory infection caused by severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) with a high rate of morbidity and mortality (Wu and McGoogan, 2020). It has been recognized as a pandemic by the World Health Organization (World Health Organization, 2021). As of February 2022, >400 million patients have been confirmed with COVID-19 globally, and >5 million people have deceased (https://coronavirus.jhu.edu/map.html) (John Hopkins University, 2022). Although COVID-19 primarily affects the respiratory system, SARS-CoV-2 infection can lead to vascular inflammation and endothelial dysfunction due to hypoxia and excessive inflammation. A consequent hypercoagulability causes pulmonary embolism with micro- and macro-thrombosis in the lung vessels (Ackermann et al., 2020; Belen-Apak and Sarıalioğlu, 2020; Varga et al., 2020), rendering it as one of the main reasons for death and poor prognosis in COVID-19 patients. Previous studies have shown that venous thrombosis is common in patients infected with SARS-CoV-2, with an estimated venous thromboembolism (VTE) incidence of 25% (Cui et al., 2020; Zhang et al., 2020). Also, thrombosis was observed in autopsy studies (Wichmann et al., 2020).

The guidelines recommend that prophylactic low-molecular weight heparin (LMWH) should be administrated to all hospitalized patients with COVID-19 to decrease mortality (Moores et al., 2020). Nonetheless, according to recent studies, despite the use of apparently adequate thrombosis prophylaxis, COVID-19 patients have a higher incidence of VTE than those with other infectious diseases (Cui et al., 2020; Zhang et al., 2020; Spiegelenberg et al., 2021). Approximately half of VTE incidence occurred at or within 24 h of admission (Lodigiani et al., 2020). Therefore, it was hypothesized that in the early stages of COVID-19, the transition from asymptomatic to severe respiratory failure might be sudden, and the anticoagulation therapy started after the advanced stage of COVID-19 may be insufficient (Aslan et al., 2021). The oral anticoagulation (OAC) treatment prior to the earliest stages of COVID-19 infection may exert a protective effect for COVID-19-related outcomes.

Several studies have evaluated the effects of chronic therapeutic OACs in reducing the risk of mortality in patients with COVID-19; however, the conclusions were controversial (Aslan et al., 2021; Buenen et al., 2021; Gülcü et al., 2021; Spiegelenberg et al., 2021), and high-quality evidence is limited. Hitherto, no systematic review and meta-analysis have been conducted to investigate the association between chronic OAC and clinical outcomes in COVID-19 patients. Furthermore, the protective effects of the different types of oral anticoagulants, direct oral anticoagulants (DOACs), or vitamin K antagonists (VKAs) in COVID-19 patients are controversial. Therefore, this systematic review and meta-analysis were conducted to investigate the benefits and hazards of chronic OACs in COVID-19 patients comprehensively and rigorously.

2 MATERIALS AND METHODS

The present study was performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement. The protocol for this review was prospectively registered in PROSPERO (CRD42021269764).

2.1 Data Sources and Search Strategy

The following electronic databases were searched comprehensively to identify all the potentially eligible articles from inception to 24 July 2021, without language restriction: PubMed, Embase, Web of Science, and Cochrane Library. In brief, the search strategy was identified based on the PICOS format, including three key concepts: 1) COVID-19, 2) oral anticoagulant, and 3) human study (**Supplementary Table S1**). The detailed search strategy is presented in **Supplementary Table S2**. For each of the concepts, the authors mapped the relevant keywords and Medical Subject Headings (MeSH). Also, the references were assessed manually to ensure that all the relevant studies were included. Two authors (MFD and STG) independently searched the databases, and any disagreements were resolved by consulting a third author (HX).

2.2 Study Selection and Outcomes

Studies were considered eligible and included if they 1) were randomized controlled trials (RCTs) or cohort studies; 2) consisted of patients with confirmed COVID-19; 3) included patients in the exposure group with chronic OAC; and 4) with outcomes including mortality or intensive care unit (ICU) admission. The exclusion criteria were as follows: 1) studies with lack of a comparison group; 2) review articles, case reports, case–control studies, cross-sectional studies, and clinical trial registries; and 3) studies with unrelated outcomes or unreported outcomes. All medical interventions after hospital admission, including anticoagulation and pharmacological treatments for COVID-19, were performed at the discretion of the medical team. The primary outcome was all-cause mortality, defined as death from any cause in COVID-19 patients. The secondary outcome was ICU admission.

2.3 Data Extraction

Information was extracted using a pre-specified form consisting of items, such as the first author's name, publication year, study design, country, patient demographics, clinical characteristics, sample size, the type of the anticoagulant, and outcomes.

The adjusted effect estimates were preferred to the unadjusted effect estimates and were extracted for reducing the confounding factors. Also, the effect estimates and tabular data obtained from

studies comprising propensity-matched groups were treated as adjusted estimates. If the reports did not provide adjusted effect estimates, we extracted crude effect estimates. The data were extracted by two authors (MFD and STG) independently and cross-checked by a third author (HX) for completeness and accuracy.

2.4 Quality Assessment

Two authors (MFD and STG) independently assessed the methodological quality of each study. The Cochrane Collaboration Risk of Bias Tool (Higgins et al., 2011) was used to identify the risk of bias of RCTs. The Newcastle–Ottawa Scale (NOS) score was determined to be nine while assessing three quality domains, selection, comparability, and outcome in the cohort studies (Wells et al., 2021). The total NOS score <5 points defined a study with a potentially high risk of bias (Wells et al., 2021). An additional reviewer (HX) was consulted to resolve any discrepancies.

2.5 Statistical Analysis

Relative risk (RR) and 95% confidence interval (CI) were calculated for the predetermined outcomes and pooled in the random-effects model due to the heterogeneity of study populations. The generic inverse variance method was used for weighing (Cumpston et al., 2019). The hazard ratio (HR) was also calculated when available. The rare case hypothesis was not used when the outcome incidence was >20%, and the odds ratio (OR) and HR were transformed to RR based on the approximation suggested by VanderWeele (2020). I^2 was used to evaluate the heterogeneity of the study: I^2 < 50%, 50.0–75.0%, and >75.0% represented as low, moderate, and substantial heterogeneity, respectively (Cumpston et al., 2019). Subgroup analysis was conducted based on different types of oral anticoagulants (DOACs or VKAs), and an interaction analysis was applied to evaluate the risk difference between DOACs and VKAs. Sensitivity analyses were performed to strengthen the robustness of the results with the leave-one-out method. Random-effects meta-regression was conducted to test the pre-specified demographic characteristics in order to explore the potential confounding factors associated with all-cause mortality, using a restricted maximum likelihood. The potential publication bias was assessed by evaluating the study effects using visual inspection of funnel plots and Egger's regression test. The meta-analyses were performed with STATA 15.1 (Stata, College Station, TX, United States).

3 RESULTS

3.1 Study Selection and Characteristics

A total of 1,310 studies were retrieved from electronic databases through a comprehensive search. After screening the titles and abstracts, 339 duplicates were removed, and 922 records were excluded. After assessing full text, 12 studies (Aslan et al., 2021; Buenen et al., 2021; Chocron et al., 2021; Covino et al., 2021; Denas et al., 2021; Fröhlich et al., 2021; Gülcü et al., 2021; Harrison et al., 2021; Iaccarino et al., 2021; Rivera-Caravaca et al., 2021; Russo et al., 2021; Spiegelenberg et al., 2021) involving 30,646 patients were selected according to the inclusion criteria. The PRISMA flowchart illustrates the results of the literature search and study selection (**Figure 1**). The characteristics of the included studies are summarized in **Table 1**. All the 12 included studies were retrospective cohort studies, of which 7 used the propensity score matching (PSM) method to minimize the influence of confounding factors between comparison groups and provided accurate results, and five studies applied the covariate adjustment. Among these, one study was conducted in the United States, while 11 studies were conducted in Europe, including Italy, Turkey, Germany, the Netherlands, and France. The patient demographics and clinical characteristics of included studies are summarized in **Table 2**. The mean age range of involved patients was 62–84 years, and the proportion of males was 45.6–63.8%. The percentage of comorbidities, such as hypertension, diabetes mellitus, congestive heart failure, pulmonary disease, and renal disease, was reported in the majority of studies. The quality of cohort studies is assessed and summarized in **Table 1**; all included studies were rated fair and good.

3.2 Association of Chronic OAC With All-Cause Mortality in COVID-19 Patients

The outcome of all-cause mortality was available in 12 studies (14 cohorts) that compared chronic OAC patients with non-OAC patients with COVID-19. The meta-analysis of studies revealed that patients with chronic OAC were not associated with a significantly lower risk of all-cause mortality than those without OAC (**Figure 2A**; RR: 0.92; 95% CI: 0.82–1.03; p = 0.165; I^2 = 65.2%). A total of seven studies provided adjusted measure as HR, and the pooled HR (**Figure 2B**; HR: 0.99, 95% CI: 0.76–1.29; p = 0.941; I^2 = 75.4%) was consistent with pooled RR. Meta-regression was conducted to test whether demographic characteristics (sample size, mean age, percentage of male, study setting, hypertension, diabetes mellitus, congestive heart failure, pulmonary disease, and renal disease) were associated with all-cause mortality; any potential confounding factors could not be identified (p > 0.05 for each characteristic, **Supplementary Table S3**). The sequential removal of each study failed to identify those with a significant influence on the results (**Supplementary Table S4**).

3.3 Association of Different Types of OACs With All-Cause Mortality

The mortality outcome was stratified by DOACs or VKAs. A total of six studies compared the risk of mortality in COVID-19 patients with chronic VKAs with those without OACs. The patients with chronic VKA therapy faced an equal risk of all-cause mortality compared to those without OACs (**Figure 3A**; RR: 0.92; 95% CI: 0.74–1.15; p = 0.463; I^2 = 45.0%). Subsequently, seven studies compared the risk of mortality in chronic DOACs patients with non-OAC patients with COVID-19, and the pooled RR showed a significantly lower all-cause mortality in chronic DOACs than non-OACs in patients with COVID-19 (**Figure 3B**;

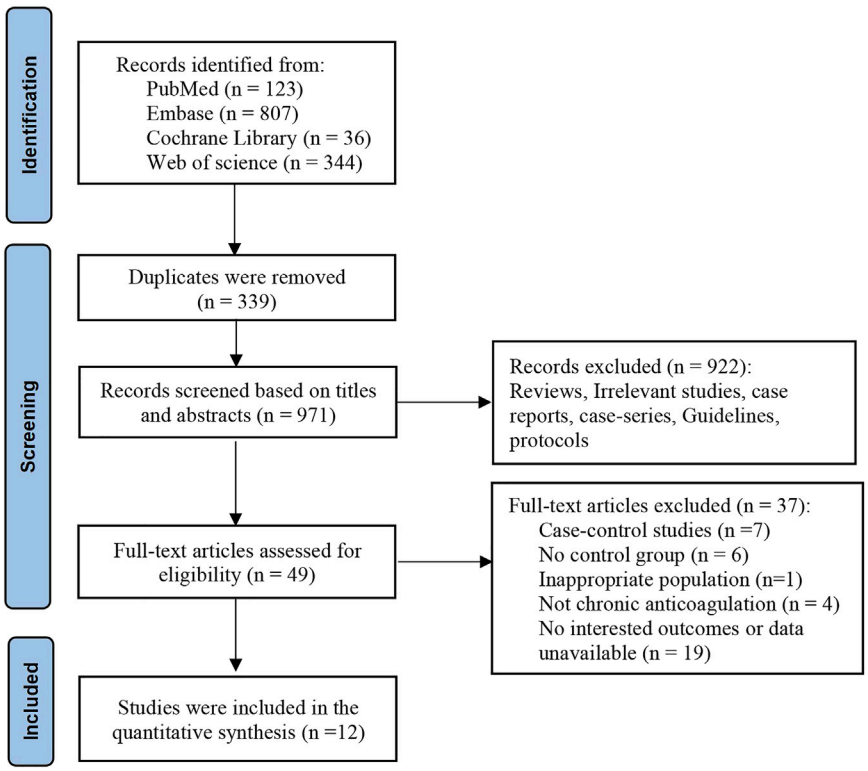

FIGURE 1 | Flow diagram for the selection of studies for the meta-analysis.

RR: 0.84; 95% CI: 0.73–0.97; p = 0.015; I^2 = 20.7%). The interaction analysis was conducted to evaluate the risk difference between VKAs and DOAC groups; however, no significant difference was detected ($P_{interaction}$ = 0.497).

3.4 Association of Chronic OAC With Intensive Care Unit Admission in COVID-19 Patients

The outcome of ICU admission was available in four studies that compared chronic OAC patients with non-OAC patients with COVID-19. We documented that patients taking chronic OAC could not ameliorate ICU admission significantly (**Figure 4**; RR: 0.65; 95% CI: 0.40–1.04; p = 0.073; I^2 = 76.5%).

3.5 Publication Bias

The visual inspection of funnel plots of primary outcomes for publication bias analyses showed a qualitatively symmetrical funnel-plot (**Supplementary Figure S1**). No significant publication bias was found, as confirmed by Egger's test (p = 0.602).

4 DISCUSSION

This study provided a comprehensive investigation of the effectiveness of chronic OAC in COVID-19. The results suggested that chronic OAC could not reduce the high risk of all-cause mortality or ICU admission in patients with COVID-19 compared to those not using OACs. Although the subgroup analysis suggested a potentially reduced risk of mortality in COVID-19 patients with chronic use of DOACs, the interaction analysis did not detect any significant difference between DOACs and VKAs. Together, the results of this study did not support that chronic OAC is beneficial in reducing the risk of mortality in patients with COVID-19.

The high rates of thromboembolic complications were observed in COVID-19 patients. The excess mortality of COVID-19 may be related to the hypercoagulability and microthrombi, leading to lung capillary occlusion (Ackermann et al., 2020; Varga et al., 2020). However, whether patients can benefit from OAC before the early stages of COVID-19 infection remains uncertain. This meta-analysis, based on 12 cohort studies involving 30,646 COVID-19 patients, is the first to provide a comprehensive overview of the effects of chronic OAC. The results demonstrated that chronic OAC does not reduce the all-cause mortality and ICU admission in COVID-19 patients. To the best of our knowledge, only one previous meta-analysis (Kamel et al., 2021) has demonstrated that there was no association between pre-admission anticoagulation and mortality in COVID-19 patients (RR: 0.84, 95% CI: 0.49–1.43; p > 0.05), which was consistent with our findings. However, the meta-analysis (Kamel et al., 2021) included only three studies (two were poor quality) with high heterogeneity (I^2 = 86%) and

TABLE 1 | Characteristics of included trials.

Author	Study design	Country	Sample size	Study setting	Type of anticoagulant (%)	Adjusted method	Reported outcomes	Quality
Aslan et al. (2021)	Retrospective cohort	Turkey	1710	Hospitalized	DOACs (100%)	CA	All-cause mortality	8
Buenen et al. (2021)	Retrospective cohort	The Netherlands	497	Emergency department	DOACs (51.8%) and VKAs (48.2%)	CA	All-cause mortality	8
Chocron et al. (2021)	Retrospective cohort	France	2,878	Hospitalized	DOACs (60.7%) and VKAs (39.3%)	PSM	All-cause mortality	7
Covino et al. (2021)	Retrospective cohort	Italy	2,666	Emergency department	DOACs (77.8%) and VKAs (22.2%)	PSM	All-cause mortality	8
Denas et al. (2021)	Retrospective cohort	Italy	4,697	Registration systems	DOACs (21.2%) and VKAs (78.8%)	PSM	ICU admission and all-cause mortality	6
Fröhlich et al. (2021)	Retrospective cohort	Germany	6,637	Registration systems	DOACs (69.5%) and VKAs (30.5%)	CA	All-cause mortality or need for non-invasive or invasive ventilation or ECMO	6
Gülcü et al. (2021)	Retrospective cohort	Turkey	5,575	Hospitalized	DOACs (84.9%) and VKAs (15.1%)	PSM	All-cause mortality	8
Harrison et al. (2021)	Retrospective cohort	America	1,026	Hospitalized	DOACs (78.8) or VKA (21.2%)	CA	All-cause mortality	7
Iaccarino et al. (2021)	Retrospective cohort	Italy	2,377	Registration systems	OAC	CA	All-cause mortality and ICU admission	6
Rivera-Caravaca et al. (2021)	Retrospective cohort	Ecuador, Germany, Italy, and Spain	1,002	Registration systems	DOACs (25.5%) and VKAs (74.5%)	PSM	All-cause mortality and several comorbidities	6
Russo et al. (2021)	Retrospective cohort	Italy	427	Hospitalized	DOACs (62.1%) and VKAs (37.9%)	PSM	All-cause mortality	8
Spiegelenberg et al. (2021)	Retrospective cohort	The Netherlands	1,154	Hospitalized	DOACs (51.6%) and VKAs (48.4%)	PSM	All-cause mortality and ICU admission	6

DOACs, direct oral anticoagulants; OAC, oral anticoagulant; VKAs, vitamin K antagonists; PSM, propensity score matching; CA, covariate adjustment; ICU, intensive care unit; ICH, intracranial hemorrhage; ECMO, extracorporeal membrane oxygenation.

TABLE 2 | Patient demographics and clinical characteristics of included studies.

Author	Mean age (years)	Male (%)	Hypertension (%)	Diabetes mellitus (%)	Pulmonary disease (%)	Congestive heart failure (%)	Renal disease (%)
Aslan et al. (2021)	62.0	49.5	42.0	27.0	6.0	3.0	NA
Buenen et al. (2021)	71.8	63.8	52.1	20.5	26.0	NA	19.5
Chocron et al. (2021)	66.6	57.9	50.5	23.5	NA	NA	NA
Covino et al. (2021)	84.0	50.0	41.8	18.5	17.4	23.4	15.2
Denas et al. (2021)	NA	50.6	64.1	19.9	NA	6.4	4.8
Gülcü et al. (2021)	64.4	50.2	49.5	26.9	13.9	5.3	2.7
Iaccarino et al. (2021)	68.2	62.7	59.0	18.0	8.0	12.0	6.0
Rivera-Caravaca et al. (2021)	82.0	59.2	82.1	31.2	18.3	4.6	15.1
Russo et al. (2021)	67.0	63.0	62.0	26.0	19.0	8.0	13.0
Spiegelenberg et al. (2021)	69.3	63.7	39.3	22.7	24.1	4.9	NA
Fröhlich et al. (2021)	66.4	50.9	58.7	22.4	11.4	12.8	16.2
Fröhlich et al. (2021)	65.6	51.3	57.0	21.9	11.0	11.9	16.1
Harrison et al. (2021)	76.3	47.5	73.8	38.3	NA	23.4	22.9
Harrison et al. (2021)	76.3	45.6	73.1	38.0	NA	20.8	21.7

NA, not available.

did not elaborate on the effects of the confounding factors that could affect the results, thereby providing weak preliminary conclusions. The evidence for the benefits and risks of an intervention can be obtained from high-quality observational studies when RCTs are lacking (Wei et al., 2018). The present meta-analysis included 12 cohort studies to build a large dataset

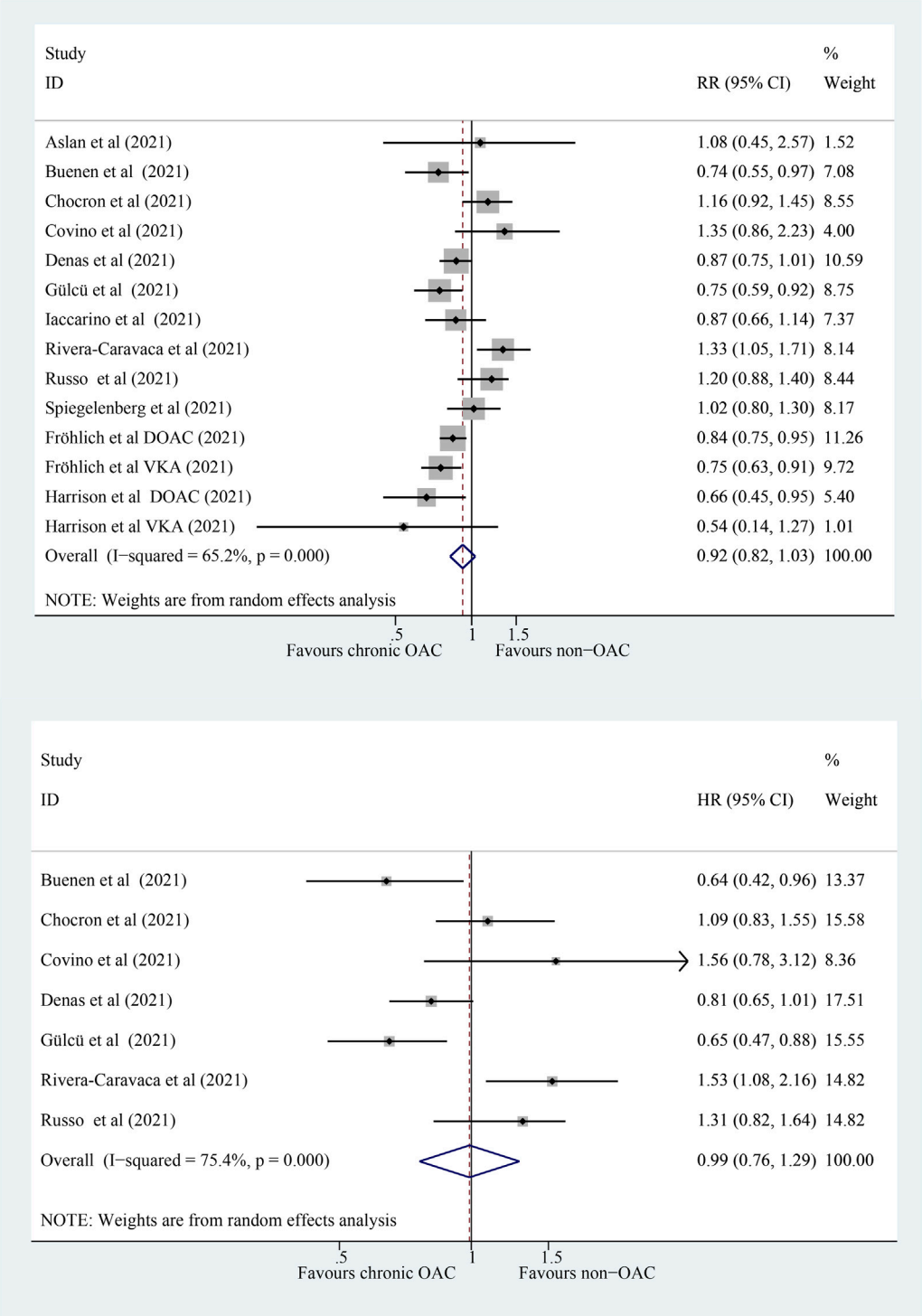

FIGURE 2 | Forest plot for association of chronic OAC with all-cause mortality in COVID-19 patients, random model. **(A)** Pooled the RR. **(B)** Pooled the adjusted HR. OAC, oral anticoagulation; COVID-19, coronavirus disease 2019; RR, relative risk; HR, hazard ratio; CI, confidence interval; DOAC, direct oral anticoagulant; VKA, vitamin K antagonist; *P*, heterogeneity test.

to provide high-quality evidence, whether chronic OAC could improve the prognosis of patients with COVID-19, and elaborated the influence of the confounding factors by the subgroup analysis and meta-regression to aid in decision-making for the treatment. In the analysis based on drug classification, each subgroup showed acceptable heterogeneity

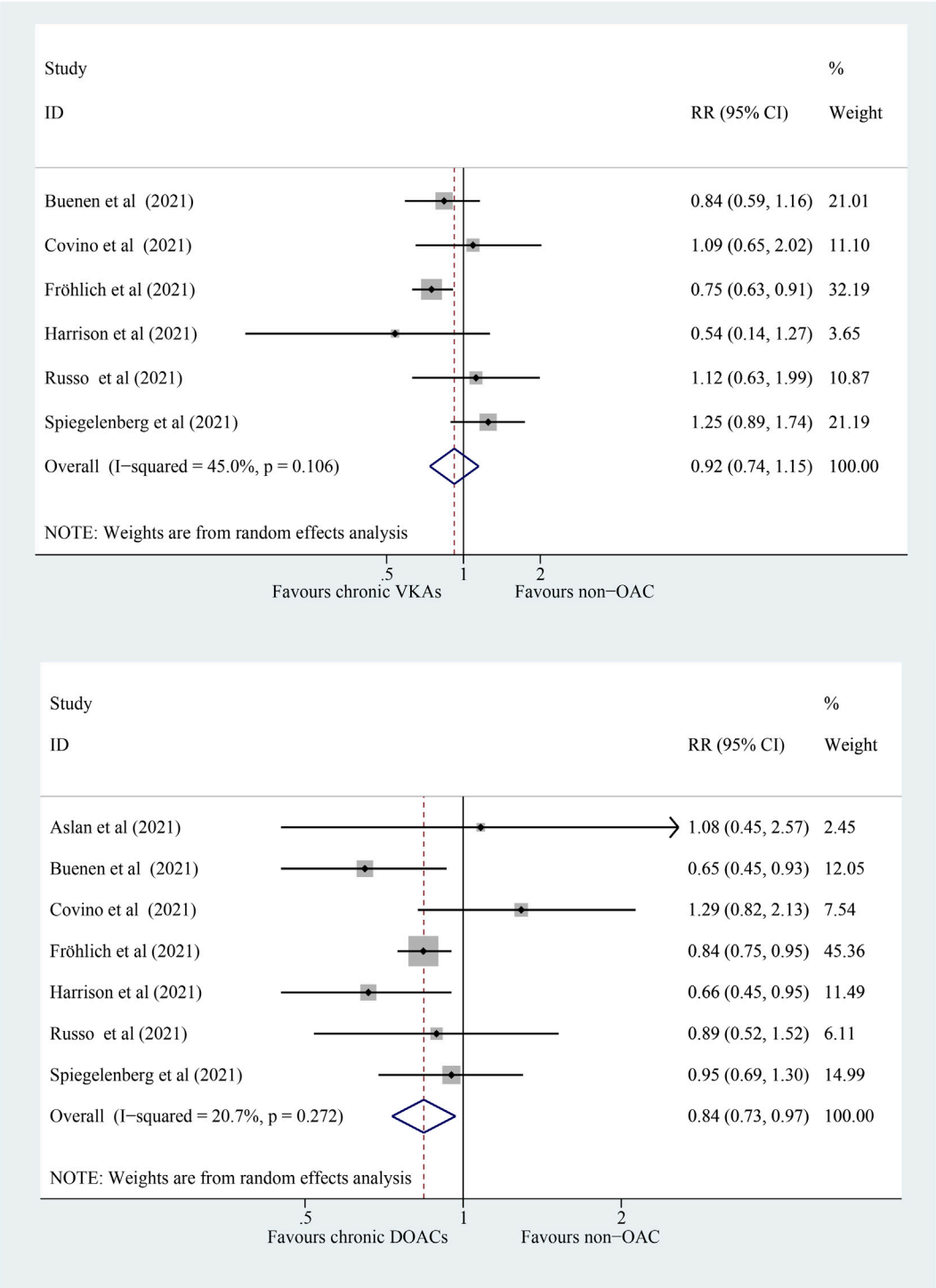

FIGURE 3 | Forest plot for the association of different types of OACs with all-cause mortality in COVID-19 patients, random model. **(A)** Chronic VKAs. **(B)** Chronic DOACs. OAC, oral anticoagulation; COVID-19, coronavirus disease 2019; RR, relative risk; CI, confidence interval; DOACs, direct oral anticoagulants; VKAs, vitamin K antagonists; *P*, heterogeneity test.

(both $I^2 < 50\%$). The type of OAC might be a source of moderate heterogeneity for the primary outcome ($I^2 = 65.2\%$), but the interaction analysis did not detect any statistical difference between DOACs and VKAs in this study. The meta-regression analysis failed to identify any potential confounding factors on all-cause mortality. However, confounding factors that might not

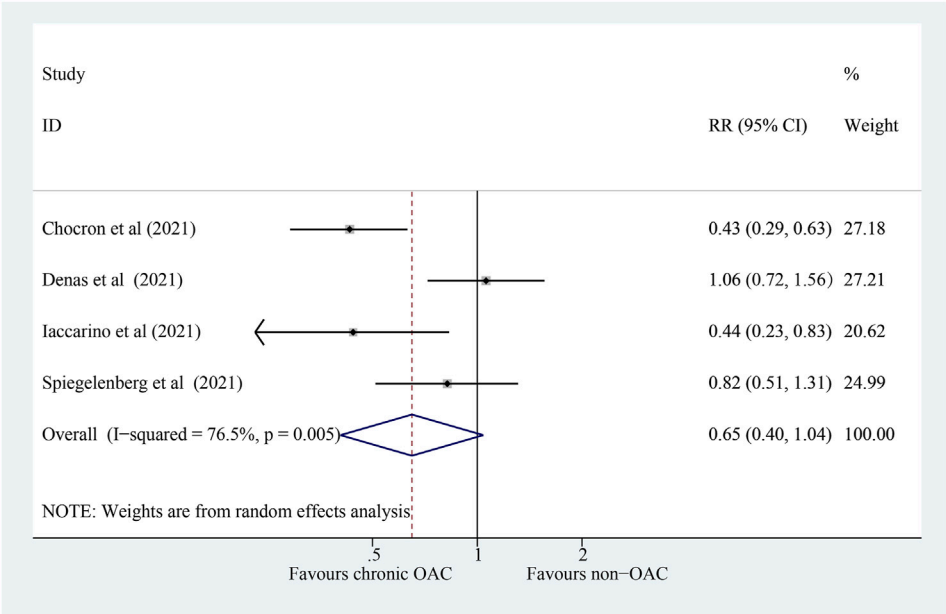

FIGURE 4 | Forest plot for the association of chronic OAC with ICU admission in COVID-19 patients, random model. OAC, oral anticoagulation; COVID-19, coronavirus disease 2019; RR, relative risk; CI, confidence interval; ICU, intensive care unit; *P*, heterogeneity test.

have been reported adequately in the included studies, such as different treatment options for COVID-19, may contribute to the unexplained heterogeneity.

Previous studies proposed that with the direct invasion of SARS-CoV-2 into endothelial cells, the extensive endotheliitis, immune complex accumulation, and type 3 hypersensitivity reaction results in diffuse vascular inflammation (Carsana et al., 2020; Aslan et al., 2021), while the hypercoagulability of COVID-19 is secondary to vascular inflammation. Furthermore, the clotting system activation and the immune-mediated inflammatory response have a complex interplay. The two processes mutually reinforce each other (Scudiero and Parodi, 2020; Hadid et al., 2021) and lead to microvascular thrombosis and death. This might explain why OAC, targeting a single pathway, cannot influence the progression of COVID-19 and eventually the mortality of the infected patients (Aslan et al., 2021; Russo et al., 2021). Thus, rather than against secondary hypercoagulability, directed against thrombogenic inflammation or vasculopathy may be a better treatment option (Flam et al., 2021).

Another study suggested that drugs that inhibit thrombin (coagulation factor IIa) and coagulation factor Xa, such as DOACs, might be promising in ameliorating disease progression and severity of COVID-19. In addition to the prevention of thrombosis, inhibition of factor IIa and factor Xa prevent a cytokine storm (Jose and Manuel, 2020), which might also have an antiviral effect *via* inhibition of SARS-CoV-2 fusion (Frydman et al., 2020). Although VKAs inhibits factors IIa and Xa, vitamin K insufficiency accelerates elastic fiber damage and thrombosis in severe COVID-19 due to impaired activation of matrix Gla protein (protects against pulmonary and vascular elastic fiber damage) and extrahepatic endothelial anticoagulant protein S, respectively, related to poor outcomes (Dofferhoff et al., 2020). However, this meta-analysis is based on real-world studies, suggesting that the impact of these mechanisms on COVID-19 mortality was negligible, and DOACs did not significantly reduce the risk of all-cause mortality compared to VKAs.

Additionally, the chronic anticoagulation indications (mostly by atrial fibrillation and VTE) of patients with COVID-19 were the per se risk factors of mortality (Mountantonakis et al., 2021). The pro-inflammatory effects, a "lethal cocktail," of COVID-19 render these patients susceptible (Rivera-Caravaca et al., 2021). Patients with anticoagulation indications should be advised to continue OAC during the COVID-19 pandemic, based on the results of this study. The OAC did not increase the mortality due to COVID-19; however, it might be beneficial to patients with anticoagulation indications and reduce the inherent high risk of thromboembolism. In addition, the recommendations from the Italian Federation of Anticoagulation Clinics (Poli et al., 2020) were consistent with those of this study, suggesting to continue OAC for patients with anticoagulation indications in the COVID-19 era. Furthermore, the switch from VKAs to DOACs should be considered in the absence of contraindication because social distancing and stay-at-home orders were implemented. The patients on VKAs were facing unique challenges, who requiring close and continued monitoring by clinic staff to optimize the benefit–risk ratio of warfarin therapy (Kish and Lekic, 2021).

The results need to be interpreted with caution. First, there is a lack of evidence that chronic OAC improves prognosis. Second, the effects of anticoagulants are influenced by genetic factors (Conti et al., 2021), other antithrombotic drugs (Valeria et al., 2021), and potential drug–drug interactions; for example, the drug–drug interactions of anticoagulants with some antiviral treatments used in COVID-19 patients. These phenomena

might increase the plasma levels of anticoagulants (Testa et al., 2020), and exposure to these molecules could cause severe bleeding events (Wu et al., 2015). Therefore, OAC during the early stages of COVID-19 infection should not be generalized to all patients in the general population. However, adhering to chronic OAC for patients with thrombotic complications diseases was essential. Several ongoing RCTs are focused on evaluating the potential effectiveness of anticoagulant therapy (including OAC) on the spectrum of COVID-19 patients (Talasaz et al., 2021). This finding needs to be substantiated further by appropriate RCTs.

5 LIMITATIONS

The present study has several limitations. First, all the included studies were observational studies. Although multiple measures, such as the adjusted effect estimates, subgroup analysis, and meta-regression, were used to reduce the risk of bias of primary analysis. The meta-analysis assembling data from observational studies had some limitations, such as among-study heterogeneity, selection bias, and various confounders. Second, moderate heterogeneity was observed in the primary outcomes, and methodological and clinical heterogeneity could be detected for all meta-analyses, especially involving retrospective studies. Moreover, no highly homogeneous studies support that chronic OAC reduces mortality in patients with COVID-19, and the conclusions of this study were credible. Third, the anticoagulation regimen during hospitalization influences the reported results. However, no standard protocol is yet available for COVID-19 treatment; all medical interventions were performed after hospital admission at the discretion of the medical team, and the results were credible under current treatment conditions. Finally, the assessment of major or non-major clinically relevant bleeding and thromboembolic complications was not performed because data were not available. Therefore, the results were considered with caution since the possibility of confounders could not be excluded.

6 CONCLUSION

The current systematic review and meta-analysis provided evidence that chronic OAC is not associated with improving the prognosis in COVID-19 patients, and the results were consistent across DOACs and VKAs subgroups. RCTs are required to produce high-quality evidence regarding the safety and efficacy of OAC on the full spectrum of COVID-19 patients.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**; further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

HX, Z-CG, and W-HG are the guarantors of the entire manuscript. M-FD and S-TG contributed to the study concept and design, data acquisition, analysis, and interpretation, and drafting and critical revision of the manuscript for intellectual content. Y-JK, B-YW, and FY contributed to the data acquisition, analysis, and interpretation. All authors contributed to the article and approved the submitted version.

FUNDING

The authors disclosed the receipt of the financial support for the research, authorship, and/or publication of this article. This study was funded by the Nanjing Medical Science and Technology Development Fund (QRX17060) and Jiangsu Pharmaceutical Association Shire Biopharmaceutical Fund (S201606).

ACKNOWLEDGMENTS

We would like to thank Prof. Dong-Jin Wang, the head of the Cardiothoracic Surgery Department of Drum Tower Hospital affiliated with Nanjing University Medical School, for support in our study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2022.781192/full#supplementary-material

REFERENCES

Ackermann, M., Verleden, S. E., Kuehnel, M., Haverich, A., Welte, T., Laenger, F., et al. (2020). Pulmonary Vascular Endothelialitis, Thrombosis, and Angiogenesis in COVID-19. *N. Engl. J. Med.* 383 (2), 120–128. doi:10.1056/NEJMoa2015432

Aslan, B., Akyüz, A., Işık, F., Çap, M., İnci, Ü., Kaya, İ., et al. (2021). The Effect of Chronic DOAC Treatment on Clinical Outcomes of Hospitalized Patients with COVID-19. *Int. J. Clin. Pract.* 75 (9), e14467. doi:10.1111/ijcp.14467

Belen-Apak, F. B., and Sarıalioğlu, F. (2020). Pulmonary Intravascular Coagulation in COVID-19: Possible Pathogenesis and Recommendations on Anticoagulant/thrombolytic Therapy. *J. Thromb. Thrombolysis* 50 (2), 278–280. doi:10.1007/s11239-020-02129-0

Buenen, A. G., Sinkeldam, M., Maas, M. L., Verdonschot, M., and Wever, P. C. (2021). Prior Use of Anticoagulation Is Associated with a Better Survival in COVID-19. *J. Thromb. Thrombolysis* 52, 1207–1211. doi:10.1007/s11239-021-02486-4

Carsana, L., Sonzogni, A., Nasr, A., Rossi, R. S., Pellegrinelli, A., Zerbi, P., et al. (2020). Pulmonary post-mortem Findings in a Series of COVID-19 Cases from

Northern Italy: a Two-centre Descriptive Study. *Lancet Infect. Dis.* 20 (10), 1135–1140. doi:10.1016/s1473-3099(20)30434-5

Chocron, R., Galand, V., Cellier, J., Gendron, N., Pommier, T., Bory, O., et al. (2021). Anticoagulation before Hospitalization Is a Potential Protective Factor for COVID-19: Insight from a French Multicenter Cohort Study. *J. Am. Heart Assoc.* 10 (8), e018624. doi:10.1161/jaha.120.018624

Conti, V., Corbi, G., Manzo, V., Sellitto, C., Iannello, F., Esposito, S., et al. (2021). The Role of Pharmacogenetics in Antithrombotic Therapy Management: New Achievements and Barriers yet to Overcome. *Curr. Med. Chem.* 28 (32), 6675–6703. doi:10.2174/0929867328666201231124715

Covino, M., De Matteis, G., Della Polla, D., Burzo, M. L., Pascale, M. M., Santoro, M., et al. (2021). Does Chronic Oral Anticoagulation Reduce In-Hospital Mortality Among COVID-19 Older Patients? *Aging Clin. Exp. Res.* 33 (8), 2335–2343. doi:10.1007/s40520-021-01924-w

Cui, S., Chen, S., Li, X., Liu, S., and Wang, F. (2020). Prevalence of Venous Thromboembolism in Patients with Severe Novel Coronavirus Pneumonia. *J. Thromb. Haemost.* 18 (6), 1421–1424. doi:10.1111/jth.14830

Cumpston, M., Li, T., Page, M. J., Chandler, J., Welch, V. A., Higgins, J. P., et al. (2019). Updated Guidance for Trusted Systematic Reviews: a New Edition of the Cochrane Handbook for Systematic Reviews of Interventions. *Cochrane Database Syst. Rev.* 10, Ed000142. doi:10.1002/14651858.Ed000142

Denas, G., Gennaro, N., Ferroni, E., Fedeli, U., Lorenzoni, G., Gregori, D., et al. (2021). Reduction in All-Cause Mortality in COVID-19 Patients on Chronic Oral Anticoagulation: A Population-Based Propensity Score Matched Study. *Int. J. Cardiol.* 329, 266–269. doi:10.1016/j.ijcard.2020.12.024

Dofferhoff, A. S. M., Piscaer, I., Schurgers, L. J., Visser, M. P. J., van den Ouweland, J. M. W., de Jong, P. A., et al. (2020). Reduced Vitamin K Status as a Potentially Modifiable Risk Factor of Severe Coronavirus Disease 2019. *Clin. Infect. Dis.* 73, e4039–e4046. doi:10.1093/cid/ciaa1258

Flam, B., Wintzell, V., Ludvigsson, J. F., Mårtensson, J., and Pasternak, B. (2021). Direct Oral Anticoagulant Use and Risk of Severe COVID-19. *J. Intern. Med.* 289 (3), 411–419. doi:10.1111/joim.13205

Fröhlich, G. M., Jeschke, E., Eichler, U., Thiele, H., Alhariri, L., Reinthaler, M., et al. (2021). Impact of Oral Anticoagulation on Clinical Outcomes of COVID-19: a Nationwide Cohort Study of Hospitalized Patients in Germany. *Clin. Res. Cardiol.* 110 (7), 1041–1050. doi:10.1007/s00392-020-01783-x

Frydman, G. H., Streiff, M. B., Connors, J. M., and Piazza, G. (2020). The Potential Role of Coagulation Factor Xa in the Pathophysiology of COVID-19: A Role for Anticoagulants as Multimodal Therapeutic Agents. *TH Open* 4 (4), e288–e299. doi:10.1055/s-0040-1718415

Gülcü, O., Aksakal, E., Aydemir, S., Doğan, R., Saraç, İ., Aydın, S. Ş., et al. (2021). Association between Previous Anticoagulant Use and Mortality Among Hospitalized Patients with COVID-19. *J. Thromb. Thrombolysis* 53, 88–95. doi:10.1007/s11239-021-02489-1

Hadid, T., Kafri, Z., and Al-Katib, A. (2021). Coagulation and Anticoagulation in COVID-19. *Blood Rev.* 47, 100761. doi:10.1016/j.blre.2020.100761

Harrison, R. F., Forte, K., Buscher, M. G., Jr., Chess, A., Patel, A., Moylan, T., et al. (2021). The Association of Preinfection Daily Oral Anticoagulation Use and All-Cause in Hospital Mortality from Novel Coronavirus 2019 at 21 Days: A Retrospective Cohort Study. *Crit. Care Explor.* 3 (1), e0324. doi:10.1097/cce.0000000000000324

Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., et al. (2011). The Cochrane Collaboration's Tool for Assessing Risk of Bias in Randomised Trials. *BMJ* 343, d5928. doi:10.1136/bmj.d5928

Iaccarino, G., Grassi, G., Borghi, C., Grassi, D., Mancusi, C., Muiesan, M. L., et al. (2021). Preexisting Oral Anticoagulant Therapy Ameliorates Prognosis in Hospitalized COVID-19 Patients. *Front. Cardiovasc. Med.* 8, 633878. doi:10.3389/fcvm.2021.633878

John Hopkins University (2022). John Hopkins University COVID-19. [Online]. Available at: https://coronavirus.jhu.edu/map.html (Accessed February 17, 2022).

Jose, R. J., and Manuel, A. (2020). COVID-19 Cytokine Storm: the Interplay between Inflammation and Coagulation. *Lancet Respir. Med.* 8 (6), e46–e47. doi:10.1016/s2213-2600(20)30216-2

Kamel, A. M., Sobhy, M., Magdy, N., Sabry, N., and Farid, S. (2021). Anticoagulation Outcomes in Hospitalized COVID-19 Patients: A Systematic Review and Meta-analysis of Case-control and Cohort Studies. *Rev. Med. Virol.* 31 (3), e2180. doi:10.1002/rmv.2180

Kish, K., and Lekic, S. (2021). Implementation of Warfarin to Direct Oral Anticoagulant Conversion Initiative in Pharmacist-run Anticoagulation Clinics during COVID-19 Pandemic. *J. Am. Coll. Clin. Pharm.* 4, 1154–1160. doi:10.1002/jac5.1470

Lodigiani, C., Iapichino, G., Carenzo, L., Cecconi, M., Ferrazzi, P., Sebastian, T., et al. (2020). Venous and Arterial Thromboembolic Complications in COVID-19 Patients Admitted to an Academic Hospital in Milan, Italy. *Thromb. Res.* 191, 9–14. doi:10.1016/j.thromres.2020.04.024

Moores, L. K., Tritschler, T., Brosnahan, S., Carrier, M., Collen, J. F., Doerschug, K., et al. (2020). Prevention, Diagnosis, and Treatment of VTE in Patients with Coronavirus Disease 2019: CHEST Guideline and Expert Panel Report. *Chest* 158 (3), 1143–1163. doi:10.1016/j.chest.2020.05.559

Mountantonakis, S. E., Saleh, M., Fishbein, J., Gandomi, A., Lesser, M., Chelico, J., et al. (2021). Atrial Fibrillation Is an Independent Predictor for In-Hospital Mortality in Patients Admitted with SARS-CoV-2 Infection. *Heart Rhythm* 18 (4), 501–507. doi:10.1016/j.hrthm.2021.01.018

Poli, D., Tosetto, A., Palareti, G., Barcellona, D., Ciampa, A., Grandone, E., et al. (2020). Managing Anticoagulation in the COVID-19 Era between Lockdown and Reopening Phases. *Intern. Emerg. Med.* 15 (5), 783–786. doi:10.1007/s11739-020-02391-3

Rivera-Caravaca, J. M., Núñez-Gil, I. J., Vivas, D., Viana-Llamas, M. C., Uribarri, A., Becerra-Muñoz, V. M., et al. (2021). Clinical Profile and Prognosis in Patients on Oral Anticoagulation before Admission for COVID-19. *Eur. J. Clin. Invest.* 51 (1), e13436. doi:10.1111/eci.13436

Russo, V., Bottino, R., D'Andrea, A., Silverio, A., Di Maio, M., Golino, P., et al. (2021). Chronic Oral Anticoagulation and Clinical Outcome in Hospitalized COVID-19 Patients. *Cardiovasc. Drugs Ther.*, 1–8. doi:10.1007/s10557-021-07194-y

Scudiero, F., and Parodi, G. (2020). Dual Antiplatelet Therapy in Patients with Acute Coronary Syndrome during the Coronavirus Disease of 2019 Pandemia: the Right Choice at the Right Time. *J. Cardiovasc. Med. (Hagerstown)* 21 (8), 535–537. doi:10.2459/jcm.0000000000001028

Spiegelenberg, J. P., Gelder, M. M. H. J., Maas, M. L., Hovens, M. M. C., Esselink, A., Dofferhoff, A. S. M., et al. (2021). Prior Use of Therapeutic Anticoagulation Does Not Protect against COVID-19 Related Clinical Outcomes in Hospitalized Patients: A Propensity Score-matched Cohort Study. *Br. J. Clin. Pharma.* 87, 4839–4847. doi:10.1111/bcp.14877

Talasaz, A. H., Sadeghipour, P., Kakavand, H., Aghakouchakzadeh, M., Kordzadeh-Kermani, E., Van Tassell, B. W., et al. (2021). Recent Randomized Trials of Antithrombotic Therapy for Patients with COVID-19: JACC State-Of-The-Art Review. *J. Am. Coll. Cardiol.* 77 (15), 1903–1921. doi:10.1016/j.jacc.2021.02.035

Testa, S., Prandoni, P., Paoletti, O., Morandini, R., Tala, M., Dellanoce, C., et al. (2020). Direct Oral Anticoagulant Plasma Levels' Striking Increase in Severe COVID-19 Respiratory Syndrome Patients Treated with Antiviral Agents: The Cremona Experience. *J. Thromb. Haemost.* 18 (6), 1320–1323. doi:10.1111/jth.14871

Valeria, C., Carmine, S., Valentina, M., Teresa, I., Maria, C., Martina, T., et al. (2021). The Need of a Multicomponent Guiding Approach to Personalize Clopidogrel Treatment. *Pharmacogenom. J.* 21 (2), 116–127. doi:10.1038/s41397-020-00189-2

VanderWeele, T. J. (2020). Optimal Approximate Conversions of Odds Ratios and hazard Ratios to Risk Ratios. *Biometrics* 76 (3), 746–752. doi:10.1111/biom.13197

Varga, Z., Flammer, A. J., Steiger, P., Haberecker, M., Andermatt, R., Zinkernagel, A. S., et al. (2020). Endothelial Cell Infection and Endotheliitis in COVID-19. *Lancet* 395 (10234), 1417–1418. doi:10.1016/s0140-6736(20)30937-5

Wei, A. H., Gu, Z. C., Zhang, C., Ding, Y. F., Liu, D., Li, J., et al. (2018). Increased Risk of Myocardial Infarction with Dabigatran Etexilate: Fact or Fiction? A Critical Meta-Analysis of over 580,000 Patients from Integrating Randomized Controlled Trials and Real-World Studies. *Int. J. Cardiol.* 267, 1–7. doi:10.1016/j.ijcard.2018.05.048

Wells, G., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., et al. (2021). The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomised Studies in Meta-Analyses [Online]. Ottawa Hospital Research Institute Research Programs. Available at: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp (Accessed August 15, 2021).

Wichmann, D., Sperhake, J. P., Lütgehetmann, M., Steurer, S., Edler, C., Heinemann, A., et al. (2020). Autopsy Findings and Venous Thromboembolism in Patients with COVID-19: A Prospective Cohort Study. *Ann. Intern. Med.* 173 (4), 268–277. doi:10.7326/m20-2003

World Health Organization (2021). WHO Announces COVID-19 Outbreak a Pandemic. [Online]. Available at: http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-COVID-19/news/news/2020/3/who-announces-COVID-19-outbreak-a-pandemic (Accessed July 24, 2021).

Wu, Z., and McGoogan, J. M. (2020). Characteristics of and Important Lessons from the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases from the Chinese Center for Disease Control and Prevention. *JAMA* 323 (13), 1239–1242. doi:10.1001/jama.2020.2648

Wu, C., Alotaibi, G. S., Alsaleh, K., Linkins, L. A., and McMurtry, M. S. (2015). Case-fatality of Recurrent Venous Thromboembolism and Major Bleeding Associated with Aspirin, Warfarin, and Direct Oral Anticoagulants for Secondary Prevention. *Thromb. Res.* 135 (2), 243–248. doi:10.1016/j.thromres.2014.10.033

Zhang, C., Shen, L., Le, K. J., Pan, M. M., Kong, L. C., Gu, Z. C., et al. (2020). Incidence of Venous Thromboembolism in Hospitalized Coronavirus Disease 2019 Patients: A Systematic Review and Meta-Analysis. *Front. Cardiovasc. Med.* 7, 151. doi:10.3389/fcvm.2020.00151

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Dai, Guo, Ke, Wang, Yu, Xu, Gu and Ge. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Public Health

SYSTEMATIC REVIEW
published: 01 April 2022
doi: 10.3389/fpubh.2022.842096

Reliability of the Evidence to Guide Decision-Making in Acupuncture for Functional Dyspepsia

Jinke Huang[1†], Jiali Liu[1†], Zhihong Liu[2], Jing Ma[1], Jinxin Ma[2], Mi Lv[1], Fengyun Wang[1] and Xudong Tang[3*]*

[1] Department of Gastroenterology, Xiyuan Hospital, China Academy of Chinese Medical Sciences, Beijing, China, [2] Department of Gastroenterology, Peking University Traditional Chinese Medicine Clinical Medical School (Xiyuan), Beijing, China, [3] China Academy of Chinese Medical Sciences, Beijing, China

Background and Aims: There has been a significant increase in the number of systematic reviews (SRs)/meta-analyses (MAs) investigating the effects of acupuncture for functional dyspepsia (FD). To systematically collate, appraise, and synthesize the current evidence, we carried out an umbrella review of SRs/MAs.

Methods: Systemic reviews/meta-analyses on acupuncture for FD were collected by searching major medical databases. The included studies were evaluated in terms of methodological quality, reporting quality, and evidence quality using the criteria from the Assessment of Multiple Systematic Reviews 2 (AMSTAR-2) tool, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement, and the Grades of Recommendation, Assessment, Development, and Evaluation (GRADE) system, respectively.

Results: Ten SRs/MAs were analyzed for this study. The methodological quality, reporting quality, and evidence quality of the included SRs/MAs were generally unsatisfactory. Lack of protocol registration, no list of excluded trials, or lack of a comprehensive search strategy were the main limitations. No high-quality evidence was found to support the effects of acupuncture for FD; the qualitative data synthesis relied on low quality trials with small sample sizes and was the main factor for evidence degradation.

Conclusions: Acupuncture seems to have a promising efficacy in the treatment of FD. It provides a new and prospective therapeutic method for FD. Although the quality of the included SRs/MAs was generally low and defects were frequent, this umbrella review highlights areas where improvement in methodology is required.

Keywords: reliability, evidence, decision-making, acupuncture, functional dyspepsia

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Seok-Jae Ko,
Kyung Hee University, South Korea
Shen Guoming,
Anhui University of Chinese Medicine, China

****Correspondence:***
Fengyun Wang
wfy811@163.com
Xudong Tang
txdly@sina.com

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Public Health Policy, a section of the journal Frontiers in Public Health

Received: *23 December 2021*
Accepted: *21 February 2022*
Published: *01 April 2022*

Citation:
Huang J, Liu J, Liu Z, Ma J, Ma J, Lv M, Wang F and Tang X (2022) Reliability of the Evidence to Guide Decision-Making in Acupuncture for Functional Dyspepsia. Front. Public Health 10:842096. doi: 10.3389/fpubh.2022.842096

INTRODUCTION

Functional dyspepsia (FD) is a functional gastrointestinal disease marked chiefly by early satiation, postprandial fullness, epigastric pain, and epigastric burning (1). The annual prevalence of FD is 9–10%, with 15% of patients having frequent (>3 attacks per week) and chronic (>3 months per year) symptoms (2). As a result of continuous medical treatment and continuous drug use,

indigestion symptoms create a huge socio-economic burden and have a negative impact on the quality of life (3, 4). It is reported that the total yearly direct cost of presentation and treatment is $805 per patient, yet the economic losses of FD on patients are much higher in combination with the indirect costs caused by resignation or reduced productivity (4). Abnormal gastric motility (delayed gastric emptying), impaired gastric adaptability to food, and visceral hypersensitivity (gastric hyperdistention sensitivity and acid hypersensitivity) are usually considered to be the main pathophysiological mechanisms of FD (5). Current pharmacological options include antidepressants, antacids, anxiolytics, prokinetics, and *Helicobacter pylori* eradication (6). However, none of the available therapies are effective in the majority of patients (6).

Acupuncture has been widely used to treat functional gastrointestinal disorders including FD. A literature search yielded that an increasing number of systematic reviews (SRs)/meta-analyses (MAs) regarding this topic have been conducted. However, not all SRs/MAs can provide reliable evidence, and low-quality SRs/MAs may instead mislead decision-makers (7). Furthermore, in an attempt to ensure validity of evidence and improve the quality of SRs/MAs, the measurement tools like the Assessment of Multiple Systematic Reviews 2 (AMSTAR-2) (8), the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (9), and the Grades of Recommendation, Assessment, Development, and Evaluation (GRADE) (10) were published in 2007, 2009, and 2004, respectively. An umbrella review is defined as an evidence-based medicine approach to the systematic review of all SRs/MAs on the same topic. Conducting an umbrella review helps to integrate the evidence more comprehensively, thereby providing higher quality evidence for clinicians. Hence, to systematically collate, evaluate, and synthesize current evidence on acupuncture for FD, we conducted this study.

METHODS

The protocol of this study was registrated in PROSPERO registry (CRD42021262020), and the methodology was performed following the criteria of the Cochrane Handbook.

Criteria for Considering Studies

The criteria for selection were as follows: (a) SRs/MAs of only enrolled randomized controlled trials (RCTs), (b) subjects were diagnosed with FD according to the Rome diagnostic criteria, (c) interventions applied to the experimental group included acupuncture alone or in combination with conventional medicine (CM), while interventions applied to the control included sham acupuncture or CM, (d) reporting at least one prespecified outcome: symptom score (SS), nepean dyspepsia index (NDI), main frequency of electrogastrogram (MF-EGG), gastric half-emptying time (GHET), plasma motilin level (PML), and effective rate (ER). The exclusion criteria were as follows: (a) repeated publications, (b) graduate dissertation, (c) lack of further data, and (d) conference abstracts.

TABLE 1 | Detailed retrieval strategy for PubMed.

Query	Search term
#1	Dyspepsia[Mesh]
#2	Dyspepsia*[Title/Abstract] OR Functional dyspepsia* [Title/Abstract] OR Indigestion* [Title/Abstract] OR non-ulcer dyspepsia* [Title/Abstract]
#3	#1 OR #2
#4	Acupuncture [Mesh]
#5	Acupuncture [Title/Abstract] OR pharmacoacupuncture[Title/Abstract] OR acupotomy[Title/Abstract] OR acupotomies[Title/Abstract] OR pharmacopuncture[Title/Abstract] OR needle[Title/Abstract] OR needling[Title/Abstract] OR dry-needling[Title/Abstract] OR body-acupuncture[Title/Abstract] OR electroacupuncture[Title/Abstract] OR electro-acupuncture[Title/Abstract] OR auricular acupuncture[Title/Abstract] OR warm needle[Title/Abstract]
#6	#4 OR #5
#7	Meta-Analysis as Topic [Mesh]
#8	Systematic review[Title/Abstract] OR Meta-Analysis[Title/Abstract] OR meta analysis[Title/Abstract] OR meta-analyses[Title/Abstract] OR metaanalysis[Title/Abstract]
#9	#7 OR #8
#10	#3 AND #6 AND #9

Search Strategy

Systematic reviews/Meta-analysis of acupuncture on FD were systematically searched in the following databases from their inception till August 2021: the Cochrane Library, PubMed, Web of Science, Embase, Chinese Biomedical Database, Chinese VIP information, and Wanfang Data. Functional dyspepsia, acupuncture, and systematic review were used as key search terms. A strategy for using PubMed search is shown in **Table 1**.

Literature Screening and Data Extraction

Literature selection and data extraction were conducted by two independent reviewers. For literature selection, titles and abstracts were screened first, and the full texts of potential literature were read for further evaluation. From the articles included, the following data were extracted: first author, publication year, sample size, interventions, outcomes, and main results. Disagreements were settled *via* discussion and consensus.

Quality Assessment

Methodological quality, reporting quality, and evidence quality of the enrolled studies were evaluated by two independent reviewers using the AMSTAR-2 tool, the PRISMA statement, and the GRADE system, respectively. The AMSTAR-2 consisted of 16 items, and the methodological quality was ranked as high, moderate, low, or critically low (8). The PRISMA statement consisted of 27 items, with each of the 16 items given a rating of yes, no, or partly yes (9). Evidence quality with GRADE was considered from five aspects (risk of bias, inconsistency, indirectness, imprecision, and publication bias) and given a

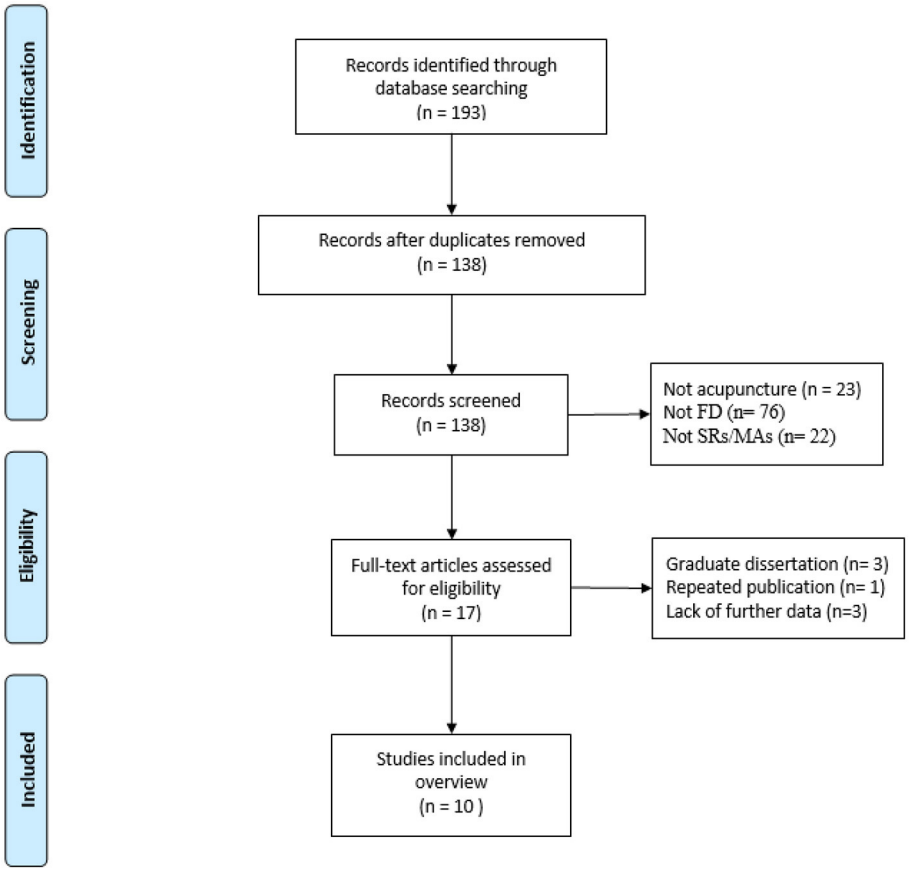

FIGURE 1 | Literature screening flowchart.

rating of high to critically low (10). Disagreements were settled *via* discussion and consensus.

RESULTS

Literature Screening

The literature search yielded 193 citations, of which 55 duplicates were removed. After screening the title and abstract, 121 citations were excluded and the remaining 17 were further evaluated through full-text reading. Finally, 10 SRs/MAs (11–20) met the inclusion criteria (**Figure 1**).

Characteristics of the Primary Studies

Then, SRs/MAs published from 2014 to 2021 were included. The interventions applied to the experimental groups were acupuncture alone or in combination with CM, while those administered to the control groups were CM or sham acupuncture. More details are presented in **Table 2**.

Methodological Quality

According to the AMSTAR-2, only one Cochrane SR/MA presented high level of methodological quality, while the other SRs/MAs presented critically low levels of methodological quality. Lack of protocol registration, no list of excluded trials, or lack of a comprehensive search strategy were the main limitations of the included SRs/MAs (**Figures 2**, **3**).

Reporting Quality

According to the PRISMA statement, only one Cochrane SR/MA fully reported information on the 27 items, while the other SRs/MAs missed serious information in item Q5 (protocol and registration), Q8 (search), and Q27 (funding). More details are shown in **Table 3**.

Evidence Quality

With the GRADE system, the evidence quality of all concerned outcome indicators ranged from "critically low" to "moderate" due to the risk of bias, inconsistency, imprecision, and publication bias. Furthermore, upgrading factors such as large effect and evidence of dose-response gradient were not applicable to the included outcome indicators. More details are shown in **Table 4**.

Meta-Analyses Outcomes of Intervention

Acupuncture vs. Sham Acupuncture

Five studies (11, 12, 14–16) compared the effects of acupuncture with sham acupuncture. Three studies (11, 12, 14, 16) reported significant improvements in SSs in the acupuncture group

TABLE 2 | A description of the characteristics.

References	Country	Trials (subjects)	Experimental intervention	Control intervention	Quality assessment	Meta-analyses	Results summary
Han (11)	China	10(1,202)	AT	Sham AT, CM	Cochrane criteria	Yes	Regarding the improvement in the symptoms of FD, the effect of acupuncture was superior to the sham-acupuncture group or the western medication group.
Mao (12)	China	7(853)	AT	Sham AT, CM	Cochrane criteria	Yes	Acupuncture was more effective than placebo in treating FD, while it was comparable to CM.
Yuan (13)	China	31(2,571)	AT	CM	Cochrane criteria	Yes	Compared with CM, acupuncture can significantly improve the effective rate, functional dyspeptic symptoms and motilin levels.
Zhou (14)	China	24(3,097)	AT, AT+CM	Sham AT, CM	Cochrane criteria	Yes	Acupuncture can effectively alleviate the symptoms of FD and improve the quality of life of patients.
Pang (15)	China	16(1,436)	AT	Sham AT, CM	Cochrane criteria	Yes	Acupuncture was superior to sham acupuncture and CM in improving symptoms and quality of life in FD patients.
Kim (16)	Korea	20(1,423)	AT, AT+CM	Sham AT, CM	Cochrane criteria	Yes	Acupuncture was superior to sham acupuncture and CM in improving symptoms and quality of life in FD patients.
Lan (17)	China	7(542)	AT	Sham AT, CM	Cochrane criteria	Yes	Whether acupuncture was more effective or safer than other treatments for the treatment of FD remains inconclusive.
Peng (18)	China	7(1,044)	AT	CM	Cochrane criteria	Yes	The therapeutic effect of acupuncture on FD was equivalent to that of CM.
Jin (19)	China	20(1,301)	AT	CM	Jadad	Yes	Acupuncture was more effective than CM in the treatment of FD, and had a lower risk of adverse events.
Wu (20)	China	16(1,088)	AT	CM	Jadad	Yes	Acupuncture was more effective than CM in the treatment of FD, and had a lower risk of adverse events.

compared to the sham acupuncture. Two studies reported ERs (11, 16), and the meta-analysis results revealed that the acupuncture group was significantly superior to the sham acupuncture group. The MF- EGG was reported in two studies (11, 12). One study (11) showed that the acupuncture group was significantly superior to the sham group, while the other (12) reported no significant difference between these two groups. Furthermore, it was reported that acupuncture was superior to sham acupuncture in NDI (15).

Acupuncture vs. Conventional Medication

Nine studies (11–14, 16–20) compared the effects of acupuncture with CM. Symptom score was reported in six studies (11–14, 16, 19), and the results of four studies (11, 13, 16, 19) revealed that acupuncture was superior to CM, while others (12, 14) showed that acupuncture was equivalent to CM. An effective rate was reported in eight studies (11–13, 16–20), and the results of five studies (13, 16, 18–20) revealed that acupuncture was superior to CM, while others (11, 12, 17) showed that acupuncture was equivalent to CM. The MF-EGG was reported in one study (11), and the pooled results showed that acupuncture was equivalent to CM. PML was reported in four studies (11–14), and the results of one study (13) revealed that acupuncture was superior to CM, while others (11, 12, 14) showed that acupuncture was equivalent to CM. Two studies reported the GHET (11, 12), and the results showed that the therapeutic effect of acupuncture group was comparable to that of CM. In addition, it was reported that acupuncture treatment was superior to CM in NDI (19).

DISCUSSION

Systematic reviews/meta-analyses, as the highest level of evidence source in evidence-based medicine, have become an important basis for evidence-based decision-making (21). However, SRs/MAs with low quality might affect the applicability of evidence. In this study, we performed an umbrella review to collate, evaluate, and synthesize the current evidence on acupuncture for FD.

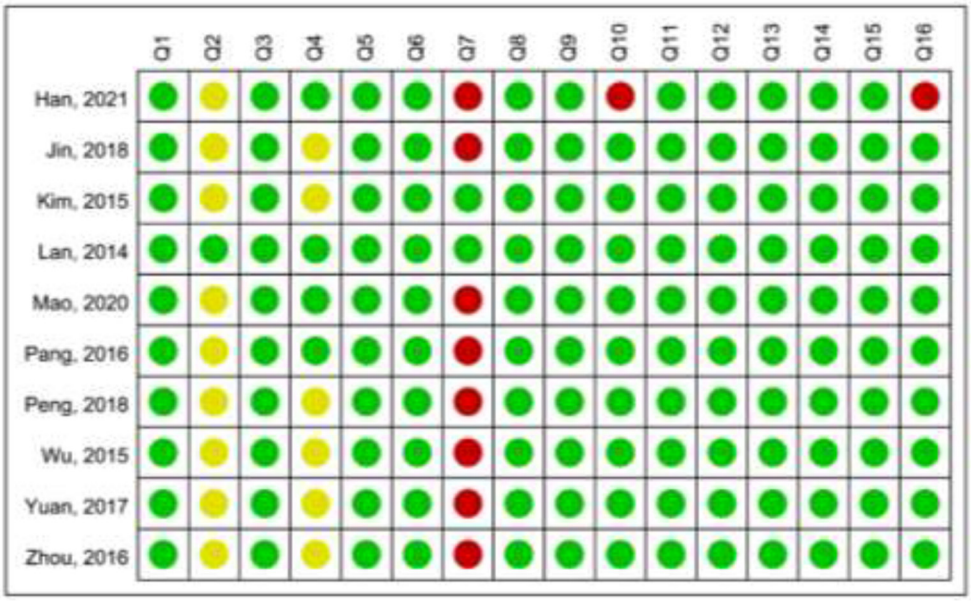

FIGURE 2 | Summary of the Assessment of Multiple Systematic Reviews 2 (AMSTAR-2).

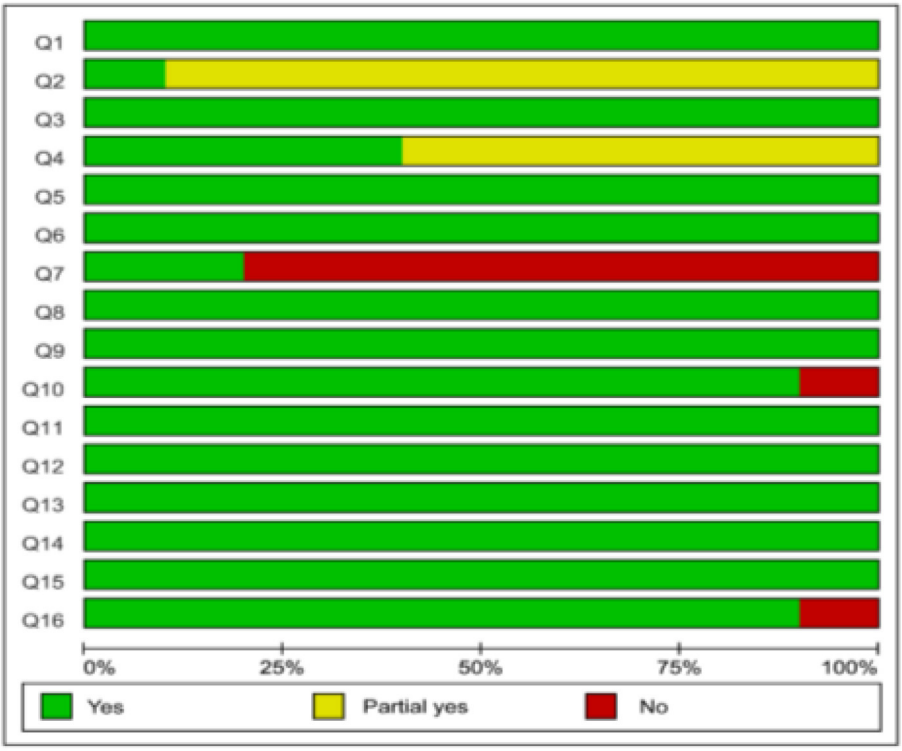

FIGURE 3 | Graphical representation of the AMSTAR-2.

Main Results

First, 10 SRs/MAs published between 2014 and 2021 were included, nine of which were published over the past 5 years, indicating an increasing interest in acupuncture for FD. Second, based on the AMSTAR-2 tool, the PRISMA statement, and the GRADE system, the methodological quality, reporting quality,

TABLE 3 | Result of the PRISMA checklist.

Section/Topic	Items	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	Compliance (%)
Title	Q1. Title	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
Abstract	Q2. Structured summary	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
Introduction	Q3. Rationale	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
	Q4. Objectives	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
Methods	Q5. Protocol and registration	N	N	N	N	N	N	Y	N	N	N	10%
	Q6. Eligibility criteria	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
	Q7. Information sources	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
	Q8. Search	Y	Y	PY	PY	Y	PY	Y	PY	PY	PY	40%
	Q9. Study selection	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
	Q10. Data collection process	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
	Q11. Data items	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
	Q12. Risk of bias in individual studies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
	Q13. Summary measures	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
	Q14. Synthesis of results	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
	Q15. Risk of bias across studies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
	Q16. Additional analyses	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
Results	Q17. Study selection	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
	Q18. Study characteristics	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
	Q19. Risk of bias within studies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
	Q20. Results of individual studies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
	Q21. Synthesis of results	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
	Q22. Risk of bias across studies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
	Q23. Additional analysis	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
Discussion	Q24. Summary of evidence	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
	Q25. Limitations	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
	Q26. Conclusions	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%
Funding	Q27. Funding	N	Y	Y	Y	Y	Y	Y	Y	Y	N	80%

and evidence quality, respectively, of the included studies did not meet the requirements of the Cochrane Handbook, which means that the conclusions of the included studies may be different from the real situation and cannot provide reliable evidence for the investigators. Third, there is considerable room for improving the methodological quality and reporting the quality of SRs/MAs. This study highlights common areas for improvement such as protocols designed and registered in advance, comprehensive literature search strategies, providing a transparent list of excluded trials, and declaring funding and conflicts of interest. Fourth, there is considerable room for improving the methodological quality of RCTs. This study highlights the common areas for improvement in RCTs, such as a strict implementation of randomization and allocation concealment and adequate reporting of blinding methods. In addition, while almost all outcome measures lead to positive

TABLE 4 | Results of evidence quality.

References	Treatments	Outcomes	Limitations	Inconsistency	Indirectness	Imprecision	Publication bias	Relative effect (95% CI)	Quality
Han (11)	AT vs. Sham AT	SS	−1	0	0	−1	−1	SMD −3.03 (−3.56, −2.50)	CL
		ER	−1	0	0	−1	−1	OR 5.09 (3.30, 7.86)	CL
		MF-ECG	−1	0	0	−1	−1	MD −14.36 (−18.31, −10.41)	CL
	AT vs. CM	SS	−1	0	0	−1	−1	MD −3.03 (−3.56, −2.50)	CL
		ER	−1	0	0	−1	−1	OR 1.33 (0.70, 2.52)	CL
		MF-ECG	−1	0	0	−1	−1	MD 0.01 (−0.03, 0.01)	CL
		PML	−1	0	0	−1	−1	SMD −0.06 (−0.23, 0.11)	CL
		GHMT	−1	0	0	−1	−1	MD 0.23 (−1.94, 2.40)	CL
Mao (12)	AT vs. Sham AT	SS	−1	0	0	−1	−1	MD −3.44 (−4.21, −2.67)	CL
		MF-ECG	−1	−1	0	−1	−1	SMD −0.69 (−3.02, 4.40)	CL
	AT vs. CM	SS	−1	−1	0	0	−1	SMD −0.18 (−0.51, 0.16)	CL
		ER	−1	0	0	0	0	RR 1.04 (0.96,1.13)	M
		PML	−1	0	0	0	0	SMD 0.93 (−0.30, 1.55)	M
		GHMT	−1	0	0	−1	−1	SMD 0.02 (−0.16, 0.20)	CL
Yuan (13)	AT vs. CM	ER	−1	0	0	0	−1	OR 3.00 (2.33, 3.87)	L
		SS	−1	−1	0	0	0	WMD −1.21 (−2.13, −0.30)	L
		PML	−1	0	0	0	0	WMD 13.99 (0.45, 27.54)	M
Zhou (14)	AT vs. Sham AT	SS	−1	−1	0	0	0	SMD −1.23 (−2.00,−0.47)	L
	AT vs. CM	SS	−1	−1	0	0	0	SMD −0.30 (−0.77,0.16)	L
		PML	−1	−1	0	0	0	SMD 0.67 (−0.07,1.42)	L
Pang (15)	AT vs. Sham AT	NDI	−1	−1	0	−1	−1	MD 20.91 (6.55,35.26)	CL
Kim (16)	AT vs. Sham AT	SS	−1	0	0	−1	−1	MD 0.54 (0.18, 0.90)	CL
		ER	−1	−1	0	−1	−1	RR 2.66 (1.85, 3.82)	CL
	AT vs. CM	SS	−1	−1	0	0	0	MD 0.54 (0.33,0.76)	L
		ER	−1	−1	0	−1	−1	RR 1.18 (1.09, 1.27)	CL
Lan (17)	AT vs. CM	ER	−1	0	0	−1	−1	RR 1.02 (0.91,1.16)	CL
Peng (18)	AT vs. CM	ER	−1	0	0	0	0	RR 1.17 (1.10,1.24)	M
Jin (19)	AT vs. CM	ER	−1	−1	0	0	0	RR 1.21 (1.15,1.27)	L
		SS	−1	−1	0	−1	−1	MD −2.10 (−3.61,−0.59)	CL
		NDI	−1	0	0	−1	−1	MD 9.94 (7.93,11.94)	CL
Wu (20)	AT vs. CM	ER	−1	0	0	0	0	RR 1.18 (1.11,1.24)	M

SS, symptom score; ER, effective rate; MF-EGG, main frequency of electrogastrogram; PML, plasma motilin level; GHMT, gastric half-emptying time; NDI, nepean dyspepsia index. CL, critically low; L, Low; M, moderate.

conclusions, it is worth noting that most authors were reluctant to draw definitive conclusions on the effects of acupuncture for FD due to the small size or low quality of the RCTs.

Impact on Further Research

This study highlights several challenges for producers of SRs/MAs that should be addressed. First, by using the AMSTAR-2 tool and the PRISMA statement, it was observed that the registration of protocols was far from being appreciated by authors of SRs/MAs. Only one Cochrane SR/MA provided a protocol registration, whereas none of the other studies provided that. It is well-known that a preregistration protocol helps to improve transparency, minimize the potential risk of bias arising from the research process, reduce duplicate work between different researchers, and keep the study up-to-date (22). Authors are advocated to register protocols in a freely open database (PROSPERO, http://www.crd.york.ac.uk/prospero) to avoid study bias. Secondly, the risk of bias in identification and literature selection was far from being paid attention to by authors of SRs/MAs. That is, authors of SRs/MAs should provide a comprehensive search strategy and be aware that an appropriate range of databases or electronic sources of published literature should be included. In addition to the database search, other methods were considered for identifying relevant literature, such as meeting reports and trial registration platforms. Furthermore, a list of excluded trials and account for exclusion help to

guarantee transparency; it can be presented by direct reference or as a supplementary file. Finally, funding sources should be fully reported as results from commercially funded research might be biased toward funders (23).

Based on the GRADE system, the risk of bias of the enrolled RCTs was the most common downgrading factor, indicating that the root cause of the reduced quality of evidence lies in the quality of the original RCTs. Only well-designed and rigorously conducted RCTs may reduce or avoid bias (24). Specific methods of randomization should be clearly described to reflect whether randomization has been successfully achieved. In addition, the information on allocation concealment and blind method should be fully reported. Due to the peculiarity of acupuncture, although blinding physicians is challenging, attempts should be made to blind patients, other caregivers, and outcome assessors to minimize the risk of bias. With the development of evidence-based acupuncture, it is hoped that researchers will continue to promote standardization and precision in the future by improving the technical operating procedures of acupuncture. Moreover, it should be realized that a large sample size and high-quality RCTs should be fundamental to high-quality evidence sources.

Limitations

This study highlights areas of improvement in the SR/MA process that may help guide future high-quality research. However, the widely used AMSTAR-2 tool, PRISMA statement, and GRADE system are all subjective evaluation tools, and the accuracy of assessor's assessments cannot be guaranteed.

CONCLUSION

Acupuncture seems to have a promising efficacy in the treatment of FD. Although the quality of the included SRs/MAs was generally low and defects were frequent, our study highlights areas where methodologies need to be improved.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

JH and JL initiated the study design and drafted the manuscript. ML, JingM, ZL, and JinxM helped with the implementation of this work. FW and XT contributed to methodology, review, and editing. All authors read and approved the final manuscript.

FUNDING

This study was supported by the National Natural Science Foundation of China [Nos. 81873297 and 81804078], the State Administration of Traditional Chinese Medicine Digestive Refractory Disease Inheritance and Innovation Team Project [No. ZYYCXTD-C-202010], and the Excellent Young Scientist Fund of China Academy of Chinese Medical Sciences [No. ZZ13-YQ-006].

REFERENCES

1. Stanghellini V, Chan FK, Hasler WL, Malagelada JR, Suzuki H, Tack J, et al. Gastroduodenal disorders. *Gastroenterology*. (2016) 150:1380–92. doi: 10.1053/j.gastro.2016.02.011
2. Kumar A, Pate J, Sawant P. Epidemiology of functional dyspepsia. *J Assoc Physicians India*. (2012) 60:9–12.
3. Aro P, Talley NJ, Agréus L, Johansson SE, Bolling-Sternevald E, Storskrubb T, et al. Functional dyspepsia impairs quality of life in the adult population. *Aliment Pharmacol Ther*. (2011) 33:1215–24. doi: 10.1111/j.1365-2036.2011.04640.x
4. Lacy BE, Weiser KT, Kennedy AT, Crowell MD, Talley NJ. Functional dyspepsia: the economic impact to patients. *Aliment Pharmacol Ther*. (2013) 38:170–7. doi: 10.1111/apt.12355
5. Chuah KH, Mahadeva S. Cultural factors influencing functional gastrointestinal disorders in the east. *J Neurogastroenterol Motil*. (2018) 24:536–43. doi: 10.5056/jnm18064
6. Masuy I, Van Oudenhove L, Tack J. Review article: treatment options for functional dyspepsia. *Aliment Pharmacol Ther*. (2019) 49:1134–72. doi: 10.1111/apt.15191
7. Huang J, Wu M, Liang S, Qin X, Shen M, Li J, et al. A critical overview of systematic reviews and meta-analyses on acupuncture for poststroke insomnia. *Evid Based Complement Alternat Med*. (2020) 2020:2032575. doi: 10.1155/2020/2032575
8. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*. (2017) 358:j4008. doi: 10.1136/bmj.j4008
9. Moher D, Liberati A, Tetzlaff J, Altman D. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*. (2009) 339:b2535. doi: 10.1136/bmj.b2535
10. Pollock A, Farmer SE, Brady MC, Langhorne P, Mead GE, Mehrholz J, et al. An algorithm was developed to assign GRADE levels of evidence to comparisons within systematic reviews. *J Clin Epidemiol*. (2016) 70:106–10. doi: 10.1016/j.jclinepi.2015.08.013
11. Han, J, Wang HC, Rong P, et al. Systematic review and meta-analysis of the therapeutic effect on functional dyspepsia treated with acupuncture and electroacupuncture. *World J Acupunct Moxibustion*. (2021) 31:44–51. doi: 10.1016/j.wjam.2020.11.015
12. Mao X, Guo S, Ni W, et al. Electroacupuncture for the treatment of functional dyspepsia: a systematic review and meta-analysis. *Medicine*. (2020) 99:e23014. doi: 10.1097/MD.0000000000023014
13. Yuan XX, Wang BY, Yang L, Li D-d, Liang H, Zhang Y-l. Efficacy of manual acupuncture on functional dyspepsia: a meta-analysis of randomized, controlled trials. *Afr J Tradit Complement Altern Med*. (2017) 14:38–55. doi: 10.21010/ajtcam.v14i6.5
14. Zhou W, Su J, Zhang H. Efficacy and safety of acupuncture for the treatment of functional dyspepsia: meta-analysis. *J Altern Complement Med*. (2016) 22:380–9. doi: 10.1089/acm.2014.0400
15. Pang B, Jiang T, Du YH, Li J, Li B, Hu YC, et al. Acupuncture for Functional Dyspepsia: What Strength Does It Have? a systematic review and meta-analysis of randomized controlled trials evid based complement. *Alternat Med*. (2016) 2016:3862916. doi: 10.1155/2016/3862916
16. Kim KN, Chung SY, Cho SH. Efficacy of acupuncture treatment for functional dyspepsia: a systematic review and meta-analysis. *Complement Ther Med*. (2015) 23:759–66. doi: 10.1016/j.ctim.2015.07.007

17. Lan L, Zeng F, Liu GJ, Ying L, Wu X, Liu M, et al. Acupuncture for functional dyspepsia. *Cochrane Database Syst Rev*. (2014) 10:CD008487. doi: 10.1002/14651858.CD008487.pub2
18. Peng JR, Tao SY, Yang CX, et al. Efficacy of acupuncture on dyspesia subtypes: a systematic review. *J Tradit Chin Med*. (2018) 45:1594–8.
19. Jin X, Cai WR, Xu PD. The effect of acupuncture and moxibustion on functional dyspepsia compared with prokinetic agents: a meta-analysis. *World Latest Medicine Information*. (2018) 18:9–12.
20. Wu XW, Ji HZ, Xu LE, et al. The effect of acupuncture and moxibustion on functional dyspepsia compared with prokinetic agents: a meta-analysis. *Chin J Integr Med.* (2015) 23:100–4.
21. Manchikanti L, Benyamin RM, Helm S, Hirsch JA. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: part 3: systematic reviews and meta-analyses of randomized trials. *Pain Physician.* (2009) 12:35–72. doi: 10.36076/ppj.2009/12/35
22. PLoS Medicine Editors. Best practice in systematic reviews: the importance of protocols and registration. *PLoS Med*. (2011) 8:e1001009. doi: 10.1371/journal.pmed.1001009
23. Schmieder RE, Neuzil P. Scientific data and transparency of conflict of interest are important, not biased editorial without facts. *JACC Cardiovasc Interv*. (2016) 9:2263. doi: 10.1016/j.jcin.2016.09.004
24. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomized trials. *BMJ.* (2010) 340:c869. doi: 10.1136/bmj.c869

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Huang, Liu, Liu, Ma, Ma, Lv, Wang and Tang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Pediatrics

ORIGINAL RESEARCH
published: 14 April 2022
doi: 10.3389/fped.2022.775378

Retrospective Analysis of the Accuracy of High-Frequency Ultrasound for Pancreaticobiliary Maljunction in Pediatrics at a Single Center

*Qiuchen Xu, Min Liu, Qiumei Wu, Wen Ling and Shan Guo**

Department of Medical Ultrasonics, Fujian Maternity and Child Health Hospital, Affiliated Hospital of Fujian Medical University, Fuzhou, China

Objective: To determine the clinical value of high-frequency ultrasonography (US) in the evaluation and diagnosis of pancreaticobiliary maljunction (PBM) among children.

Methods: The clinical subjects consisted of 31 pediatric patients who were diagnosed with PBM from January 2015 to May 2021 in Fujian Provincial Maternity and Children's Hospital. The primary outcomes included diagnosis accuracy, imaging characteristics of each type of PBM based on JSPBM, time length of operation, and cost of service. Secondary outcomes were the serum amylase and bilirubin levels.

Results: The diagnostic accuracy of US was 90.3% and comparable to the other imaging methods—MRCP (82.6%), IOC (79.2%), and ERCP (100%), respectively. The time length of operation and direct cost were significantly lower than other imaging pathways. Stenotic type (A) is associated with a high internal diameter of CBD, and dilated channel type (C) presents increased internal diameter and length of CC as well as internal diameter of PD. There were higher levels of the serum bilirubin seen in type A and of serum amylase in type C compared with others.

Conclusion: High-frequency US is a safe, cost-effective, and non-invasive imaging tool for the diagnosis and evaluation of PBM in pediatrics.

Keywords: ultrasonography, pancreaticobiliary maljunction (PBM), diagnostic imaging, congenital biliary dilatation, magnetic resonance cholangiopancreatography, endoscopic retrograde cholangiopancreatography, intraoperative cholangiography

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Hong Han,
Fudan University, China
Lijun Tang,
Wake Forest University, United States

****Correspondence:***
Shan Guo
shanguo21@yahoo.com

Specialty section:
This article was submitted to Pediatric Gastroenterology, Hepatology and Nutrition, a section of the journal Frontiers in Pediatrics

Received: *13 September 2021*
Accepted: *08 February 2022*
Published: *14 April 2022*

Citation:
Xu Q, Liu M, Wu Q, Ling W and Guo S (2022) Retrospective Analysis of the Accuracy of High-Frequency Ultrasound for Pancreaticobiliary Maljunction in Pediatrics at a Single Center. Front. Pediatr. 10:775378. doi: 10.3389/fped.2022.775378

INTRODUCTION

Pancreaticobiliary maljunction (PBM) is a congenital anomaly in which a junction of the pancreatic and bile ducts is anatomically outside the duodenal wall, forming a relatively long common channel above the sphincter of *Oddi*. Half a century ago, it was widely recognized as an abnormal union of biliopancreatic ducts or anomalous junction of the pancreaticobiliary ductal system (1). According to the presence or absence of the bile duct dilatation, PBM can be classified as PBM with congenital biliary dilatation and PBM without biliary dilatation (maximal diameter of the bile duct $\leq$ 10 mm).

In PBM, the regurgitation between the pancreatobiliary and biliopancreatic tract systems and premalignant changes in the gallbladder can occur due to loss of the regulation of the sphincter of *Oddi* (1–3). Patients with PBM have been found to have a high incidence of gallbladder and biliary tract cancers (4, 5). According to the report by a Japanese study (6), biliary cancer was seen in 21.6% of adult patients with congenital biliary dilatation and in 42.4% of patients with PBM without biliary dilatation. The cancer was located in the bile duct and gallbladder in between 32.1 and 62.3% of patients with dilatation and between 7.3 and 88.1% of patients with PBM without biliary dilatation. Moreover, these cancers presented 15–20 years earlier in patients with PBM than in patients without PBM. The incidence of biliary tract cancers in pediatric patients (younger than 15 years of age) with PBM is unknown, whereas a few case series/reports were published (7).

The main symptoms of PBM include abdominal pain, vomiting, jaundice, fever, and hyperamylasemia. Compared with patients with PBM without biliary tract dilatation, a larger proportion of patients with congenital biliary dilatation are symptomatic and related to the formation of the dilated duct. Many patients with neonatal/infant onset have been classified as showing cystic subtype in bile duct, in which main symptoms are jaundice and an abdominal mass. Those with the early onset version of the disease have a bile duct of a fusiform/cylindrical type and also present abdominal pain mostly due to increased intra-pancreatic/biliary pressure or due to the common channel being blocked by protein plugs. Generally, patients with non-dilated PBM often show symptoms in adulthood (8–10). Surgery is recommended for patients with PBM irrespective of the presence or absence of symptoms (7).

Abdominal ultrasonography (US) has conventionally been recommended as a primary imaging modality in the care of pediatric patients because of not only lower cost than other techniques, but also it being safety and being child-friendly. High-frequency ultrasound generally refers to the ultrasound probe frequency of more than 10 MHz. Because of its shorter wavelengths, it is absorbed easily for abdominal structures and, hence, is a reliable and non-invasive technique that is being used along with pediatric examination for assessment, diagnosis, and management. However, informative findings about the value of high-frequency US in the diagnosis of PBM have been limited. Although the development of imaging modalities including MRCP, ERCP, and IOC provided advantages for pediatric radiologists in early diagnosis of PBM, escalating medical costs is becoming more and more a concern in worldwide pediatric care clinics, suggesting a primary demand for pediatrics is reducing cost. We retrospectively analyzed our pediatric high-frequency US practice with diagnosis accuracy in clinical diagnosis of PBM. This study also explored the time length of operation and direct cost of different imaging tools in pediatric practice.

Abbreviations: PBM, Pancreaticobiliary maljunction; US, Ultrasonography; JSPBM, Japanese Study Group on Pancreaticobiliary Maljunction; MRCP, Magnetic resonance cholangiopancreatography; ERCP, Endoscopic retrograde cholangiopancreatography; IOC, Intraoperative cholangiography; CBD, Common bile duct; PD, Pancreatic duct; CC, Common duct; PBJ, Pancreaticobiliary junction; DU, Duodenum.

MATERIALS AND METHODS

Patients

This study was approved by the Institutional Clinical Research Ethics Committee of Fujian Provincial Maternity and Children's Hospital. Written informed consent was obtained from the parents or guardians of all participants.

In this retrospective study, all pediatric patients with clinical features of biliary obstructive disease were included in the study. The US findings among 31 children (age < 10 years) diagnosed with PBM between January 2015 and May 2021 at our Medical Center were analyzed. To confirm the diagnosis, we used the gold standard imaging tools—MRCP, IOC, and ERCP, which results were evaluated by radiologist blinded to the US findings. The morphologic characteristics of PBM were described according to the last classification proposed by JSPBM, which is based on the formation of the pancreaticobiliary junction (11). The four types are stenotic type (A) in which the distal common bile duct with stenosis joins the common channel, non-stenotic type (B) in which the distal common bile duct without stenosis joins the common channel, dilated channel type (C) in which the common channel is dilated, and complex type (D). The length of operation and the direct cost analysis in medical servings as high-frequency, MRCP, ERCP, and IOC were calculated on the basis of the updated standards issued by the national health and family planning administration.

The primary outcomes were diagnosis accuracy, imaging characteristics of each type of PBM, time length of operation, and cost of service. Secondary outcomes were the serum amylase and bilirubin levels associated with PBM types. The serum bilirubin level was determined on the basis of an orthodox diazo method, in which bilirubin reacts with diazotized sulfanilic acid to produce azobilirubin (purplein color). Intensity of color is directly proportional to the amount of bilirubin in the serum. To quantify the activity of amylase, the iodometric (amyloclastic) method was used to measure the rate at which substrate (starch) is reacted away. Iodine readily reacts with starch to produce a purple color, in which intensity is quantitatively associated with of the activity of amylase.

High-Frequency US Technique

The high-frequency US facility was performed by the Voluson S8 scanner (GE Medical Systems, Waukesha, WI, USA) equipped with 9-MHz/11-MHz linear array transducers. All young patients underwent US imaging in the supine position within the sonography department at the hospital center. There was no additional force applied other than the weight of the probe to avoid any extra pain in the patients. During the operation, we applied a specific technique of graded compression to eliminate the disturbing impact of bowel gas. It allowed the use of a high-frequency probe for the assessment of the morphological abnormality by evaluating its reaction on compression. The compression should be applied slowly and gently to reduce abdominal pain. Four to five vertically oriented and overlapping

lanes in the peritoneal cavity were needed to obtain better image quality using probes. The maximum inner diameters and lengths of the common bile duct, pancreatic duct, and common channel were reported by three experienced pediatric sonographers. The existence of a protein plug within the ductal system was observed. The electronic clinical records, sonography/radiology reports, and laboratory data of the children were analyzed for evidence of accuracy, sensitivity, and specificity in accordance with the other imaging diagnosis.

Statistical Analysis

Data for continuous variables were analyzed for normality by Kolmogorov–Smirnov test and presented as mean ± SD. The *chi*-square or Fisher's exact tests were used to compare categorical variables and presented as mean and frequency with percentages. The normally distributed continuous variables were compared by unpaired *t*-test, whereas the others were tested by Mann–Whitney U test. Significant differences were defined as those with p-value < 0.05. All analyses were performed using SPSS statistical software (version 22).

RESULTS

Clinical Characteristics

Of the 31 patients included in this study, there were 24 girls and 7 boys. The age distribution and clinical symptoms is shown in **Table 1**. Abdominal pain was the most common complaint (77.4%), followed by vomiting (32.3%), jaundice (22.6%), and fever (9.7%). One patient had chromosomal abnormality (*Down*'s syndrome) confirmed by karyotype test.

TABLE 1 | Clinical characteristics of pediatrics with PBM.

		***N* = 31**
Gender	Girl	24 (77.4%)
	Boy	7 (22.6%)
Age (overall range)		2 days−9 years
Clinical manifestation		
Abdominal pain		24 (77.4%)
Vomiting		10 (32.3%)
Jaundice		7 (22.6%)
Fever		3 (9.7%)
Down's syndrome		1 (3.2%)

The Primary Outcomes

Among 31 cases of PBM, the vision of high-frequency US was read thoroughly by three sonographers. Further imaging with MRCP, IOC, or ERCP was considered to confirm the features suggestive of PBM. The overall diagnosis rate, time length of operation, and direct cost of service are shown in **Table 2**.

Ultrasound accurately diagnosed PBM in 28 of 31 cases (90.3%) included in the study. Ultrasound was unable to diagnose a specific cause for three cases where false visualization in the common channel and common bile duct was read because of excessive intestinal gas. MRCP examination was performed in 23 cases, out of which 19 cases were accurately diagnosed PBM and four cases were false negative due to the very young age (2 days) with dilated CBD. IOC scan confirmed PBM in 19 of 24 children, whereas five were false negative due to failure to read common ducts (four cases) and being out of time as administration of the contrast agent drained into the duodenum (one case). ERCP was performed in nine cases, all of which diagnosed PBM accurately.

The snapshot of the time taken in imaging operation in diagnosis of PBM was summarized in **Table 2**. The high-frequency US procedure needed about half an hour in our clinical department. Over three times of US running operation time were needed in MRCP and ERCP examinations. IOC also needed a longer operation time compared with US measurement. Half of parents or guardians were concerned about the challenging issues from the medical imaging. Very few participants reported dissatisfactions or complaints just before/after the high-frequency US procedure.

Total per-patient direct costs in the medical center were \$34 ± 3.3 for high-frequency US and \$270 ± 32 for MRCP. IOC and ERCP services costed \$90 ± 10.2 and \$388 ± 61, respectively. We calculated that, as a result of administrating a high-frequency US examination, 11 (35.5%) MRCP and 13 (41.9%) IOC cases were avoided. The potential cost savings for use of high-frequency US in pediatric PBM instead of MRCP, IOC, and ERCP would be \$236, 56, and 354 per examination.

Ultrasonic Characteristics of PBM in Children

On the basis of the new JSPBM's classification, the individual morphologic features of PBM were described as four types (11)

TABLE 2 | Overall diagnosis rates, time length of operation, and cost of service in pediatrics with PBM.

	US			**MRCP**			**IOC**			**ERCP**		
	***N* = 31**			***N* = 23**			***N* = 24**			***N* = 9**		
	DA	**SE**	**SP**	**DA**	**SE^**	**SP^**	**DA**	**SE#**	**SP#**	**DA**	**SE^**	**SP^**
Diagnostic rate of PBM (%)	90.3	90.9	88.9	82.6	50–52.9	77.8–100	79.2	78.9	100.0	100.0	50–52.9	77.8–100
Time length of operation (min)		30 ± 4*			106 ± 19			176 ± 48			129 ± 31	
Cost (\$/case)		34 ± 3.3*			270 ± 32			90 ± 10.2			388 ± 61	

*DA, diagnosis accuracy; SE, sensitivity; SP, specificity; *p < 0.05 versus the other three tools. Based on the current sample studied, no expected SE/SP was available in groups of MRCP, IOC and ERCP. The data were shown according to the literatures about ^MRCP and ERCP applications in bile duct dilatation/occlusion and pancreatic duct occlusion/strictures (12), and #IOC in bile duct injury (13).*

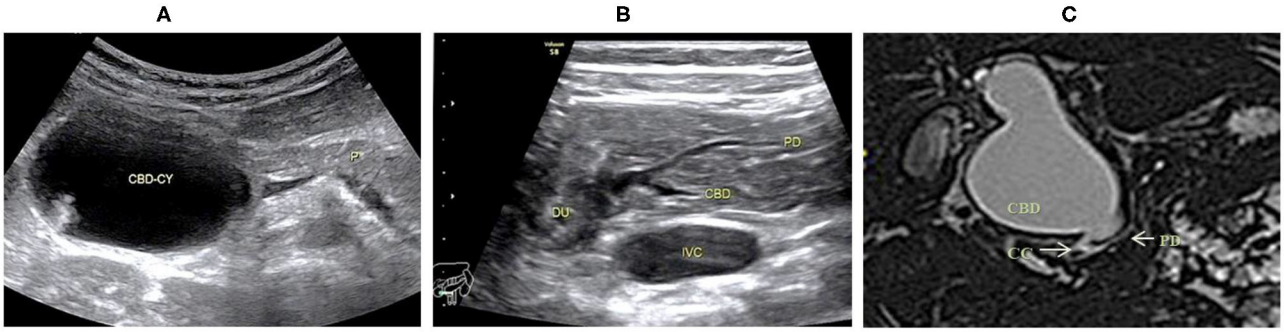

FIGURE 1 | Images of a representative type A of PBM. **(A)** US transverse scan showed the dilated CBD containing the protein plugs; **(B)** US transverse scan showed the PD joining the terminal stenotic CBD within pancreas to form a common channel which drains into the duodenum; **(C)** MRCP indicated the dilated CBD with a narrow distal end joining the common channel.

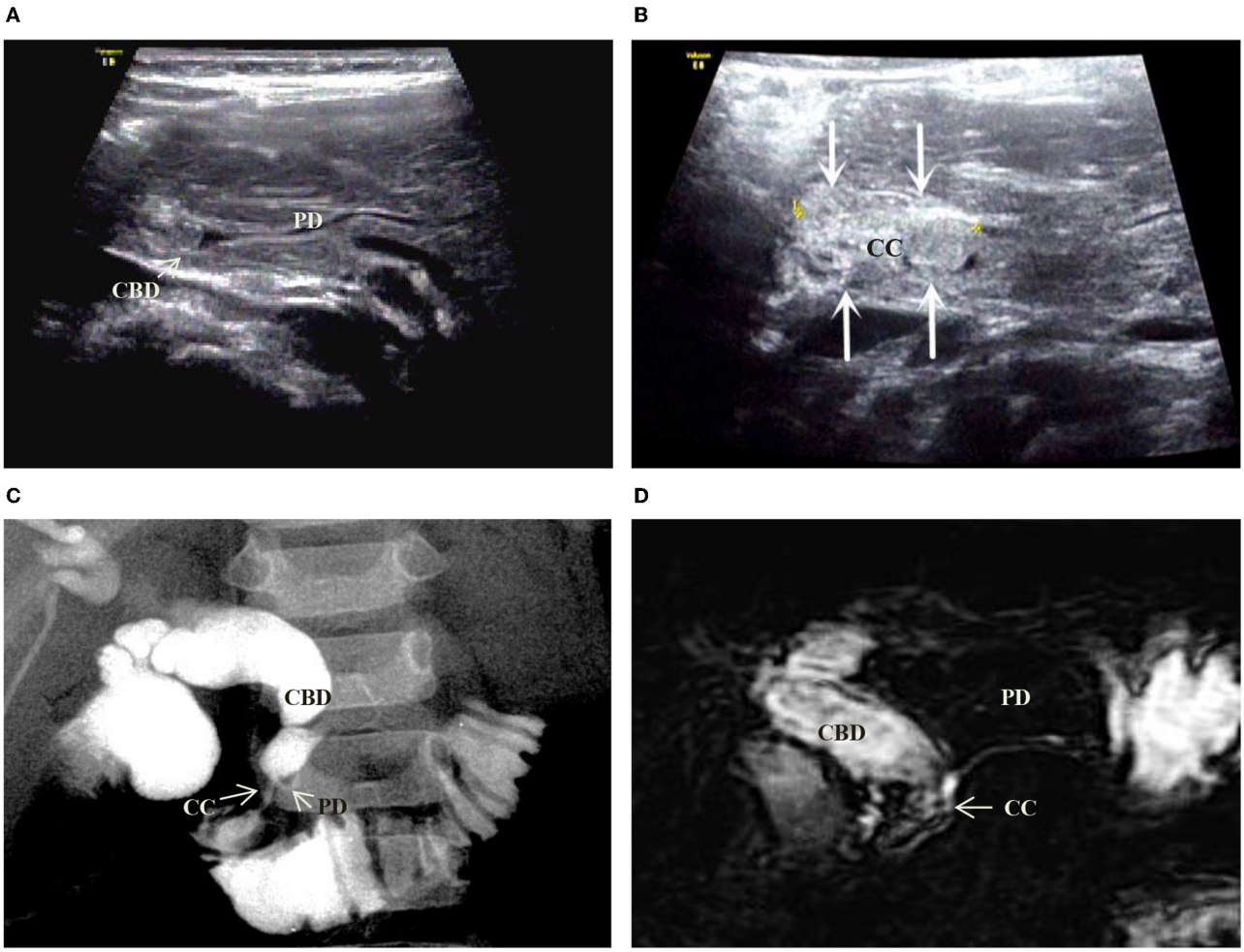

FIGURE 2 | Images of a representative type B of PBM. **(A)** US transverse scan showed the PD (longitudinal section) joining the CBD which contains the protein plugs. **(B)** US oblique scan showed the common channel (between the arrows) is dilated and filled with protein plugs. **(C)** IOC indicated the PD and the dilated CBD joining at a right angle to form a common channel which inserts into the duodenum. **(D)** MRCP indicated the same features of PD and CBD as that in US and IOC. No significant narrowness in the distal CBD.

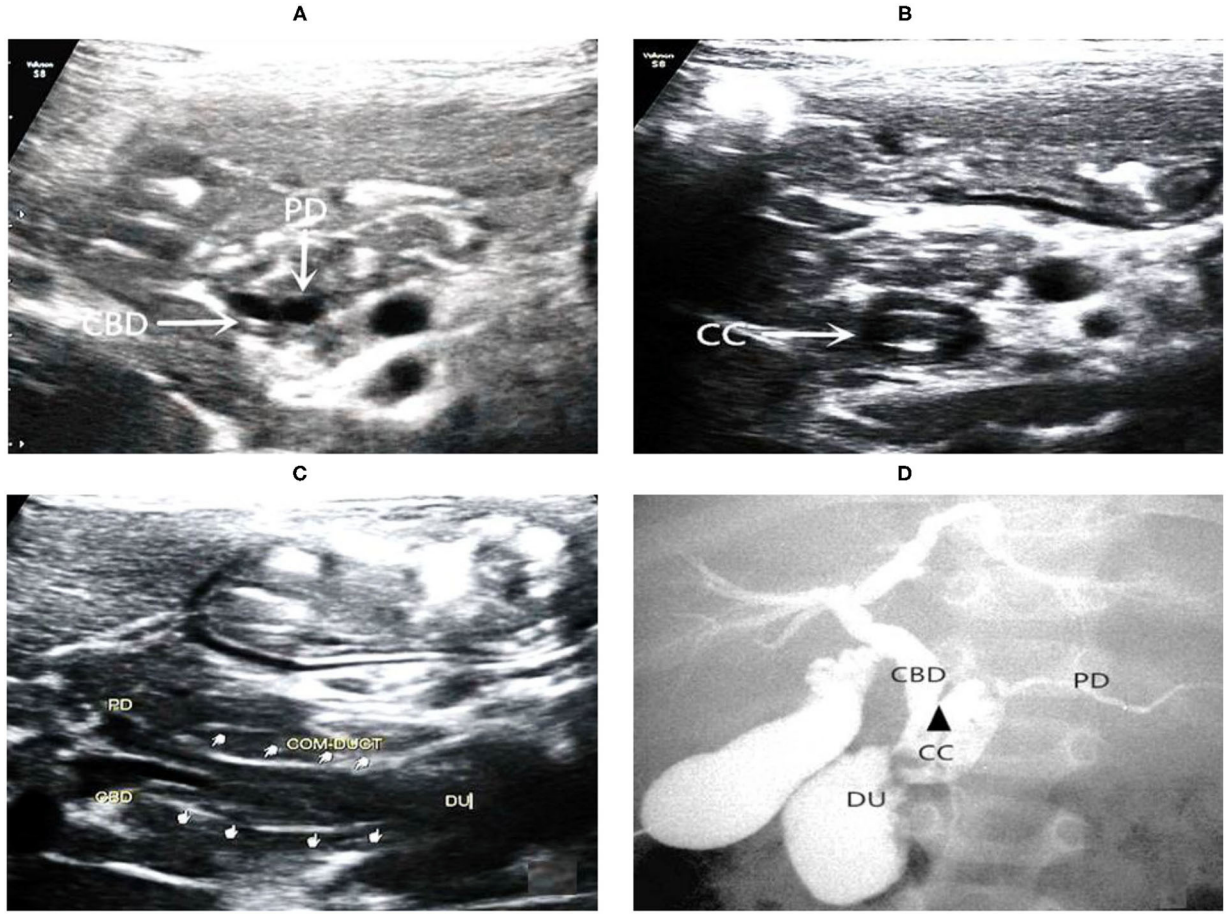

FIGURE 3 | Images of a representative type C of PBM. **(A)** US transverse scan showed the axial section of both the terminal of the CBD and the PD; **(B)** US transverse scan showed the axial section of common channel and indicated protein plugs within; **(C)** US oblique scan showed the terminal CBD and the PD parallel each other and join at the end into the common channel which drains into the duodenum; **(D)** IOC confirmed dilations of the common channel and the CBD. There is an acute angle (arrow head) of interjunction of the PD and CBD.

in presenting our ultrasound report: type A (stenotic)—the distal CBD with stenosis before joining with the CC, and an angle of slightly $< 90^{\circ}$ usually observed in the conjunction between PD and CBD at a site distant from the papilla of *Vater*; type B (non-stenotic)—the distal CBD emerging but without stenosis pattern before joining with the CC, in addition, an approximately right angle customarily seen in the conjunction between PD and CBD at a site distant from the papilla; type C (dilated channel)—the dilated CC significantly occurred; type D (complex)—a complicated configuration in PBJ contexture including the PD, the terminal portion of the CBD, and the dilated accessory PD, and one end opening as a small nipple in the duodenum and the other end converged with the CBD downward into the CC. **Figures 1–4** show the representative imagines of A, B, C, and D types of PBM, respectively.

Comparison of Morphological Classification

Among the 31 cases, 14 cases were stenotic type (A), 11 cased were non-stenotic (B), and five were dilated channel (C). The complex type (D) rarely occurred and only one case was found in our medical center. **Table 3** shows the comparison among A, B, and C types in terms of gender, age, the maximum internal diameter of CBD, the internal diameter of CC, the length of CC, the internal diameter of the main PD, and the thickness of the gallbladder. Type A causes a significant increase in the internal diameter of CBD. Type C appears significant increases in internal diameter and length of CC and internal diameter of PD.

The Secondary Outcomes

The serum bilirubin/amylase concentrations were shown in **Table 3**. Type A presented a higher level of serum bilirubin than other types, and the serum amylase level showed a significant higher in type C. All 31 children underwent extrahepatic bile duct resection and hepaticojejunostomy with Roux-en-Y anastomosis. No difference in biliary amylase levels was seen among the types (**Table 3**). The detective rates of protein plug were indicated individually among three types (**Table 3**), and the case of Type D was pointed out a positive image of protein plug in the US. The result showed that the overall detective rate of protein plug

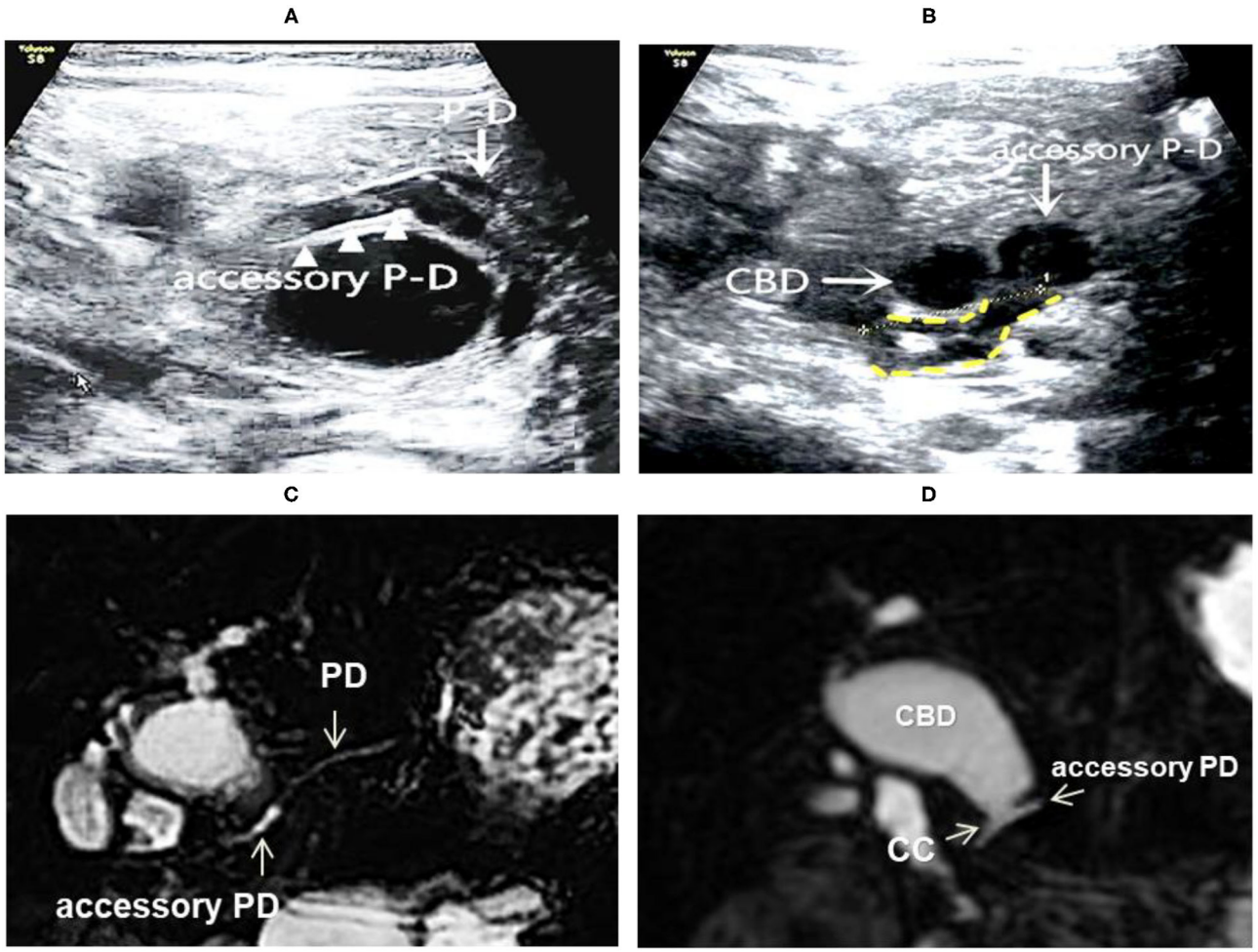

FIGURE 4 | Images of a representative type C of PBM. **(A)** US transverses scan showing the main PD merging into the dilated accessory PD; **(B)** US transverse scan showing the terminal CBD and the accessory PD joining downward to form the common channel (shown by yellow dotted lines); **(C)** MRCP indicated the main PD imports to the accessary PD; **(D)** MRCP indicated the accessary PD joins with the distal CBD to form the common channel, suggesting a type D.

TABLE 3 | Comparison of clinical data among the types (A, B, and C) of PBM.

PBM type	A	B	C
Girl/Boy	11/3	9/2	3/2
Age (range)	2 days−5 years	2–9 years	1.17–4.33 years
Internal diameter of CBD (mm)	26.43 ± 9.17*	12.66 ± 7.23	18 ± 8.17#
Internal diameter of CC (mm)	2.07 ± 0.87	4.23 ± 1.72	7.00 ± 2.11*
Length of CC (mm)	10.78 ± 2.25	11.82 ± 3.10	17.4 ± 3.73*
Internal diameter of PD (mm)	1.29 ± 0.45	1.74 ± 0.67	2.06 ± 0.79*
Thickness of gallbladder (mm)	2.27 ± 0.94	2.58 ± 0.95	2.48 ± 0.53
Serum bilirubin (μmol/L)	57.28 ± 23.19*	17.58 ± 7.87	15.26 ± 7.71
Serum amylase (U/L)	213.87 ± 48.24	245.17 ± 78.18	491.78 ± 151.30*
Bile amulase (U/L)	6,888.83 ± 2,250.88	7,247.03 ± 1,675.53	9,761.77 ± 2,200.40
Detective rate of protein plug (%)	85.7 (12/14)	100 (11/11)	100 (5/5)

A, stenotic type; B, non-stenotic type; C, dilated channel type; CBD, common bile duct; CC, common duct.
$^{*}p < 0.05$ vs. the other two types.
$^{\#}P < 0.05$ vs. type B.

was 93.5%. The surgical treatment confirmed protein plugs in the distal segment of CBD and/or CC reported by preoperative US.

DISCUSSION

In the present study, we mainly aim to describe the current application of high-frequency US in the diagnosis of PBM and associated morphological findings in our pediatric center. Because of its higher resolution (16–158 μm) compared with ordinary ultrasound, high-frequency US was recognized as the ultrasound biomedical microscope (14, 15). In our investigation, its imaging qualities allow us clearly obtain the structure and thickness of ducts, channel, and gallbladder. Pediatric sonographers can diagnose the PBM on the basis of these results and can evaluate the follow-up treatment (data not shown). More importantly, our study suggests that high-frequency US is non-invasive, convenient, practical, and safe; therefore, it should gain more application in the field of pediatric PBM.

PBM is characterized by an abnormally long common channel and/or an abnormal union between the pancreatic and bile ducts. In 2020, the JSPBM proposed a new PBM classification, which is based on the formation of the PBJ especially the presence of dilatation in ductal structure (11). The four types classified by JSPBM provide us an actionable direction in early diagnosis of PBM with or without biliary dilatation by high-frequency US. Although the imaging methods used to demonstrate pancreatic and bile ducts remain direct cholangiography (ERCP, percutaneous trans-hepatic cholangiography, or IOC), MRCP, or three-dimensional drip infusion cholangiography computed tomography (1), clinical features of pediatric patients conceivably make quite difficult to support prevailing practice with the above golden standard tools. For example, ERCP requires general anesthesia and has potentially serious complications such as pancreatitis, hemorrhage, and bowel perforation. A recent study of 116 children demonstrated a morbidity of 3.4%, with a procedural failure rate of 2.5% (mainly due to the inability of major papilla catheterization) (16). Our practice suggests that ERCP operation is more difficult than sonography among infants and toddlers because the parents/guardians poorly accept ERCP operation for their children especially those <5 years. MRCP is a non-invasive, radiation-free, and safe procedure in examination of the pancreatic and liver parenchyma in addition to the ductal structures. However, the main disadvantage is limited spatial resolution especially in screening a PD $<$ 1 mm, compared with high-frequency US. In our study, the range of average internal diameter of the PD in children with PBM was 1.29 $\pm$ 0.45 to 2.06 $\pm$ 0.79 mm; additionally, MRCP hardly pinpointed a distinct exposure of pancreas in one case, overall resulting in a poor visualization and a lower accuracy (82.6%) than high-frequency US. Our findings are similar to other reports, which showed a 40–80% accuracy rate in diagnosis of PBM by MRCP (17–19). CT offers a higher resolution in abdominal contents than MRCP, but it has not been widely used in children due to exposure to ionizing radiation and usually iodinated contrast media, both of which have known safety concerns (20). MRI provides excellent evaluation of organ characteristics without exposure to ionizing radiation, but it has the difficulties in rendering a young patient sedation for the examination and the potential toxic influence of gadolinium-based contrast agents (21).

The abdominal wall of young children has thinner musculature and less abdominal adipose tissue than old children and adults; the solid organs are comparatively larger in young children than in adults, which indicates that more surface area is exposed, making the bile duct and pancreas accessible for high-frequency US to achieve. Although a low resolution has been associated with a depth of its scanning organs, high-frequency US did fulfill its detection among our young children (averaged 5 years old). Furthermore, high-frequency US is a non-invasive, radiation-free, high-resolution, and real-time dynamic technique that does not require sedation or have complications. One of the advantages in preoperative diagnosis of PBM by high-frequency US is that the scanning image not only can discern the abnormal confluence of the PD and the CBD but also further reveal the CC with slight rotation of the probe and continued examination to the extent possible. In addition, high-frequency US provides a satisfying measurement of the inner diameters of the ductal structures and visualizes protein plugs within, if present in the current study. For example, the overall detection rate of protein plugs is 95.2% (85.7–100%) in our study, which is higher than one reported for cholangiography and MRCP images (9.3–40%) (22, 23). The resulting difference displays that protein plugs are translucent in cholangiography and MRCP images but are detectable in high-frequency US examination, in which it presents acoustic images with homogeneous and moderately echoic characteristics. In the process of PBM diagnosis, because cholangiography or MRCP accurately showed the filling of contrast agent or liquid in the duct or channel, the protein plugs inside of it can falsely result in the illusion of lumen stenosis. In contrast, the high resolution of high-frequency US can truly clarify the inner diameter of the CBD/CC even with the presence of protein plugs.

Our study presents 90.9% in sensitivity by high-frequency US. The rate is lower than other report (24) with 100% in diagnosis of PBM. The difference could be possibly due to younger age in our study. Our results demonstrate a specificity of 88.9% on high-frequency US, which is similar with other studies with 88.4% in detecting the benign etiology of obstruction (25). There are a few potential reasons for false-negative ultrasound detection. One is that we could be able to overlook the inflamed PBJ. Other is the existence of air-filled dilated bowel loops that may hide ductal structure from view. Another pitfall is some demonstration of the part of ducts/channels within the scanning field because it is obscured or confused by multiple-sized protein plugs with inflamed tissues.

The fact that the US of abdominal organs is more susceptible to gastrointestinal gas was recognized in the investigation and would be most likely the one of reasons for false-negative detection by high-frequency US. Because of excessive intestinal gas in two cases, ultrasound visualization was limited to accessibility of the common channel and common bile duct, therefore a false reading occurring. To reduce the influence of

the GI gas, some gentle compression should be recommended to pull way the bowel gas during the operation.

Our study indicated the financial and time advantages of using high-frequency US in pediatric PBM diagnosis, particularly in reducing the direct serving cost and cutting down time in imaging operation compared with the amount of other imaging tools performed. It is suggested that high-frequency US can be the recommended imaging modality of choice in pediatric PBM diagnosis.

Our study showed that the serum amylase concentration of patients with type C PBM was higher than other types of PBM. The possible explanation is that the dilated CC may have higher inner pressure against the pancreatic fluid upstream, or a smaller intersection angle of PD with CC is quite sharp to potentially obstruct the fluid from pancreas. However, further studies need to confirm the certain mechanisms on it.

Limitations

The present study was limited by its retrospective nature and small sample size, e.g., type D. Some cases showed that PBJ region has poor visualizations on US, resulting in low-quality imagines.

CONCLUSION

High-frequency US is a safe and cost-effective diagnostic tool for suspected cases of PBM in children. Apart from the diagnosis, it can provide valuable information about the type of PBM and the ductal structure as well as protein plugs.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institutional Clinical Research Ethics Committee of Fujian Provincial Maternity and Children's Hospital: https://www.natureindex.com/institution-outputs/china/fujian-provincial-maternity-and-children-s-hospital/5df1dfe3051f67680f677c40. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

ML, QW, and WL: data collection. QX: draft manuscript preparation. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. Kamisawa T, Ando H, Hamada Y, Fujii H, Koshinaga T, Urushihara N, et al. Diagnostic criteria for pancreaticobiliary maljunction 2013. *J Hepatobiliary Pancreat Sci.* (2014) 21:159–61. doi: 10.1002/jhbp.57
2. Ono S, Fumino S, Iwai N. Diagnosis and treatment of pancreaticobiliary maljunction in children. *Surg Today.* (2011) 41:601–5. doi: 10.1007/s00595-010-4492-9
3. Guo WL, Huang SG, Wang J, Sheng M, Fang L. Imaging findings in 75 pediatric patients with pancreaticobiliary maljunction: a retrospective case study. *Pediatr Surg Int.* (2012) 28:983–8. doi: 10.1007/s00383-012-3159-6
4. Kamisawa T, Kuruma S, Chiba K, Tabata T, Koizumi S, Kikuyama M. Biliary carcinogenesis in pancreaticobiliary maljunction. *J Gastroenterol.* (2017) 52:158–63. doi: 10.1007/s00535-016-1268-z
5. Ragot E, Mabrut JY, Ouaïssi M, Sauvanet A, Dokmak S, Nuzzo G, et al. Pancreaticobiliary maljunctions in European patients with bile duct cysts: results of the multicenter study of the French Surgical Association (AFC). *World J Surg.* (2017) 41:538–45. doi: 10.1007/s00268-016-3684-x
6. Kamisawa T, Ando H, Shimada M, Hamada Y, Itoi T, Takayashiki T, et al. Recent advances and problems in the management of pancreaticobiliary maljunction: feedback from the guidelines committee. *J Hepatobiliary Pancreat Sci.* (2014) 21:87–92. doi: 10.1002/jhbp.8
7. Le Roy B, Gagnière J, Filaire L, Fontarensky M, Hordonneau C, Buc E. Pancreaticobiliary maljunction and choledochal cysts: from embryogenesis to therapeutics aspects. *Surg Radiol Anat.* (2016) 38:1053–60. doi: 10.1007/s00276-016-1669-y
8. Tashiro S, Imaizumi T, Ohkawa H, Okada A, Katoh T, Kawaharada Y, et al. Pancreaticobiliary maljunction: retrospective and nationwide survey in Japan. *J Hepato Biliary Pancreatic Surg.* (2003) 10:345–51. doi: 10.1007/s00534-002-0741-7
9. Iwai N, Fumino S, Tsuda T, Ono S, Kimura O, Deguchi E. Surgical treatment for anomalous arrangement of the pancreaticobiliary duct with nondilatation of the common bile duct. *J Pediatr Surg.* (2004) 39:1794–6. doi: 10.1016/j.jpedsurg.2004.08.010
10. Ohuchida J, Chijiiwa K, Hiyoshi M, Kobayashi K, Konomi H, Tanaka M. Long-term results of treatment for pancreaticobiliary maljunction without bile duct dilatation. *Arch Surg.* (2006) 141:1066–70. doi: 10.1001/archsurg.141.11.1066
11. Ono A, Arizono S, Isoda H, Togashi K. Imaging of pancreaticobiliary maljunction. *Radiographics.* (2020) 40:378–92. doi: 10.1148/rg.2020190108
12. Hurter D, De Vries C, Potgieter P, Barry R, Botha F, Joubert G. Accuracy of MRCP compared to ERCP in the diagnosis of bile duct disorders. *SA J Radiology.* (2008) 12:14–22.
13. Alvarez FA, De Santibañes M, Palavecino M, Sánchez Clariá R, Mazza O, Arbues G, et al. Impact of routine intraoperative cholangiography during laparoscopic cholecystectomy on bile duct injury. *J Bri Surg.* (2014) 101:677–84.
14. Jasaitiene D, Valiukeviciene S, Linkeviciute G, Raisutis R, Jasiuniene E, Kazys R. Principles of high-frequency ultrasonography for investigation of skin pathology. *J Eur Acad Dermatol Venereol.* (2011) 25:375–82. doi: 10.1111/j.1468-3083.2010.03837.x
15. Petrella LI, de Azevedo Valle H, Issa PR, Martins CJ, Machado JC, Pereira WC. Statistical analysis of high frequency ultrasonic backscattered signals from basal cell carcinomas. *Ultrasound Med Biol.* (2012) 38:1811–9. doi: 10.1016/j.ultrasmedbio.2012.06.001
16. Morine Y, Shimada M, Takamatsu H, Araida T, Endo I, Kubota M, et al. Clinical features of pancreaticobiliary maljunction: update analysis of 2nd Japan-nationwide survey. *J Hepato Biliary Pancreatic Sci.* (2013) 20:472–80. doi: 10.1007/s00534-013-0606-2
17. Varadarajulu S, Wilcox CM, Hawes RH, Cotton PB. Technical outcomes and complications of ERCP in children. *Gastrointest Endosc.* (2004) 60:367–71. doi: 10.1016/S0016-5107(04)01721-3
18. Kim MJ, Han SJ, Yoon CS, Kim JH, Oh JT, Chung KS, et al. Using MR cholangiopancreatography to reveal anomalous pancreaticobiliary ductal

union in infants and children with choledochal cysts. *Am J Roentgenol.* (2002) 179:209–14. doi: 10.2214/ajr.179.1.1790209
19. Matos C, Nicaise N, Deviere J, Cassart M, Metens T, Struyven J, et al. Choledochal cysts: comparison of findings at MR cholangiopancreatography and endoscopic retrograde cholangiopancreatography in eight patients. *Radiology.* (1998) 209:443–8. doi: 10.1148/radiology.209.2.9807571
20. Doria AS, Moineddin R, Kellenberger CJ, Epelman M, Beyene J, Schuh S, et al. US or CT for diagnosis of appendicitis in children and adults? A meta-analysis. *Radiology.* (2006) 241:83–94. doi: 10.1148/radiol.2411050913
21. Malviya S, Voepel-Lewis T, Eldevik OP, Rockwell DT, Wong JH, Tait AR. Sedation and general anaesthesia in children undergoing MRI and CT: adverse events and outcomes. *Br J Anaesth.* (2000) 84:743–8. doi: 10.1093/oxfordjournals.bja.a013586
22. Chapuy S, Gorincour G, Roquelaure B, Aschero A, Paris M, Lambot K, et al. Sonographic diagnosis of a common pancreaticobiliary channel in children. *Pediatr Radiol.* (2006) 36:1300–5. doi: 10.1007/s00247-006-0322-z
23. Kaneko K, Ando H, Ito T, Watanabe Y, Seo T, Harada T, et al. Protein plugs cause symptoms in patients with choledochal cysts. *Am J Gastroenterol.* (1997) 92:1018–21.
24. Huang SG, Guo WL, Wang J, Sheng M, Lan XH, Fang L. Factors interfering with delineation on MRCP of pancreaticobiliary maljunction in paediatric patients. *PLoS ONE.* (2016) 11:e0154178. doi: 10.1371/journal.pone.0154178
25. Kushnir VM, Wani SB, Fowler K, Menias C, Varma R, Narra V, et al. Sensitivity of endoscopic ultrasound, multidetector computer tomography and magnetic resonance cholangiopancreatography in the diagnosis of pancreas divisum: a tertiary center experience. *Pancreas.* (2013) 42:436. doi: 10.1097/MPA.0b013e31826c711a

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Xu, Liu, Wu, Ling and Guo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Public Health

SYSTEMATIC REVIEW
published: 21 April 2022
doi: 10.3389/fpubh.2022.801981

Economic Burden and Quality of Life of Hepatocellular Carcinoma in Greater China: A Systematic Review

Huimin Zou[1†], Meng Li[1†], Qing Lei[1], Zejun Luo[1], Yan Xue[1], Dongning Yao[1], Yunfeng Lai[1,2], Carolina Oi Lam Ung[1,3] and Hao Hu[1,3]**

[1] State Key Laboratory of Quality Research in Chinese Medicine, Institute of Chinese Medical Sciences, University of Macau, Macau, Macao SAR, China, [2] School of Public Health and Management, Guangzhou University of Chinese Medicine, Guangzhou, China, [3] Department of Public Health and Medicinal Administration, Faculty of Health Sciences, University of Macau, Macau, Macao SAR, China

Background: Hepatocellular carcinoma (HCC) accounts for more than 85%-90% of primary liver cancer globally, and approximately 45% of deaths from HCC occur in greater China. This disease poses a significant economic burden for patients, payers and society and significantly affects patients' quality of life (QoL). However, such impact of HCC in greater China has not been well characterized. This review was conducted to analyze the current evidence about the economic and humanistic impact of HCC in greater China for informing national disease management and identifying clinical gaps yet to be resolved.

Methods: A systematic search literature using seven databases (Web of Science, PubMed, Medline, Cochrane Central, China National Knowledge Infrastructure, Wanfang, and Weipu) was performed to identify interventional and observational studies that reported the impact of HCC on cost or QoL and published before April 6, 2021. The focus population included adult patients with HCC in greater China. This review excluded any studies that focused on any specific treatment. Study quality was assessed using the Effective Public Health Practice Project tool.

Results: Of 39,930 studies retrieved, 27 were deemed eligible for inclusion. The methodologies, perspectives and data sources used in studies were heterogeneous. In greater China, while few studies reported the health expenditures of HCC patients and investigations about economic burden at national level was lacking, the significant economic impact of HCC on patients and their families had been reported. Health-related costs increased as the disease deteriorated. Additionally, HCC also has a negative impact on the QoL of patients, mostly in terms of physical, cognitive, social functioning and severe symptoms.

Conclusions: HCC has brought significant economic and QoL burden to patients in greater China. Both physical and psychological factors predicted QoL in patients with HCC in greater China. Future studies should explore the disease-related economic effects on Chinese patients and their families, the effects of physical and psychological factors on QoL and the relationships of physical and psychological factors in the region.

Systematic Review Registration: www.crd.york.ac.uk/prospero/display_record.php?RecordID=278421, PROSPERO: CRD42021278421

Keywords: hepatocellular carcinoma, economic burden, quality of life, systematic review, China

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Sok King Ong,
Ministry of Health, Brunei
Hon-Yi Shi,
Kaohsiung Medical University, Taiwan

****Correspondence:***
Carolina Oi Lam Ung
carolinaung@um.edu.mo
Hao Hu
haohu@um.edu.mo

† These authors share first authorship

Specialty section:
This article was submitted to Health Economics, a section of the journal Frontiers in Public Health

Received: *26 October 2021*
Accepted: *11 March 2022*
Published: *21 April 2022*

Citation:
Zou H, Li M, Lei Q, Luo Z, Xue Y, Yao D, Lai Y, Ung COL and Hu H (2022) Economic Burden and Quality of Life of Hepatocellular Carcinoma in Greater China: A Systematic Review. Front. Public Health 10:801981. doi: 10.3389/fpubh.2022.801981

INTRODUCTION

Hepatocellular carcinoma (HCC) is the major pathological classification of primary liver cancer, accounting for 85%-90% liver cancer globally (1). In terms of global incidence, HCC ranks 7th, and there are about 850 thousand new cases of HCC worldwide every year (1, 2). The incidence of HCC increases with age and is 2.7 times higher in males than in females (3). HCC is a highly fatal cancer, causing approximately 800 thousand deaths every year (4). It is the fourth leading cause of cancer-related deaths, second only to lung cancer, colorectal cancer, and gastric cancer (5). The high prevalence of HCC poses increasing demand for treatment and hospitalization, which bring huge disease burden for both patients and the society (6). In greater China, HCC is the fourth highest incidence of malignant tumors, with 466 thousand new cases each year accounting for about 50% of the total global new cases (7, 8); there are about 444 thousand deaths due to HCC each year accounting for approximately 45% of the total deaths from HCC worldwide.

HCC can be induced by multiple etiologies and risk factors such as Hepatitis B Virus (HBV) infection, Hepatitis C Virus (HCV) infection, alcohol abuse, metabolic liver disease (such as non-alcoholic fatty liver disease) and exposure to dietary toxins (such as aflatoxin and aristolochic acid) (9). Among them, chronic hepatitis caused by viral infection accounts for more than 70% of HCC in the world (10). Chronic hepatitis B (HBV) and hepatitis C (HCV), as well as aflatoxin, are classified as Group 1 carcinogen by the International Agency for Research on Cancer (IARC) (11, 12). Recent studies suggested that patients with diabetes, obesity, smoking, drug-induced liver damage and men over the age of 40 also carry a higher risk of HCC (13, 14). HCC may be developed by the additive effects of multiple complex pathogeneses.

Most HCC cases were discovered at a late stage, and patients suffered from poor prognosis and short survival due to limited treatment options during advanced stage (15–17). The treatment of HCC can be divided into two types: surgical and non-surgical. Surgical treatment includes curative resection, liver transplantation, etc.; non-surgical treatment mainly includes radiofrequency ablation, trans-arterial chemoembolization (TACE), radioembolization, and systemic targeted agent, etc (7, 18). The utilization of medical, clinical treatments and work loss due to the disease brings a huge negative impact on the economic status and quality of life (QoL) for patients with HCC and their families.

Research showed that the overall financial charges for HCC hospitalization, especially for the costs related to disease progression, increased substantially in the past decades (6, 19, 20). The total inpatient expenses related to HCC increased from 1.0 billion USD in 2005 to 2.0 billion USD in 2009 in USA (19). In terms of QoL, extrahepatic manifestations and complications largely affected physical, social, emotional, and functional well-being of patients with HCC (21, 22). Understanding the economic impact and QoL affected by HCC is essential to determine the factual disease burden on patients and their families. Despite that greater China has a high prevalence rate of HCC in the world and has developed relevant policies to reduce patients' burden, the evidence about the economic burden and the QoL in patients affected by HCC in greater China has not been well characterized.

Therefore, a systematic review of the published literature on the economic burden and QoL in patients with HCC in greater China was conducted to provide evidence for public health strategies of HCC management and identify clinical gaps to inform future research of HCC in greater China.

TABLE 1 | Eligibility criteria.

Inclusion criteria	
Patient population	Adults (age ≥ 18 years) with any stage of HCC in greater China
Intervention	All interventions except for any specific treatment
	No intervention
Comparator	All interventions except for any specific treatment Baseline
	No comparator
Outcomes	Direct costs Indirect costs
	Health insurance payments QoL measures based on clinical instruments
Study type	Interventional studies
	Observational studies
Language	English Chinese
Exclusion criteria	
Publication type	Commentaries
	Conference abstracts
	Reviews Editorials
	Letters
	Case reports
	Protocols
Design	*In vitro* studies Animal studies
Language restrictions	Not English or Chinese

QoL, quality of life.

METHODS

The reporting of this systematic review was performed according to the PRISMA statement (23). The protocol of this systematic review was registered with PROSPERO (CRD42021278421).

Eligibility Criteria

Eligibility criteria presented in **Table 1** were specified according to the PICOS approach (24). Interventional and observational studies published in English or Chinese were included if they reported direct and indirect costs or QoL measures based on clinical instruments. Both prospective and retrospective studies were included. The duration of follow-up or the year of publication was not restricted. The focus population included adult patients suffering from any stage of HCC in greater China. This review excluded any studies that focused on any specific treatment. Commentaries, conference abstracts, reviews, editorials, letters, case reports, protocols, *in vitro* studies, animal studies, and data-uncompleted publications were also excluded. Included studies were further grouped into economic burden studies and QoL studies.

Data Sources and Search Strategies

Literature search was conducted comprehensively in four English databases (Web of Science, PubMed, Medline, and Cochrane Central) and three Chinese databases (China National Knowledge Infrastructure, Wanfang and Weipu). Additional references of relevant reviews or included articles were also manually screened. Articles published from inception to Apr 6, 2021, were included. The details regarding the search strategy are provided in **Supplementary Table 1**. Endnote X9 software was used to categorize and file all the references.

Study Selection

Firstly, seven authors (HZ, ML, QL, ZL, YX, DY, and YL) screened retrieved articles and excluded irrelevant results according to titles and abstracts. Secondly, six authors (HZ, ML, QL, ZL, YX, and YL) conducted full-text review and obtained appropriate studies. They independently checked the included articles. The disagreements were resolved by discussion or consultation with two other authors (HH and COLU).

Data Extraction

We used a standardized form to extract prespecified data elements from each study, including first author, year of publication, year of investigation, region, study design, age range or mean age if available, gender ratio (male/female), disease stage at diagnosis, sample source, sample size, duration of follow-up for the longitudinal studies, and main outcomes (economic burden and QoL). Two authors extracted data from obtained studies (ML and HZ), and disagreements were resolved by discussion or consultation with two other authors (HH and COLU). Excel 2013 was used to extract data and record.

Quality Assessment

Study quality was assessed using the Effective Public Health Practice Project (EPHPP) tool, which included the following components: selection bias, study design, confounders, blinding, data collection methods, withdrawals and dropouts, intervention integrity, and data analysis. Two authors independently performed all quality ratings (HZ and ML), and disagreements were resolved by discussion or consultation with two other authors (HH and COLU).

RESULTS

Study Selection

We identified a total of 34,274 titles and abstracts by electronic searches, of which 6,684 were selected for full-text review (**Figure 1**). Twenty-seven publications were finally included reporting 6 studies in the economic burden and 21 in the QoL correlated with HCC.

Assessment of Study Quality

The quality of the 6 economic burden studies was heterogeneous (**Table 2**). While 1 study relying on cancer registry claims data was regarded as having "strong" global rating (27), 2 studies based on hospital databases, social insurance information system, and national health insurance database were rated "moderate" (26, 29), and another 3 studies based on survey and hospital electronic health records were rated "weak" (25, 28, 30). A discrepancy occurred and resolved in raters, and the reasons were differences in interpretation of criteria.

Six of 21 QoL studies were considered as having "strong" global rating, while 6 were considered as "moderate" and 9 were considered as "weak". Data collection method was rated "strong" in all cases, since validated instruments were used for all evaluations. Studies with longer follow-up periods had major challenges in dropouts (46–48), which caused "weak" ratings in component of withdrawals and dropouts. Contrarily, study design was rated "weak" in cross-sectional studies (32, 34, 35, 42, 43, 45, 51) based on the definition of the EPHPP dictionary. Thirteen studies were rated "weak" in controlling for confounders because < 60% of relevant confounders were controlled or control of confounders not described (31, 32, 34–39, 42, 43, 45, 48, 51). A discrepancy occurred and resolved in raters, and the reasons were differences in interpretation of criteria.

Economic Burden of HCC

Study Characteristics

HCC had a long-term and poor prognosis and the costs of HCC posed a huge economic burden for patients, household, and society worldwide. Six studies were included in this review after the data-extraction process (**Figure 1**): 3 were cross-sectional studies, 2 were cohort studies, and 1 was a case-control study. The characteristics of economic burden studies are summarized in **Table 3**.

Direct Costs

Medical Costs

Direct cost was divided into direct medical cost and direct nonmedical cost. Direct medical costs were generally collected through the hospital billing system (physician services, examination, medicine costs hospitalization charges) (52). Che et al. conducted a cross-sectional study to investigate the economic burden of patients at different stages of HBV-related liver diseases from December 2012 to June 2013 in Yunnan Province (25). Information of inpatients and outpatients were consecutively collected from medical records database and questionnaire in this study. A total of 131 patients with HCC were included whose monthly household and individual income was 4,000 CNY and 2,000 CNY respectively. For the financial burden of patients with HCC, the average annual direct costs were 33,044 CNY, while the average medical cost was 31,092 CNY.

Another hospital-based cross-sectional survey study covering 39 hospitals and 21 project sites in 13 provinces was carried out to investigate the economic burden of liver cancer in mainland China from 2012-2014 (28). Demographic and societal information, expenditures, and time loss due to clinical visits of 1,196 patients with HCC were investigated with the use of a structured questionnaire. The overall expenditure was 55,529 CNY, whereas medical expenditure was 50,937 CNY. The total expenditure of a newly diagnosed course of patients with HCC was 48,150 CNY, of which the reimbursed medical expenditures accounted for 48.1%. The out-of-pocket expenditure was estimated to be 25,896 CNY, which accounted for nearly 50% of the previous year's household income (56,536 CNY). In this

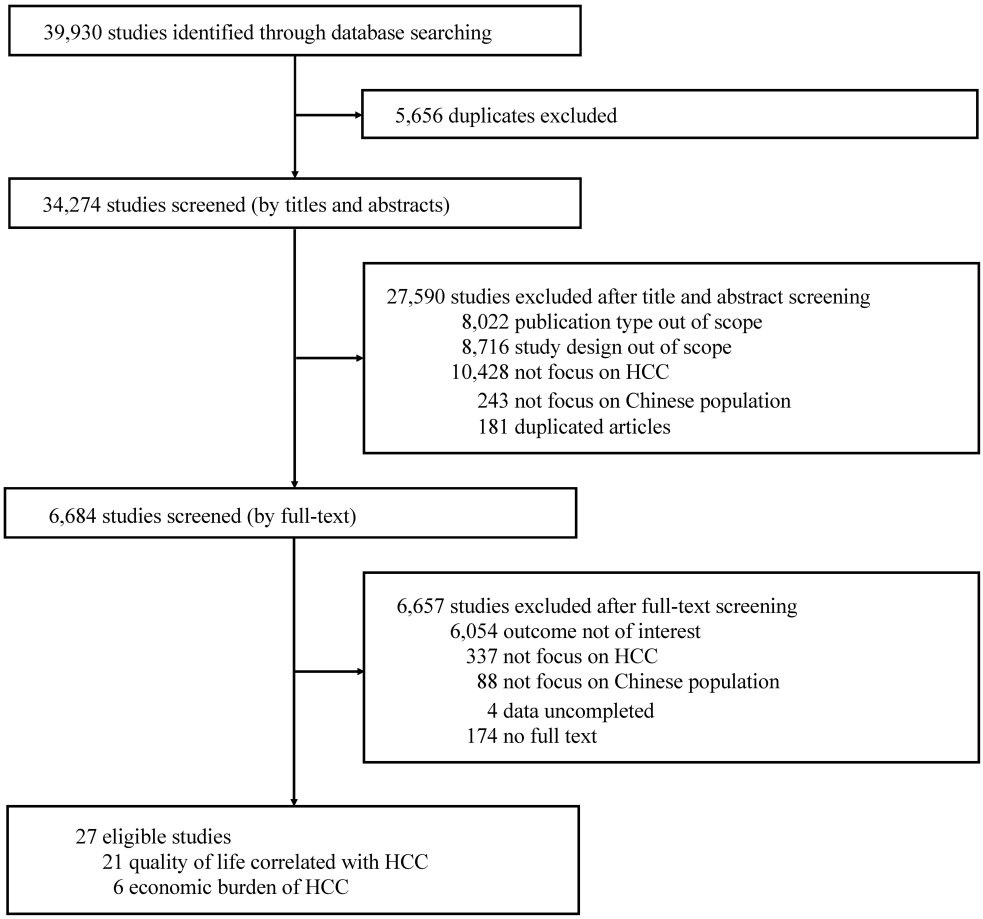

FIGURE 1 | PRISMA flowchart of study selection.

study, 49.1% of patients with HCC considered the economic burden was overwhelming for them compared to only 19.7% of other patient groups who reported the same concerns ($P < 0.001$).

One retrospective cohort study conducted in Beijing and Guangzhou collected the data from the information system of community health center and a structured household questionnaire survey to investigate the economic impact of chronic HBV-related diseases and HCC was one of the conditions under investigation (26). In this study, the annual direct economic burden which referred to the economic resources for healthcare services utilized were 6,084 USD and 5,472 USD in Beijing and Guangzhou. Ran et al. analyzed the top 10 diseases in the total hospitalization cost of a tertiary hospital in Beijing between 2017–2018 (30). The results of the study showed that the average direct medical cost per visit for HCC was 29,896.82 CNY.

Nguang et al. retrospectively investigated the HCC-related burden in Taiwan (29). Data from national health insurance program research, the Longitudinal Health Insurance Database 2000 (LHID 2000), was extracted for analysis. From 1997 to 2012, 5,522 patients were newly diagnosed with HCC. The total healthcare expenditure for all patients with HCC was 92 million USD of which 58% was due to hospital cost, and 42% was due to outpatient and emergency department services. The average total cost per patient was 16,711 (±21,350) USD, of which inpatient care cost was 9,721 (±11,811) USD and outpatient cost was 6,989 (±14,726) USD.

Another prospective study was conducted to determine the lifetime cancer-related medical costs of patients with HCC in Taiwan by Lang et al. (27). The data of 2,592 patients with HCC was extracted from cancer registry statistics of patients with HCC and the claims data of Taipei Veterans General Hospital (TVGH) during 1999–2002. The costs of 3 HCC disease phases including the initial treatment phase (3 month period after diagnosis), continuous care phase and terminal phase (final 6 months prior to death) were applied to estimate the lifetime cost (53). For patients who survived < 1 year (Group 1), all costs were calculated and the average cost was 206,573 TWD per person (6,259.7 USD, 2002). For patients in survived ≥ 1 year but died before the end of 2002 (Group 2), the data was used to derive the terminal cost and the cost was 237,032 TWD (7,182.7 USD, 2002). For patients who were still surviving by the end of 2002 (Group 3), the data was used to calculate the initial and monthly continuing costs. The cost was found to be 140,403

TABLE 2 | Quality assessment of identified studies.

Study	Outcome	Component rating						Global rating
		Selection bias	Study design	Confounders	Blinding	Data collection methods	Withdrawals and dropouts	
Che et al. (25)	Cost	Moderate	Weak	Weak	Moderate	Weak	NA	Weak
Hu et al. (26)	Cost	Moderate	Moderate	Weak	Moderate	Strong	NA	Moderate
Lang et al. (27)	Cost	Moderate	Moderate	Strong	Moderate	Strong	NA	Strong
Lei et al. (28)	Cost	Moderate	Weak	Weak	Moderate	Strong	NA	Weak
Nguang et al. (29)	Cost	Moderate	Moderate	Strong	Moderate	Strong	Weak	Moderate
Ran et al. (30)	Cost	Moderate	Weak	Weak	Moderate	Strong	NA	Weak
Fan et al. (31)	QoL	Moderate	Weak	Weak	Moderate	Strong	NA	Moderate
Fan et al. (32)	QoL	Moderate	Weak	Weak	Moderate	Strong	NA	Weak
Fielding et al. (33)	QoL	Moderate	Moderate	Strong	Moderate	Strong	Strong	Strong
Hsu et al. (34)	QoL	Moderate	Weak	Weak	Moderate	Strong	NA	Weak
Huang et al. (35)	QoL	Moderate	Weak	Weak	Moderate	Strong	NA	Weak
Li et al. (36)	QoL	Moderate	Moderate	Weak	Moderate	Strong	Strong	Moderate
Li et al. (37)	QoL	Moderate	Moderate	Weak	Moderate	Strong	Strong	Moderate
Li et al. (38)	QoL	Moderate	Moderate	Weak	Moderate	Strong	Strong	Moderate
Li et al. (39)	QoL	Moderate	Moderate	Weak	Moderate	Strong	Strong	Moderate
Jie et al. (40)	QoL	Moderate	Moderate	Strong	Moderate	Strong	Strong	Strong
Jie et al. (41)	QoL	Moderate	Moderate	Strong	Moderate	Strong	Strong	Strong
Qiao et al. (42)	QoL	Moderate	Weak	Weak	Moderate	Strong	NA	Weak
Qiu et al. (43)	QoL	Moderate	Weak	Weak	Moderate	Strong	NA	Weak
Shun et al. (44)	QoL	Moderate	Moderate	Strong	Moderate	Strong	Strong	Strong
Wei et al. (45)	QoL	Moderate	Weak	Weak	Moderate	Strong	NA	Weak
Wong et al. (46)	QoL	Moderate	Moderate	Strong	Moderate	Strong	Weak	Moderate
Wong et al. (47)	QoL	Moderate	Moderate	Strong	Moderate	Strong	Weak	Moderate
Yang et al. (48)	QoL	Moderate	Moderate	Weak	Moderate	Strong	Weak	Weak
Yeo et al. (49)	QoL	Strong	Moderate	Strong	Moderate	Strong	Strong	Strong
Zhang et al. (50)	QoL	Moderate	Moderate	Strong	Moderate	Strong	Strong	Strong
Zheng et al. (51)	QoL	Moderate	Weak	Weak	Moderate	Strong	NA	Weak

QoL, quality of life; NA, not applicable.

TABLE 3 | Characteristics of selected healthcare economic burden studies.

Study	Perspective	Region, year of pricing	Study design	Sample source	Sample size	Main findings/Cost categories and costs
Che et al. (25)	Patient	Yunnan, 2013	Cross-sectional study	Hospital-based	131 (of 940 HBV infected patients)	(1) Hospital costs accounted for the majority of total direct medical expenses. (2) Median estimated monthly household income: 4,000 CNY (IQR: 1,950, 9,750); Individual incomes: 2,000 CNY (IQR: 900, 3,000). (3) Insurance scheme [n (%)]: • No: 2 (1.5); UEM: 48 (36.6); URM: 27 (20.6); NRCM: 54 (41.2). (4) Estimated average annual direct costs for HCC patients was 33,044 CNY. (5) The percent of decrease in total direct expense (Total direct costs/household income per year median) after reimbursement was 36.1% in HCC group (33,044–17,273, 61.9–38.8%). (6) After reimbursement it could save 24% of patients not undergoing catastrophic health expenditure health in HCC groups.
Hu et al. (26)	Patient	Beijing and Guangzhou	Retrospective cohort Study	Hospital-based	596 CHB-related disease	(1) The total annual cost per patient for HCC was 6,615 USD in Beijing and 6,054 USD in Guangzhou. (2) Beijing: direct economic burden: 6,084 USD, indirect Economic Burden: 531 USD. (3) Guangzhou: direct economic burden: 5,472 USD, indirect Economic Burden: 582 USD.
Lang et al. (27)	Payer	Taiwan, 2002	Retrospective case-control study	Hospital-based	2873	(1) Group 1 (survived <1 year): 895 patients, average medical care costs: 206,573 TWD (6,259.7 USD, 2002), year 2002 value). (2) Group 2 (survived ≥ 1 year but died before the end of 2002): 735 patients, 237,032 TWD (7,182.7 USD, 2002). (3) Group 3 (still surviving by the end of 2002): 858 patients for initial phase, 140,403 TWD (4,254.6 USD, 2002); 1011 cases for continuous phase, 8,687 TWD (263.2 USD, 2002) per month. (4) For the average HCC patient, the 10-year lifetime cost was 418,554 TWD (12,683.4 USD, 2002).
Lei et al. (28)	Patient	13 Provinces, 2014	Cross-sectional study	Hospital-based	1196 (of 2223 liver cancer)	(1) HCC is the major pathological type of liver cancer. (2) Medical expenditure of HCC was 50,937 CNY, Non-medical expenditure was 4,592 CNY. Overall expenditure was 55,529 CNY. (3) Economic impact of overall expenditure on patients' family: • Expenditure of newly diagnosed course/Annual expenditure of illness of HCC: 48,150 CNY; • Self-reported predicted reimbursement ratio of HCC: 48.1%; • Out-of-pocket expenditure of HCC: 25,896 CNY; • Previous year household income: 56,536 CNY. (4) Self-reported degree of economic pressure of HCC: • Not at all (5.2%); • Somewhat but manageable (15.2%); • Heavy (30.5%); • Overwhelmed (49.1%).
Nguang et al. (29)	Payer	Taiwan, 2016	Retrospective cohort study	Hospital-based	5522	(1) Total healthcare expenditure (insurance payments) for treating HCC patients was approximately 92 million USD (92,269,551), including 53.4 million USD (58%) for hospital care and 38.7 million USD (42%) for outpatient and emergency department services. (2) Underlying comorbid conditions, liver transplants, hepatectomy, and trans arterial chemoembolization were associated with increased total cost, with liver transplants having the greatest impact over time.
Ran et al. (30)	Patient	Beijing, 2018	Cross-sectional study	Hospital-based	4174 hospitalizations	(1) Total hospitalization expenses: 158,320,500 CNY, average hospitalization expenses per time: 29,896.82 CNY (IQR:14,306.36, 50,567.79). (2) HCC was the second most expensive disease in this large tertiary hospital.

UEM, the basic health insurance for urban employees medical scheme; URM, the basic health insurance for urban residents medical scheme; NRCM, the new rural cooperative medical scheme.

TWD (4,254.6 USD, 2002) at the initial phase, 8,687 TWD per month at the continuous phase (263.2 USD, 2002), and 418,554 TWD (12,683.4 USD, 2002) for a 10-year lifetime.

Non-medical Costs

Three studies explored the direct non-medical cost of patients with HCC by specific questionnaires. The non-medical costs in these studies normally referred to additional meals, additional nutrition supplements, transportation, accommodation and informal nursing worker fee due to illness. Che et al. reported the non-medical cost of patients with HCC in Yunnan province was 565 CNY annually (25). The average non-medical expenditure in 13 provinces in mainland China was 4,592 CNY per case (28).

Indirect Costs

The indirect costs were estimated from the social perspective in two studies in which the work loss of both patients and their families due to HCC were considered. Hu et al. reported that the average indirect economic burden of HCC in Beijing and Guangzhou were 531 USD and 582 USD (26). It was discovered by Lei et al. that the working day loss of patients and their caregivers due to HCC diagnosis was about 70.5 days, including 41.4 days for patients with HCC and 29.1 days for accompanying persons (28).

Health Insurance Payments

The impact of health insurance on patients' economic burden was also explored by two studies. Che et al. found almost all patients in Yunnan province were covered by at least one of the health insurance schemes (25): namely the New Rural Cooperative Medical Scheme (41.2%), the Urban Employee Basic Medical Insurance (36.6%), and the Urban Resident Basic Medical Insurance Scheme (20.6%). After reimbursement, the median percentage of decrease in annual total direct expense was 36.1% (11,928 CNY) in patients with HCC. In terms of the catastrophic health expenditure, health insurance could save more than 24% of patients with HCC not undergoing catastrophic health expenditure. Ran et al. also analyzed patients with medical insurance (30), and the results showed that during 2017–2018, medical insurance payments covered a total of 12,617.5 thousand CNY for patients with HCC.

QoL Correlated With HCC

Study Characteristics

Of the 27 studies identified, 21 included outcomes on QoL effects: 11 of these were prospective cohort studies, 8 were cross-sectional studies, 1 was case-control study, and 1 was interrupted time series. The 21 studies encompassed a total of 3,060 HCC patients and were conducted in Shanghai ($n = 1$), Chongqing ($n = 3$), Hebei ($n = 1$), Guangxi ($n = 1$), Jiangxi ($n = 1$), Yunnan ($n = 1$), Hong Kong ($n = 8$), and Taiwan ($n = 5$). Summarized of characteristics of QoL studies are provided in **Table 4**.

Generic, Disease-Specific and Supplementary Measures of QoL in Chinese Patients

The European Organization for Research and Treatment for Cancer Quality of Life Questionnaire Core-30 (EORTC QLQ-C30) ($n = 11$) and the Functional Assessment of Cancer Treatment-General (FACT-G) ($n = 3$) were the 2 most commonly used measures to assess generic QoL. The EORTC QLQ-C30 reported 5 functional domains (physical, role, emotional, cognitive and social functioning), 8 symptom domains, 1 financial problem and 1 global health status (54). The FACT-G included 4 subscales for physical, social/family, emotional and functional well-being (55).

The European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core-18 (EORTC QLQ-HCC18) ($n = 8$), the Functional Assessment of Cancer Treatment-Hepatobiliary (FACT-Hep) ($n = 2$), and the quality of life-liver cancer (QOL-LC) ($n = 1$) were the 3 standardized instruments to assess disease-specific QoL. The EORTC QLQ-HCC18 assessed 8 symptom domains for fatigue, body image, jaundice, nutrition, pain, fever, abdominal swelling and sex life, which was primarily proposed for patients with HCC (56). The FACT-Hep was used widely for liver cancer-specific assessment, consisting of FACT-G and newly validated hepatobiliary subscale (57). The Chinese versions of both of these tools have demonstrated good validity, reliability, and responsiveness (48, 58). For Chinese patients diagnosed with liver cancer, the QOL-LC was specifically developed, and the disease-specific items were evaluated in terms of the symptom/side effect subscale (59).

In addition to conventional assessment instruments mentioned above, 2 original studies reported the development of supplementary measures for Chinese patients with HCC, which included QLQ-C30, QLQ-HCC18 index scores and Mini Nutritional Assessment (MNA) (34, 36). C30 and HCC18 index scores could represent all domains of QLQ-C30 and HCC18, respectively (36). The mathematical formulas were presented as follows:

$$\begin{aligned}\textit{C30 index score} = {} & [(100 - \textit{global QoL}) + (100 - \textit{physical functioning}) \\ & + (100 - \textit{role fuctioning}) + (100 - \textit{emotional functioning}) \\ & + (100 - \textit{cognitive functioning}) + (100 - \textit{social functioning}) \\ & + \textit{fatigue} + \textit{nausea and vomiting} + \textit{pain} + \textit{dyspnoea} + \textit{insomnia} \\ & + \textit{appetite loss} + \textit{constipation} + \textit{diarrhea} + \textit{financial difficulties}]/15; \\ \textit{HCC18 index score} = {} & (\textit{fatigue} + \textit{body image} + \textit{jaundice} + \textit{nutrition} + \textit{pain} \\ & + \textit{fever} + \textit{abdominal swelling} + \textit{sex life})/8 \text{ (36).}\end{aligned}$$

The index scores had more advantages in clinical practice, since they converted complex QoL data to more straightforward tools. The MNA, a nutritional assessment tool, performed better than global functional status and global QoL of the EORTC QLQ-C30 to predict functional status and QoL in Chinese patients with HCC (34).

QoL in Chinese Patients With HCC Compared With Benign Liver Diseases and the General Population

HCC Compared With Benign Liver Diseases

Patients with HCC had lower physical functioning (QOL-LC score: 42 ± 10), symptoms/side effects (41 ± 7), social functioning (28 ± 10) and global QoL (150 ± 24) than patients with benign liver diseases ($P < 0.05$), but showed no difference in psychological functioning (39 ± 10 vs. 40 ± 10; $P > 0.05$) and self-assessment (70 ± 15 vs. 75 ± 15; $P > 0.05$) (50).

TABLE 4 | Characteristics of selected QoL studies.

Study	Region	Study design	Sample source	Sample size	Measure(s)	QoL data
Measures of QoL in Chinese patients						
Hsu et al. (34)	Taiwan	Cross-sectional study	Hospital-based	300	MNA-LF, MNA-SF, EORTC QLQ-C30 V3.0	(1) Both the long-form and short-form of the MNA performed better than GQL and GFS in predicting quality of life and functional status of patients with HCC. (2) The MNA is suitable for identifying the risk of deteriorating quality of life or functional status, in addition to identifying the risk of malnutrition, in patients with HCC.
Li et al. (36)	Hong Kong	Prospective cohort study	Hospital-based	472	EORTC QLQ-C30, EORTC QLQ-HCC18, C30 and HCC18 index-scores	(1) In multivariate analysis, independent prognostic HRQOL variables for OS were QLQ-C30 pain (HR 1.346 [1.092–1.661], $P = 0.0055$), QLQ-C30 physical functioning (HR 0.652 [0.495–0.860), $P = 0.0024$); QLQ-HCC18 pain (HR 1.382 [1.089–1.754], $P = 0.0077$) and QLQ-HCC18 fatigue (HR 1.441 [1.132–1.833], $P = 0.0030$). (2) C30 index-score (HR 2.143 [1.616–2.841], $P < 0.0001$) and HCC18 index-score (HR 1.957 [1.411–2.715], $P < 0.0001$) were highly significant factors for OS. (3) The median OS of patients with C30 index-score of 0–20, 21–40, 41–60, 61–100 were 16.4, 7.3, 3.1, 1.8 months respectively ($P < 0.0001$); while for HCC18 index-score: 16.4, 6.0, 2.8, 1.8 months respectively ($P < 0.0001$).
Yang et al. (48)	Yunnan	Interrupted time series	Hospital-based	114	EORTC QLQ-HCC18, FACT-Hep	(1) The internal consistency Cronbach's α were > 0.60 for all domains (exception of Jaundice 0.38), and all test-retest reliability coefficients were > 0.80. (2) Four out of eight domains had statistically significant changes with effect size standardized response mean ranging from 0.31 to 0.73.
HCC compared with benign liver diseases and the general population						
Fan et al. (32)	Taiwan	Cross-sectional study	Hospital-based	286	EORTC QLQ-C30, EORTC QLQ-HCC18, Brief IPQ, Jalowiec coping scale	(1) Patients with HCC had worse global QoL, physical, role, cognitive and social functioning, but better emotional functioning than the general population. (2) Cognitive representation was significant predictors of global QoL, physical functioning and emotional functioning. (3) Cognitive representation mediated the relationships between physical variables and global QoL, physical functioning and emotional functioning, but coping only mediated the relationship between cognitive representation and global QoL.
Wei et al. (45)	Guangxi	Cross-sectional study	Hospital-based	63	CD-RISC, EORTC QLQ-C30 V3.0, EORTC QLQ-HCC18	(1) The total health status score of QLQ-C30 was (56.61 ± 27.24) points, with lower quality of life scores and higher clinical symptoms scores. (2) Emotional functioning, social functioning, and general health status were positively correlated with the total score of psychological resilience ($r = 0.382$, $P < 0.01$; $r = 0.324$, $P < 0.01$; $r = 0.383$, $P < 0.01$). (3) Symptoms of fatigue, pain, fatigue, and nutritional changes were negatively correlated with the total score of psychological resilience ($r = -0.303$, $P < 0.05$; $r = -0.286$, $P < 0.05$; $r = -0.360$, $P < 0.01$; $r = -0.259$, $P < 0.05$).
Zhang et al. (50)	Hebei	Case-control study	Hospital-based	81 HCC 44 benign liver diseases	QOL-LC V2.0	(1) The scores of physical functioning (42 ± 10), symptom/side effect (41 ± 7), social functioning (28 ± 10) and total QoL (150 ± 24) of HCC group were lower than that of control group ($P < 0.05$). (2) The scores of physical functioning, symptom/side effect are negatively correlated with the way of case finding, TNM stage, Child-Pugh class, respectively. (3) The score of psychological functioning is positively correlated with age. (4) The score of social functioning is positively correlated with age, education level, income, and mode of payment, and negatively correlated with the way of case finding.
Physical factors and symptoms associated with QoL						
Fielding et al. (33)	Hong Kong	Prospective cohort study	Hospital-based	176 liver cancer 358 lung cancer	FACT-G, Visual analog: eating ability, eating appetite, eating enjoyment, self-care ability, and current health perception	(1) No association between QoL and survival in patients with liver cancer. (2) Less advanced cancer stage and better appetite were associated significantly with longer survival in patients with liver cancer.

(Continued)

TABLE 4 | Continued

Study	Region	Study design	Sample source	Sample size	Measure(s)	QoL data
Li et al. (37)	Hong Kong	Prospective cohort study	Hospital-based	445	EORTC QLQ-C30, EORTC QLQ-HCC18, C30 and HCC18 index-scores	(1) Significant correlations were found between IL-8 levels and EORTC QLQ-C30, QLQ-HCC18, C30, and HCC18 index-scores. (2) The strongest correlated factors were those reflective of constitutional symptoms, namely QLQ-C30 "appetite loss" (with Pearson's correlation coefficient, $r = 0.322$, $P < 0.0001$); QLQ-C30 "fatigue" ($r = 0.311$, $P < 0.0001$); QLQ-C30 "role functioning" ($r = -0.305$, $P < 0.0001$); QLQ-HCC18 "nutrition" ($r = 0.317$, $P < 0.0001$); and QLQ-HCC18 "fatigue" ($r = 0.306$, $P < 0.0001$). (3) Moderate but significant correlations were also observed with HCC18 index score ($r = 0.321$, $P < 0.0001$), and C30 index score ($r = 0.306$, $P < 0.0001$). (4) QoL factors were also significantly correlated with mIBI.
Li et al. (38)	Hong Kong	Prospective cohort study	Hospital-based	445	EORTC QLQ-C30, EORCT QLQ-HCC18, C30 and HCC18 index-scores	(1) Higher inflammatory states were significantly correlated with worse QoL. (2) For CRP and CRP/alb ratio, the QoL factors with higher correlations included C30 and HCC18 index-scores, certain QLQ-C30 domains and items ("physical functioning", "role functioning", "fatigue", "pain", "appetite loss") and QLQ-HCC18 items ("fatigue", "body image", "nutrition" and "abdominal swelling"), where the Pearson's correlation coefficients were up to 0.416. (3) Multivariate analyses indicated that worse QoL factors were significantly correlated with worse scores in GPS, IBI and PI.
Li et al. (39)	Hong Kong	Prospective cohort study	Hospital-based	472	EORTC QLQ-C30, EORCT QLQ-HCC18, C30 and HCC18 index-scores	(1) After adjusting for clinical variables, significant correlations were found between QoL (QLQ-C30 and QLQ-HCC18) and dichotomized liver function variables (including Child-Pugh class, ALBI grade and the presence of ascites). (2) It was demonstrated that QoL had significant and potentially clinically important correlations with continuous liver function variables (albumin, bilirubin, ALP and albumin-to-ALP ratio), with the highest Spearman's rank correlation coefficient (rho) exceeding 0.4. (3) HCC18 and C30 index scores were also significantly correlated with these liver function variables. HCC18 index score, which had rho up to 0.37, generally performed better than C30 index score, which had rho up to 0.33.
Qiao et al. (42)	Shanghai	Cross-sectional study	Hospital-based	140	FACT-Hep	(1) The mean FACT-Hep scores were reduced significantly from TNM Stage I to Stage II, Stage IIIA, Stage IIIB group (687 ± 39.69 vs. 547 ± 42.57 vs. 387 ± 51.24 vs. 177 ± 71.44, $P = 0.001$). (2) Regarding the physical and emotional well-being subscales, scores decreased gradually from Stage I to Stage IIIB ($P = 0.002$ vs. Stage I; $P = 0.032$ vs. Stage II; $P = 0.033$ vs. Stage IIIA). (3) Mean FACT-Hep scores varied by Child-Pugh class, especially in the subscales of physical well-being, functional well-being and the hepatobiliary cancer ($P = 0.001$; $P = 0.036$; $P = 0.032$ vs. Child A). (4) For the social and family well-being subscale, only Child C scores were significantly lower as compared with Child A scores ($P = 0.035$). (5) For the subscales of functional well-being and hepatobiliary cancer, there were significant differences for Child A, B and C ($P = 0.002$ vs. Child A).
Wong et al. (46)	Hong Kong	Prospective cohort study	Hospital-based	253 liver cancer 334 lung cancer	FACT-G, MISS-Cog, ChPSQ-9, The single-item visual analog to assess eating appetite, optimism, and depression	(1) There were no differences in QoL, patient satisfaction, and psychosocial measures between the 2 cancer groups. (2) The patients' informational support from medical staff did not predict QoL, but all psychosocial factors emerged as a covariate of the satisfaction in predicting QoL; after controlling for sociodemographic and psychosocial variables, only satisfaction predicted QoL.
Wong et al. (47)	Hong Kong	Prospective cohort study	Hospital-based	253 liver cancer 250 breast cancer 334 lung cancer 242 nasopharyngeal cancer	FACT-G, Eating function: eating ability, eating appetite, eating enjoyment, Pain rating (visual analog), Depression (single item)	(1) Patients with liver cancer reported lower scores in total QoL, physical, functional, emotional, eating appetite, and depression than patients with nasopharyngeal cancer. (2) After controlling for sociodemographic and medical variables, pain, depression, and eating function significantly predicted overall QoL, physical, and functional well-being over time (all cancer). (3) Patients with liver cancer had a slightly decreased score in eating appetite, ability, and enjoyment.

(Continued)

TABLE 4 | Continued

Study	Region	Study design	Sample source	Sample size	Measure(s)	QoL data
Yeo et al. (49)	Hong Kong	Prospective cohort study	Hospital-based	233	EORTC QLQ-C30	(1) Significant independent predictors of shorter survival were advanced Okuda staging ($P = 0.0030$; HR = 2.058), high baseline total bilirubin ($P = 0.0008$; HR = 1.013) and worse QoL score in the appetite score domain ($P = 0.0028$; HR for 10 point increase = 1.070). (2) Patients who were entered into the chemotherapy trial ($P = 0.0002$; HR = 0.503), those who scored better in the physical functioning domain ($P = 0.0034$; HR for 10 point decrease = 0.911) and the role functioning domain ($P = 0.0383$; HR for 10 point decrease = 0.944) of the QoL questionnaire, were associated with longer survival.
Demographic characteristics and psychological factors associated with QoL						
Fan et al. (31)	Taiwan	Cross-sectional study	Hospital-based	33	Semistructured interview	(1) The impact of disease: HCC was associated with physical symptoms and psychosocial stress, as well as positive changes. (2) Illness perceptions: patients perceived HCC as a long-term and chronic disease that could not be cured but might be controlled. (3) Coping strategies: these included focusing on managing HCC and its symptoms, emotional responses, and leading a normal life.
Huang et al. (35)	Taiwan	Cross-sectional study	Hospital-based	77	BFI-T, PSQI-Taiwan Form, Depression subscale of the HADS	(1) Fatigue, sleep disturbance, and depression are positively interrelated and co-occur in patients with HCC. (2) Depression completely mediates the effects of sleep disturbance on fatigue.
Jie et al. (40)	Chongqing	Prospective cohort study	Hospital-based	218	EORTC QLQ-C30, Brief IPQ	(1) When comparing the patients in the disclosed group with the patients who were uninformed, the patients in the disclosed group had higher scores for global QoL at discharge ($P = 0.013$) and higher scores on understanding of their illness regarding illness perceptions ($P = 0.022$). (2) When comparing the patients in the 'autonomy-satisfied' group with the patients whose desire for disclosure was not satisfied, the patients in the autonomy-satisfied group had better emotional functioning and better global QoL at discharge ($P < 0.001$ and $P = 0.001$, respectively). (3) Additionally, the patients in the autonomy-satisfied group had higher scores for personal control ($P = 0.009$) and lower scores for emotional reaction ($P = 0.007$) regarding illness perceptions, even after controlling for other confounding factors.
Jie et al. (41)	Chongqing	Prospective cohort study	Hospital-based	300	PCL-C, PTGI, EORTC QLQ-C30	(1) Compared with the uninformed group, patients in the disclosed group had lower scores for PTSS ($P < 0.001$), higher scores for PTG ($P < 0.001$), better emotional functioning ($P < 0.001$), and better global QoL ($P = 0.006$) at 1 month after discharge.
Qiu et al. (43)	Chongqing	Cross-sectional study	Hospital-based	220	EROTC QLQ-C30, Brief IPQ, SCSQ, SSRS	(1) The mean score of quality of life was obviously higher in the male patients than the female patients ($P < 0.001$). (2) In regard to illness perceptions, personal control ($\beta = 1.707$, $P = 0.003$), identity of symptoms ($\beta = -1.315$, $P = 0.016$) and illness comprehensibility ($\beta = 1.489$, $P = 0.014$) were significantly correlated to the global QoL.
Shun et al. (44)	Taiwan	Prospective cohort study	Hospital-based	104	SDS, HADS, SCNS-SF34	(1) Overall symptom distress decreased monthly, with the highest level before discharge. Compared with the elderly group, the young group had a significantly higher level of symptom distress ($P = 0.024$) but had a lower level of this 2 months after discharge. (2) Patients in the young group who were male ($\beta = -14.24$, $P = 0.012$) and single/widowed ($\beta = -19.17$, $P = 0.033$) and with higher levels of education ($\beta = 1.38$, $P = 0.020$), stage C ($\beta = 30.21$, $P = 0.006$), good functional status ($\beta = 0.78$, $P = 0.024$), and higher levels of symptom distress ($\beta = 2.67$, $P < 0.0001$) and anxiety ($\beta = 6.03$, $P < 0.0001$) had higher levels of overall unmet supportive care needs. (3) For the elderly group, those patients with recent diagnosis status ($\beta = -14.08$, $P = 0.008$), portal vein thrombosis ($\beta = 31.32$, $P = 0.011$), and higher levels of symptom distress ($\beta = 2.09$, $P = 0.003$), anxiety ($\beta = 4.60$, $P < 0.0001$), and depression ($\beta = 2.71$, $P = 0.007$) had higher levels of overall unmet care needs.

(Continued)

TABLE 4 | Continued

Study	Region	Study design	Sample source	Sample size	Measure(s)	QoL data
Zheng et al. (51)	Jiangxi	Cross-sectional study	Hospital-based	166	EORTC QLQ-HCC18, HCC18 index-score, CES-D, RSES, LOT-R	(1) Scores of CES-D, RSES, and LOT-R were (17.34 ± 2.25), (29.59 ± 4.67), and (29.78 ± 3.14), respectively. (2) EORTC QLQ-HCC18, CES-D, RSES, and LOT-R were correlated with scores in various fields ($P < 0.05$). (3) In the influence path of QoL on depressive symptoms, RSES regulation belonged to a complete mediation model and LOT-R regulation belonged to partial intermediary model.

TABLE 5 | The pooled data and norms of the EORTC QLQ-C30 subscales.

	Pooled data	Norms, general population[a]	Norms, heterogeneous cancer[a]
N	1,588	7,802	23,553
Global quality of life	57.1 (5.5)	71.2 (22.4)	61.3 (24.2)
Physical functioning	77.6 (6.7)	89.8 (16.2)	76.7 (23.2)
Role functioning	81.2 (8.2)	84.7 (25.4)	70.5 (32.8)
Emotional functioning	74.9 (3.6)	76.3 (22.8)	71.4 (24.2)
Cognitive functioning	76.4 (5.4)	86.1 (20)	82.6 (21.9)
Social functioning	68.5 (7.4)	87.5 (22.9)	75.0 (29.1)
Fatigue	36.4 (4.4)	24.1 (24)	34.6 (27.8)
Nausea/vomiting	8.9 (2.5)	3.7 (11.7)	9.1 (19)
Pain	26.4 (9.5)	20.9 (27.6)	27.0 (29.9)
Dyspnoea	20.3 (8.2)	11.8 (22.8)	21.0 (28.4)
Insomnia	27.5 (9.4)	21.8 (29.7)	28.9 (31.9)
Appetite loss	20.8 (8.3)	6.7 (18.3)	21.1 (31.3)
Constipation	13.4 (3.8)	6.7 (18.4)	17.5 (28.4)
Diarrhea	12.7 (3.2)	7.0 (18)	9.0 (20.3)
Financial difficulties	40.8 (10.5)	9.5 (23.3)	16.3 (28.1)

Means and standard deviations shown.
[a]*Data from Scott et al. (60).*

HCC Compared With the General Population

Patients with HCC had worse QoL than the general population, especially in global QoL (32, 45), physical (32, 45), role (32), cognitive (32), social functioning (32, 45), pain (45), insomnia (45), appetite loss (45), and financial difficulties (45). Conversely, patients reported better scores in emotional functioning ($P < 0.05$ for all) (32).

The EORTC QLQ-C30 was applied in 11 of 21 studies to measure QoL, but means and standard deviations were only provided in 6 studies (34, 36, 41, 43, 45, 49). We compared the pooled means and standard deviations in baseline of the 6 studies (total HCC cases, 1,588) with the norms which provided by EORTC Group for the general population and patients with heterogeneous cancer (**Table 5**) (60). Patients with HCC had worse global QoL ($t = -6.78$; $P < 0.01$), physical functioning ($t = -4.79$; $P < 0.01$), cognitive functioning ($t = -4.766$; $P < 0.01$), social functioning ($t = -6.798$; $P < 0.01$), fatigue ($t = 5.588$, $P < 0.05$), appetite loss ($t = 3.792$; $P < 0.05$) and financial difficulties ($t = 5.149$; $P < 0.05$) than the general population. Additionally, patients with HCC had worse cognitive functioning ($t = -3.049$; $P < 0.05$), but better role functioning ($t = 3.44$; $P < 0.05$) and emotional functioning ($t = 2.546$; $P < 0.05$) than cancer patients.

Physical Factors and Symptoms Associated With QoL in Chinese Patients

Physical Factors Associated With QoL

Better liver function was significantly correlated with better QoL. Patients with better Child–Pugh classification had better QoL, mainly in terms of higher physical well-being, functional well-being and fewer hepatobiliary symptoms ($P < 0.05$) (39, 42, 50). Similarly, patients with better ALBI grade or the

presence of ascites had higher physical functioning or worse abdominal swelling respectively ($P < 0.001$) (39). In addition to dichotomized liver function variables, lower serum bilirubin levels, higher albumin levels, lower alkaline phosphatase (ALP) levels, better albumin-to-ALP ratios were correlated with better QoL (39).

Tumor stages were negatively related to QoL. The mean FACT-Hep scores of Chinese patients with HCC were markedly decreased with the increasing TNM Stages (Stage I was 687 ± 39.69 vs. Stage II 547 ± 42.57 vs. Stage IIIA 387 ± 51.24 vs. Stage IIIB 177 ± 71.44; $P = 0.001$) (42). The most impaired subscales of QoL were physical, emotional well-being (measured by FACT-Hep; $P < 0.05$) and symptom/side effect (measured by QOL-LC; $P < 0.05$) (42, 50).

There were also significant negative correlations between inflammatory states and QoL. Patients with higher IL-8 levels suffered worse QoL, especially in constitutional symptoms, namely worse role functioning, appetite loss, fatigue, and nutrition ($P < 0.0001$) (37). For C-reactive protein (CRP) and CRP-to-albumin ratio, the QoL factors with higher negative correlations included physical and role functioning, as well as with higher positive correlations included fatigue, pain, appetite loss, body image, nutrition, and abdominal swelling ($P < 0.0001$ for all) (38). Worse QoL factors were also significantly associated with worse scores in Inflammation-Based Index, Glasgow Prognostic Score, and Prognostic index (37, 38).

Furthermore, QoL was considered as a significant predictor for survival time (36, 49). Yeo et al. indicated that significant independent predictors for survival were QLQ-C30 physical functioning (HR 0.911 [0.856-0.969]; $P = 0.0034$), role functioning (HR 0.944 [0.894-0.996]; $P = 0.0383$) and appetite loss (HR 1.070 [1.023–1.118]; $P = 0.0028$) in patients with unresectable HCC (49), whereas the results from Li et al. reported that independent prognostic QoL factors for survival consisted of QLQ-C30 physical functioning (HR 0.652 [0.495–0.860]; $P = 0.0024$), pain (HR 1.346 [1.092–1.661]; $P = 0.0055$), QLQ-HCC18 fatigue (HR 1.441 [1.132–1.833]; $P = 0.0030$), pain (HR 1.382 [1.089–1.754]; $P = 0.0077$) in patients affected by any stage of HCC (36).

Symptoms Associated With QoL

Severe symptoms were significantly associated with worse QoL, including appetite loss and pain (46). It was proved that worse eating appetite ($P < 0.001$) and higher reported pain ($P < 0.001$) predicted lower FACT-G total scores (worse QoL) (46). Moreover, eating ability and performance status were correlated positively with patient QoL (32, 47).

Demographic Characteristics and Psychological Factors Associated With QoL in Chinese Patients

Demographic Characteristics Associated With QoL

Older patients had better QoL (44, 50), mainly in terms of psychological functioning and social functioning ($P < 0.05$) (50). The female gender was associated with worse QoL (QLQ-C30 score for female patients was 41.36 ± 24.06 vs. 55.74 ± 22.03 for male patients; $P < 0.001$) (43). Higher education levels were correlated with better social functioning ($P < 0.05$) (50). Furthermore, income and mode of payment were correlated positively with social functioning, while the way of case finding was correlated negatively with physical and social functioning and correlated positively with symptoms/side effects ($P < 0.05$) (50).

Psychological Factors Associated With QoL

Depression was correlated negatively with QoL (47). One study showed positive correlations among depression, sleep disturbance and fatigue. These symptoms occurred together in HCC patients. Depression had complete mediating effects for the impact of sleep disturbance on fatigue (35). Meanwhile, a recent study showed that the EORTC QLQ-HCC18 and Center for Epidemiological Survey, Depression Scale (CES-D) for HCC patients were correlated with scores in various fields ($P < 0.05$), and self-esteem and optimism completely and partially mediated the effects of QoL on depressive symptoms respectively (51), which suggested that psychological intervention to strengthen the patient′s self-esteem and optimism could be helpful to reduce the symptoms of depression in patients.

Moreover, QoL and depression were significantly associated with care needs. This was supported by evidence that patients with good functional status and higher levels of symptom distress, anxiety and depression had higher levels of overall unmet care needs ($P < 0.05$) (44).

Psychological resilience and QoL were positively interrelated. Emotional, social functioning, and global QoL were positively correlated with psychological resilience ($r = 0.382$, $P < 0.01$; $r = 0.324$, $P < 0.01$; $r = 0.383$, $P < 0.01$) (45). Symptoms of fatigue, pain, fatigue, and nutritional changes were negatively correlated with psychological resilience ($r = -0.303$, $P < 0.05$; $r = -0.286$, $P < 0.05$; $r = -0.360$, $P < 0.01$; $r = -0.259$, $P < 0.05$) (45). Clinically, improving psychological resilience of HCC patients could promote symptom management and improve QoL.

Illness perception and coping were significantly correlated with QoL. Patients who had positive illness perceptions had better QoL, while patients who made more emotion-oriented coping reported worse QoL (31, 32). Furthermore, personal control ($\beta = 1.707$, $P = 0.003$), identity of symptoms ($\beta = -1.315$, $P = 0.016$) and illness comprehensibility ($\beta = 1.489$, $P = 0.014$) regarding illness perceptions were greatly correlated with global QoL (43).

Additionally, diagnosis disclosure and patient autonomy had impacts on illness perceptions and QoL. Patients who were aware of cancer diagnosis had better illness comprehensibility ($P = 0.022$) and global QoL ($P = 0.013$) (40). Autonomy-satisfied patients had lower emotional reaction ($P = 0.007$) and better personal control ($P = 0.009$), and had better global QoL ($P = 0.001$), emotional functioning ($P < 0.001$) (40). Consequently, the desires of patients for autonomy in regard to their diagnosis disclosure should be satisfied. Meanwhile, diagnosis disclosure also correlated to less severe posttraumatic stress symptoms ($P < 0.001$), greater posttraumatic growth ($P < 0.001$), better global QoL ($P = 0.006$), and better emotional functioning ($P < 0.001$) (41). Besides, patients who satisfied with medical services had better QoL (46).

DISCUSSION

This systematic review provided a landscape of the burden caused by HCC in greater China in terms of economic and QoL dimensions from database inception to Apr 6, 2021. The economic impacts on patients with HCC have been receiving rising attention worldwide in the past decades (8, 61–63). Nevertheless, few studies have reported the economic burden of patients with HCC in greater China, and only 6 studies have been identified and included in the present review. As far as the study region was concerned, no publication investigated the economic burden of patients with HCC at a national level. Despite the 2 studies that analyzed the expenditure of HCC in one province or city (26, 28), other studies only focused on a single area, including Taiwan (27, 29), Yunnan province (25) and Beijing (30).

The research methodologies and data sources employed in the included studies were heterogeneous. Two cross-sectional designs and one cohort design were adopted using questionnaires to explore more specific individual expenditures (25, 26, 28). The data of 3 other studies were drawn from the expenditure data of different electronic databases, including the national health insurance database, cancer registry claims data and hospital health records. Such research designs only provided the information of patients' direct medical costs. Non-medical costs and indirect costs, such as supportive services, loss of productivity and family burden, were not considered individually. Although direct hospitalization costs caused huge economic burden for patients with HCC, numerous studies have shown that in addition to the direct medical costs, the economic burden on the family caused by the disease was also extremely significant (64, 65). Therefore, the studies that analyzed the hospitalization data from database could not fully explain the other costs that HCC patients and their families had to bear. In this review, 3 studies that employed questionnaire design reported the economic burden of patients in greater China. One cohort study analyzed both the direct medical costs and indirect costs, and 2 cross-sectional studies investigated the medical costs, non-medical costs and family/household health expenditure of patients with HCC. In 2009, Hu et al. reported that the annual economic burden (direct and indirect cost) of patients with HCC in Beijing and Guangzhou were 6,615 USD (45,180.5 CNY, 2009), and 6,054 USD (41,348.8 CNY, 2009), where the direct cost accounted for 92.0% [6,084 USD (41,553.72 CNY, 2009)] and 90.4% [5,472 USD (37,373.76 CNY, 2009)] (26). The study conducted by Che et al. in Yunnan province between 2012 and 2013 showed that the average annual direct costs for patients with HCC were 33,044 CNY of which the total direct costs/household income per year was 61.9%. These results revealed the big economic impact of HCC on patients and their families (25). Another research reported that the average total expenditure per HCC patient was 55,529 CNY during 2012-2014 based on the 13 provinces in mainland China, including 50,937 CNY of medical expenditure and 4,592 CNY of non-medical expenditure (28). Based on current studies in greater China, HCC imposed a substantial economic burden on patients. The results also showed that health insurance reimbursement could alleviate patients' health care financial burden with high levels of household economic status. The patients who lived in first-line cities might have a higher economic burden than patients who lived in other regions, and the cost of patients with HCC from hospitals and communities also varied greatly, prompting the need for further investigation in the future.

Almost all studies in this review reported the annual direct medical expenditure of HCC. Studies in the US reported that the median overall direct costs for per patient with HCC were 176,456 USD (66), and the annual estimates of lost productivity were 3,553 USD per patient (63). In Australia, the total health expenditure attributed to HCC in 2019–2020 was estimated to be 31,775 USD per patient with HCC, while the total cost of HCC inpatient admissions was 26,438 USD per patient (67). The cost of HCC in Japan was calculated to be 607.2 billion JPY in 2014 (68). It was difficult to directly compare the costs of HCC among different studies due to different regions, health systems, calculation methodologies, the currency and year of calculating prices. Since greater China still had a relatively lower average household income than these developed countries, the economic burden of HCC patients in greater China was still large.

Concerning the QoL burden, the EORTC QLQ-C30 and the FACT-G were the most commonly used measures to assess cancer-specific QoL in patients diagnosed as HCC. The EORTC QLQ-HCC18 and the FACT-Hep supplemented the EORTC QLQ-C30 and the FACT-G, respectively, as disease-specific modules for the assessment of patients with HCC. The Chinese versions of these 4 measures showed good validity and reliability (48, 58, 69, 70), and were suitable for use in Chinese patients. The QOL-LC was also appropriate for Chinese patients due to its cultural sensitivity. We further identified 3 supplementary measures of QoL for use involving Chinese patients. C30 and HCC18 index scores were developed to summarize and analyse raw QoL data to render QLQ-C30 and HCC18 more communicable and meaningful (36). The MNA was suitable to identify the malnutrition and deteriorating QoL for Chinese patients (34). These tools could serve potentially as standardized measures for QoL research regarding patients with HCC in the future.

Compared with medical indices such as functional index, survival rate and mortality rate, the scores of QoL could be more informative as it comprehensively covered all aspects of performance and health status (71). The present review suggested that compared with the general population, Chinese patients with HCC had poor QoL in terms of physical, cognitive functioning and severe symptoms, consistent with the previous results based on the patients with HCC across the globe (72). However, Chinese patients also had worse social relationships, different from the global result (72). Additionally, Chinese patients with HCC had poorer QoL than patients who had benign liver diseases, especially social and physical aspects. Side effects of treatments and severe symptoms might cause impaired physical functioning, especially weight loss, pain, digestive problem, diarrhea and fever.

The physical and psychosocial predictors for QoL in Chinese patients with HCC were identified in this review. Regarding the physical variables, patients who had a better liver function, early tumor stage, less severe inflammatory states and

symptoms, and better eating ability and performance status had better QoL. Psychosocial factors such as better psychosocial resilience, positive illness perceptions, diagnosis disclosure, patient autonomy, and satisfaction of medical services predicted better QoL, whereas depression and emotion-oriented coping indicated worse QoL. The findings suggested that physicians and patients could improve these physical and psychosocial factors to promote QoL correlated with HCC in daily clinical practice.

This systematic review provided the most up-to-date and in-depth overview of the available data about the cost and QoL of HCC in greater China. It revealed the substantial economic and QoL burden of HCC in the region. The findings can contribute to gauge whether health resource allocations match the burden of HCC in greater China and supply useful information for evaluating the population's health status. The burden of HCC measures can also provide significant insights into clinical gaps, leading to further research on the causes of the HCC burden and their resolutions. In addition, as we make great efforts to measure and improve our knowledge of the HCC burden, we may have an opportunity to address the existing methodological challenges to investigate this field further.

However, several limitations of this review can be addressed in future studies. Firstly, limited research provided information on HCC disease staging or progression related to the frequency of clinical services and hospitalizations. More comprehensive health-related costs may be demonstrated by taking into account further disease information. Secondly, differences in the results reported might be a consequence of differences in methodologies and tools for measurement. Thirdly, only a few studies provided the baseline means and standard deviations of QoL data, resulting in limitations regarding pooled data and conduction of meta-analysis. Fourthly, norms of EORTC for the general population were based on individuals from multiple nations except China and might not be representative of Chinese populations. However, a growing body of research is now exploring the applicability of the applications of EORTC in greater China, and there will be more evident in the future to verify reliability and validity of EORTC in Chinese populations. Lastly, living standards in greater China have improved substantially over the past few decades, and improvements in socioeconomic structure, health system, clinical treatment, patient education, and employment support may also affect the disease economic costs and QoL for HCC patients. The specific and quantitative relationship between living standards and disease burden requires further investigation.

CONCLUSION

In summary, HCC has bought a significant economic burden to patients and their families in greater China. HCC also negatively impacted QoL in Chinese patients, especially in physical, cognitive, social functioning and severe symptoms. As the data about the economic and QoL burden of HCC in greater China is lacking, more investigations are needed to provide a deeper interpretation of the disease burden of patients with HCC in greater China in recent years. Moreover, further studies should focus on the disease-related economic impacts on both patients and their families, the effects of physical factors and psychological factors on QoL, the interactions between physical factors and psychological factors, and other predictors that can improve patients' QoL with HCC in greater China.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

HZ, CU, and HH conceived the research project and coordinated the contributors. HZ, ML, QL, ZL, YX, DY, and YL participated in the study selection. HZ and ML extracted, analyzed and interpreted the data, and drafted the manuscript. HH and CU made critical revisions to the manuscript and supervised the study. All the authors reviewed and approved the final vision.

FUNDING

This research was partially supported by a grant from the University of Macau (MYRG2020-00230-ICMS).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2022.801981/full#supplementary-material

REFERENCES

1. Llovet JM, Villanueva A, Lachenmayer A, Finn RS. Advances in targeted therapies for hepatocellular carcinoma in the genomic era. *Nat Rev Clin Oncol.* (2015) 12:408–24. doi: 10.1038/nrclinonc.2015.103
2. Ding X-X, Zhu Q-G, Zhang S-M, Guan L, Li T, Zhang L, et al. Precision medicine for hepatocellular carcinoma: driver mutations and targeted therapy. *Oncotarget.* (2017) 8:55715. doi: 10.18632/oncotarget.18382
3. Yuen MF HJ, Chutaputti A. Hepatocellular carcinoma in the Asia pacific region. *J Gastroenterol Hepatol.* (2009) 24:346–53. doi: 10.1111/j.1440-1746.2009.05784.x
4. Organisation WH. *News Room, Fact Sheets, Cancer*. World Health Organisation. (2018). Available online at: http://www.who.int/news-room/fact-sheets/detail/cancer (accessed Feb 1, 2018).
5. Kim H-S, El-Serag HB. The epidemiology of hepatocellular carcinoma in the USA. *Curr Gastroenterol Rep*. (2019) 21:1–8. doi: 10.1007/s11894-019-0681-x
6. Jinjuvadia R, Salami A, Lenhart A, Jinjuvadia K, Liangpunsakul S, Salgia R. Hepatocellular carcinoma: a decade of hospitalizations and

financial burden in the United States. *Am J Med Sci.* (2017) 354:362–9. doi: 10.1016/j.amjms.2017.05.016
7. Ma J, Wang X, Song ZF, Qiao L, Zhao YQ. Comparison of AASLD, EASL & JSH guidelines and Chinese criteria for diagnosis and treatment of hepatocellular carcinoma. *J Cancer Control Treat.* (2019) 32:1031–8.
8. Sayiner M, Golabi P, Younossi ZM. Disease burden of hepatocellular carcinoma: a global perspective. *Dig Dis Sci.* (2019) 64:910–7. doi: 10.1007/s10620-019-05537-2
9. Yang JD, Hainaut P, Gores GJ, Amadou A, Plymoth A, Roberts LR, et al. global view of hepatocellular carcinoma: trends, risk, prevention and management. *Nat Rev Gastroenterol Hepatol.* (2019) 16:589–604. doi: 10.1038/s41575-019-0186-y
10. Kanda T, Goto T, Hirotsu Y, Moriyama M, Omata M. Molecular mechanisms driving progression of liver cirrhosis towards hepatocellular carcinoma in chronic hepatitis B and C infections: a review. *Int J Mol Sci.* (2019) 20:1358. doi: 10.3390/ijms20061358
11. de Martel C, Georges D, Bray F, Ferlay J, Clifford GM. Global burden of cancer attributable to infections in 2018: a worldwide incidence analysis. *Lancet Glob Health.* (2020) 8:e180–90. doi: 10.1016/S2214-109X(19)30488-7
12. Org C. *Known and Probable Human Carcinogens*. Kennesaw, GA: American Cancer Society (2019).
13. Liver EAFTSOT, EASL. Clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol.* (2018) 69:182–236. doi: 10.1016/j.jhep.2018.03.019
14. Zhang B-H, Yang B-H, Tang Z-Y. Randomized controlled trial of screening for hepatocellular carcinoma. *J Cancer Res Clin Oncol.* (2004) 130:417–22. doi: 10.1007/s00432-004-0552-0
15. Dhanasekaran R, Bandoh S, Roberts LR. Molecular pathogenesis of hepatocellular carcinoma and impact of therapeutic advances. *F1000Res.* (2016) 5:879. doi: 10.12688/f1000research.6946.1
16. Bruix J, Sherman M. Management of hepatocellular carcinoma. *Hepatology.* (2005) 42:1208–36. doi: 10.1002/hep.20933
17. El-Serag HB, Marrero JA, Rudolph L, Reddy KR. Diagnosis and treatment of hepatocellular carcinoma. *Gastroenterology.* (2008) 134:1752–63. doi: 10.1053/j.gastro.2008.02.090
18. Raza A, Sood GK. Hepatocellular carcinoma review: current treatment, and evidence-based medicine. *World J Gastroenterol.* (2014) 20:4115. doi: 10.3748/wjg.v20.i15.4115
19. Mishra A, Otgonsuren M, Venkatesan C, Afendy M, Erario M, Younossi ZM. The inpatient economic and mortality impact of hepatocellular carcinoma from 2005 to 2009: analysis of the US nationwide inpatient sample. *Liver Int.* (2013) 33:1281–6. doi: 10.1111/liv.12201
20. El Khoury A, Klimack W, Wallace C, Razavi H. Economic burden of hepatitis C-associated diseases in the United States. *J Viral Hepat.* (2012) 19:153–60. doi: 10.1111/j.1365-2893.2011.01563.x
21. Bianchi G, Loguercio C, Sgarbi D, Abbiati R, Brunetti N, De Simone T, et al. Reduced quality of life of patients with hepatocellular carcinoma. *Dig Liver Dis.* (2003) 35:46–54. doi: 10.1016/S1590-8658(02)00011-7
22. Gandhi S, Khubchandani S, Iyer R. Quality of life and hepatocellular carcinoma. *J Gastrointest Oncol.* (2014) 5:296. doi: 10.3978/j.issn.2078-6891.2014.046
23. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med.* (2009) 151:264–9. doi: 10.7326/0003-4819-151-4-200908180-00135
24. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ.* (2009) 339:b2700. doi: 10.1136/bmj.b2700
25. Che Y, Chongsuvivatwong V, Li L, Sriplung H, Wang Y, You J, et al. Financial burden on the families of patients with hepatitis B virus-related liver diseases and the role of public health insurance in Yunnan province of China. *Public Health.* (2016) 130:13–20. doi: 10.1016/j.puhe.2015.03.015
26. Hu M, Chen W. Assessment of total economic burden of chronic hepatitis B (CHB)-related diseases in Beijing and Guangzhou, China. *Value Health.* (2009) 12(Suppl. 3:S89–92. doi: 10.1111/j.1524-4733.2009.00636.x
27. Lang H-C, Wu J-C, Yen S-H, Lan C-F, Wu S-L. The lifetime cost of hepatocellular carcinoma. *Appl Health Econ Health Policy.* (2008) 6:55–65. doi: 10.2165/00148365-200806010-00005
28. Lei H, Lei L, Shi J, Wu Y, Liang L, Huang H, et al. No expenditure difference among patients with liver cancer at stage I– IV: findings from a multicenter cross-sectional study in China. *Chin J Cancer.* (2020) 32:516. doi: 10.21147/j.issn.1000-9604.2020.04.09
29. Nguang S-H, Wu C-K, Liang C-M, Tai W-C, Yang S-C, Ku M-K, et al. Treatment and cost of hepatocellular carcinoma: a population-based cohort study in Taiwan. *Int J Environ Res Public Health.* (2018) 15:2655. doi: 10.3390/ijerph15122655
30. Ran TY, Sun LH. Study on the direct medical expenses and cost composition of inpatient diseases in a large tertiary a hospital. *Chin Med J.* (2019) 20:67–71.
31. Fan S-Y, Eiser C. Illness experience in patients with hepatocellular carcinoma: an interpretative phenomenological analysis study. *Eur J Gastroenterol Hepatol.* (2012) 24:203–8. doi: 10.1097/MEG.0b013e32834ec184
32. Fan SY, Eiser C, Ho MC, Lin CY. Health-related quality of life in patients with hepatocellular carcinoma: the mediation effects of illness perceptions and coping. *Psycho-Oncol.* (2013) 22:1353–60. doi: 10.1002/pon.3146
33. Fielding R, Wong WS. Quality of life as a predictor of cancer survival among Chinese liver and lung cancer patients. *Eur J Cancer.* (2007) 43:1723–30. doi: 10.1016/j.ejca.2007.05.002
34. Hsu W-C, Tsai AC, Chan S-C, Wang P-M, Chung N-N. Mini-nutritional assessment predicts functional status and quality of life of patients with hepatocellular carcinoma in Taiwan. *Nutr Cancer.* (2012) 64:543–9. doi: 10.1080/01635581.2012.675620
35. Huang T-W, Lin C-C. The mediating effects of depression on sleep disturbance and fatigue: symptom clusters in patients with hepatocellular carcinoma. *Cancer Nurs.* (2009) 32:398–403. doi: 10.1097/NCC.0b013e3181ac6248
36. Li L, Mo FK, Chan SL, Hui EP, Tang NS, Koh J, et al. Prognostic values of EORTC QLQ-C30 and QLQ-HCC18 index-scores in patients with hepatocellular carcinoma - clinical application of health-related quality-of-life data. *BMC Cancer.* (2017) 17:8. doi: 10.1186/s12885-016-2995-5
37. Li L, Chan SL, Mo F, Hui EP, Koh J, Chan AK, et al. Correlations of health-related quality of life with serum inflammatory indicators IL-8 and mIBI in patients with hepatocellular carcinoma. *Cancer Manag Res.* (2019) 11:2719. doi: 10.2147/CMAR.S178482
38. Li L, Chan SL, Mo F, Hui EP, Koh J, Chan AK, et al. Status of inflammation in relation to health related quality of life in hepatocellular carcinoma patients. *Qual Life Res.* (2019) 28:2597–607. doi: 10.1007/s11136-019-02190-0
39. Li L, Mo F, Hui EP, Chan SL, Koh J, Tang NL, et al. The association of liver function and quality of life of patients with liver cancer. *BMC Gastroenterol.* (2019) 19:1–12. doi: 10.1186/s12876-019-0984-2
40. Jie B, Qiu Y, Feng ZZ, Zhu SN. Impact of disclosure of diagnosis and patient autonomy on quality of life and illness perceptions in Chinese patients with liver cancer. *Psycho-Oncol.* (2016) 25:927–32. doi: 10.1002/pon.4036
41. Jie B, Zhou Y-H, Qiu Y, Yang G-Y, Feng Z-Z. Impact of the disclosure of diagnosis on posttraumatic stress and growth and quality of life in Chinese patients with hepatocellular carcinoma. *Support Care Cancer.* (2020) 28:3371–9. doi: 10.1007/s00520-019-05141-0
42. Qiao C-X, Zhai X-F, Ling C-Q, Lang Q-B, Dong H-J, Liu Q, et al. Health-related quality of life evaluated by tumor node metastasis staging system in patients with hepatocellular carcinoma. *World J Gastroenterol.* (2012) 18:2689. doi: 10.3748/wjg.v18.i21.2689
43. Qiu Y, Yang J, Ma KS, Jie B, Shu WQ. Correlation of illness perceptions with quality of life in early-and intermediate-stage hepatocellular carcinoma patients. *J Amry Med Univ.* (2017) 39:1150–4.
44. Shun S-C, Lai Y-H, Hung H, Chen C-H, Liang J-D, Chou Y-J. The role of age in change in unmet supportive care needs in hepatocellular carcinoma patients during transition from hospital to home. *Cancer Nurs.* (2017) 40:245–54. doi: 10.1097/NCC.0000000000000403
45. Wei YQ, Li AH, Chen Y, Zhao XH, Tang J, Wei QL, et al. Analysis of psychological resilience and its effect on quality of life and clinical symptoms in hospitalized patients with hepatocellular carcinoma. *Chin Med.* (2019) 26:206–10.

46. Wong WS, Fielding R. The association between patient satisfaction and quality of life in Chinese lung and liver cancer patients. *Med Care.* (2008) 46:293–302. doi: 10.1097/MLR.0b013e31815b9785
47. Wong WS, Fielding R. Eating ability predicts subsequent quality of life in Chinese patients with breast, liver, lung, or nasopharyngeal carcinoma: a longitudinal analysis. *Acta Oncol.* (2008) 47:71–80. doi: 10.1080/02841860701441814
48. Yang Z, Wan C, Li W, Cun Y, Meng Q, Ding Y, et al. Development and validation of the simplified chinese version of EORTC QLQ-HCC18 for patients with hepatocellular carcinoma. *Cancer Invest.* (2015) 33:340–6. doi: 10.3109/07357907.2015.1036280
49. Yeo W, Mo F, Koh J, Chan A, Leung T, Hui P, et al. Quality of life is predictive of survival in patients with unresectable hepatocellular carcinoma. *Ann Oncol.* (2006) 17:1083–9. doi: 10.1093/annonc/mdl065
50. Zhang B, Wang YB, Liu SP, Chen MH. Evaluation of quality of life and its influencing factors of primary liver cancer before treatment. *J China Med Univ.* (2009) 31:2069–70.
51. Zheng Q, Wang K, Zhou Y, Lei YF, Huang YY, Feng LM. Mediating effect of self?esteem and optimism on effect of health related quality of life and depression in patients with hepatocellular carcinoma. *Chin Nurs Res.* (2021) 35:602–7.
52. Eisenberg JM. Clinical economics: a guide to the economic analysis of clinical practices. *Jama.* (1989) 262:2879–86. doi: 10.1001/jama.1989.03430200123038
53. Brown ML, Riley GF, Schussler N, Etzioni R. Estimating health care costs related to cancer treatment from SEER-Medicare data. *Med Care.* (2002) 40(Suppl. 8):IV104–17. doi: 10.1097/00005650-200208001-00014
54. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst.* (1993) 85:365–76. doi: 10.1093/jnci/85.5.365
55. Cella DF, Tulsky DS, Gray G, Sarafian B, Linn E, Bonomi A, et al. The functional assessment of cancer therapy scale: development and validation of the general measure. *J Clin Oncol.* (1993) 11:570–9. doi: 10.1200/JCO.1993.11.3.570
56. Blazeby JM, Currie E, Zee BC, Chie WC, Poon RT, Garden OJ, et al. Development of a questionnaire module to supplement the EORTC QLQ-C30 to assess quality of life in patients with hepatocellular carcinoma, the EORTC QLQ-HCC18. *Eur J Cancer.* (2004) 40:2439–44. doi: 10.1016/j.ejca.2004.06.033
57. Heffernan N, Cella D, Webster K, Odom L, Martone M, Passik S, et al. Measuring health-related quality of life in patients with hepatobiliary cancers: the functional assessment of cancer therapy-hepatobiliary questionnaire. *J Clin Oncol.* (2002) 20:2229–39. doi: 10.1200/JCO.2002.07.093
58. Zhu ZC, Lang QB, Chen Z, Li DT, Ling CQ. Evaluation of Chinese version of the Functional Assessment of Cancer Therapy-Hepatobiliary questionnaire. *Zhong Xi Yi Jie He Xue Bao.* (2008) 6:341–5. doi: 10.3736/jcim20080403
59. Wan CH, Fang JQ, Zhang CZ, Lin Y, Luo YM. Development and evaluation of a quality of life scale for patients of liver cancer. *Chin J Behav Med Sci.* (1998) 7:170–3.
60. Scott NW, Fayers PM, Aaronson NK, Bottomley A, Graeff A, Groenvold M, et al. *EORTC QLQ-C30 Reference Values*. Brussels: EORTC Quality of Life Group (2008).
61. Nguyen V, Law M, Dore G. Hepatitis B-related hepatocellular carcinoma: epidemiological characteristics and disease burden. *J Viral Hepat.* (2009) 16:453–63. doi: 10.1111/j.1365-2893.2009.01117.x
62. Sartorius K, Sartorius B, Aldous C, Govender P, Madiba T. Global and country underestimation of hepatocellular carcinoma (HCC) in 2012 and its implications. *Cancer Epidemiol.* (2015) 39:284–90. doi: 10.1016/j.canep.2015.04.006
63. Lang K, Danchenko N, Gondek K, Shah S, Thompson D. The burden of illness associated with hepatocellular carcinoma in the United States. *J Hepatol.* (2009) 50:89–99. doi: 10.1016/j.jhep.2008.07.029
64. Council G, Assembly G, Medal HLB, Medal LB, Lecture LWDG, Centers T, et al. *Global Burden Of Liver Disease: A True Burden on Health Sciences and Economies!!* (2012).
65. Kohn CG, Singh P, Korytowsky B, Caranfa JT, Miller JD, Sill BE, et al. Humanistic and economic burden of hepatocellular carcinoma: systematic literature review. *Am J Manag Care.* (2019) 25:SP61–73.
66. Tapper EB, Catana AM, Sethi N, Mansuri D, Sethi S, Vong A, et al. Direct costs of care for hepatocellular carcinoma in patients with hepatitis C cirrhosis. *Cancer.* (2016) 122:852–8. doi: 10.1002/cncr.29855
67. Deloitte. *The Social and Economic Cost of Primary Liver Cancer Hepatocellular Carcinoma (HCC) in Australia* (2021).
68. Matsumoto K, Wu Y, Kitazawa T, Fujita S, Seto K, Hasegawa T. Cost of illness of hepatocellular carcinoma in Japan: a time trend and future projections. *PLoS ONE.* (2018) 13:e0199188. doi: 10.1371/journal.pone.0199188
69. Wan C, Meng Q, Yang Z, Tu X, Feng C, Tang X, et al. Validation of the simplified Chinese version of EORTC QLQ-C30 from the measurements of five types of inpatients with cancer. *Ann Oncol.* (2008) 19:2053–60. doi: 10.1093/annonc/mdn417
70. Yu CL, Fielding R, Chan CL, Tse VK, Choi PH, Lau WH, et al. Measuring quality of life of Chinese cancer patients: a validation of the Chinese version of the Functional Assessment of Cancer Therapy-General (FACT-G) scale. *Cancer.* (2000) 88:1715–27
71. Bonnetain F, Paoletti X, Collette S, Doffoel M, Bouche O, Raoul JL, et al. Quality of life as a prognostic factor of overall survival in patients with advanced hepatocellular carcinoma: results from two French clinical trials. *Qual Life Res.* (2008) 17:831–43. doi: 10.1007/s11136-008-9365-y
72. Fan SY, Eiser C, Ho MC. Health-related quality of life in patients with hepatocellular carcinoma: a systematic review. *Clin Gastroenterol Hepatol.* (2010) 8:559–64 e1–10. doi: 10.1016/j.cgh.2010.03.008

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Zou, Li, Lei, Luo, Xue, Yao, Lai, Ung and Hu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Public Health

SYSTEMATIC REVIEW
published: 28 April 2022
doi: 10.3389/fpubh.2022.858587

The Dose-Related Efficacy of Acupuncture on Endometrial Receptivity in Infertile Women: A Systematic Review and Meta-Analysis

Xiaoyan Zheng[1,2], Siyi Yu[1], Liying Liu[1], Han Yang[1], Fangge Wang[2], Hongmei Yang[2], Xingyu Lv[2] and Jie Yang[1,2]**

[1] Acupuncture and Tuina School, Chengdu University of Traditional Chinese Medicine, Chengdu, China, [2] Clinical Research Center for Acupuncture and Moxibustion in Sichuan Province, Chengdu Xi'nan Gynecological Hospital, Chengdu, China

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Qinyu Zhao,
Shandong University of Traditional Chinese Medicine, China
Yujia Wei,
Mayo Clinic, United States

***Correspondence:**
Xingyu Lv
lvxy@jxr-fertility.com
Jie Yang
jenny_yang_jie@126.com

Specialty section:
This article was submitted to Family Medicine and Primary Care, a section of the journal Frontiers in Public Health

Received: 20 January 2022
Accepted: 15 March 2022
Published: 28 April 2022

Citation:
Zheng X, Yu S, Liu L, Yang H, Wang F, Yang H, Lv X and Yang J (2022) The Dose-Related Efficacy of Acupuncture on Endometrial Receptivity in Infertile Women: A Systematic Review and Meta-Analysis. Front. Public Health 10:858587. doi: 10.3389/fpubh.2022.858587

Background: Progress has been achieved by using acupuncture widely for poor endometrial receptivity (PER). However, different acupuncture dosages may lead to controversy over efficacy.

Objective: To evaluate the evidence-based conclusions of dose-related acupuncture on infertile women with PER.

Method: References were retrieved from nine databases from inception to 26 February 2022. This meta-analysis included randomized controlled trials (RCTs) that investigated the dose-related efficacy of acupuncture for PER with outcomes of endometrium receptivity (ER) parameters by transvaginal sonography (TVS) and the subsequent pregnancy outcomes in three acupuncture-dose groups: the high-dosage group (three menstrual cycles), the moderate-dosage group (one menstrual cycle), and the low-dosage group (two or four days). Since there remained sufficient heterogeneity among the three subsets, we prespecified seven subgroup variables (four clinical and three methodological) to investigate the heterogeneities.

Results: A total of 14 RCTs (1,564 women) of moderate or low overall quality were included. The results were different when the dosage of acupuncture was restricted. For the moderate or high-dosage group, CPR and part of ER parameters were improved in the acupuncture group (i.e., CPR: OR = 2.00, 95% CI [1.24, 3.22], $p = 0.004$, $I^2 = 0\%$ in one menstrual cycle; OR = 2.49, 95%CI [1.67, 3.72], $p < 0.05$, $I^2 = 0\%$ in three menstrual cycles). However, for the low-dosage group, no statistical difference was observed in CPR (OR = 0.07, 95% CI [−0.10, 0.23], $p = 0.44$, $I^2 = 82\%$) and a part of the ER parameters. In subgroup analysis, four subgroup variables (the routine treatment, risk of performance bias, duration of acupuncture treatment, and the age of participants) could explain some of the heterogeneities across all trials.

Conclusion: The finding indicated that the trend of relatively more acupuncture dosage showed better effects for poor endometrial receptivity among PER women. It remains a potential heterogeneity in our studies. Further high-quality trials with a homogeneity trial design need to be conducted.

Keywords: dose-related, acupuncture, endometrial receptivity, meta-analysis, heterogeneity analysis

INTRODUCTION

The incidence of infertility has begun to increase annually (1, 2), and it has gradually become the third most common disease worldwide (3). Potential embryonic development, optimal endometrial receptivity (ER), and synchronization of the embryo and endometrium play critical roles in a successful pregnancy (4). ER refers to the endometrium's ability to allow the blastocyst to attach and grow. Impaired ER will reduce the synchronization and lead to infertility. Approximately 2/3 of embryo implantation (ET) failure is closely related to poor endometrial receptivity [PER; (5, 6)]. However, little progress has been achieved for PER over three decades after the introduction of *in vitro* fertilization (IVF), which has achieved great improvement in embryo quality (4).

Acupuncture has been widely used to treat infertility for a long period. Some potential mechanisms have been postulated to explain the role of acupuncture in the IVF procedure. Firstly, for the stimulation of ovary induction, electro-acupuncture can alter several different neuroendocrinological factors, such as β-endorphin, which can mediate the hypothalamus-pituitary-gonadal (HPG) and-adrenal axes (HPA) and regulate the menstrual cycle, ovulation, and fertility (7). Secondly, electro-acupuncture can circulate the blood flow of the uterus, reduce the resistance of uterine arteries (8), and increase ovarian blood flow through the ovarian sympathetic nerves (9). Thirdly, acupuncture may have the efficacy of meditating the immune response for the achievement and maintenance of a successful pregnancy (10). Besides, acupuncture can reduce the anxiety level (11) and regulate serum cortisol (CORT) and PRL (12), which simulate estradiol (E2) and progestin (P) independently. However, the dosage of the acupuncture intervention varies from 2 (13) to 18 sessions (14) (lasting for three menstrual cycles), and the efficacy of acupuncture remains a contention worldwide. In 2002, Paulus et al. (13) conducted the first randomized controlled trial (RCT) which evaluated the efficacy of acupuncture for women undergoing *in vitro* fertilization and embryo transfer (IVF-ET) with two sessions of acupuncture (25 min before ET and after ET) and found out that acupuncture could significantly improve clinical pregnancy rate (CPR), which was 42.5% in the acupuncture group compared to 26.3% in the control group. However, many trials (15–18) which attempted to repeat this outcome with the same intervention (2 sessions of acupuncture) have not been successful. Moreover, Ajeena et al. (19) found out that six sessions of transcutaneous electrical nerve stimulation (TENS), which is a kind of electro-acupoint stimulation, are beneficial to increase endometrial thickness in healthy women at their childbearing age. For the duration of three menstrual cycles of acupuncture, Zhuang et al. (20) found that CPR and EMT were improved significantly, while Shuai et al. (14) did not find out the same outcome in EMT. Furthermore, Belinda (21) confirmed that infertility is a complex medical issue, often associated with either significant previous gynecological issues/pathology and/or with advanced-age patients (>30 y or <40 y). Infertile women who receive longer-term treatment may benefit from the effects of regular acupuncture on their other health issues. But it did not mention the optimal dosages of acupuncture.

Therefore, the purpose of this systematic review is to evaluate the efficacy of acupuncture in improving the ER for infertile women with PER and to identify the optimal acupuncture plan (in terms of duration of acupuncture intervention, and optimal intervention measurement). The finding can be used to obtain more vigorous evidence-based clinical practice.

METHOD

Protocol and Registration

The systematic review protocol has been registered on the prospective international register of systematic review (PROSPERO: registration number is CRD42020206790) (https://www.crd.york.ac.uk/PROSPERO/).

All contents and report details were strictly referred to as Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) (22), as shown in **Supplementary Material S1**.

Literature Search Strategy

References were retrieved from nine databases: four English databases (i.e., PubMed, Embase, Cochrane Library, and Web of Science), and five Chinese databases, i.e., SinoMed (formerly Chinese Biomedical Database), Chinese National Knowledge Infrastructure (CNKI), Wanfang Data, China Biomedical Literature Database, and China Science Journal Database (VIP database) from inception to 25 February 2022. The search strategy was based on the guidance of the Cochrane handbook. The language was limited to English and Chinese. The search strategies are shown in **Supplementary Material S2**.

In addition, to reduce the publication bias, we searched for lists of relevant references. Clinical trial registries, i.e., Menstrual Disorders and Subfertility Group (MDSG) Specialized Register, Cochrane Central Register of Controlled Trials (CENTRAL), World Health Organization International Clinical Trials Registry Platform, Chinese clinical registry, and Clinical Trials. Gov, and we manually searched key journals and meetings such as the European Society for Human Reproduction and Embryology (ESHRE) and American Society of Reproduction Medicine

(ARSM) for relevant articles, including relevant journals and conferences abstracts, by connecting with the coordinator.

Inclusion and Exclusion Criteria

Study Participants

In our study, the starting points are those patients who suffered from PER where the uterine factor is routinely evaluated by transvaginal sonography (TVS) without anatomic abnormality. Repeated implantation failure (RIF) (23) and polycystic ovarian syndrome (PCOS) (24) were included considering impaired ER can be found in the categories of patients.

Study Intervention and Comparison

We collected randomized controlled trials (RCTs) that compared verum acupuncture with placebo acupuncture, or no adjuvant treatment in the review. In a broad sense, verum acupuncture included auricular acupuncture, electro-acupuncture, manual acupuncture, and transcutaneous electrical acupoint stimulation (TEAS) with or without moxibustion. Placebo acupuncture included placebo acupuncture devices, shallow acupuncture, non-acupoints or non-therapeutic acupoints, and mock electrical stimulation. No adjuvant treatment only included western routine treatment.

Study Outcomes Measures

The primary outcome was the clinical pregnancy: the presence of at least one intrauterine gestational sac or fetal heartbeat confirmed by ultrasound 4–6 weeks after embryo transfer or ovulation, and at least one of the following outcomes was extracted as the secondary outcomes by TVs:

i) EMT: It refers to the distance between the endometrium's anterior and posterior walls, including the uterine cavity gap. EMT is one of the most frequently employed indirect predictors of ER. The pregnancy outcome was significantly higher in women with $7 < \text{EMT} \leq 14$ mm (25). However, EMT is a conflicting indicator in predicting the pregnancy outcome alone (26).
ii) Endometrium pattern (EMP), according to Gonen's criterion (27): Type A, trilinear or multilayered endometrium, strong echo in the outer and moderate parts, hypoechoic or dark areas in the inner layer, and unmistakable linear echo in the uterine cavity; Type B, weak trilinear, isolated echo in the moderate, inconspicuous echo in the moderate uterine cavity; and Type C, strong echo, no intrauterine midline echo. Type A with "triple-line" appears important in improving the pregnancy rate (25).
iii) Blood flow indicators include the resistance index (RI), pulse index (PI), endometrial vascular index (VI), flow index (FI) of the uterine artery, and endometrial blood flow: peak systolic velocity/ end-diastolic blood velocity (S/D) (28). Reducing uterine vascular RI and improving uterine blood flow can improve the implantation rate (29).
iv) Live birth: The live birth is newborns (>28 weeks of gestation) who were delivered with signs of life. Live birth rate (LBR is a ratio of live birth and clinical pregnancy).

We excluded RCTs that extracted EMT as the only outcome after acupuncture treatment because the utilization of EMT as a tool to decide on IVF cycle cancellation is not justified based on the current meta-analysis (30). The Chinese herb utilization was excluded to evaluate the efficacy of acupuncture precisely.

Selection of Studies and Data Extraction

Two reviewers (XYZ and HMY) independently selected the studies, extracted data, and downloaded the citations into Note Express software Version 2.6.1 (Aegean Sea software company, Beijing, China) for data management and eliminated the duplicate research by software. Any discrepancies were resolved through further discussion with the third reviewer (FGW).

For trials to be eligible, intervention dosage (duration and frequency) for acupuncture treatment is extracted in detail. We categorized the acupuncture dosages into three groups: low, moderate, and high. Moderate-dosage group was defined as "one menstrual cycle" of acupuncture duration [including the acupuncture conducted during the controlled ovarian hyperstimulation (COH) procedure in IVF or during frozen-thawed embryo transfer (FET)]. For the duration of acupuncture that was less than "one menstrual cycle," we extracted studies in the "low-dosage group," while for acupuncture duration that was more than "one menstrual cycle," we extracted the studies in the "high-dosage group."

Assessment of Included Studies

Assessment of Bias

The risks of bias of RCTs were assessed using the Cochrane Collaboration's tool (31). The criteria consist of seven items: i) selection bias (random sequence generation and allocation concealment); ii) performance bias (blinding of participants and personnel); iii) detection bias (blinding of outcome assessment); iv) attrition bias (incomplete outcome data); v) reporting bias (selective reporting), and vi) other bias. Each study was evaluated as high, moderate, low, or very low for each evidence.

The Quality of Evidence

We used the Grades of Recommendation, Assessment, Development, and Evaluation [GRADE; (32)] score to assess the quality of each piece of evidence. The quality was classified into high, moderate, or very low, and five reasons to possibly rare down the quality of each evidence are as follows: i) limitation in study design or execution (risk of bias); ii) inconsistency of results, inconsistent results; iii) evidence of indirectness; iv) imprecision; v) publication bias. Each study was evaluated as high, low, or unclear risk of bias for each item.

Data Synthesis and Analysis

We used the RevMan5.4.1 provided by the Cochrane Collaboration to analyze data. For dichotomous data (CPR, LBR, and EMP), we expressed the results for each study as the odds ratio (OR) with a 95% confidence interval (CIs). For continuous data (EMT, RI, PI, and S/D), we expressed the results as the difference or standardized mean difference (SMD) with 95% CI. If data could be synthesized, we used descriptive analysis to solve this problem. For pooled data, we

used I-square (I^2) statistics, which indicates the proportion of variability across trials not explained by sampling variation alone, and the Cochran Q-test for heterogeneity assessing (31). A fixed-effects model was performed for low heterogeneity ($I^2 < 50\%$ statistics or the Cochran Q-test, $p > 0.05$), otherwise, random-effects were performed. We used network construction between acupoints-associated and duration-associated to show their relationship visually.

To find out the dose-related efficacy of acupuncture, we divided the extracted data into three groups based on the different dosages of acupuncture intervention separately (2 or 4 days only, one menstrual cycle, or three menstrual cycles). We assessed the likelihood of publication bias by constructing funnel plots and the Eggers' test.

Subgroup Analyses

For pooled data, characteristics varied from different studies. We used subgroup analysis instead of meta-regression to find the source of heterogeneity, mainly considering the insufficient number of included studies. Subgroup analysis was based on four clinical characteristics: i) acupuncture duration; ii) intervention of different routine treatment (IVF-ET, FET, or clomiphene citrate (CC)/letrozole (LE) for ovulation induction); iii)acupuncture used alone or with other intervention; iv) mean age of participants (≥35 years or not) and three methodological characteristics: i) risk of blinding of participants and personnel (low risk or high risk); ii) risk of random sequence generation (low risk or high risk); and iii) risk of allocation concealment (low risk or high risk).

For each subgroup analysis, we performed a single covariate weighted random effects with 95% confidence interval in Stata version 16 (*StataCorp)* to investigate whether differences in effects of adjuvant acupuncture between the covariate's two subgroups were statistically significant. For each single covariate subgroup analysis, we calculated the interaction p-value of the test and the percentage of the heterogeneity explained by the covariate (I^2).

RESULTS

Selection of Studies

Figure 1 shows the details of the study selection process.

The PRISMA flow diagram (**Figure 1**) shows the details of the study selection proces. Of the 101 identified trials, 76 studies were excluded after screening full-text (6 studies were not available for full-text. Six studies were not RCTs, seven studies had no outcomes for interests, 34 studies had the only outcome of EMT, which cannot represent ER in our opinion, and 34 other studies were also excluded because of low quality or repeated publication). Thus, 14 RCTs with a total of 1,564 participants met the inclusion criteria. Twelve (14, 16, 20, 33–41) studies were conducted in China [ten studies (14, 20, 33, 34, 36–41) were from the mainland, China and two were (16, 35) from Taiwan, China], and two (13, 42) were from Germany.

Trial Characteristics

Table 1 shows the characteristics of RCTs included in the review. The differences in trial eligibility criteria were that six trials (14, 20, 34, 36, 39, 41) included women with RIF and one trial (37) included women with PCOS. The duration of the acupuncture sessions differed among trials (**Table 1**). Three studies (13, 16, 42) conducted two sessions around ET and one study (35) conducted 4 sessions during COH. Some women received acupuncture for almost one menstrual cycle [acupuncture was performed for 1 month before FET in one study (38) and during FET in three studies (36, 39, 40)]. Meanwhile, the other six trials (14, 20, 33, 34, 37, 41) lasted for three menstrual cycles before embryo transfer. Six studies (13, 16, 33, 35, 38, 42) conducted IVF-ET of routine treatment, seven studies (14, 20, 34, 36, 39–41) were performed in FET, and one study (37) administered letrozole (LE) and human chorionic gonadotropin (HCG) for the induction of ovulation in routine treatment.

For the frequency of the collected studies, one study (41) did not mention the frequency of acupuncture, three studies (33, 34, 40) conducted acupuncture every day, three studies (36, 37, 39) were conducted every other day, two studies (20, 38) were conducted three times a week, while one study (14) conducted for six times per cycle. However, all of the collected studies did not mention the total dosage of acupuncture treatments, except that three studies (13, 16, 42) were conducted only for 2 days (25 min before and after ET), and one study (35) was conducted for 4 days in total.

Besides, the acupoints-associated and duration-associated network construction chart (**Figure 2**) shows that various acupoints were selected in treating PER. Among the whole trails, SP6, EX-CA1, RE4, ST36, and LR3 were the top five acupoints, and three menstrual cycles were the most frequently used duration for the whole treatment.

The Methodological Quality of Included Trials

Risk Bias of Included Studies

A summary of the risks of bias is presented in **Table 2** and **Figure 3**. Three trials (33, 35, 41) did not mention random sequence generation. Eight studies (13, 14, 16, 20, 33, 35, 41, 42) mentioned allocation concealment, which used opaque sealed envelopes subjects to assign included subjects randomly. As presented in **Table 1**, only one study (16) utilized Streitberger control. Zhao et al. (39) used shallow needles (depth <5 mm) placed on the non-acupoints. Dieterle et al. (42) used an actual needling procedure on acupoints that were designed not to affect fertility, Wang et al. (41) conducted sham acupuncture (acupoint pressing stimulated epidermal irritation slightly without de qi), two studies (14, 40) utilized mock TEAS, and others used no intervention as the control group. For 4 studies (20, 37, 38, 40), some randomized women were recruited from the beginning but did not complete the treatment (i.e., there was no embryo transfer or they did not complete the ER tests), and missing data were not reported which would increase the attrition bias. Four studies (33, 36, 39, 41) remained with other unclear biases.

GRADE for Quality of Each Evidence

Fourteen reviews (13, 14, 16, 20, 33–42) included 24 outcomes that were related to the efficacy of acupuncture for PER. There was low or very low evidence to indicate that acupuncture might improve the CPR or ER parameters when acupuncture

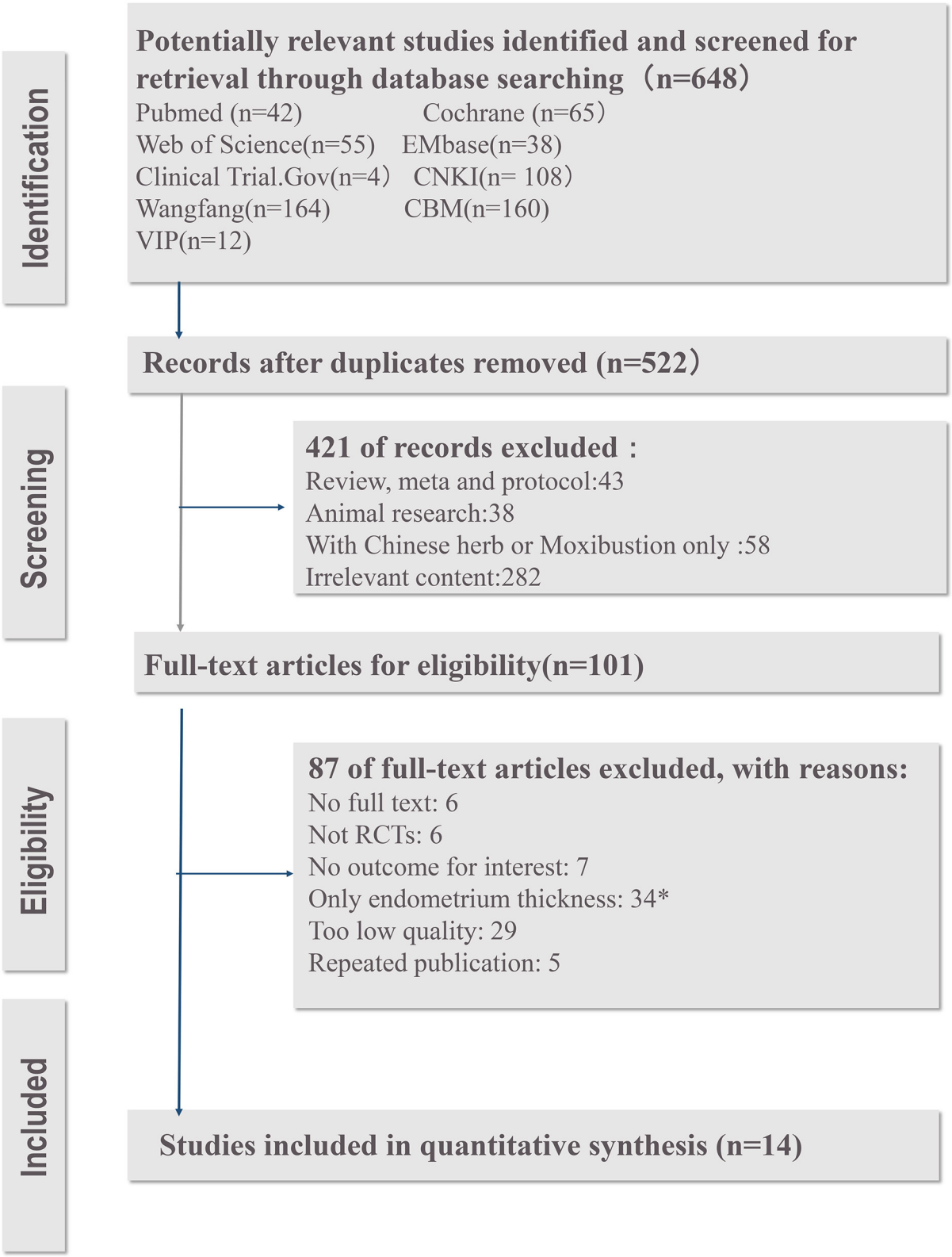

FIGURE 1 | The PRISMA flow diagram of study screening process. *The studies were excluded because that the utilization of EMT as a tool to decide on IVF cycle cancelation is not justified based on the current meta-analysis.

TABLE 1 | Characteristics of included trials.

References	*N**	Study population	Mean age	Acupuncture group	Control group†	Starting time of intervention§ And the duration of intervention	Outcomes	Routine treatment
Chen and Hau (33)	T:57 C:57	i)age: 24–35y ii)IVF-ET iii)regular menstrual cycles	T:31 ± 3 C:31 ± 3	**Acupuncture& Moxibustion (*n* = 57)** i)Acupuncture: [Needles stimulated manually by rotating, lifting, and thrusting the handle of the needle to maintain de qi sensation, both during initial insertion and after 10 min] ii)Moxibustion: [Light the moxibustion strip to navel ¶ until the patient feel warm for 30 min]	**No adjuvant treatment (*n* = 57)**	**Ac time:** during COH **Frequency**: every day **Duration:**3 menstrual cycles	EMT, EMP, S/D, PI, RI	IVF-ET
Paulus et al. (13)	T:80 C:80	i)IVF-ET	T: 32.1 ± 3.9 C: 32.8 ± 4.1	**Acupuncture& auricular acupuncture (*n* = 80)** i) Acupuncture: [soreness, numbness, or distention around the point = Deqi sensation) occurred during the initial insertion.]	**No adjuvant treatment (*n* = 80)**	**Ac time:** around ET **Frequency:** 25 min before and after embryo transfer **Duration:**2 days	EMT, PI, CPR	IVF-ET
Ho et al. (35)	T:30 T:26	i)IVF-ET	T: 35.5 ± 4.5 C: 34.0 ± 5.2	**Electro-acupuncture (*n* = 30)** [The needles were twirled by hand to evoke a needle reaction; this often resulted in soreness, numbness, and distension around the point, The needles were then attached to an electrical stimulator at a low frequency of 10 Hz for 30 min]	**No adjuvant treatment (*n* = 14)‡**	**Ac time:** during COH (from day 2 of the study to the day before oocyte retrieval) **Frequency:** Twice a week for 2 weeks **Duration:**4 days	PI, CPR	IVF-ET
Dieterle et al. (42)	T:116 C:109	i)infertility	T: 35.1 ± 3.8 C:34.7 ± 4.0	**Acupuncture & auricular acupuncture (*n* = 116)** i)Acupuncture: [soreness, numbness, or distention around the point = Deqi sensation) occurred during the initial insertion.] ii) auricular acupuncture: [a special Chinese medical drug (the seed of Caryophyllaceae) was placed on the patient's ears. The seeds remained in place for 2 days and were pressed twice daily for 10 min. 3 days after ET, all patients received a second acupuncture treatment.]	**Placebo acupuncture & placebo auricular acupuncture (*n* = 109)** i)Placebo acupuncture: [the same needle reaction was utilized as the acupuncture group with acupoints designed not to influence fertility] ii) placebo auricular acupuncture: [the same needle reaction was utilized as the acupuncture group with acupoints designed not to influence fertility]	**Ac time:** after ET **Frequency:** immediately after ET and again 3 days later **Duration:** 2 days	EMT, CPR	IVF-ET
Zhong ** et al. (43)	T:51 C:52	i)age <35y ii)IVF-ET iii)normal BMI	T:31.37 ± 2.91 C:33.19 ± 2.57	**TEAS (*n* = 50)** g [TEAS (*HANS, Beijing, China*) stimulated via waves at 2 Hz frequency, 8–25 mA, which was the level of maximal tolerance without discomfort.]	**Mock TEAS (*n* = 50)††** [Ineffective electrical stimulation]	**Ac time:** during FET (from day 2 or 3 of each cycle to 1 day before ET) **Frequency:** every day **Duration**:1 menstrual cycle	EMT, EMP, RI, PI, EMB, S/D	FET

(Continued)

TABLE 1 | Continued

References	*N**	Study population	Mean age	Acupuncture group	Control group†	Starting time of intervention§ And the duration of intervention	Outcomes	Routine treatment
Wang et al. (41)	T:30 C:30	i)infertility ii)RIF	T:35.0 ± 3.71 C:34.7 ± 3.18	**Acupuncture& Moxibustion (*n* = 30)** i) Acupuncture: [Needles manipulated to obtain de qi for 30min] ii)Moxibustion: [Burned moxa stuck at the top of needles]	**Sham acupuncture& Placebo moxibustion (*n* = 30)** i)Sham acupuncture: [Acupoint pressing stimulated epidermal irritation slightly without de qi.] ii)Placebo moxibustion: [Electromagnetic waves bake electric lamp produced the sensation of heat only.]	**Ac time:** three courses before FET (one course: 10 days before the period starting treatment to the day before the next menstruation) **Frequency:** not mentioned **Duration**:3 menstrual cycles	EMP, RI, PI, CPR	FET
Shuai et al. (14)	T:34 C:34	i)age: 25–40 y ii)regular menstrual cycles iii)RIF	T: 29.47 ± 3.24 C: 29.65 ± 2.60	**TEAS (*n* = 34):** [TEAS (*LH202H HANS, Huawei Co Ltd, Beijing, China*) using dispersed-dense waves at 2 Hz frequency. The intensity was set to approximately 10–20 mA]	**Mock TEAS (*n* = 34)** [TEAS electrodes applied to the same sites and received intermittent 2 Hz (10 s on and 20 s off) TEAS at an intensity of 5 mA.] ‖	**Ac time:** three menstrual cycles before the scheduled FET **Frequency:** six times per cycle **Duration:** 3 cycles (18 treatment sessions in all).	EMT, EMP, CPR, LBR	FET
Zhao et al. (39)	T:38 C:34	i)age <38y ii)RIF	T:32.57 ± 4.25 C:33.71 ± 4.22	**Electro-acupuncture (*n* = 38):** [Needles manipulated to obtain de qi and connected electric-wire, the density wave (2/80 Hz) was set to the maximum comfort intensity that the patient could tolerate]	**Shallow acupuncture (*n* = 34):** [Shallow needles (depth <5 mm) placed on the non-acupoints, and not manipulated to achieve "de qi" sensation. Electrode wires connected with the shallow needles without electrical current]	**Ac time:** during FET (from day 2 or 3 of each cycle to 1 day before ET) **Frequency:** every other day **Duration:**1 menstrual cycle	EMT, EMV, FI, RI, VFI, CPR	FET
Zhuang (20)	T:36 C:36	i)age:24–45 ii)RIF	T: 34.2 ± 5. 31 C: 34.26 ± 5.30	**Acupuncture & Moxibustion & Cupping (*n* = 34)** ‖ i)Acupuncture: [Needles manipulated to obtain de qi for 30min] ii)Moxibustion: [Burned moxa stuck at the top of needles] iii)Flash Cupping: [Flash cupping stimulated at abdominal acupoints until flushing skin during the follicular phase and luteal phase.]	**No adjuvant treatment (*n* = 35)** ‖	**Ac time:** three menstrual cycles before the scheduled FET **Frequency: three** times a week **Duration:**3 menstrual cycles	EMT, EMB, RI, PI, CPR, LBR	FET
Ma and Zhang (36)	T:35 C:35	i)age:25–40y ii)RIF iii)BMI: 18.5~23.9 kg /m^2 iv) more than two high-grade embryos remained	T:30.04 ± 2.98 C: 30.55 ± 3.71	**Acupuncture (*n* = 35)** [Needles stimulated manually by rotating, lifting, and thrusting to obtain de qi until the maximum comfort intensity that the patient could tolerate, different acupoints selected before and after embryo transfer.]	**No adjuvant treatment (*n* = 35)**	**Ac time:** during FET and immediately after ET **Frequency**: every other day **Duration:**1 menstrual cycle	EMT, RI, PI, CPR	FET

(Continued)

TABLE 1 | Continued

References	N*	Study population	Mean age	Acupuncture group	Control group†	Starting time of intervention§ And the duration of intervention	Outcomes	Routine treatment
Chen and Hau (33)	T:25 C:31	i) age:22–40y ii)RIF iii)EMP:C type	T:35.64 ± 3.73 C:36.0±2.98	**Acupuncture & Moxibustion (*n* = 25)** i)Acupuncture: [Needles manipulated to obtain de qi for 30 min] ii)Moxibustion: [Burned moxa stuck at the top of needles]	**No adjuvant treatment (*n* = 31)**	**Ac time:** three menstrual cycles **Frequency**: every day **Duration:**3 menstrual cycles	EMT, EMP, RI, PI, CPR	FET
So (16)	T:185 C:185	i) normal uterine cavity	A: 35.64±3.73 C: 36.0±2.98	**Acupuncture (*n* = 185)** [Needle reaction (soreness, numbness, or distension around the puncture sites or sometimes propagate along the corresponding meridians which termed the DeQi sensation) was elicited during the initial insertion]	**Placebo acupuncture (*n* = 185)** [The Streitberger's placebo needles were used, which was blunt. When it was pushed forward against, the skin, the needle slid into the handle and the whole needle appeared shortened.]	**Ac time:** around ET **Frequency:** 25 min before and after embryo transfer **Duration**:2 days	VI, FI, VFI, SEVI, SEVFI, SEFI, CPR, LBR	IVF-ET
Lin et al. (37)	T:35 C:35	i) age:20∼40y ii)PCOS (based on Rotterdam criteria, 2003)	T: 27.74 ± 3.07 C: 26.71 ± 2.98	**Electro-acupuncture & Moxibustion(n=32)** i)Electro-acupuncture: [Filiform needles were used at different acupoints with continuous wave, 2 Hz in frequency and tolerable current intensity based on different phase of menstrual cycle.] ii)Moxibustion: [ginger-isolated moxibustion: the moxa cone was placed on the ginger and ignited with red skin in the local area, once every two days, 30 min each time]	**No adjuvant treatment (n=35)**	**Ac time:** during COH **Frequency:** every two days **Duration**:3 menstrual cycles.	EMT, RI, PI, S/D, CPR	letrozole and HCG
Zhong et al. (38)	T:32 C:32	i)infertility ii)age:35–42y iii)underwent IVF-ET	T: 36.52 ± 2.11 C: 36.19 ± 1.95	**Electro-acupuncture (*n* = 30)‡‡** [Needles manipulated to obtain de qi and connected electrode wires, dilatational wave, maximum comfort intensity that the patient could tolerate for 30min, every other day.	**No adjuvant treatment (*n* = 31)‡‡**	**Ac time:** 1 menstrual cycle before the scheduled IVF-ET **Frequency: three** times a week **Duration:1** cycle	EMT, EMP, PI, RI, S/D, CPR	IVF-ET

**Number randomized; † For the control group trials, the procedure was given the same as the acupuncture group; § Based on the start time of the first acupuncture session: in menstruation or after menstruation; ¶ Navel is an acupoint of Ren meridian based on Traditional Chinese Medicine (TCM); ‡12 patients in the control group were excluded from the study: four of them failed ovarian stimulation and the other eight participants were declined to participate further after randomization; ** Three groups were set in the trial, TESA vs. Mock TESA vs. no adjuvant control group, we collected the first group and the second group in our study; †† One patient in acupuncture group quitted and two patients in control group fell out due to incomplete data; ‖ Two patients in the acupuncture group interrupted treatment because of out of work and one patient in the control group embryo transferred advanced; ‡‡ In acupuncture group, one patient dropped out and one was excluded due to insufficient data. In the control group, one patient dopped out.*

*T, the treatment group; C, the control group; COH, control ovarian hyperstimulation; Ac, acupuncture; EMT, endometrial thickness; EMP, endometrial pattern; S/D, peak systolic velocity/ end-diastolic blood velocity; EMB, endometrial blood; RI, resistive index; PI, pulse index; IVF-ET, in vitro fertilization and embryo transfer; ET, embryo transfer; CPR, clinical pregnancy rate; TEAS, transcutaneous electrical acupoint stimulation; RIF, repeated implantation failures; HCG, human chorionic gonadotropin; PCOS, polycystic ovarian syndrome, ‡§, ¶, ‖, **, ††, ‡‡.*

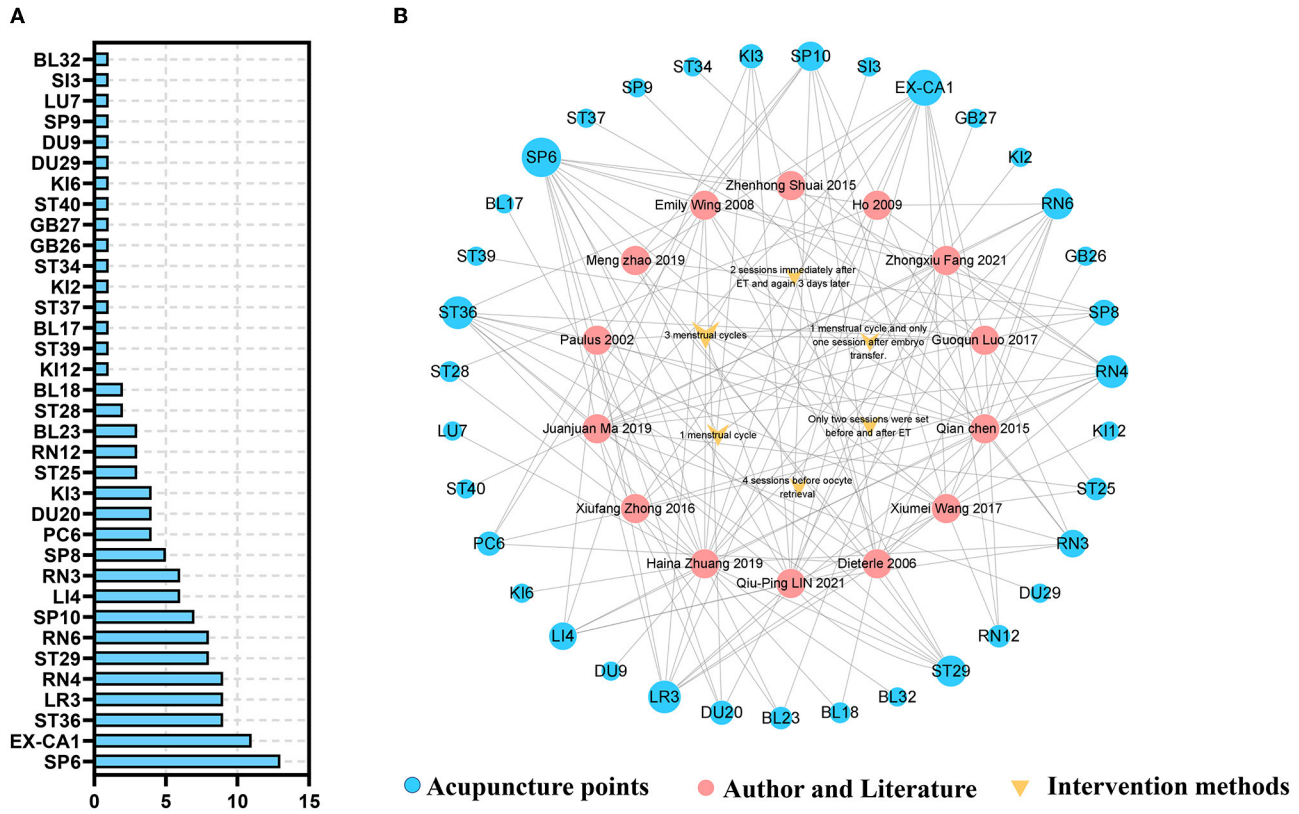

FIGURE 2 | The acupoints and acupuncture duration in included studies. **(A)** The frequency of acupoints in included studies. **(B)** The acupoints-associated and duration-associated network construction (The size of each circle or triangle represent the frequency in the studies).

TABLE 2 | Risk of bias summary of included trials.

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
So (16)	Low risk	Low risk	Low risk	Unclear risk	Low risk	Low risk	Low risk
Dieterle et al. (42)	Low risk	Low risk	Low risk	Unclear risk	Low risk	Low risk	Low risk
Ho et al. (35)	Unclear risk	Low risk	Low risk	Unclear risk	Low risk	Low risk	Low risk
Paulus et al. (13)	Low risk	Low risk	High risk	Unclear risk	Low risk	Low risk	Low risk
Chen and Hau (33)	Unclear risk	Low risk	High risk	Unclear risk	Low risk	Low risk	Unclear risk
Shuai et al. (14)	Low risk	Low risk	Low risk	Unclear risk	Low risk	Low risk	Low risk
Zhong et al. (43)	Low risk	Unclear risk	Low risk	Unclear risk	High risk‡	Low risk	Low risk
Wang et al. (41)	Unclear risk	Low risk	Low risk	Unclear risk	Low risk	Low risk	Unclear risk §
Luo et al. (34)	Low risk	Unclear risk	High risk	Unclear risk	Low risk	Low risk	Low risk
Zhao et al. (39)	Low risk	Unclear risk	High risk	Unclear risk	Low risk	Low risk	Unclear risk¶
Ma and Zhang (36)	Low risk	Unclear risk	High risk	Unclear risk	Low risk	Low risk	Unclear risk
Zhuang (20)	Low risk	Low risk	High risk	Low risk	High risk‡	Low risk	Low risk
Lin et al. (37)	Low risk	Unclear risk	High risk	Unclear risk	High risk‡	Low risk	Low risk
Zhong et al. (38)	Low risk	Unclear risk	High risk	Unclear risk	High risk‡	Low risk	Low risk

‡In these trials, one to three participants fell out due to interrupt treatment and did not finish the whole treatment, but the authors did not synthesis the intention to treat analysis; § This trail did not mention about the procedure of IVF-ET, and the measurement of endometrium thickness or endometrium pattern; ¶ In this trial, the measurements of ER were not mentioned. We considered the outcomes of ER, especially about the ultrasound procedure might cause measurement bias. But the outcome of CPR would not be expected to cause an important bias due to the measurements. Thus, we thought the bias is not clear.

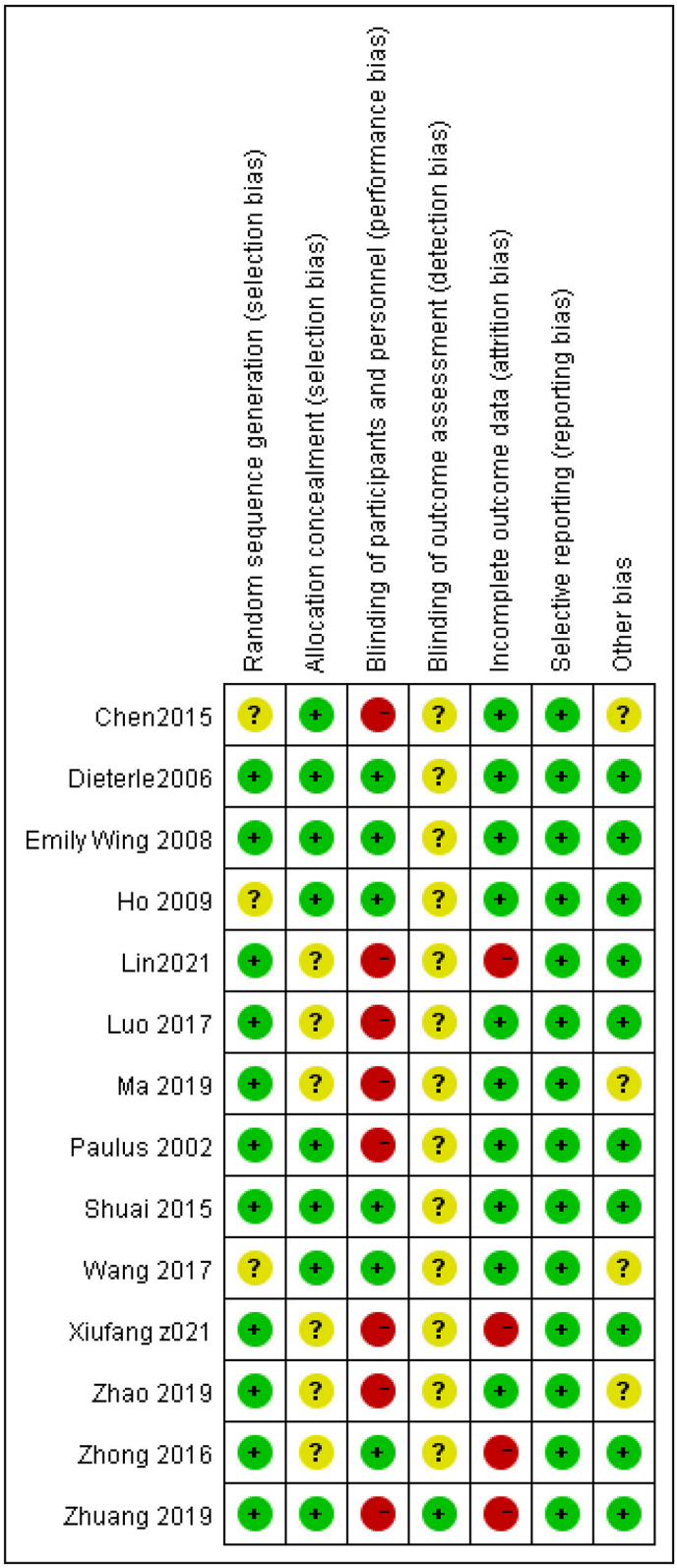

FIGURE 3 | Risk of bias summary of included trials.

TABLE 3 | Quality of evidence-based on GRADE.

Outcome indicators	No. of studies	No. of participants		Effect [95% CI]	I^2	Quality of evidence
		Intervention group	Control group			
CPR						
Low dosage group	4	411	388	1.42 [0.63, 3.20]	82%	⊕◯◯◯ Very low (b, d, e)
Moderate dosage group	4	153	150	2.00 [1.24, 3.22]	0%	⊕⊕◯◯ Low (a, e)
High dosage group	6	212	220	2.49 [1.67, 3.72]	0%	⊕◯◯◯ Low (a, e)
Total	**14**	**776**	**758**	**1.97 [1.97, 2.79]**	**56%**	⊕◯◯◯ Very low (a, b, e)
EMP						
Moderate dosage group	3	115	116	1.41 [0.25, 7.90]	85%	⊕◯◯◯ Very low (a, b, d)
High dosage group	6	212	220	3.25 [2.05, 5.15]	10%	⊕⊕◯◯ Low (a, e)
Total	**9**	**327**	**336**	**2.48 [2.26, 4.90]**	**71%**	⊕◯◯◯ Very low (a, b, e)
LBR						
Low dosage group	1	185	185	0.68 [0.44, 1.05]	-	⊕◯◯◯ Very low (d, e)
High dosage group	2	68	69	2.96 [1.42, 6.16]	0%	⊕⊕◯◯ Low (a, e)
Total	**3**	**253**	**254**	**1.68 [0.54, 5.27]**	**83%**	⊕◯◯◯ Very low (a, b, d, e)
EMT						
Low dosage group	2	196	189	−0.06 [−0.54, 0.41]	81%	⊕◯◯◯ Very low (b, d, e)
Moderate dosage group	4	153	150	0.75 [0.08, 1.42]	87%	⊕◯◯◯ Very low (a, b, e)
High dosage group	5	182	190	0.51 [0.04, 0.98]	80%	⊕◯◯◯ Very low (a, b, e)
Total	**11**	**531**	**529**	**0.48 [0.13, 0.83]**	**87%**	⊕⊕◯◯ Low (a, b)
RI						
Moderate dosage group	4	153	150	−0.74 [−1.44,−0.04]	88%	⊕◯◯◯ Very low (a, b, e)
High dosage group	5	185	188	−0.94 [−1.41, −0.48]	77%	⊕◯◯◯ Very low (a, b, e)
Total	**9**	**338**	**338**	**−0.86 [−1.23, −0.48]**	**82%**	⊕◯◯◯ Very low (a, b, e)
PI						
Low dosage group	3	140	108	−0.13 [−0.45, 0.20]	23%	⊕⊕◯◯ Low (d, e)
Moderate dosage group	3	115	116	−1.02 [−1.64, −0.40]	79%	⊕◯◯◯ Very low (a, b, e)

(Continued)

TABLE 3 | Continued

Outcome indicators	No. of studies	No. of participants		Effect [95% CI]	I^2	Quality of evidence
		Intervention group	Control group			
High dosage group	5	185	188	−2.35 [−3.59, −1.11]	96%	⊕○○○ Very low (a, b, e)
Total	**11**	**440**	**412**	**−1.33 [-1.93, −0.72]**	**93%**	⊕○○○ Very low (a, b, e)
S/D						
Moderate dosage group	2	80	81	−2.24 [−3.86, −0.62]	93%	⊕○○○ Very low (a, b, e)
High dosage group	2	89	90	−1.61 [−3.87, 0.65]	97%	⊕○○○ Very low (a, b, e)
Total	**4**	**169**	**171**	**−1.91 [−3.08, −0.75]**	**95%**	⊕○○○ Very low (a, b, e)

a: Download one level for serious risk of bias: failure to develop and apply appropriate eligibility criteria (inclusion of control population), flawed measurement of both exposure and outcome, failure to adequately control confounding, or Incomplete or inadequately short follow-up. b: Downgraded one level for serious inconsistent: inconsistency refers to an unexplained heterogeneity of results, which includes the wide variance of point estimates across studies, minimal or no overlap of confidence intervals (CI), and statistical criteria, including tests of heterogeneity which test the null hypothesis that all studies have the same underlying magnitude of effect, have a low p-value (p < 0.05), indicating to reject the null hypothesis. c: Downgraded one level for serious indirectness: including differences in the population (applicability), differences in interventions (applicability), differences in outcomes measures (surrogate outcomes), and indirect Comparisons. d: Downgraded one level for serious imprecision: dichotomous outcomes and continuous outcomes were considered separately, including that, optimal information size criterion was not med, or 95% CI overlaps no effect. e: Downgraded one level when publication bias is suspected.

was performed in low-, moderate-, or high-dosage groups. The qualities of the evidence are shown in **Table 3**.

Efficacy Analysis

The RCTs included in this study varied in study design, especially in ways of intervention of treatment group, such as, manual acupuncture (MA) alone (16, 36), electro-acupuncture (EA) alone (35, 38, 39), MA or EA with moxibustion (20, 33, 34, 37, 41), MA with auricular acupuncture (13, 42), and TEAS (14, 40). Meanwhile, the duration of acupuncture differed for trails, such as 2 or 4 days of acupuncture, one or three menstrual cycles. Finally, we categorized these trails according to the duration of acupuncture treatment in three groups: the high-dosage group [three menstrual cycles; (14, 20, 33, 34, 37, 41)], the moderate-dosage group [one menstrual cycle; (36, 38–40)], and the low-dosage group [2 or 4 days; (13, 16, 35, 42)].

Comparisons of Endometrial Receptivity Outcomes by the Duration of Acupuncture

Intervention With Low Dosage

For the primary outcome, CPR was from four trials [(13, 16, 35, 42); $n = 815$], but the statistical between the studies was found to be significant ($I^2 = 82\%$, $p = 0.007$). A random-effect model was performed, and no statistical significance was found ($p = 0.39$, OR = 1.42, 95%CI [0.63, 3.20] (**Figure 4A**).

For the secondary outcomes, EMT was available in two trials (43, 44; $n = 385$), and there was significant heterogeneity between these trials ($I^2 = 81\%$, $P = 0.02$). Pooling the results of these two trials into the random-effects showed no distinct EMT between the treatment group and the control group ($P = 0.79$, SMD = −0.06 95%CI [−0.54, 0.41] (**Figure 5A**). Only two trials (36, 44; $n = 160$) collected PI, and Ho et al. (35) measured the left and right uterine arteries that were separated into two groups for analysis. No statistical significance was performed ($p = 0.44$, SMD = −0.13, 95%CI [−0.45, 0.20] (**Figure 6A**), and mild heterogeneity was found between the studies ($I^2 = 23\%$, $p = 0.27$) in the random-effect model. LBR was from only one study (16), and the measurement of acupuncture with the control group did not show a statistically significant benefit ($P = 0.08$, OR = 0.68, 95%CI [0.44, 1.05] (**Figure 7A**). EMP and RI were not measured with 2 or 4 days of acupuncture.

Intervention With Moderate Dosage

Four studies [(36, 38–40); $n = 309$] conducted acupuncture treatment for one menstrual cycle, and CPR data were found in all of these trials. There was significantly higher CPR of acupuncture vs. the control group ($P = 0.004$, OR = 2.00, 95%CI [1.24, 3.22] without heterogeneity ($I^2 = 0\%$, $P = 0.74$; **Figure 4B**).

There were also significant benefits in EMT, PI, RI, S/D. A total of four trials [(36, 38, 43, 44); $n = 309$] included EMT, the acupuncture treatment significantly increased the endometrial thickness ($p = 0.03$, SMD = 0.75, 95%CI [0.08, 1.42]) with significant heterogeneity ($I^2 = 87\%$, $p < 0.01$) (**Figure 5B**); three trails(37, 39, 41; $n = 237$) included PI, acupuncture can significantly decreased PI ($p = 0.01$, SMD = −1.02, 95%CI [−1.64, −0.40] with high heterogeneity ($I^2 = 79\%$, $p = 0.008$) (**Figure 6B**); four trails (37, 39–41; $n = 309$) conducted RI, acupuncture can significantly reduce the RI ($p = 0.04$, SMD = −0.74, 95%CI [−1.44, −0.04] with high heterogeneity ($I^2 = 88\%$, $p < 0.01$) (**Figure 8A**); and two trails (39, 41; $n = 167$) conducted S/D, acupuncture can significantly reduce S/D ($P = 0.007$, SMD = −2.24, 95%CI [−3.86, −0.62] with significant heterogeneity ($I^2 = 79\%$, $P < 0.01$) (**Figure 9A**). No significant difference was found in EMP [three trails (36, 38, 40), $n = 237$, P

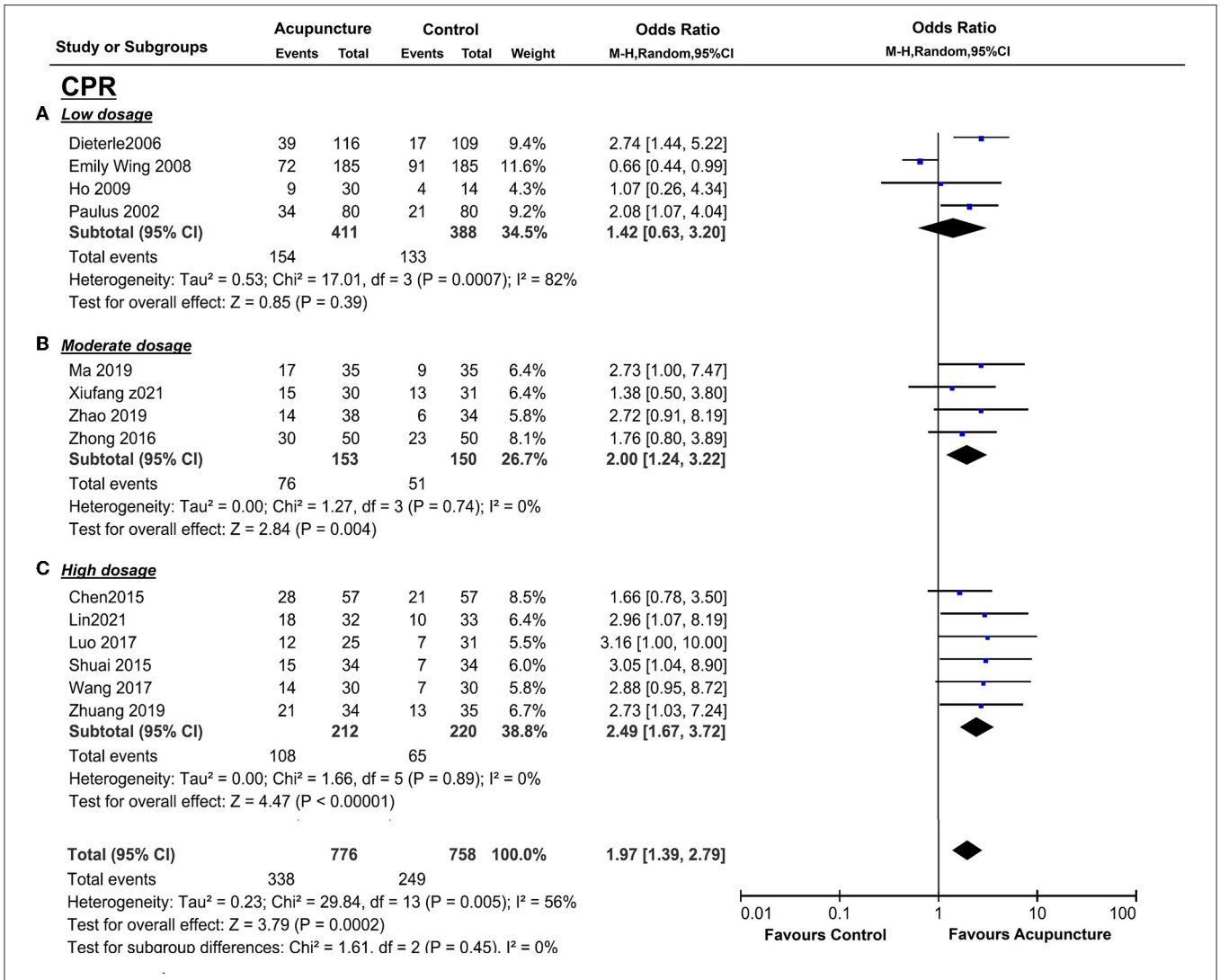

FIGURE 4 | Effects of acupuncture on CPR based on intervention of low dosage **(A)**, moderate dosage **(B)**, and high dosage **(C)**.

= 0.070, OR = 1.41, 95%CI [0.25, 7.9]; **Figure 10A**]. LBR was not measured with acupuncture treatment for one menstrual cycle before embryo transfer.

Intervention With High Dosage

The pooled indicators of CRR, EMT, PI, EMP, RI, and LBR showed significant differences between the acupuncture group and the control group, while S/D showed non-significant differences. A total of six trails [(14, 20, 33, 34, 37, 41); $n = 440$] evaluated the CPR, acupuncture can significantly improve CPR (P<0.05, OR=2.49, 95%CI [1.67, 3.72] without heterogeneity (I^2 = 0 %, $P = 0.89$) (**Figure 4C**). For ER parameters, EMT was significantly thicker in the acupuncture group from five trials (14, 20, 33, 34, 37) ($n = 380$, $P = 0.03$, SMD = 0.51, 95%CI [0.04, 0.98] with high heterogeneity ($I^2 = 80\%$, $P < 0.01$; **Figure 5C**); PI: four trails (20, 33, 37, 41), $n = 316$, $P < 0.05$, SMD = −2.35, 95%CI [− 3.59, −1.11] with high heterogeneity ($I^2 = 96\%$, $P <$ 0.01; **Figure 6C**); RI: four trails (20, 33, 37, 41), $n = 316$, $P < 0.05$, SMD = −0.94,95%CI [−1.41, −0.48] with high heterogeneity (I^2 = 77%, $P < 0.01$; **Figure 8B**); EMP was significantly improved in 6 trails (14, 20, 33, 34, 37, 41), $n = 440$, $P < 0.05$, OR = 3.25, 95%CI [2.05, 5.15] with small heterogeneity ($I^2 = 10\%$, $P = 0.35$) (**Figure 10B**). For LBR, it was significantly increased in two trails (14, 20), $n = 140$, $P = 0.004$, OR-2.96, 95%CI [1.42, 6.16]) with non-significant heterogeneity ($I^2 = 0\%$, $P =$ 0.82) (**Figure 7B**). However, there was no significant difference between the acupuncture treatment group and the control group in S/D [two trails (33, 37), $n = 184$, $P = 0.16$, SMD = −1.61, 95%CI (−3.78, 0.65)] with high heterogeneity ($I^2 = 97\%$, $p <$ 0.01; **Figure 9B**).

Subgroup Analysis and Investigation of Heterogeneity

However, the duration of acupuncture subgroup analysis did not significantly explain this statistical heterogeneity [interaction p

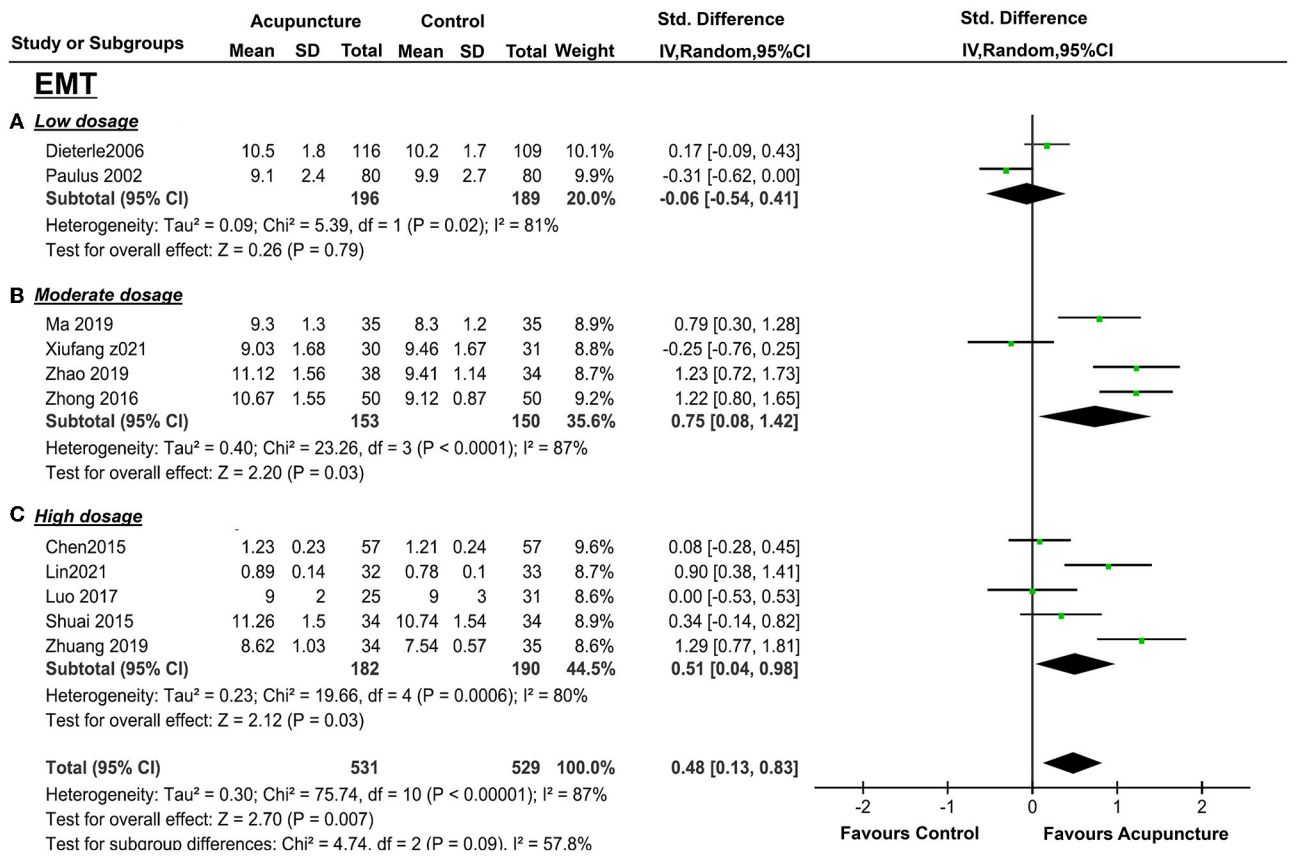

FIGURE 5 | Effects of acupuncture on EMT based on intervention of low dosage **(A)**, moderate dosage **(B)**, and high dosage **(C)**.

= 0.53 for clinical pregnancy outcome and interaction $p = 0.09$ for endometrium thickness, respectively (**Figure 4**)]. We used subgroup analysis to find out the heterogeneities of the outcomes. With a wide range of study designs, we chose seven variables to assess heterogeneity, according to four clinical characteristics and three methodological characteristics. The heterogeneity analysis was only conducted in EMT, EMP, RI, and PI due to the amounts of included studies.

Generally, 4 subgroup variables (routine treatment, risk of performance bias, duration of acupuncture treatment, and the age of participants) could explain some of the heterogeneities across all trials (**Figure 11**). Routine treatment was significantly different with subgroup analysis in EMT, RI, PI (EMT: interaction $p = 0.003$; RI: interaction $p = 0.03$; PI: interaction $p < 0.05$, respectively), risk of performance bias (blinding of participants and personnel) was significantly different with subgroup analysis in RI (interaction $P = 0.003$). Subgroup analysis in PI was significantly different when restricted in the duration of acupuncture treatment ($P = 0.004$), and EMT was almost statistically significant when restricted in the age of participants (interaction P=0.05). That is, heterogeneity is reduced when restricted to variables (i.e., the maximum I^2 percentage change of the subgroup was 43%).

Furthermore, considering the controlled intervention of different comparisons is the heterogeneity, and we made a further subgroup analysis in different dosage groups according to the different way of controlled group: no intervention or placebo treatment. As shown in **Supplementary Material S3**, though we found converse results in the re-subgroup analysis, the heterogeneities remained high among the groups.

Publication Bias Analysis

The funnel plot (see **Figure 12**) of the primary outcome CPR showed a small positive studies effect, with the smaller studies showing more benefit being conducted than larger studies, and also positive results were easily published by researchers. The Eggers'test $p = 0.003$, which means there is public bias in the collected studies.

DISCUSSIONS

Summary of the Main Results

To the best of our knowledge, this is the first comprehensive review addressing the pooled analysis of whether dose-related acupuncture therapeutic efficacy improves ER for infertile women. Finally, we collected 14 trials in our review, including four (36–39) updated relevant studies.

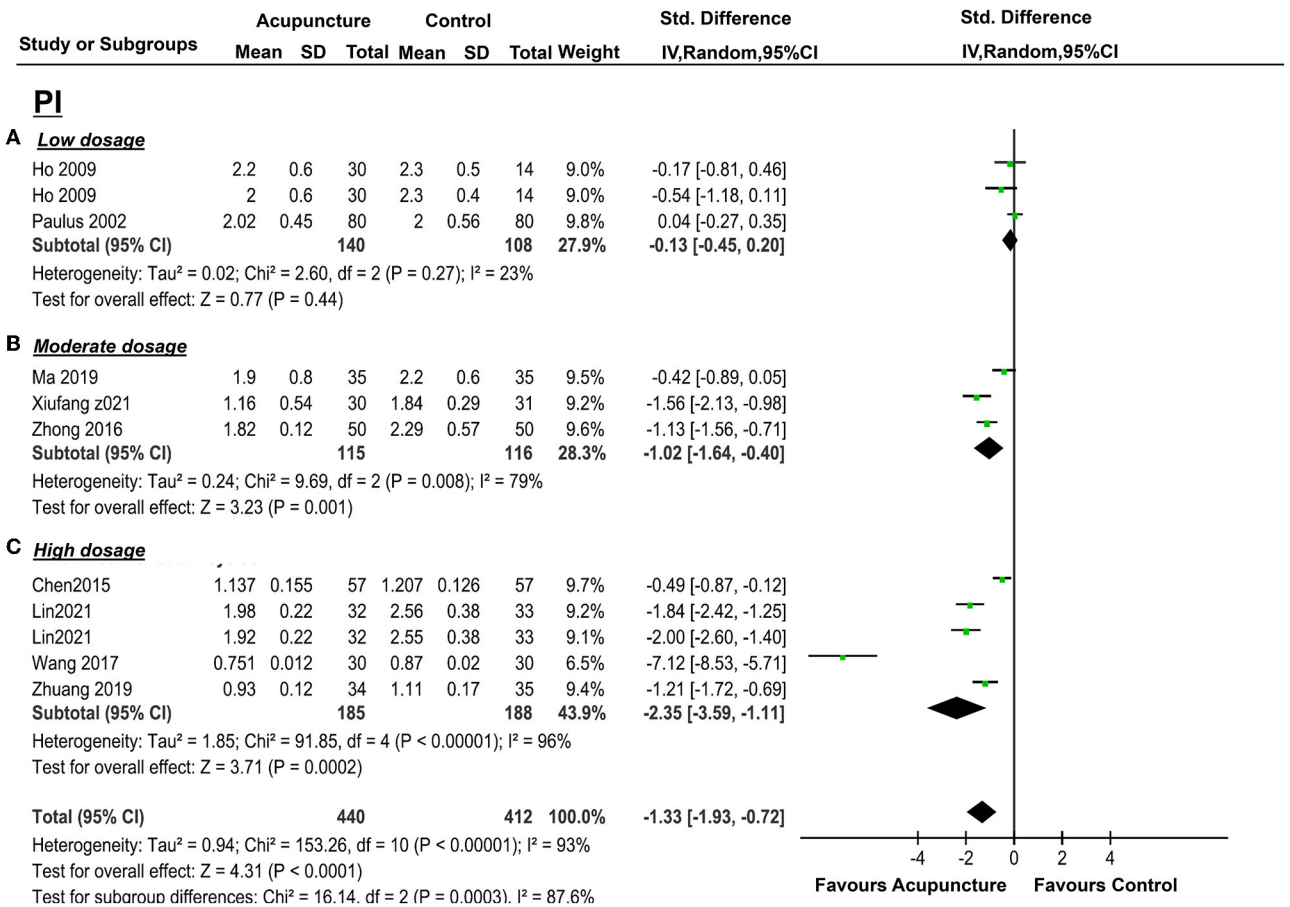

FIGURE 6 | Effects of acupuncture on PI based on intervention of low dosage **(A)**, moderate dosage **(B)**, and high dosage **(C)**.

Generally, acupuncture can significantly improve the CPR and the ER parameters (EMT, EMP, RI, PI, and S/D) with moderate or high heterogeneity (I^2: 56–95%), whereas the LBR was not statistically significant between the two groups with high heterogeneity ($I^2 = 83\%$), also. For the primary outcome, CPR was significantly increased in the moderate and high acupuncture groups without heterogeneity, whereas in the low-dosage group, CPR failed to show a significant improvement with considerable heterogeneity ($I^2 = 82\%$) (**Table 3**). For the ER parameters, EMP was significantly improved in the high-dosage group with a small heterogeneity ($I^2 = 10\%$) and was not statistically significant in the moderate group with considerable heterogeneity ($I^2 = 85\%$) (**Table 3**). Though the total heterogeneity of EMP was moderate ($I^2 = 71\%$), we failed to find the sources of the heterogeneity (**Figure 11A**). The decrease of PI showed a significant difference in the high-dosage and moderate group between the acupuncture group vs. the controlled one with considerable heterogeneity ($I^2 = 96\%$ and $I^2 = 79\%$, respectively), whereas, in the low-dosage group, no significant difference was found with considerable heterogeneity (**Table 3**). The total heterogeneity of PI was high ($I^2 = 93\%$) in our analysis, and not only the duration of acupuncture but also the different routine treatments for infertile women were attributed to the source of heterogeneity (interaction $P < 0.01$, < 0.01, respectively) (**Figure 11B**). EMT was significantly improved in the moderate- and high-dosage groups with significant heterogeneity ($I^2 = 87\%$, $I^2 = 80\%$, respectively), and was not statistically different in the low-dosage group with considerate heterogeneity ($I^2 = 81\%$). The total heterogeneity of EMT was high ($I^2 = 87\%$) too, and we found out that the different routine treatment was attributed to the source of heterogeneity (interaction $P = 0.003$). The heterogeneity was decreased and was 65% in the IVF-ET group and 84% in the FET group (**Figure 11C**). Though the significant decreases of RI were shown in the moderate- and high-dosage acupuncture groups, the heterogeneities were significant too ($I^2 = 88\%$, $I^2 = 77\%$, respectively) (**Table 3**). In our subgroup analysis, the routine treatment and the method of the compared intervention (placebo acupuncture was utilized or not) may be the sources of heterogeneity (**Figure 11D**). LBR was statistically improved in the high-dosage group while it was non-significant in the low-dosage group. S/D was significantly decreased in both moderate and high-dosage groups. However, the indicators of S/D and LBR showed high heterogeneities in the studies ($I^2 = 95$, $I^2 = 83\%$, respectively) (**Table 3**). Considering the small

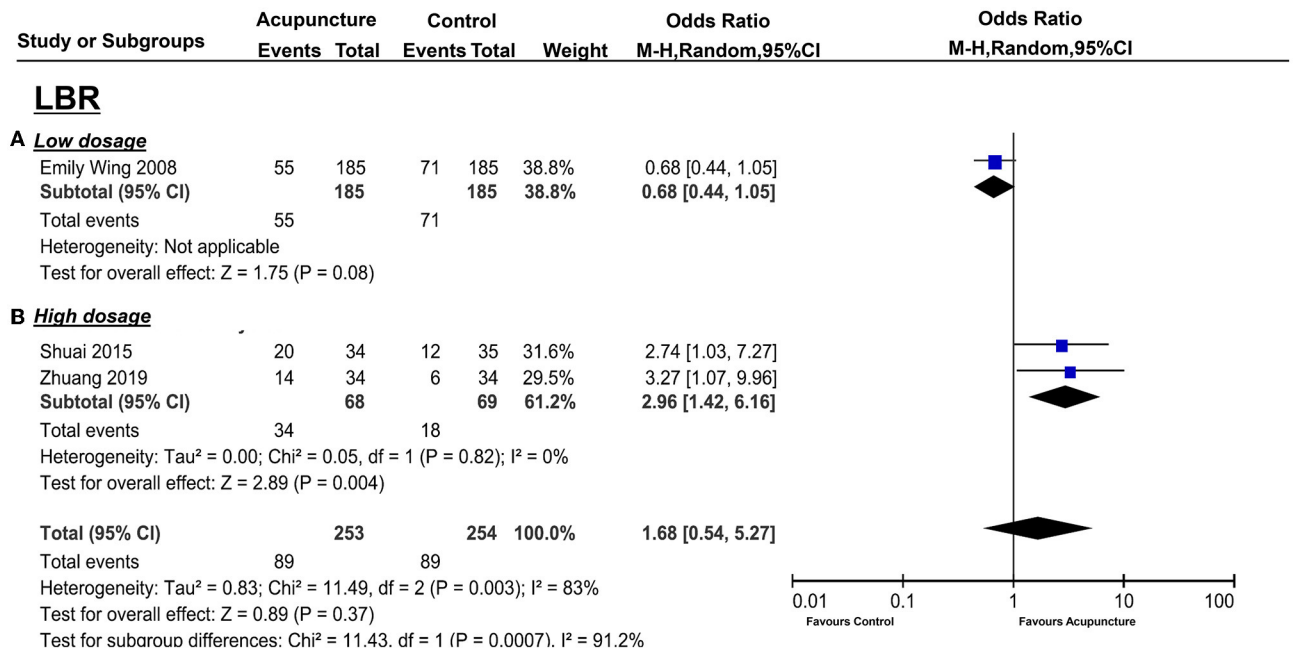

FIGURE 7 | Effects of acupuncture on LBR based on intervention of low dosage **(A)** and high dosage **(B)**.

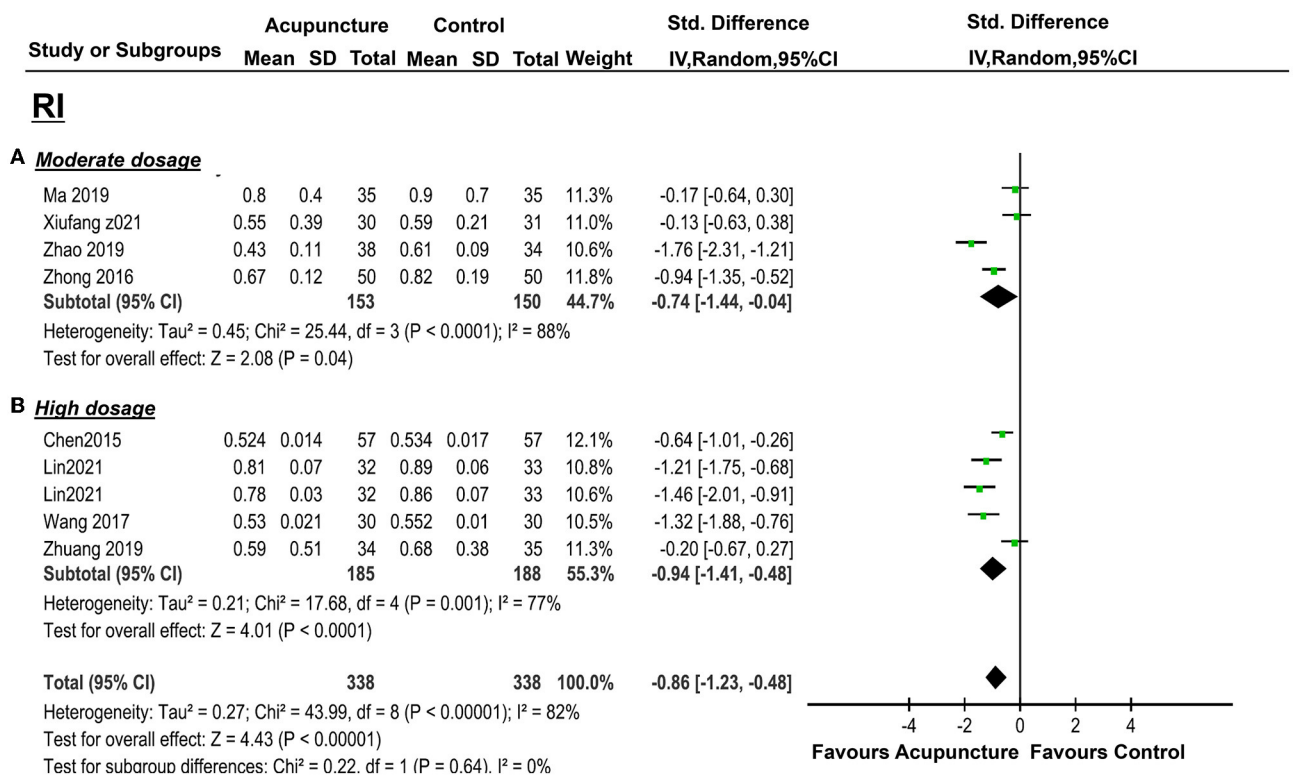

FIGURE 8 | Effects of acupuncture on RI based on intervention of moderate dosage **(A)** and high dosage **(B)**.

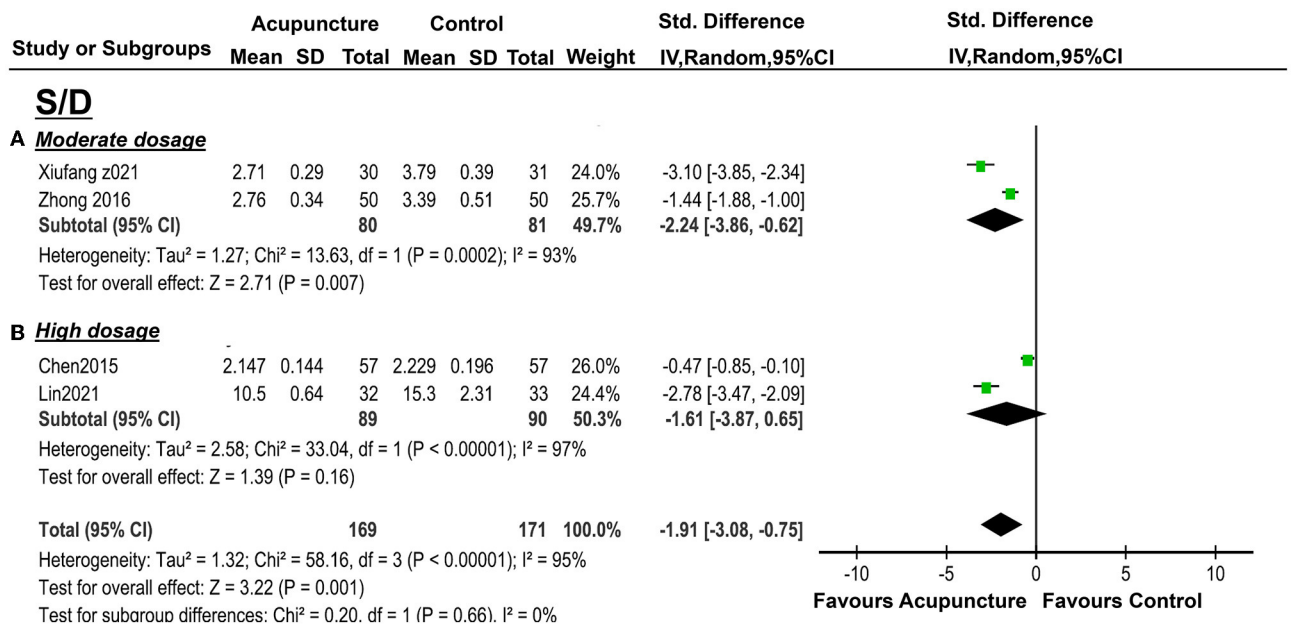

FIGURE 9 | Effects of acupuncture on S/D based on intervention of moderate dosage **(A)** and high dosage **(B)**.

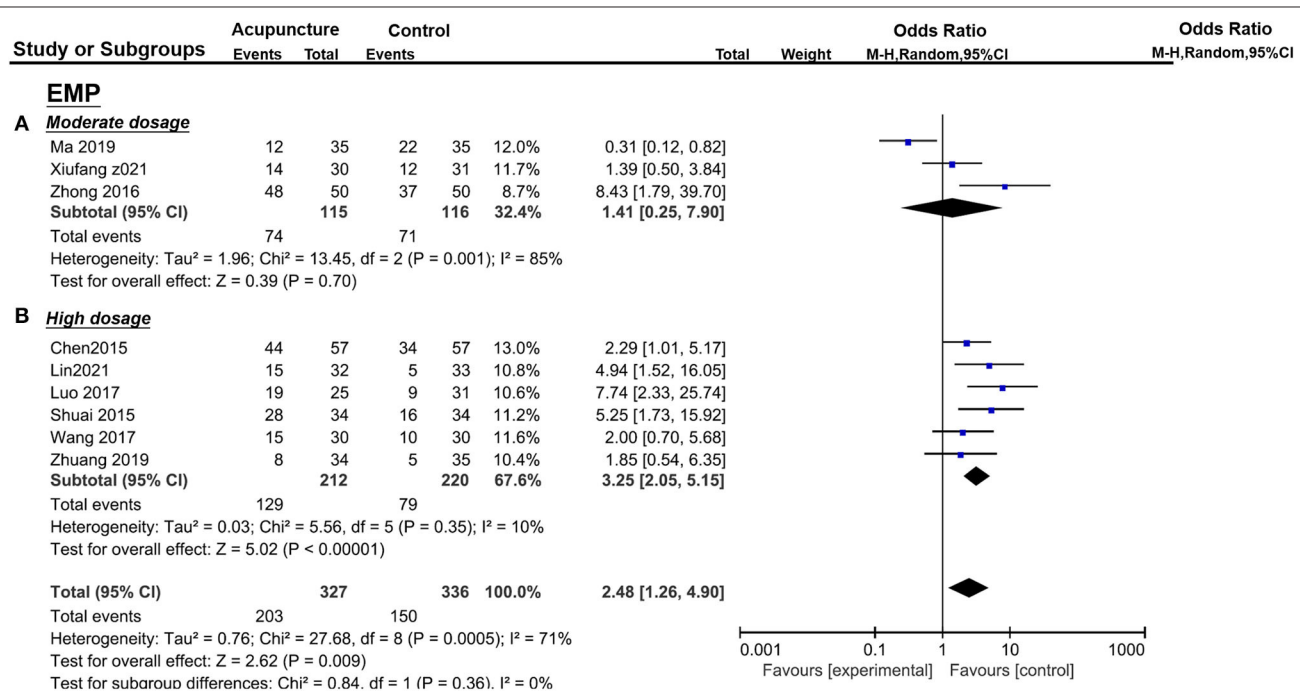

FIGURE 10 | Effects of acupuncture on EMP based on intervention of moderate dosage **(A)** and high dosage **(B)**.

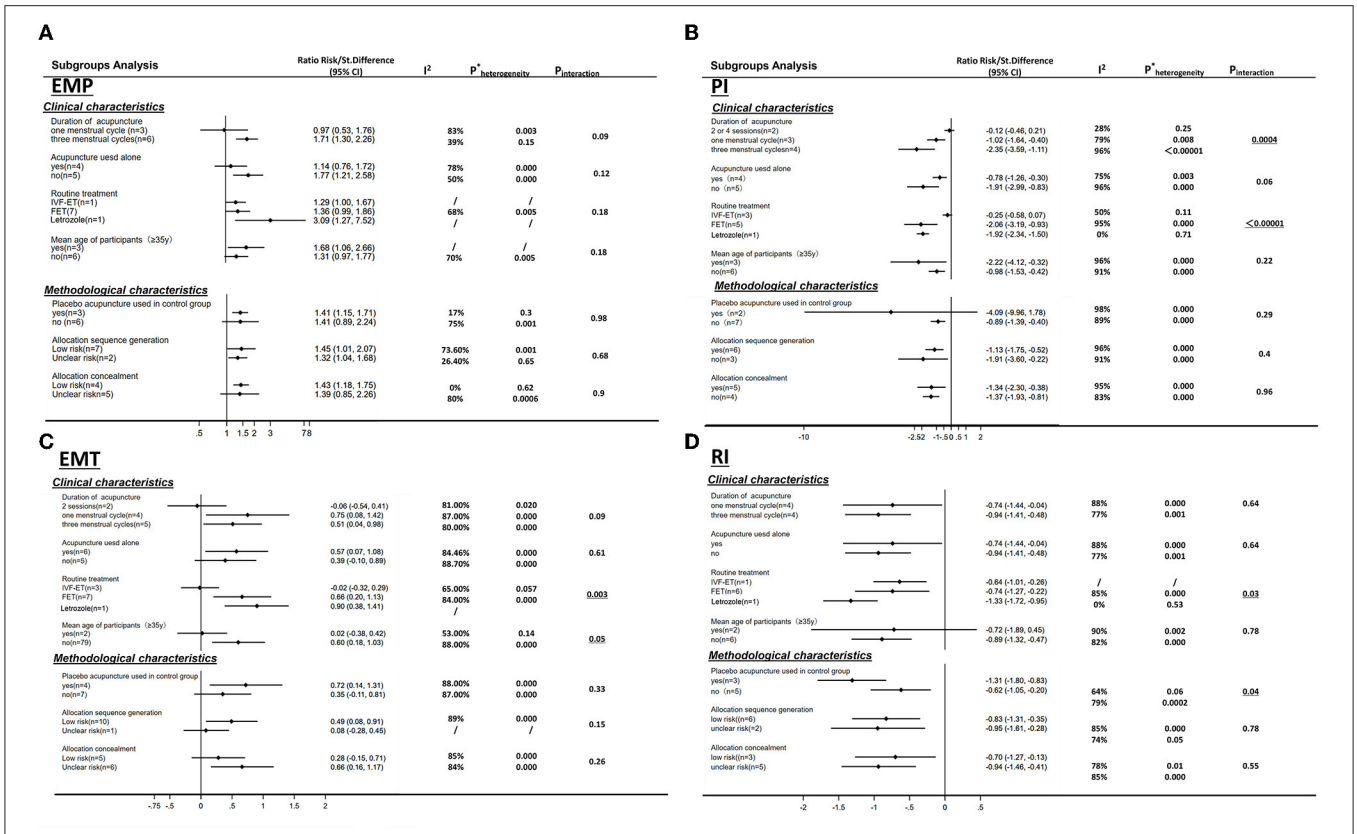

FIGURE 11 | Heterogeneity analysis of four endometrial receptivity characteristics by seven variables. **(A)** endometrial pattern; **(B)** pulse index; **(C)** endometrial thickness; **(D)** resistance index. Seven variables including four clinical characteristics. Acupuncture duration, intervention of different routine treatment (IVF-ET, FET or clomiphene for ovulation induction), acupuncture used alone or with other interventions and mean age of participants (≥35y or not) and three methodological characteristics blinding of participants (sham/placebo or no adjuvant treatment); random sequence generation (adequate or not); allocation concealment (adequate or not) (*:0.000 of $P_{heterogeneity}$ means <0.00001).

numbers of included studies that measured the LBR and the S/D, the heterogeneity analysis was not performed. Above all, comparisons were evaluated as a very low to moderate level of evidence, even very low and low, mostly based on the GRADE (**Table 3**).

Thus, the therapeutic effect of acupuncture in different dosage groups is relatively weak so far. Though it is difficult to draw a definitive conclusion that acupuncture is more effective, we discussed the findings based on our subgroup analysis under the method of study design.

Analyses Based on Acupuncture Dosage

Above all, the evaluation among the outcomes from the included studies and acupuncture dosage resulted in different therapeutic effects. Though significant heterogeneities were not found, the trend of relatively more dosage showed better effects in CPR, EMT, PI (moderate or high dosage showed significantly improved while the low dosage did not), and EMP (high dosage was significantly improved while the moderate did not; **Table 3**).

Our findings of the relationship between acupuncture therapeutic efficacy and the dosage were consistent with the previous studies. White et al.(45) thought the dose may be affected by the state of the patient (e.g., nervous, immune, and endocrine systems); different doses may be required for different conditions. For the analgesia of acupuncture, a system review (46) evaluated whether the effect of acupuncture is dose-related for symptom management in knee osteoarthritis. Based on the consideration of the number of points needled during each treatment; de qi response; frequency of treatment per week; and several treatment sessions, the study found out that the effect may be associated with the dose of acupuncture, with a higher dosage (the number of points needled ≥9; de qi response; frequency ≥2 sessions a week; the total number of treatment sessions ≥8) related to better treatment outcomes in terms of relief of pain and dysfunction in patients with knee osteoarthritis. Gao (47) utilized functional MRI (fMRI) in evaluating acupuncture efficacy in treating migraine. Four weeks of acupuncture treatments strengthened the brain connection to improve therapeutic efficacy in migraine when compared with 2 weeks. For the acupuncture modulation in default mode network (DMN) of the brain with fMRI, Lin et al. (48) suggested that enhancing the acupuncture dose could potentially be applied as a means of modulating brain activity. For improving the outcome of women who underwent IVF, Magarelli et al.(12) suggested that

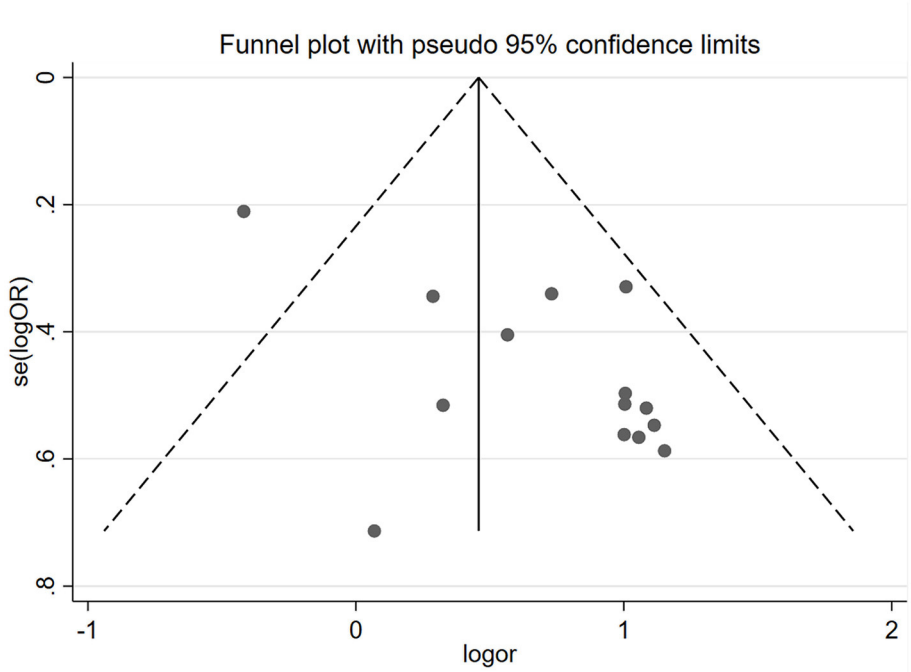

FIGURE 12 | Funnel plot of trials meeting inclusion criteria (CPR). The Egger'test $P = 0.003$.

patients receiving acupuncture weeks before the stimulation of IVF (usually 4 weeks at two treatments per week until the day of retrieval) and in the times before egg retrieval and ET may benefit more than those patients treated just pre-/post-ET. In a systematic and meta-review, Zheng et al. (49) reported that the pooled biochemical pregnancy rate (BPR) was significantly higher in the acupuncture group than in the controls (OR = 2.07, $P = 0.02$). The CPR, LBR, and IR results tended to be higher when acupuncture treatment was conducted during the COH time (from the starting time of the menstrual cycle to the oocyte retrieval). Whereas, the pooled BPR, CPR, OPR, LBR, IR, and MR showed no significant differences when acupuncture was performed only around the day of ET (i.e., CPR: $p = 0.48$).

Infertility for women with PER is a complex medical issue with multiple factors, often associated with previous gynecological issues/pathology (e.g., fibroids, endometriosis, autoimmune disease, etc.). Although patients may benefit from two sessions of acupuncture treatment (13), many studies (15–17, 33) failed to repeat this outcome. Thus, chronic conditions of long-term duration need larger dosages of acupuncture, especially in patients with multiple health issues.

Why S/D is lower in three-cycle treatments than in one cycle? It was found that there were more stressful and uncomfortable feelings during the IVF process in the acupuncture group (50). Also, the S/D is a parameter of the uterine artery. The S/D of women will change under too much stress and discomfort. Besides, how many acupuncture sessions should be performed for participants who suffered PER is necessary to reconsider because of the financial burden and time cost for PER (51).

Analyses Based on the Cause of Heterogeneity

Heterogeneity was expected due to the variables of the routine treatment, acupuncture treatment duration, performance bias (blinding of participants and personnel), and the age of participants.

For routine treatment for infertile women, to the best of our knowledge, IVF-ET is a procedure in which fresh embryos would be transferred after ovarian stimulation, while FET is frozen embryo transfer. Ovarian stimulation for IVF would affect the luteal phase and the window of implantation (WOI) by altering the endocrinological environment of the endometrium. That is, ovarian stimulation will reduce ER in IVF-ET (52), whereas FET in subsequent cycles avoids possible adverse effects without ovarian stimulation. For women with PCOS, LE has been used as the traditional medication to induce ovulation for being superior to CC in ER (53). But it is uncertain whether LE for PCOS will improve the EMT compared to the programmed COH cycle (54).

For blinding of participants and personnel, placebo acupuncture treatment is not inert in numerous studies (55), and the possibility that some of the sham interventions may have acupuncture-specific efficacy (56, 57). Even in Madsen's review (58), a greater effect of acupuncture with penetrative placebo needles compared with the non-penetrative placebo needles was found ($P = 0.04$). Furthermore, the device or method of different placebo interventions is possible resource of heterogeneity. To the best of our knowledge, four different types of placebo acupuncture: i) placebo acupuncture device (with a blunt needle tip, pressing the skin without penetrating),

including Streitberger's needle [(15, 16, 59); the needle tip is blunt but a pricking sensation is felt by the patient, simulating the puncturing of the skin] and Park needle [(60, 61); this device has a retractable needle shaft and a blunt tip. Acupuncturists were instructed to lightly place the sham needle on the surface of the skin with no manipulation of the needle to minimize any physiological effect]; ii)shallow acupuncture (39), which stimulates at the real acupoints (sham acupuncture was standardized as minimal-depth needling without stimulation "DeQi"); iii) non-acupoints (17) or non-therapeutic acupoints (42) utilization in the control group, and iv)Mock electrical stimulation [(14, 62); the technique has been shown to constitute a successful placebo treatment in functional brain networks (61). However, the difference of placebo efficacy among the above placebo intervention was not found in an accepted way (63). Thus, the potential heterogeneity may be also related to the placebo intervention.

For the age of participants, one of the major factors related to the decline in fertility with age is the aging uterus. However, the underlying factors that decrease endometrial receptivity in older women are still unclear.

Though different interventions in the treatment group (acupuncture was used alone or not) showed no significant heterogeneity, the previous meta-analysis showed the intervention in the treatment group accounts for some heterogeneity (64). Further analysis should be conducted. Furthermore, other potential variables, such as the way of measurement of three-dimensional ultrasound, may cause heterogeneity and care should be taken when the therapeutic efficacy of acupuncture treatment is evaluated.

Research Strengths

To reduce the risk of bias, we only included the studies of randomized trials. We assessed the GRADE score to determine the strength of evidence. We excluded the RCTs with the outcome of EMT only for a more reliable result. To strengthen the efficacy of our findings, we explored the cause of heterogeneity associated with the clinical characteristics and methodological characteristics.

Research Limitations

The review included moderate or low-quality studies. In our review, the duration of acupuncture was assessed. However, based on the study design of acupuncture trial guideline, Standards for Reporting Interventions in Controlled Trials of Acupuncture (STRICTA), and the frequency and starting time of acupuncture should be also considered in acupuncture dosage evaluation. Also, the variables subgroup analysis did not decrease the heterogeneity, which indicated that there were still potential factors in the collected studies. Therefore, a standardized individualized acupuncture program with a homogeneity trial design needed to be performed to evaluate the therapeutic efficacy of acupuncture, such as enough fixed acupuncture treatment dosage. The latest RCTs or negative unpublished trials were unlikely to be included in our analysis, so there was some publication bias.

CONCLUSION

Overall, the trend of relatively higher acupuncture dosage showed better effects for poor endometrial receptivity, though, it is difficult to draw a definitive conclusion with the considered heterogeneities among the outcomes and insufficiently high quality of the evidence. More high-quality studies with a homogeneity trial design need to be conducted to explore the efficacy of acupuncture for improving PER for infertile women. Also, there should be more consensus and dimensionality in the assessment of the measurement of the three-dimensional ultrasound.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

XZ conceived and designed the study and wrote the paper. HaY and LL contributed to the literature search and extracted data and conducted the format and tables. XZ and HoY performed the statistical analysis. FW and SY participated in critical dialogue and revised the manuscript. Moreover, all authors have approved the final manuscript for submission.

FUNDING

This study was funded by a program of the Science and Technology foundation of Sichuan Province (grant number 2020JDJQ0051).

ACKNOWLEDGMENTS

We gratefully thank Hang Zhou Ph.D. (School of Basic Medical Sciences, Chengdu University of Traditional Chinese Medicine) for assistance with the evidence-based methodological suggestion.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2022.858587/full#supplementary-material

REFERENCES

1. Harper JC, Aittomäki K, Borry P, Cornel MC, De Wert G, Dondorp W, et al. Recent developments in genetics and medically assisted reproduction: from research to clinical applications. *Europ J Human Gen: EJHG*. (2018) 26:12–33. doi: 10.1038/s41431-017-0016-z
2. Roussos-Ross D, Rhoton-Vlasak AS, Baker KM, Arkerson BJ, Graham G. Case-based care for pre-existing or new-onset mood disorders in patients undergoing infertility therapy. *J Assisted Reprod Gen*. (2018) 35:1371–6. doi: 10.1007/s10815-018-1222-5
3. Kazemijaliseh H, Ramezani Tehrani F, Behboudi-Gandevani S, Hosseinpanah F, Khalili D, Azizi F. The prevalence and causes of primary infertility in Iran: a population-based study. *Glob J Health Sci*. (2015) 7:226–32. doi: 10.5539/gjhs.v7n6p226
4. Diedrich K, Fauser BC, Devroey P, Griesinger G, Evian Annual Reproduction (EVAR) Workshop Group. The role of the endometrium and embryo in human implantation. Hum Reprod Update. (2007) 13:365–77. doi: 10.1093/humupd/dmm011
5. Ledee-Bataille N, Olivennes F, Frydman R, Taupin JL, Kadoch J, Chaouat G. Concentration of leukemia inhibitory factor (LIF) in uterine flushing fluid is highly predictive of embryo implantation. *Fertil Steril*. (2002) 78:S110-S. doi: 10.1016/S0015-0282(02)03669-5
6. Franasiak JM, Forman EJ, Hong KH, Werner MD, Upham KM, Treff NR, et al. The nature of aneuploidy with increasing age of the female partner: a review of 15,169 consecutive trophectoderm biopsies evaluated with comprehensive chromosomal screening. *Fertil Steril*. (2014) 101:656–63. doi: 10.1016/j.fertnstert.2013.11.004
7. Ferin M, Wiele RV. Endogenous opioid peptides and the control of the menstrual cycle. *Europ J Obstetrics, Gynecol, Reprodu Biol*. (1984) 18:365–73. doi: 10.1016/0028-2243(84)90059-5
8. Stener-Victorin E, Waldenstrom U, Andersson SA, Wikland M. Reduction of blood flow impedance in the uterine arteries of infertile women with electro-acupuncture. *Hum Reprod*. (1996) 11:1314–7. doi: 10.1093/oxfordjournals.humrep.a019378
9. Stener-Victorin E, Lundeberg T, Cajander S, Aloe L, Manni L, Waldenström U, et al. Steroid-induced polycystic ovaries in rats: effect of electro-acupuncture on concentrations of endothelin-1 and nerve growth factor (NGF), and expression of NGF mRNA in the ovaries, the adrenal glands, and the central nervous system. *Reproductive Biol Endocrinol: RB&E*. (2003) 1:33. doi: 10.1186/1477-7827-1-3
10. Huang DM, Huang GY, Lu FE, Stefan D, Andreas N, Robert G. Acupuncture for infertility: is it an effective therapy? *Chin J Integr Med*. (2011) 17:386–95. doi: 10.1007/s11655-011-0611-8
11. Guven PG, Cayir Y, Borekci B. Effectiveness of acupuncture on pregnancy success rates for women undergoing *in vitro* fertilization: a randomized controlled trial. *Taiwan J Obstet Gynecol*. (2020) 59:282–6. doi: 10.1016/j.tjog.2020.01.018
12. Magarelli PC, Cridennda DK, Cohen M. Changes in serum cortisol and prolactin associated with acupuncture during controlled ovarian hyperstimulation in women undergoing *in vitro* fertilization-embryo transfer treatment. *Fertil Steril*. (2009) 92:1870–9. doi: 10.1016/j.fertnstert.2008.10.067
13. Paulus WE, Zhang M, Strehler E, El-Danasouri I, Sterzik K. Influence of acupuncture on the pregnancy rate in patients who undergo assisted reproduction therapy. *Fertil Steril*. (2002) 77:3273. doi: 10.1016/S0015-0282(01)03273-3
14. Shuai Z, Lian F, Li P, Yang W. Effect of transcutaneous electrical acupuncture point stimulation on endometrial receptivity in women undergoing frozen-thawed embryo transfer: a single-blind prospective randomised controlled trial. *Acupunct Med*. (2015) 33:9–15. doi: 10.1136/acupmed-2014-010572
15. Andersen D, Løssl K, Andersen AN, Fürbringer J, Bach H, Simonsen J, et al. Acupuncture on the day of embryo transfer: a randomized controlled trial of 635 patients. *Reproductive BioMed Online*. (2010) 21:29. doi: 10.1016/j.rbmo.2010.03.029
16. So EW, Ng EH, Wong YY, Lau EY, Yeung WS, Ho PC. A randomized double blind comparison of real and placebo acupuncture in IVF treatment. *Human Reprod*. (2009) 24:341–8. doi: 10.1093/humrep/den380
17. Moy I, Milad MP, Barnes R, Confino E, Kazer RR, Zhang X. Randomized controlled trial: effects of acupuncture on pregnancy rates in women undergoing *in vitro* fertilization. *Fertil Steril*. (2011) 95:24. doi: 10.1016/j.fertnstert.2010.05.024
18. Domar AD, Meshay I, Kelliher J, Alper M, Powers RD. The impact of acupuncture on *in vitro* fertilization outcome. *Fertil Steril*. (2009) 91:723–6. doi: 10.1016/j.fertnstert.2008.01.018
19. Ajeena IM, Al-Haris N, Al-Kefae HN, Goloom SA, Al-Abudi K. Impact of transcutaneous electrical nerve stimulation (TENS) on endometrial thickness in the healthy women, a step to improve the female fertility. *J Contemp MedSci*. (2018) 4:16–9.
20. Zhuang H. *Clinical Study on the Improvement of Endometrial Recepbility of Freeze-thaw Embryo Transfer Cycle in RIF Patients with Kidney Deficiency and Liver Depression by Staged Acupuncture Therapy Based on the Theory of'Chong Meridian is the Sea of Blood'*. Kunming: University of Yunnan TCM (2019).
21. Anderson B, Rosenthal L. Acupuncture and *in vitro* fertilization: critique of the evidence and application to clinical practice. *Complemen Theraps Clin Prac*. (2013) 19:02. doi: 10.1016/j.ctcp.2012.11.002
22. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ (Clin Res ed)*. (2021) 372:n71. doi: 10.1136/bmj.n71
23. Saxtorph MH, Hallager T, Persson G, Petersen KB, Eriksen JO, Larsen LG, et al. Assessing endometrial receptivity after recurrent implantation failure: a prospective controlled cohort study. *Reprod Biomed Online*. (2020) 41:998–1006. doi: 10.1016/j.rbmo.2020.08.015
24. Zhao J, Chen Q, Xue X. An update on the progress of endometrial receptivity in women with polycystic ovary syndrome. *Reprod Sci*. (2021). doi: 10.1007/s43032-021-00641-z [Online ahead of print].
25. Zhao J, Zhang Q, Li Y. The effect of endometrial thickness and pattern measured by ultrasonography on pregnancy outcomes during IVF-ET cycles. *Reprod Biol Endocrinol*. (2012) 10:100. doi: 10.1186/1477-7827-10-100
26. Galliano D, Bellver J, Díaz-García C, Simón C, Pellicer A. ART and uterine pathology: how relevant is the maternal side for implantation? *Hum Reprod Update*. (2015) 21:13–38. doi: 10.1093/humupd/dmu047
27. Gonen Y, Casper R. Prediction of implantation by the sonographic appearance of the endometrium during controlled ovarian stimulation for *in vitro* fertilization (IVF). *J in vitro Fertil Embryo Trans: IVF*. (1990) 7:146–52. doi: 10.1007/BF01135678
28. Yang JH, Wu MY, Chen CD, Jiang MC, Ho HN, Yang YS. Association of endometrial blood flow as determined by a modified colour Doppler technique with subsequent outcome of in-vitro fertilization. *Human Reprod*. (1999) 14:1606–10. doi: 10.1093/humrep/14.6.1606
29. Weckstein LN, Jacobson A, Galen D, Hampton K, Hammel J. Low-dose aspirin for oocyte donation recipients with a thin endometrium: prospective, randomized study. *Fertil Steril*. (1997) 68:927–30. doi: 10.1016/S0015-0282(97)00330-0
30. Kasius A, Smit JG, Torrance HL, Eijkemans MJ, Mol BW, Opmeer BC, et al. Endometrial thickness and pregnancy rates after IVF: a systematic review and meta-analysis. *Human Reprod Update*. (2014) 20:530–41. doi: 10.1093/humupd/dmu011
31. Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. (editors). *Cochrane Handbook for Systematic Reviews of Interventions Version 6.1*. Hoboken: A John Wiley & Sons, Ltd. (2021).
32. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ (Clini Res ed)*. (2008) 336:924–6. doi: 10.1136/bmj.39489.470347.AD
33. Chen Q, Hau C. Impacts on pregnancy outcome treated with acupuncture and moxibustion in IVF-ET patients. Zhongguo zhen jiu [Chinese Acupuncture & Moxibustion]. (2015) 35:313–7. doi: 10.7669/j.issn.0253-357X.2015.03.0159
34. Luo G, Su W, Ma W, Situ Y, Xie C, Yang M. The effect of warm needling pretreatment for endometrial receptivity of frozen embryo transfer.

Zhongguo zhen jiu [Chinese acupuncture & moxibustion]. (2017) 37:831–35. doi: 10.13703/j.0255-2930.2017.08.009
35. Ho M, Huang LC, Chang YY, Chen HY, Chang WC, Yang TC, et al. Electroacupuncture reduces uterine artery blood flow impedance in infertile women. *Taiwanese J Obst Gynecol*. (2009) 48:148–51. doi: 10.1016/S1028-4559(09)60276-X
36. Ma J, Zhang Q. Clinical study of acupuncture on improving endometrial receptivity, emotion of anxiety and depression,and pregnancy outcome in IVF-ET repeated implantation failure patients with syndrome of kidney deficiency and liver stagnation. *Academic J Shanghai Univ Traditional Chinese Med*. (2019) 33:1339–1344. doi: 10.16306/j.1008-861x.2021.02.006
37. Lin QP, Liu Y, Xu JB, Yang J, You XM, Zhang JX. Effect of electroacupuncture combined with ginger-isolated moxibustion on endometrial receptivity in infertile patients with polycystic ovarian syndrome. *World J Acupunct - Moxibustion*. (2021) 32:15–20. doi: 10.1016/j.wjam.2022
38. Zhong X, Ren C, Chi L, Zhang W, Yan H, Liu X, et al. Effects of electroacupuncture intervention on endometrial receptivity and pregnancy outcome in patients undergoing IVF-ET with syndrome of kidney deficiency and phlegm stasis. *Aca J Shanghai Unive Traditional Chinese Med*. (2021) 35:28–31+37. doi: 10.13460/.issn.1005-0957.2018.12.1339
39. Zhao M, Liu L, Quan X, Lan Y, Tian X, Ji Z, et al. Effect of electroacupuncture combined with infrared irradiation on endometrial receptivity, embryo transfer and pregnancy in patients with repeated implantation failure. *Progress Mod Biomed*. (2019) 19:4549–52+35.
40. Zhong X, Li K, Jiang X, Tong G. Effect of TEAS therapy on endometrial receptivity of IVF patients with kidney Qi deficiency. *J Liaoning Univ Tradi Chinese*. (2016) 18:188–91. doi: 10.13194/j.issn.1673-842x.2016.10.059
41. Wang X, Tian L, Yuan H, Liu H, Fang L, Zhou J. Curative observation of warming acupuncture treatment for the impact on receptivity of endometrium of embryo transfer patients. *J Sichuan Trad Chinese Med*. (2017) 35:184–7. doi: 10.13703/j.0255-2930.2017.35.002
42. Dieterle S, Ying G, Hatzmann W, Neuer A. Effect of acupuncture on the outcome of *in vitro* fertilization and intracytoplasmic sperm injection: a randomized, prospective, controlled clinical study. *Fertil Steril*. (2006) 85:1347–51. doi: 10.1016/j.fertnstert.2005.09.062
43. Zhong KL, Jiang X, Tong G. Effect of TEAS Therapy on endometrial receptivity of IVF patients with kidney Qi deficiency. *J Liaoning University Trad Chinese*. (2016) 18:188–91. doi: 10.13194/j.issn.167-842x.2016.10.059s
44. Zha LL, Quan X, Lan Y, Tian X, Ji Z, et al. Effect of Electro-acupuncture combined with infrared irradiation on endometrial receptivity, embryo transfer and pregnancy in patients with repeated implantation failure. *Progress Mod Biomed*. (2019) 19:4549–52+35. doi: 10.13241/j.cnki.pmb.2019.23.035
45. White A, Cummings M, Barlas P, Cardini F, Filshie J, Foster NE, et al. Defining an adequate dose of acupuncture using a neurophysiological approach–a narrative review of the literature. *Acupuncture Med: J Br Med Acupuncture Soc*. (2008) 26:111–20. doi: 10.1136/aim.26.2.111
46. Sun N, Tu JF, Lin LL, Li YT, Yang JW, Shi GX, et al. Correlation between acupuncture dose and effectiveness in the treatment of knee osteoarthritis: a systematic review. *Acupuncture Med: J Br Med Acupuncture Soc*. (2019) 37:261–7. doi: 10.1136/acupmed-2017-011608
47. Gao Y. Dynamic *Impact of the Specific Acupoints of the Shaoyang Meridian on Brain Function of Migraineurs:a fMRI Study*. Chengdu: Chengdu University of TCM (2012).
48. Lin YJ, Kung YY, Kuo WJ, Niddam DM, Chou CC, Cheng CM, et al. Effect of acupuncture 'dose' on modulation of the default mode network of the brain. *Acupunct Med*. (2016) 34:425–32. doi: 10.1136/acupmed-2016-011071
49. Zheng CH, Guang YH, Zhang MM, Wang W. Effects of acupuncture on pregnancy rates in women undergoing *in vitro* fertilization: a systematic review and meta-analysis. *Fertil Steril*. (2012) 97:599–611. doi: 10.1016/j.fertnstert.2011.12.007
50. Stener-Victorin E, Waldenström U, Nilsson L, Wikland M, Janson PO. A prospective randomized study of electro-acupuncture vs. alfentanil as anaesthesia during oocyte aspiration in *in-vitro* fertilization. *Human Reprod*. (1999) 14:2480–4. doi: 10.1093/humrep/14.10.2480
51. Lensen S, Hammarberg K, Polyakov A, Wilkinson J, Whyte S, Peate M, et al. How common is add-on use and how do patients decide whether to use them? A national survey of IVF patients. *Human Reprod*. (2021) 36:1854–61. doi: 10.1093/humrep/deab098
52. Bourgain C, Devroey P. The endometrium in stimulated cycles for IVF. *Hum Reprod Update*. (2003) 9:515–22. doi: 10.1093/humupd/dmg045
53. Wang L, Lv S, Li F, Bai E, Yang X. Letrozole vs. clomiphene citrate and natural cycle: endometrial receptivity during implantation window in women with polycystic ovary syndrome. *Front Endocrinol*. (2020) 11:532692. doi: 10.3389/fendo.2020.532692
54. Glujovsky D, Pesce R, Fiszbajn G, Sueldo C, Hart RJ, Ciapponi A. Endometrial preparation for women undergoing embryo transfer with frozen embryos or embryos derived from donor oocytes. *Cochrane Database Sys Rev*. (2020) 10:Cd006359. doi: 10.1002/14651858.CD006359.pub3
55. Zhang CS, Tan HY, Zhang GS, Zhang AL, Xue CC, Xie YM. Placebo devices as effective control methods in acupuncture clinical trials: a systematic review. *PloS ONE*. (2015) 10:e0140825. doi: 10.1371/journal.pone.0140825
56. Manheimer E. Selecting a control for *in vitro* fertilization and acupuncture randomized controlled trials (RCTs): how sham controls may unnecessarily complicate the RCT evidence base. *Fertil Steril*. (2011) 95:2456–61. doi: 10.1016/j.fertnstert.2011.04.040
57. Manheimer E, van der Windt D, Cheng K, Stafford K, Liu J, Tierney J, et al. The effects of acupuncture on rates of clinical pregnancy among women undergoing *in vitro* fertilization: a systematic review and meta-analysis. *Human Reprod Update*. (2013) 19:696–713. doi: 10.1093/humupd/dmt026
58. Madsen M, Gøtzsche P, Hróbjartsson A. Acupuncture treatment for pain: systematic review of randomised clinical trials with acupuncture, placebo acupuncture, and no acupuncture groups. *BMJ (Clin Res ed)*. (2009) 338:a3115. doi: 10.1136/bmj.a3115
59. Streitberger K, Kleinhenz J. Introducing a placebo needle into acupuncture research. *Lancet*. (1998) 352:364–5. doi: 10.1016/S0140-6736(97)10471-8
60. Park J, White A, Stevinson C, Ernst E, James M. Validating a new non-penetrating sham acupuncture device: two randomised controlled trials. *Acupuncture Med*. (2003) 22:13–7. doi: 10.1136/aim.20.4.168
61. Smith CA, De Lacey S, Chapman M, Ratcliffe J, Norman RJ, Johnson NP, et al. Effect of acupuncture vs. sham acupuncture on live births among women undergoing *in vitro* fertilization: a randomized clinical trial. *Jama*. (2018) 319:1990–8. doi: 10.1001/jama.2018.5336
62. Xiufang K, Zhong L, Xiaomei J, Guoqing T. Effect of TEAS therapy on endometrial receptivity of IVF patients with kidney Qi deficiency. *J Liaoning Univ Tradi Chinese*. (2016) 18:188–91.
63. Zhu D, Gao Y, Chang J, Kong J. Placebo acupuncture devices: considerations for acupuncture research. Evidence-based complementary and alternative medicine: *eCAM*. (2013) 2013:628907. doi: 10.1155/2013/628907
64. Zhong Y, Zeng F, Liu W, Ma J, Guan Y, Song Y. Acupuncture in improving endometrial receptivity: a systematic review and meta-analysis. *BMC Complement Alt Med*. (2019) 19:61. doi: 10.1186/s12906-019-2472-1

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Zheng, Yu, Liu, Yang, Wang, Yang, Lv and Yang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Pharmacology

SYSTEMATIC REVIEW
published: 17 May 2022
doi: 10.3389/fphar.2022.805966

Vancomycin, Daptomycin, Antistaphylococcal β-Lactam, and Trimethoprim-Sulfamethoxazole Monotherapy and Combination Therapy in the Management of Methicillin-Resistant *Staphylococcus aureus*: A Network Meta-Analysis

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Hasan Ejaz,
Al Jouf University, Saudi Arabia
Hua Zhou,
Zhejiang University, China

***Correspondence:**
Xiaonan Xu
xuxiaonanlzu@163.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Drugs Outcomes Research and Policies, a section of the journal Frontiers in Pharmacology

Received: 31 October 2021
Accepted: 29 March 2022
Published: 17 May 2022

Citation:
Xu X, Lu N, Song P, Zhou M, Li Y, Wang Z and Gao X (2022) Vancomycin, Daptomycin, Antistaphylococcal β-Lactam, and Trimethoprim-Sulfamethoxazole Monotherapy and Combination Therapy in the Management of Methicillin-Resistant Staphylococcus aureus: A Network Meta-Analysis. Front. Pharmacol. 13:805966. doi: 10.3389/fphar.2022.805966

Xiaonan Xu[1], Ni Lu[2†], Pan Song[3†], Mingzhen Zhou[2], Yuanxiao Li[1], Zirui Wang[2] and Xin Gao[2]*

[1]Department of Pediatrics, Second Hospital of Lanzhou University, Lanzhou, China, [2]The Clinical Medical College of Lanzhou University, Lanzhou, China, [3]Department of Urology, Institution of Urology, West China Hospital of Sichuan University, Chengdu, China

Objective: The aim was to evaluate the efficacy and safety of vancomycin or daptomycin (VAN/DAP), antistaphylococcal β-lactam (ASBL), trimethoprim-sulfamethoxazole (TMP-SMX), and combination therapy of VAN/DAP + ASBL in the management of methicillin-resistant *Staphylococcus* aureus (MRSA).

Methods: Databases including PubMed, Cochrane Library, Embase database, and google scholar were searched on 1 September 2021. The randomized control trials (RCTs) and comparable clinical studies of VAN/DAP, VAN/DAP + ASBL, ASBL, and TMP-SMX in the management of MRSA were identified. A network meta-analysis was conducted with STATA 14.0.

Results: Seven RCTs and two matched cohorts with 1,048 patients were included in the analysis. The pooled results showed that VAN/DAP + ASBL had a significantly lower rate of persistent bacteremia >3 days than VAN/DAP alone [OR:0.46, 95%CI (0.26, 0.81), $p < 0.001$]. No obvious differences were observed in the outcomes of all-cause mortality, relapsed bacteremia, microbiological treatment failure, embolic or metastatic infection, and total adverse events. However, the ranking results showed that VAN/DAP + ASBL had slightly better efficacy (all-cause mortality, persistent bacteremia >3 days, duration of bacteremia, microbiological treatment failure, and relapsed bacteremia) but slightly higher adverse events than VAN/DAP alone. No obvious differences in the comparisons of VAN/DAP vs. ASBL, and VAN/DAP vs TMP-SMX in the analyzed outcomes. The ranking results revealed that ASBL and TMP-SMX did not have better efficacy or lower adverse events compared with the treatment of VAN/DAP.

Conclusion: The efficacy of VAN/DAP + ASBL was slightly but not significantly better than VAN/DAP alone in the management of MRSA.

Keywords: vancomycin, daptomycin, β-lactams, network meta-analysis, MRSA

INTRODUCTION

Antibiotic resistance is a major global health problem, and methicillin-resistant *Staphylococcus aureus* is a serious threat among Gram-positive drug-resistant bacteria. (Vestergaard et al., 2019)According to its original source, MRSA can be divided into hospital-acquired MRSA and community-acquired MRSA. In China, the proportion of MRSA obtained by hospitals has reached 50.4%. (Shang et al., 2016)The proportion of methicillin-resistant *Staphylococcus aureus* in Europe decreased from 26.6% in 2007 to 16.8% in 2015, but it is estimated that the incidence of MRSA infection in the European Union / European Economic area actually increased by 1.28 times over the same period. (Borg and Camilleri, 2021)It has high morbidity and mortality and can cause metastatic or complex infections such as infective endocarditis or septicemia. (Hassoun et al., 2017)It seriously endangers the life and health of patients and causes huge economic losses and heavy burdens to the global medical system.

Currently, the standard treatment for MRSA bacteremia is vancomycin or daptomycin (VAN/DAP) (Liu et al., 2011). However, their efficacy is limited and there are many disadvantages such as poor tissue permeability and slow killing time (Rybak, 2006). Some alternative medicines such as trimethoprim-sulfamethoxazole (TMP-SMX) and ceftaroline have been evaluated in the treatments of MRSA in recent years (Campbell et al., 2012; Holmes et al., 2015). Another strategy for treating MRSA infection is to combine the standard regimen with other drugs. Many *in vitro* data studies have demonstrated that VAN/DAP in combination with antistaphylococcal β-lactam (ASBL) has a synergistic effect on MRSA strains, which can increase the speed of bacterial killing (Holmes et al., 2015). Whether the combination of VAN/DAP with β-lactam has obvious better outcomes than VAN/DAP is still inconclusive.

Network meta-analysis is a new meta-analysis method that uses the evidence of comprehensive randomized trials to evaluate the relative effectiveness of two or more interventions (Cipriani et al., 2013). The purpose of this study is to evaluate the efficacy and safety of VAN/DAP, VAN/DAP + ASBL, ASBL, and TMP-SMX in the management of MRSA.

METHODS

Search Strategy and Study Selection

A systematic review was performed based on the guidelines for the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement (Page et al., 2021). Four electronic databases including PubMed, Cochrane Library, Web of Science, and Embase database were searched for the clinical studies of VAN/DAP in the management of MRSA from inception to 1 September 2021. MeSH terms and related synonyms including "*staphylococcus aureus*," "*bacteremia*", "*Staphylococcus aureus bacteremia*", "*vancomycin*," "*vancocine*," "*vancomicina*," "*daptomycin*," "*cubicin*," and "*LY146032*" were combined and searched. No language limitation was used while searching the databases. All records we retrieved were imported into Endnote X9, and duplicate records were removed by using this software and manually. Two researchers assessed the relevance of each article by reading the title and abstract individually. According to the selection criteria, full-text scanning was conducted to evaluate the eligibility of each related study. Any discrepancies and inconsistencies during the entire evaluation process of the two researchers were resolved by discussion or inquiring with a third researcher.

Selection Criteria

The trials were included when they met the following criteria: (Vestergaard et al., 2019) the article types were RCTs or clinical studies with comparable basic characteristics; (Shang et al., 2016) patients were diagnosed with MRSA by one or more positive blood cultures for *S. aureus* within two calendar days; (Borg and Camilleri, 2021) the treatment in the control group was using VAN or DAP; treatment in the intervention group was using one of the VAN/DAP + ASBL, ASBL, or TMP-SMX; ASBL was a class of drugs but not specific drugs.

Studies were excluded if they met the following conditions: (Vestergaard et al., 2019) significant difference existed in the basic characteristic of two group of patients; (Shang et al., 2016) patients were associated with other diseases such as the liver and kidney insufficiency; (Borg and Camilleri, 2021) patients were infected with MSSA rather than MRSA; (Hassoun et al., 2017) some other treatments besides those we mentioned above were added during the treatment; (Liu et al., 2011) the study did not provide raw data in the article and we could not obtain it by contacting the corresponding author; (Liu et al., 2011) the outcomes were rare and no other study had similar outcomes; we were unable to conduct a combination of analyses.

Data Collection and Quality Assessment

For each included study, two independent researchers collected all available data from the article. The information including the first author's name, year of publication, journal, study design, treatments in experimental and control groups, number of patients in each group, and reported outcomes were collected and collated in Microsoft Excel 2016. The important outcomes which were reported in three or more studies and measured in the same way were pooled and defined as the outcomes of this study. The primary outcomes of this study were all-cause mortality and total adverse events, and the secondary outcomes included duration of bacteremia days, microbiological treatment failure, relapsed bacteremia, and embolic or metastatic infection.

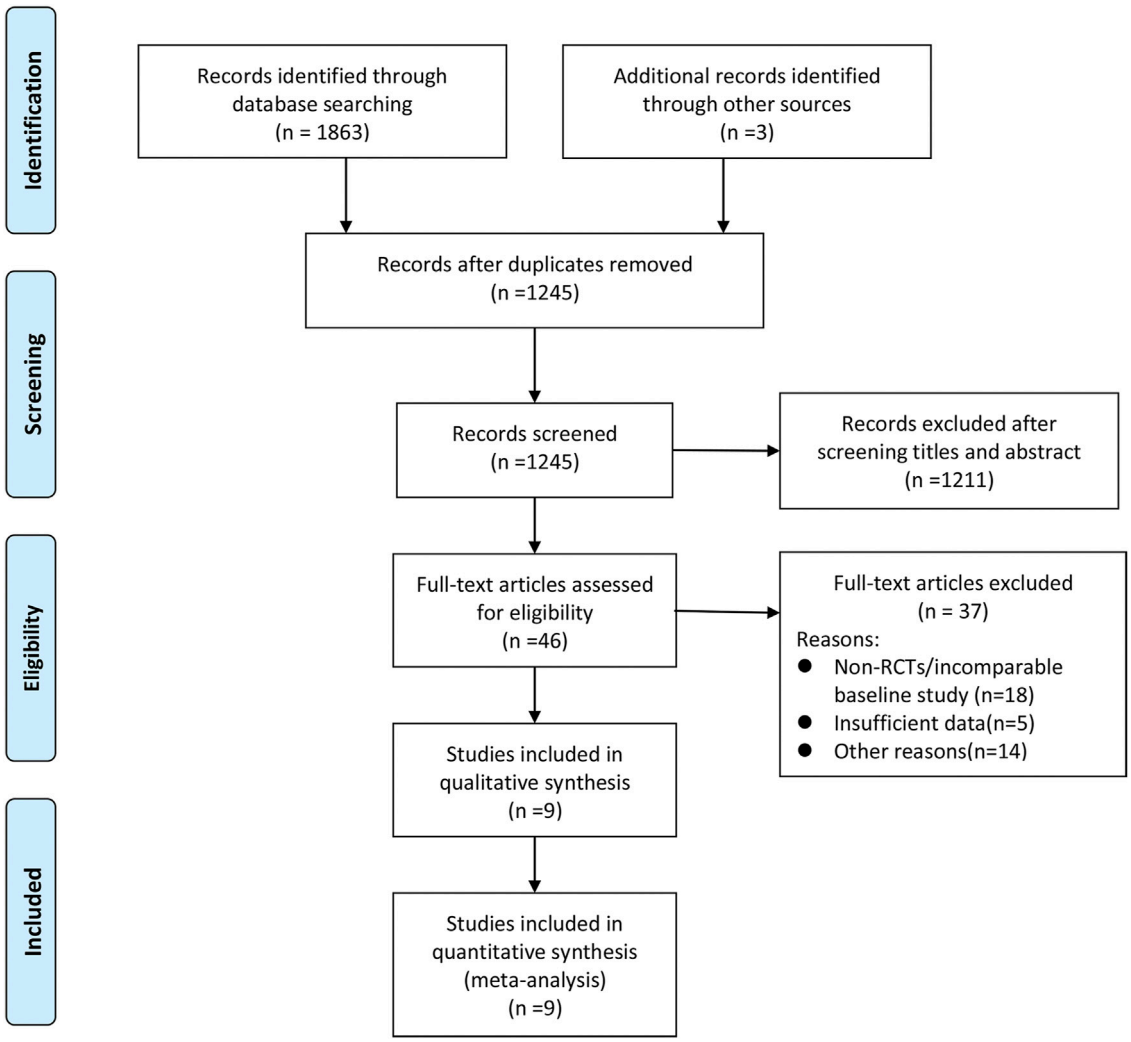

FIGURE 1 | PRISMA flow diagram of the study selection process for network meta-analysis.

The quality of the included studies was assessed by two independent reviewers based on the tools provided in version 5.1.3 of the Cochrane Handbook. Each trial was evaluated and scored according to the following criteria: random sequence generation, the blindness of participants and result evaluators, incomplete result data reporting, allocation concealment, selective data result reporting, and other deviations. The disagreement between the two reviewers was resolved through discussion or inquiry with the third reviewer.

Data Synthesis and Analysis

Odds ratio (OR) and 95% confidence interval (95% CI) were used for the statistical analysis of binary variables, and weighted mean difference (WMD) and 95% CI were used for the analysis of continuous variables. The random effect model was adopted in the analysis because inconsistency might exist between different trials and treatments. The comparisons between every two interventions for every outcome were conducted. The ranking analysis which could rank the best to worst interventions for every outcome was introduced to distinguish the best treatment. The consistency analysis was evaluated by the node-splitting analysis and the publication bias was accessed by funnel plots. All analyses were conducted with the software of STATA version 14.0. The packages of "network," "mvmeta," and "netfunnel" were used during the analyses . $p < 0.05$ was regarded as a statistical difference.

RESULTS

Baseline Characteristics of Included Studies

A total of 1866 records were retrieved and screened first by reading the titles and abstracts. The remaining 46 articles were evaluated by full-text scanning. Nine studies (Fowler et al.,

TABLE 1 | Characteristics of the included randomized controlled trials in network meta-analysis.

Authors	Year	Journal	Study design	Experimental group	Control group	Samples (E/C)	Main outcomes
Tong et al. (2020)	2020	JAMA	RCT	VAN/DAP + ASBL (flucloxacillin, cloxacillin, or cefazolin)	VAN/DAP	174/178	All-cause mortality; persistent bacteremia; relapsed bacteremia; microbiological treatment failure; adverse events
Geriak et al. (Geriak et al. (2019)	2019	Clinical Infectious Diseases	RCT	VAN/DAP + ASBL (ceftaroline)	VAN/DAP	17/23	All-cause mortality; persistent bacteremia; median duration of bacteremia days; adverse events
Davis et al. (2016)	2015	Antimicrob Agents Chemother	RCT	VAN/DAP + ASBL (flucloxacillin)	VAN/DAP	31/29	All-cause mortality; median duration of bacteremia days; relapsed bacteremia; embolic or metastatic infection; adverse events
McCreary et al. (2020)	2019	Open Forum Infectious Diseases	MatchedCohort	VAN/DAP + ASBL (ceftaroline)	VAN/DAP	58/113	All-cause mortality; median duration of bacteremia days; microbiological treatment failure;
Dilworth et al. (2014)	2014	Antimicrob Agents Chemotherapy	MatchedCohort	VAN/DAP + ASBL (ceftaroline)	VAN/DAP	50/30	All-cause mortality; microbiological eradication
Hamed et al. (2020)	2020	Future Microbiol	RCT	VAN/DAP	ASBL (ceftobiprole)	195/195	All-cause mortality; microbiological eradication; relapsed bacteremia; embolic or metastatic infection; adverse events
Pericàs et al. (2018)	2018	Clinical Microbiology and Infection	RCT	VAN/DAP	ASBL (imipenem)	7/8	All-cause mortality; persistent bacteremia; adverse events
Fowler et al. (2006)	2006	New England Journal of Medicine	RCT	VAN/DAP	ASBL (nafcillin, oxacillin, or flucloxacillin)	120/116	The median duration of bacteremia days; adverse events
Paul et al. (2015)	2015	BMJ	RCT	TMP-SMX	VAN/DAP	53/51	All-cause mortality; persistent bacteremia; microbiological treatment failure; adverse events

E/C, Experimental group versus control group; JAMA, Journal of the American medical association; BMJ, British medical journal; RCT, randomized controlled trial; VAN/DAP, vancomycin or daptomycin; ASBL, antistaphylococcal β-lactam; TMP-SMX, Trimethoprim-sulfamethoxazole.

2006; Dilworth et al., 2014; Paul et al., 2015; Davis et al., 2016; Pericàs et al., 2018; Geriak et al., 2019; Hamed et al., 2020; McCreary et al., 2020; Tong et al., 2020) with 1,048 patients were finally enrolled into our analysis. The whole process of the literature search and screening is shown in **Figure 1**. The baseline characteristics of the included studies are presented in **Table 1**.

The included studies were seven RCTs and two matched cohorts with comparable basic characteristics. Most of the RCTs were randomized groups and described adequate randomization in detail. The detailed information about allocation sequence concealment was well described in most of the included RCTs. Attrition bias and reporting bias were well-performed. The risk-of-bias assessment is summarized in **Figure 2**.

Network Meta-Analysis of All-Cause Mortality

Seven studies with 951 patients described the outcome of all-cause mortality. No significant differences were detected among the four treatments. With VAN/DAP as the reference, the ORs and 95%CIs of VAN/DAP + ASBL, ASBL, and TMP-SMX were 0.78(0.32, 1.92), 1.34 (0.43, 4.18), and 2.36 (0.60, 9.27), individually. These results are presented in **Figure 3A**. Ranking results for these interventions showed that VAN/DAP + ASBL was the best and was slightly better than other treatments. ASBL was the worst for this outcome (**Figure 4A**).

Network Meta-Analysis of Persistent Bacteremia

Four studies with 489 patients reported the number of patients with persistent bacteremia for more than 3 days. Results showed that VAN/DAP + ASBL had a significantly lower rate of persistent bacteremia longer than 3 days than VAN/DAP [OR = 0.51, 95%CI (0.29, 0.91), $p < 0.001$]. With VAN/DAP as the reference, the OR and 95% CI of ASBL and TMP-SMX were 1.09 (0.20, 5.94) and 0.86 (0.34, 2.14). All results are presented in **Figure 3B**. The ranking results revealed that VAN/DAP + ASBL ranked 1st and was the best treatment in this outcome (**Figure 4B**).

Six studies described the duration of bacteremia days. There were no significant differences among VAN/DAP, VAN/DAP + ASBL, and TMP-SMX. With VAN/DAP as the reference, the SMD and 95%CI of VAN/DAP + ASBL and TMP-SMX were 0.46 (0.26, 0.81) and 0.86(0.34,2.14). The ranking results revealed that VAN/DAP + ASBL had the shortest duration of bacteremia days and was slightly better than the other treatments. All these results are shown in **Figure 4C**.

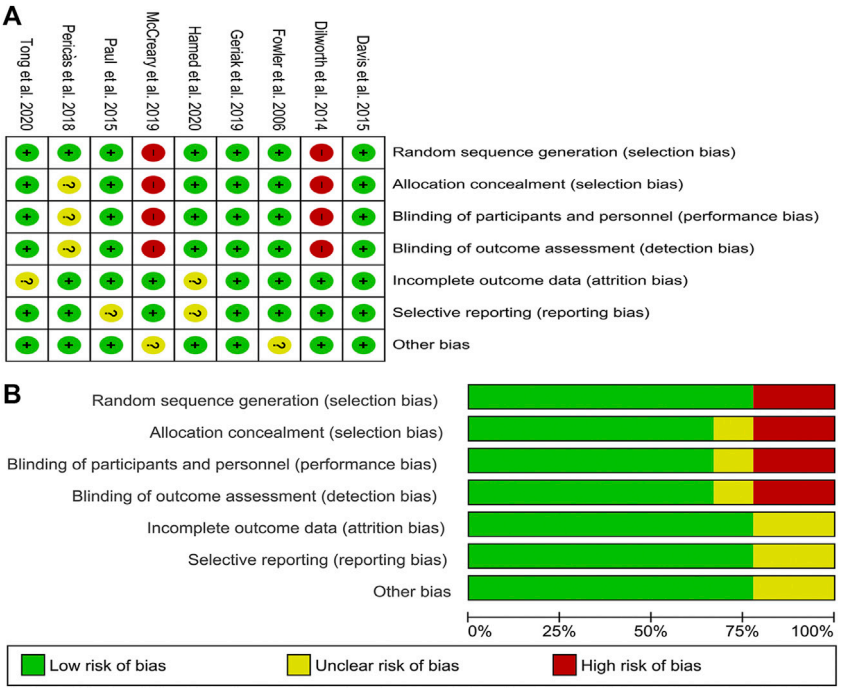

FIGURE 2 | Risk of bias graph and summary of the included studies **(A)** reviewers' judgments about each risk of bias item for eligible studies and **(B)** the judgments about each risk of bias item presented as percentages across all eligible studies.

Network Meta-Analysis of Relapsed Bacteremia

Six trials with 879 patients were analyzed. Results suggested that there were no significant differences among VAN/DAP, VAN/DAP + ASBL, and ASBL. With VAN/DAP as the reference, the OR and 95%CI of VAN/DAP + ASBL and ASBL were 0.85(0.47, 1.55) and 0.54(0.25, 1.17). The ranking results showed that ASBL was the best, and VAN/DAP + ASBL ranked 2nd. All these results are shown in **Figure 4D**.

Network Meta-Analysis of Microbiological Treatment Failure

Five trials with 604 patients were pooled in this analysis. No significant differences were found among VAN/DAP, VAN/DAP + ASBL, ASBLA, and TMP-SMX. With VAN/DAP as the reference, the OR and 95%CI of VAN/DAP + ASBL, ASBL, and TMP-SMX were 0.52 (0.13, 2.15), 1.32(0.32, 5.47), and 1.97 (0.28, 14.11), respectively. These results are presented in **Figure 3E**. The ranking results revealed that VAN/DAP + ASBL was the best. VAN/DAP ranked 2nd (**Figure 4E**).

Network Meta-Analysis of Embolic or Metastatic Infection

Three studies reported embolic or metastatic infection, recruiting 158 patients. The pooled estimates showed that the differences among VAN/DAP + ASBL, VAN/DAP, and ASBL were not significant. With VAN/DAP as the reference, the OR and 95% CI of VAN/DAP + ASBL and ASBL were 0.28(0.03, 2.92) and 0.43(0.03, 5.34). The ranking results showed that VAN/DAP + ASBL was the best. These results are presented in **Figure 4F**.

Network Meta-Analysis of Adverse Events

Seven trials described the total number of adverse events and were combined. Results showed that there were no significant differences in the incidence of any adverse events among the four treatments. With VAN/DAP as the reference, the OR and 95%CI of VAN/DAP + ASBL, ASBL, and TMP-SMX were 1.71 (0.55, 5.35), 1.54 (0.31, 7.53), and 1.44(0.30, 6.78), respectively. These results are shown in **Figure 3G**. The ranking results revealed that VAN/DAP ranked 1st, TMP-SMX ranked 2nd, VAN/DAP + ASBL ranked 3rd, and ASBL was the worst (**Figure 4G**).

Consistency Analysis

The node-splitting analysis was adopted to evaluate inconsistency by comparing the differences between direct and indirect evidence. The publication bias was evaluated by funnel plots. No obvious inconsistency and publication bias were detected in all outcomes.

DISCUSSION

From the comparisons of VAN/DAP and the combination therapy of VAN/DAP + ASBL, we found that (Vestergaard

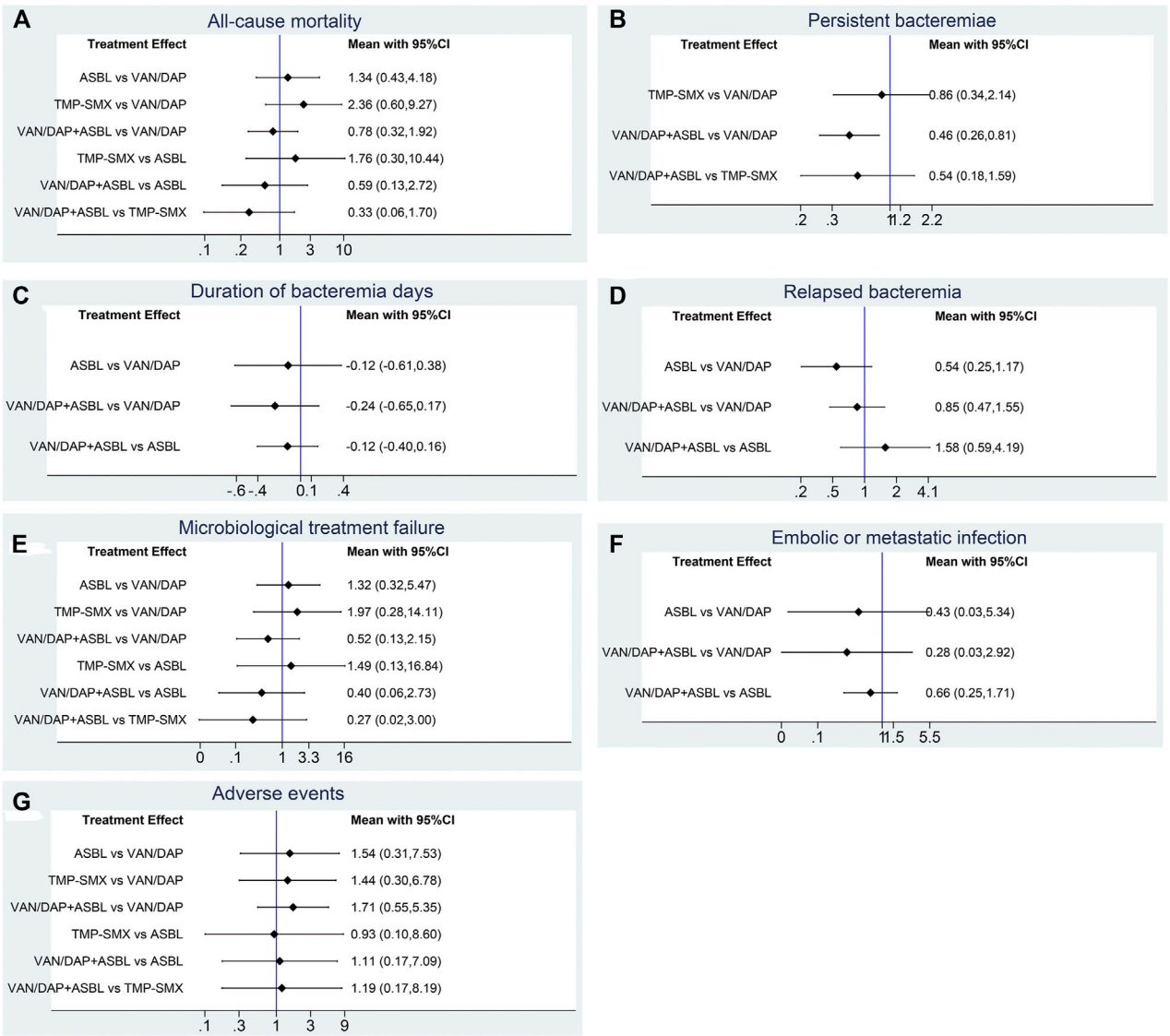

FIGURE 3 | Forest plots for the outcomes. **(A)** All-cause mortality; **(B)** Persistent bacteremia; **(C)** Duration of bacteremia days; **(D)** Relapsed bacteremia ; **(E)** Microbiological treatment failure; **(F)** Embolic or metastatic infection; **(G)** Adverse events. VAN/DAP, Vancomycin or daptomycin; ASBL, antistaphylococcal β-lactam; TMP-SMX, trimethoprim-sulfamethoxazole.

et al., 2019) VAN/DAP + ASBL significantly reduced the number of patients with persistent bacteremia >3 days. (Shang et al., 2016) No obvious difference was observed between VAN/DAP + ASBL and VAN/DAP in the outcomes of all-cause mortality, relapsed bacteremia, microbiological treatment failure, embolic or metastatic infection, and adverse events. (Borg and Camilleri, 2021) The ranking results revealed that VAN/DAP + ASBL was slightly better than VAN/DAP alone in all-cause mortality, persistent bacteremia >3 days, duration of bacteremia, microbiological treatment failure, and relapsed bacteremia.

Some previous studies reported similar outcomes. Atalla and Mylonakis (2020) pointed out that the blood culture bacteria clearance rate of the combined group (VAN/DAP + ASBL) was faster than that of the standard group (VAN/DAP). Hagihara et al. (2012) reported that the combination of vancomycin and ASBL in simulated human exposure increased the killing rate of these MRSA isolates and resulted in a greater overall antibacterial effect. Holubar suggested that the combination of vancomycin or daptomycin with β-lactam antibiotics was associated with a shorter duration of bacteremia, but there was no significant clinical benefit. (Holubar et al., 2020) However, Wang pointed out that the combination therapy has certain clinical benefits, but it may be related to the use of a certain type of beta-lactam drugs in the experimental group, such as ceftaroline, and perhaps ceftaroline itself has anti-MRSA effects. (Wang et al., 2020) Some scholars believe that the clinical effect of

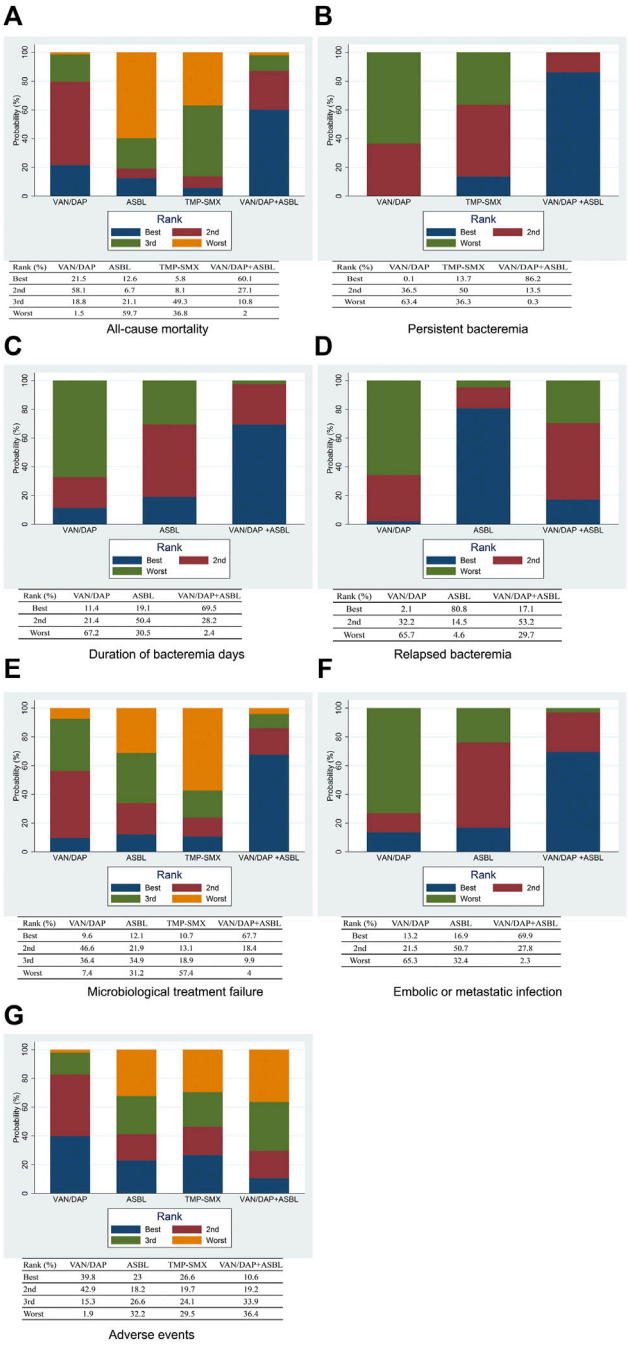

FIGURE 4 | The ranking plots for the efficacy outcomes **(A)** All-cause mortality; **(B)** persistent bacteremia; **(C)** Duration of bacteremia days; **(D)** Relapsed bacteremia; **(E)** Microbiological treatment failure; **(F)** Embolic or metastatic infection; **(G)** Adverse events. VAN/DAP, Vancomycin or daptomycin; ASBL, antistaphylococcal β-lactam; TMP-SMX, trimethoprim-sulfamethoxazole.

combined therapy is related to the length of treatment time and the concentration of IL-10 in the blood. (Geriak et al., 2019; Johnson et al., 2021) MRSA is naturally resistant to β-lactam drugs, resistance is usually conferred by the acquisition of a nonnative gene encoding a penicillin-binding protein (PBP2a), with a significantly lower affinity for β-lactams. This resistance allows cell-wall biosynthesis, the target of β-lactams, to continue even in the presence of typically inhibitory concentrations of antibiotics (Peacock and Paterson, 2015), and the significant changes in the cell-wall composition required for vancomycin resistance make these strains resistant to β-lactams antibiotics are more sensitive (Domenech et al., 2005). *In vitro* studies have demonstrated synergistic interactions between vancomycin and a wide variety of β-lactams against *S. aureus* (Fox et al., 2006; Ortwine et al., 2013). The proposed mechanism for vancomycin and β-lactam synergy is related to cell-wall thinning via the addition of a β-lactam, which increases vancomycin binding to target sites during cell wall synthesis (Sarkar et al., 2017). A study found that adding ASBL to VAN/DAP therapy for at least 24 h within 72 h could increase the odds of clinical success (Alosaimy et al., 2020).

As for the results of TMP-SMX and ASBL in the treatment of MRSA. Even though no significant differences were detected with VAN/DAP, the ranking results showed that ASBL and TMP-SMX were less efficient than VAN/DAP in most of the outcomes. Some previous studies had reported similar results. Tissot-Dupont et al. (2019) reported that TMP-SMX did not achieve noninferiority compared with vancomycin among patients with invasive MRSA infections. There were also some different conclusions. Eliakim-Raz et al. (2017) pointed out that TMP-SMX had an efficacy profile superior to that of vancomycin in the treatment of MRSA pneumonia, with lower rates of overall mortality and the same safety profile. As for ASBL, it is often adopted in the treatment of MSSA. It was reported that Cefazolin and Ertapenem could rapidly clear persistent MSSA (Ulloa et al., 2020). Recently, it was also explored in the management of MRSA. Ceftaroline and ceftobiprole are the first β-lactams with antiMRSA activity (Lee et al., 2018). Many *in vitro* studies (Espedido et al., 2015; Sader et al., 2015) had demonstrated the potent *in vitro* activity of Ceftaroline against MRSA isolates causing bacteremia. In a matched clinical cohort study, Arshad et al. (2017) compared Ceftaroline fosamil with VAN and DAP for MRSA bacteremia. They found that Ceftaroline fosamil, especially as salvage therapy, had a longer duration of bacteremia but comparable clinical outcomes with VAN/DAP for MRSA bacteremia patients. These results were somewhat similar to our results.

Our meta-analysis had certain limitations. First, although the quality of the studies we included was relatively high, the number of studies included in our analysis was insufficient, which may affect the accuracy of the results. Second, some of the outcome indicators included in the study were not perfect which made it impossible to combine and analyze these results. It might cause incomplete analysis results or omission of some results. Third, the follow-up time of included studies was quite different. Some of the included studies only followed-up for a short period which might influence the pooled results. Therefore, more high-quality studies are needed in the future to evacuate the results of our study.

CONCLUSION

VAN/DAP + ASBL had slightly but not significantly better efficacy than VAN/DAP alone in the treatment of MRSA. ASBL and TMP-SMX didn't have better efficacy or lower adverse events than VAN/DAP.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

XX, NL, and PS: conception and design, collection and assembly of data, data analysis and interpretation, manuscript writing, final approval of the manuscript. MZ, ZW, XG, and YL: data analysis and interpretation, manuscript writing, final approval of the manuscript.

FUNDING

This work was funded by the 2020 Science and Technology Plan of Chengguan District, Lanzhou City, Grant ID: 2020SHFZ0041.

REFERENCES

Alosaimy, S., Sabagha, N. L., Lagnf, A. M., Zasowski, E. J., Morrisette, T., Jorgensen, S. C. J., et al. (2020). Monotherapy with Vancomycin or Daptomycin versus Combination Therapy with β-Lactams in the Treatment of Methicillin-Resistant Staphylococcus Aureus Bloodstream Infections: A Retrospective Cohort Analysis. *Infect. Dis. Ther.* 9 (2), 325–339. doi:10.1007/s40121-020-00292-8

Arshad, S., Huang, V., Hartman, P., Perri, M. B., Moreno, D., and Zervos, M. J. (2017). Ceftaroline Fosamil Monotherapy for Methicillin-Resistant Staphylococcus Aureus Bacteremia: A Comparative Clinical Outcomes Study. *Int. J. Infect. Dis.* 57, 27–31. doi:10.1016/j.ijid.2017.01.019

Atalla, E., and Mylonakis, E. (2020). In MRSA Bacteremia, Adding a β-lactam to Usual Care Did Not Improve a Composite Outcome at 90 Days. *Ann. Intern. Med.* 172 (12), Jc67. doi:10.7326/acpj202006160-067

Borg, M. A., and Camilleri, L. (2021). What Is Driving the Epidemiology of Methicillin-Resistant Staphylococcus Aureus Infections in Europe? *Microb. Drug Resist.* 27 (7), 889–894. doi:10.1089/mdr.2020.0259

Campbell, M. L., Marchaim, D., Pogue, J. M., Sunkara, B., Bheemreddy, S., Bathina, P., et al. (2012). Treatment of Methicillin-Resistant *Staphylococcus aureus* Infections with a Minimal Inhibitory Concentration of 2 μg/mL to Vancomycin: Old (Trimethoprim/sulfamethoxazole) versus New (Daptomycin or Linezolid) Agents. *Ann. Pharmacother.* 46 (12), 1587–1597. doi:10.1345/aph.1R211

Cipriani, A., Higgins, J. P., Geddes, J. R., and Salanti, G. (2013). Conceptual and Technical Challenges in Network Meta-Analysis. *Ann. Intern. Med.* 159 (2), 130–137. doi:10.7326/0003-4819-159-2-201307160-00008

Davis, J. S., Sud, A., O'Sullivan, M. V. N., Robinson, J. O., Ferguson, P. E., Foo, H., et al. (2016). Combination of Vancomycin and β-Lactam Therapy for Methicillin-Resistant *Staphylococcus aureus* Bacteremia: A Pilot Multicenter Randomized Controlled Trial. *Clin. Infect. Dis.* 62 (2), 173–180. doi:10.1093/cid/civ808

Dilworth, T. J., Ibrahim, O., Hall, P., Sliwinski, J., Walraven, C., and Mercier, R. C. (2014). β-Lactams Enhance Vancomycin Activity against Methicillin-Resistant *Staphylococcus aureus* Bacteremia Compared to Vancomycin Alone. *Antimicrob. Agents Chemother.* 58 (1), 102–109. doi:10.1128/AAC.01204-13

Domenech, A., Ribes, S., Cabellos, C., Taberner, F., Tubau, F., Domínguez, M. A., et al. (2005). Experimental Study on the Efficacy of Combinations of Glycopeptides and Beta-Lactams against Staphylococcus Aureus with Reduced Susceptibility to Glycopeptides. *J. Antimicrob. Chemother.* 56 (4), 709–716. doi:10.1093/jac/dki294

Eliakim-Raz, N., Hellerman, M., Yahav, D., Cohen, J., Margalit, I., Fisher, S., et al. (2017). Trimethoprim/Sulfamethoxazole versus Vancomycin in the Treatment of Healthcare/Ventilator-Associated Mrsa Pneumonia: A Case-Control Study. *J. Antimicrob. Chemother.* 72 (3), 882–887. doi:10.1093/jac/dkw510

Espedido, B. A., Jensen, S. O., and van Hal, S. J. (2015). Ceftaroline Fosamil Salvage Therapy: An Option for Reduced-Vancomycin-Susceptible Mrsa Bacteraemia. *J. Antimicrob. Chemother.* 70 (3), 797–801. doi:10.1093/jac/dku455

Fowler, V. G., Boucher, H. W., Corey, G. R., Abrutyn, E., Karchmer, A. W., Rupp, M. E., et al. (2006). Daptomycin versus Standard Therapy for Bacteremia and Endocarditis Caused by Staphylococcus Aureus. *N. Engl. J. Med.* 355 (7), 653–665. doi:10.1056/NEJMoa053783

Fox, P. M., Lampen, R. J., Stumpf, K. S., Archer, G. L., and Climo, M. W. (2006). Successful Therapy of Experimental Endocarditis Caused by Vancomycin-Resistant Staphylococcus Aureus with a Combination of Vancomycin and Beta-Lactam Antibiotics. *Antimicrob. Agents Chemother.* 50 (9), 2951–2956. doi:10.1128/aac.00232-06

Geriak, M., Haddad, F., Rizvi, K., Rose, W., Kullar, R., LaPlante, K., et al. (2019). Clinical Data on Daptomycin Plus Ceftaroline versus Standard of Care Monotherapy in the Treatment of Methicillin-Resistant Staphylococcus Aureus Bacteremia. *Antimicrob. Agents Chemother.* 63 (5), 63. doi:10.1128/AAC.02483-18

Hagihara, M., Wiskirchen, D. E., Kuti, J. L., and Nicolau, D. P. (2012). *In Vitro* pharmacodynamics of Vancomycin and Cefazolin Alone and in Combination against Methicillin-Resistant *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 56 (1), 202–207. doi:10.1128/aac.05473-11

Hamed, K., Engelhardt, M., Jones, M. E., Saulay, M., Holland, T. L., Seifert, H., et al. (2020). Ceftobiprole versus Daptomycin in *Staphylococcus aureus* Bacteremia: a Novel Protocol for a Double-Blind, Phase III Trial. *Future Microbiol.* 15, 35–48. doi:10.2217/fmb-2019-0332

Hassoun, A., Linden, P. K., and Friedman, B. (2017). Incidence, Prevalence, and Management of Mrsa Bacteremia across Patient Populations-A Review of Recent Developments in Mrsa Management and Treatment. *Crit. Care* 21 (1), 211. doi:10.1186/s13054-017-1801-3

Holmes, N. E., Tong, S. Y., Davis, J. S., and van Hal, S. J. (2015). Treatment of Methicillin-Resistant Staphylococcus Aureus: Vancomycin and beyond. *Semin. Respir. Crit. Care Med.* 36 (1), 17–30. doi:10.1055/s-0034-1397040

Holubar, M., Meng, L., Alegria, W., and Deresinski, S. (2020). Bacteremia Due to Methicillin-Resistant Staphylococcus Aureus: An Update on New Therapeutic Approaches. *Infect. Dis. Clin. North. Am.* 34 (4), 849–861. doi:10.1016/j.idc.2020.04.003

Johnson, T. M., Molina, K. C., Miller, M. A., Kiser, T. H., Huang, M., and Mueller, S. W. (2021). Combination Ceftaroline and Daptomycin Salvage Therapy for Complicated Methicillin-Resistant Staphylococcus Aureus Bacteraemia Compared with Standard of Care. *Int. J. Antimicrob. Agents* 57 (4), 106310. doi:10.1016/j.ijantimicag.2021.106310

Lee, A. S., de Lencastre, H., Garau, J., Kluytmans, J., Malhotra-Kumar, S., Peschel, A., et al. (2018). Methicillin-Resistant Staphylococcus Aureus. *Nat. Rev. Dis. Primers* 4, 18033. doi:10.1038/nrdp.2018.33

Liu, C., Bayer, A., Cosgrove, S. E., Daum, R. S., Fridkin, S. K., Gorwitz, R. J., et al. (2011). Clinical Practice Guidelines by the Infectious Diseases Society of America for the Treatment of Methicillin-Resistant Staphylococcus Aureus Infections in Adults and Children. *Clin. Infect. Dis.* 52 (3), e18–55. doi:10.1093/cid/ciq146

McCreary, E. K., Kullar, R., Geriak, M., Zasowski, E. J., Rizvi, K., Schulz, L. T., et al. (2020). Multicenter Cohort of Patients with Methicillin-Resistant *Staphylococcus aureus* Bacteremia Receiving Daptomycin Plus Ceftaroline Compared with Other MRSA Treatments. *Open Forum Infect. Dis.* 7 (1), ofz538. doi:10.1093/ofid/ofz538

Ortwine, J. K., Werth, B. J., Sakoulas, G., and Rybak, M. J. (2013). Reduced Glycopeptide and Lipopeptide Susceptibility in Staphylococcus Aureus and the

"Seesaw Effect": Taking Advantage of the Back Door Left Open? *Drug Resist. Updat* 16 (3-5), 73–79. doi:10.1016/j.drup.2013.10.002

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The Prisma 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *BMJ* 372, n71. doi:10.1136/bmj.n71

Paul, M., Bishara, J., Yahav, D., Goldberg, E., Neuberger, A., Ghanem-Zoubi, N., et al. (2015). Trimethoprim-Sulfamethoxazole versus Vancomycin for Severe Infections Caused by Meticillin Resistant Staphylococcus Aureus: Randomised Controlled Trial. *BMJ* 350, h2219. doi:10.1136/bmj.h2219

Peacock, S. J., and Paterson, G. K. (2015). Mechanisms of Methicillin Resistance in Staphylococcus Aureus. *Annu. Rev. Biochem.* 84, 577–601. doi:10.1146/annurev-biochem-060614-034516

Pericàs, J. M., Moreno, A., Almela, M., García-de-la-Mària, C., Marco, F., Muñoz, P., et al. (2018). Efficacy and Safety of Fosfomycin Plus Imipenem versus Vancomycin for Complicated Bacteraemia and Endocarditis Due to Methicillin-Resistant Staphylococcus Aureus: A Randomized Clinical Trial. *Clin. Microbiol. Infect.* 24 (6), 673–676. doi:10.1016/j.cmi.2018.01.010

Rybak, M. J. (2006). The Pharmacokinetic and Pharmacodynamic Properties of Vancomycin. *Clin. Infect. Dis.* 42 (Suppl. 1), S35–S39. doi:10.1086/491712

Sader, H. S., Farrell, D. J., Flamm, R. K., and Jones, R. N. (2015). Activity of Ceftaroline and Comparator Agents Tested against Staphylococcus Aureus from Patients with Bloodstream Infections in Us Medical Centres (2009-13). *J. Antimicrob. Chemother.* 70 (7), 2053–2056. doi:10.1093/jac/dkv076

Sarkar, P., Yarlagadda, V., Ghosh, C., and Haldar, J. (2017). A Review on Cell Wall Synthesis Inhibitors with an Emphasis on Glycopeptide Antibiotics. *Medchemcomm* 8 (3), 516–533. doi:10.1039/c6md00585c

Shang, W., Hu, Q., Yuan, W., Cheng, H., Yang, J., Hu, Z., et al. (2016). Comparative Fitness and Determinants for the Characteristic Drug Resistance of St239-Mrsa-Iii-T030 and St239-Mrsa-Iii-T037 Strains Isolated in China. *Microb. Drug Resist.* 22 (3), 185–192. doi:10.1089/mdr.2015.0226

Tissot-Dupont, H., Gouriet, F., Oliver, L., Jamme, M., Casalta, J. P., Jimeno, M. T., et al. (2019). High-Dose Trimethoprim-Sulfamethoxazole and Clindamycin for Staphylococcus Aureus Endocarditis. *Int. J. Antimicrob. Agents* 54 (2), 143–148. doi:10.1016/j.ijantimicag.2019.06.006

Tong, S. Y. C., Lye, D. C., Yahav, D., Sud, A., Robinson, J. O., Nelson, J., et al. (2020). Effect of Vancomycin or Daptomycin with vs without an Antistaphylococcal β-Lactam on Mortality, Bacteremia, Relapse, or Treatment Failure in Patients with MRSA Bacteremia: A Randomized Clinical Trial. *JAMA* 323 (6), 527–537. doi:10.1001/jama.2020.0103

Ulloa, E. R., Singh, K. V., Geriak, M., Haddad, F., Murray, B. E., Nizet, V., et al. (2020). Cefazolin and Ertapenem Salvage Therapy Rapidly Clears Persistent Methicillin-Susceptible Staphylococcus Aureus Bacteremia. *Clin. Infect. Dis.* 71 (6), 1413–1418. doi:10.1093/cid/ciz995

Vestergaard, M., Frees, D., and Ingmer, H. (2019). Antibiotic Resistance and the Mrsa Problem. *Microbiol. Spectr.* 7 (2). doi:10.1128/microbiolspec.GPP3-0057-2018

Wang, C., Ye, C., Liao, L., Wang, Z., Hu, Y., Deng, C., et al. (2020). Adjuvant β-Lactam Therapy Combined with Vancomycin or Daptomycin for Methicillin-Resistant *Staphylococcus aureus* Bacteremia: a Systematic Review and Meta-Analysis. *Antimicrob. Agents Chemother.* 64 (11). doi:10.1128/aac.01377-20

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or any claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Xu, Lu, Song, Zhou, Li, Wang and Gao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Medicine

STUDY PROTOCOL
published: 30 May 2022
doi: 10.3389/fmed.2022.761419

Efficacy of the Integrative Acupuncture and Moxibustion Treatment in Patients With Major Depressive Disorder: The Study Protocol for a Multicenter, Single-Blinded, Randomized Trial in China

Yuan Zhang[1†], Yamin Liu[2†], Baile Ning[1†], Luda Yan[2†], Lihua Wu[2], Delong Zhang[3], Changhong Li[4], Wenwei Ouyang[5], Shengyong Su[6], Shuo Jiang[7], Guangcai Zhang[8], Junfeng Xu[9], Zhen Wang[10], Zhong Zheng[11], Dong Zheng[12], Shan Chen[1], Lu Sun[13] and Wenbin Fu[1]**

[1] Department of Acupuncture and Moxibustion, The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, China, [2] Shenzhen Bao'an Traditional Chinese Medicine Hospital, Guangzhou University of Chinese Medicine, Shenzhen, China, [3] School of Psychology, South China Normal University, Guangzhou, China, [4] College of Teacher Education, Guangdong University of Education, Guangzhou, China, [5] Key Unit of Methodology in Clinical Research, The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, China, [6] The First Affiliated Hospital of Guangxi Chinese Medical University, Nanning, China, [7] The First Affiliated Hospital of Zhejiang Chinese Medical University, Hangzhou, China, [8] Hainan Provincial Hospital of Chinese Medicine, Haikou, China, [9] The First Affiliated Hospital of Tianjin Chinese Medical University, Tianjin, China, [10] The Second Affiliated Hospital of Anhui Chinese Medical University, Hefei, China, [11] Sleep Medical Center, West China Hospital of Sichuan University, Chengdu, China, [12] Brain Hospital Affiliated Guangzhou Medical University, Guangzhou, China, [13] Department of Psychosomatic Medicine, Guangdong Provincial Hospital of Chinese Medicine, Guangzhou, China

Introduction: Antidepressants are the front-line treatments for major depressive disorder (MDD), but remain unsatisfactory in outcome. An increasing number of patients are interested in acupuncture and moxibustion treatment as complementary therapies. This study aims to evaluate the efficacy and safety of integrative acupuncture and moxibustion (iAM) treatment in patients with MDD.

Methods and Analysis: This multicenter, single-blind, 2 × 2 factorial randomized trial will enroll 592 patients with MDD of moderate severity from nine hospitals. All patients will be randomized, in a ratio of 2:2:2:1, through a computerized central randomization system, into four groups (the combined, iAM-only, sertraline-only, and placebo groups). Participants will undergo a 12-week intervention with either 50 mg of sertraline or a placebo once a day and active/sham iAM treatment three times per week. The primary outcome is depression severity, assessed using the Hamilton Depression Scale-17. The secondary outcomes include self-rated depression severity, anxiety, and sleep quality.

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Liyuan Tao,
Peking University Third Hospital,
China
Mustafa Bayraktar,
Atatürk University, Turkey
Shyam Odeti,
Carilion Clinic, United States

***Correspondence:**
Wenbin Fu
fuwenbin@gzucm.edu.cn
Lu Sun
sunlu197202@126.com

[†] These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Family Medicine and Primary Care, a section of the journal Frontiers in Medicine

Received: 19 August 2021
Accepted: 05 May 2022
Published: 30 May 2022

Citation:
Zhang Y, Liu Y, Ning B, Yan L, Wu L, Zhang D, Li C, Ouyang W, Su S, Jiang S, Zhang G, Xu J, Wang Z, Zheng Z, Zheng D, Chen S, Sun L and Fu W (2022) Efficacy of the Integrative Acupuncture and Moxibustion Treatment in Patients With Major Depressive Disorder: The Study Protocol for a Multicenter, Single-Blinded, Randomized Trial in China. Front. Med. 9:761419. doi: 10.3389/fmed.2022.761419

The primary and secondary outcomes will be measured at weeks 0, 4, 8, 12, and the 8th week posttreatment. Safety will be evaluated through liver and kidney function tests conducted before and after treatment and through monitoring of daily adverse events. An intent-to-treat principle will be followed for the outcome analyses.

Conclusion: This trial will provide sufficient evidence to ascertain whether iAM is effective and safe for treating MDD and provides a suitable combination strategy for treating MDD.

Clinical Trial Registration: [www.chictr.org.cn], identifier [ChiCTR2100042841].

Keywords: acupuncture, moxibustion, auricular acupuncture, integrative treatment, depressive disorders, sertraline

INTRODUCTION

Depressive disorders are common worldwide, and their prevalence reaches up to 4.4 and 4.2% of the global and Chinese populations, respectively (1). Major depressive disorder (MDD) is the most prevalent subtype and a leading cause of global disease burden (2). Estimates show that by 2030, this disorder is likely to rank high in the measure of total global disability-adjusted life years (DALY) (3). If left untreated, MDD can lead to suicide in 2.5% of patients (4). Thus, it is recognized as a major public health issue, having a substantial impact on individuals, families, and society (5, 6).

MDD is characterized by a lasting dysphoric mood, loss of interest or enjoyment, as well as fatigue or reduced energy. Patients with MDD often experience anxiety symptoms, sleep disturbances, cognitive impairment, and even other somatic symptoms (7–9). The treatment is currently based on symptom control, achieving remission, and restoring patients' function to baseline levels (10). The guideline-recommended treatments include a range of psychological interventions, such as behavioral activation, cognitive behavioral therapy, counseling, interpersonal psychotherapy, and pharmacological interventions such as first- and second-generation antidepressants. Despite the vast range of available treatments, a survey based on 17 countries has found that untreated rates of mental disorders are very high in low- and middle-income countries, ranging from 76 to 85% (11). A Chinese report illustrated that the number of patients with depressive symptoms, who received treatment for MDD was < 10% in four provinces (12).

The reasons for low treatment rates are inadequate medical resources, the social stigma associated with mental disorders, concerns about adverse drug reactions, and over-reliance on medication (12). These factors adversely affect patients, minimize their chances and motivation to receive any treatment, and are barriers to treatment. Due to the lack of professional psychologists, particularly in low- and middle-income countries, antidepressants are more frequently prescribed than psychological interventions (13). Second-generation antidepressants, such as selective serotonin reuptake inhibitors, are usually recommended as the front-line treatment for MDD based on the clinical guidelines (13–15). However, there are still debates regarding the effectiveness of medication because of limitations, such as delayed onset (16, 17), inadequate response in up to 50% of patients (4, 18), and lack of adherence in 30% (19). These limitations are additional barriers to effective treatment. Interestingly, patients nowadays show a preference for complementary therapies (12, 20).

Acupuncture is a therapy representative of Traditional Chinese Medicine (TCM) and is a common clinical practice in China (21). Internationally, acupuncture is used to treat depression and is recommended for primary care in the United Kingdom (22, 23). In 2016, the American College of Physicians recommended it as a complementary and alternative medicine in the clinical practice guidelines for adult patients with MDD (15). Currently, various styles of acupuncture are used in clinical practice, ranging from traditional/classical to modern, such as abdominal, auricular, and electro-acupuncture. The advantages of these different styles may vary, and two or more techniques may be applied jointly in clinical practice for diseases with complex symptoms (24). Moxibustion is another treatment that involves burning moxa and producing heat to stimulate acupoints and is mostly used with acupuncture to enhance the treatment's effect (25). Although MDD is not a simple disease consisting of a single symptom, most previous trials on treatment and acupuncture have only focused on the outcomes of a single-acupuncture therapy, which is deficient in terms of integrative intensity and treatment principles. Although there is an increasing accumulation of randomized trials on acupuncture and depression, these studies have methodological limitations and do not provide high-quality experimental evidence (26).

Since 2000, we have studied the effect of acupuncture on depression and formed an integrative acupuncture and moxibustion (iAM) treatment protocol called "Shugan Tiaoshen" (27). Specifically, this treatment protocol uses acupuncture and moxibustion to regulate the "Liver Qi" and spirit, which are associated with managing mood and energy in TCM theory. Our previous pilot studies (28, 29) showed that this integrative therapy could improve depressive symptoms and sleep quality in people with mild to moderate depression. Therefore, this protocol was designed as a large-sample multicenter 2 × 2 factorial randomized trial, which aimed to investigate the effectiveness and safety of this integrative therapy for patients with moderate MDD, as well as identify a suitable treatment strategy.

MATERIALS AND METHODS

Design and Setting

This study is being conducted at nine clinical hospitals in China, covering the northern, southern, eastern, and western parts of the country. The primary sponsor is the Second Affiliated Hospital of Guangzhou Chinese Medical University (Guangdong Provincial Hospital of Chinese Medicine, GPHCM). Ethical approval was obtained from the Ethics Committee of the GPHCM (No. BF2020-186). This trial was registered at the Chinese Clinical Trial Registry (ChiCTR2100042841), and all patients will be required to provide written informed consent to participate. The study is being conducted in compliance with the local regulations and international principles established in the Declaration of Helsinki. This protocol follows the guidelines of the Standards for Reporting Interventions in Clinical Trials of Acupuncture (30) and the Standard Protocol Items: Recommendations for Interventional Trials (31).

A total of 592 patients with MDD with moderate symptoms will potentially be recruited and randomly assigned per a 2 (active/sham iAM) × 2 (sertraline/placebo) factorial design with a ratio of 2:2:2:1 into active iAM and sertraline (hereafter referred to as combined treatment group), active iAM and placebo (iAM-only group), sham iAM and sertraline (sertraline-only group), and sham iAM and placebo (placebo group).

Eligibility Criteria

The inclusion criteria for this study are as follows: (1) diagnosis of MDD according to the Diagnostic and Statistical Manual of Mental Disorders 5th Edition criteria and the Standard for TCM Diseases and Syndromes Therapeutic Results, published by the Chinese TCM Authority (ZY/T001.1-94); (2) age of 18–65 years, with an educational level higher than junior middle school; and (3) with a written consent form. Patientes will be exclude if they meet the following criteria: (1) Hamilton Depression Scale-17 (HAMD-17) scores $\leq$17 or $>$24; (2) suicidal tendency; (3) treatment with antidepressants over the last 6 weeks before enrolment; (4) history of psychiatric diseases (schizophrenia, bipolar disorder, substance abuse, and others); (5) brain organic diseases, severe somatic diseases, patient or strong family history of epilepsy; (6) skin lesions or diseases, severe diabetes, tumor, and significant organ dysfunction or severe internal diseases; (7) pregnant, intending to become pregnant, or lactating.

Recruitment and Consent

Participants will be recruited from the nine centers via local advertising, newspapers, and the internet, and interested individuals will be instructed to contact research assistants by phone or email to make an appointment. During the first visit, psychological interviews will be conducted, evaluated, and recorded according to the inclusion and exclusion criteria to ensure each individual's eligibility. Eligible participants will be given detailed information about the trial and consent forms. Then, research assistants will obtain a signed form from those willing to participate. The study flow diagram is shown in **Figure 1**.

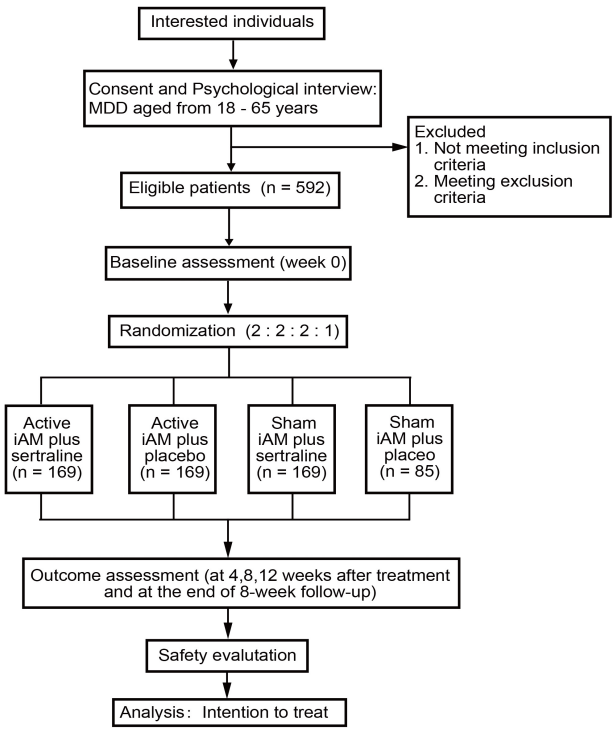

FIGURE 1 | Randomized trial flow diagram.

Randomization and Allocation

The allocation will be concealed using a central randomization method. A 2:2:2:1 permuted block sequence will be generated by the central randomization system of the GPHCM, which divides the eligible participants into either the combined treatment, iAM-only, sertraline-only, or placebo groups. The entire allocation process will be performed by a specific staff member.

Blinding

The participants, data managers, and statisticians will be blinded to the treatment allocations, which will not be revealed until the end of the study. However, this cannot be done for acupuncturists, because acupuncturists must provide the proper iAM treatment, active or sham, according to the patients' allocations. Participants will be given the corresponding labeled bottle filled with either medication or placebo pills, which will be identical in appearance. Each participant will be treated with acupuncture and moxibustion in a separate room to ensure adequate privacy. Due to the nature of auricular acupuncture, acupuncturists rather than participants will be instructed to remove the needles. In addition, acupuncturists and researchers will be instructed not to communicate with the participants about the possibility of their allocations.

Intervention

The participants will receive active/sham integrative acupuncture and moxibustion treatment three times a week, with each

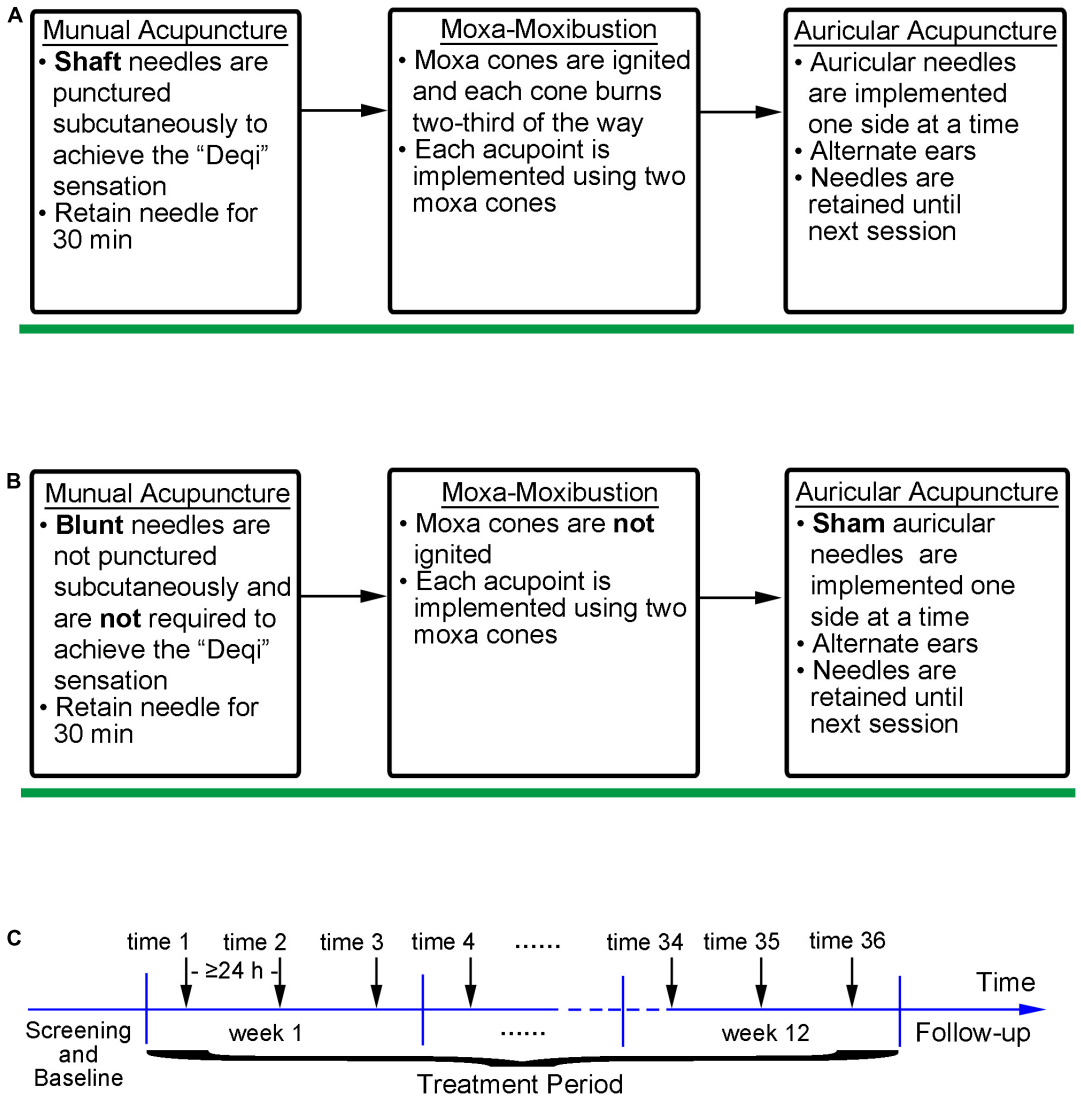

FIGURE 2 | The iAM implementation and intervention processes. **(A)** Implementation process of active iAM intervention; **(B)** Implementation process of sham iAM intervention; **(C)** Intervention frequency of the iAM.

interval > 24 h, for a total of 36 times for 12 consecutive weeks. The integrative treatment process is illustrated in **Figure 2**.

The Integrative Acupuncture-Moxibustion Treatment and Sham Method

The iAM treatment protocol comprises three interventions: manual acupuncture, moxibustion, and auricular acupuncture (**Figure 2**). For the manual acupuncture, patients will be asked to lie in the supine position, wearing eye masks for a better curative effect. After skin disinfection with 75% alcohol, acupuncture will be performed at the following acupoints: Baihui (GV20), Yintang (GV29), Touwei (ST8), Jiuwei (CV15), Zhongwan (CV12), Qihai (CV6), Taichong (LR3), Hegu (LI4), and Sanyinjiao (SP6) using tube needles (Hwato Suzhou Medical Instruments, Suzhou, China). The locations of the acupoints are shown in **Figure 3**. Needle lifting, thrusting, and twisting will be used to achieve the "Deqi" sensation, which is believed to indicate effective needling based on the TCM theory (32). The needles will be retained *in situ* for 30 min and then removed. The same procedure will be performed for the sham conditions, but without percutaneous puncture using blunt-tipped needles, meaning that the "Deqi" sensation will not be achieved. Detailed information about the active/sham acupuncture method and needle sizes are shown in **Table 1** and **Figure 4**.

After acupuncture, moxibustion will be performed (**Figure 2**). Patients will be asked to lie in the prostrate position, wearing eye masks. The alternate groups of selected acupoints are Feishu (BL13), Geshu (BL17), Danshu (BL19), and Yongquan (KI1) in one group, and Pohu (BL42), Geguan (BL46), Yanggang (BL48), and Yongquan (KI1) in the other (**Figure 5**). Wanhua oil will be smeared evenly on the skin to fix moxa cones and prevent heating. A moxa cone, 2 mm in

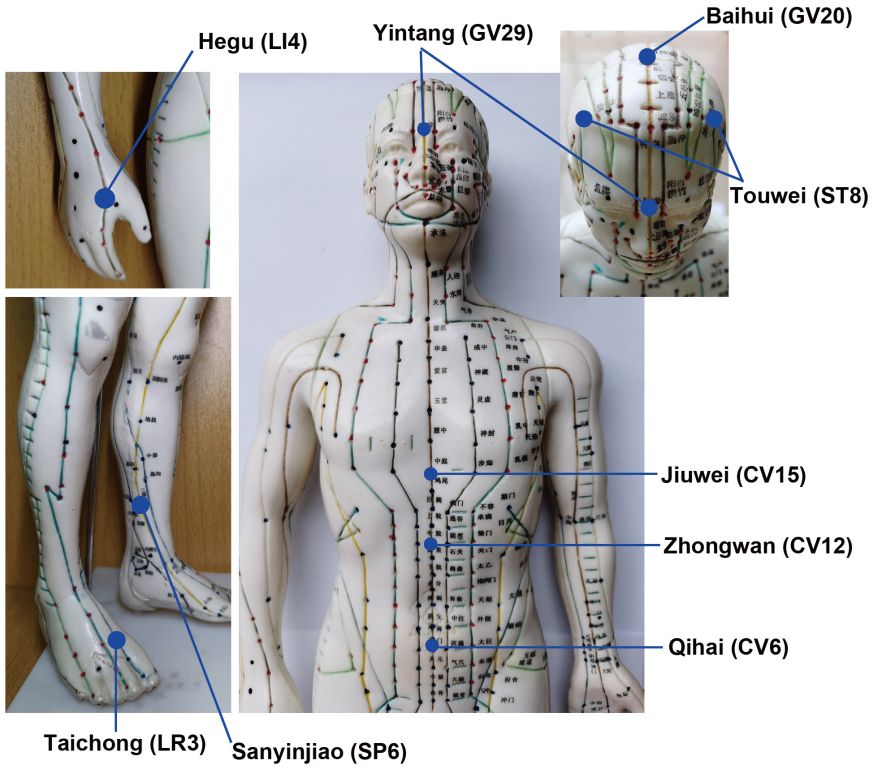

FIGURE 3 | Acupuncture acupoints and locations.

TABLE 1 | Details regarding active acupuncture and sham methods.

No.	Acupoint	Active acupuncture		Sham acupuncture	
		Needing method	Needles sizes	Needing method	Needle size
1	Baihui (GV20)	Punctured perpendicularly reaching the periosteum	0.25 × 25 mm	All needle slightly pressed but did not penetrate the skin	0.40 × 25 mm
2	Yintang (GV29)	Punctured perpendicularly reaching the periosteum	0.25 × 25 mm		0.40 × 25 mm
3	Touwei (ST8)	Punctured perpendicularly reaching the periosteum	0.25 × 25 mm		0.40 × 25 mm
4	Jiuwei (CV15)	Punctured perpendicularly 0.3 cun	0.22 × 40 mm		0.40 × 40 mm
5	Zhongwan (CV12)	Punctured perpendicularly 1 cun	0.22 × 40 mm		0.40 × 40 mm
6	Qihai (CV6)	Punctured perpendicularly 1 cun	0.22 × 40 mm		0.40 × 40 mm
7	Taichong (LR3)	Punctured perpendicularly 0.5 cun	0.25 × 25 mm		0.40 × 25 mm
8	Hegu (LI4)	Punctured perpendicularly 0.5 cun	0.25 × 25 mm		0.40 × 25 mm
9	Sanyinjiao (SP6)	Punctured perpendicularly 0.5 cun	0.25 × 25 mm		0.40 × 25 mm

1 cun = 25 mm.

diameter and 3 mm in height, will be placed and ignited using a joss stick (**Figure 6**). It will be removed when two-thirds of it have burned off. Two moxa cones will be applied at each point. In sham moxibustion, the moxa cones will not be ignited.

Finally, auricular acupuncture will be performed. The acupoints of the heart (CO15), liver (CO12), and kidney (CO10) will be selected for the intervention (**Figure 7A**). Auricular needles (33) (0.2 mm in diameter, 0.6 mm in length; Seirin Co Ltd., Shimize-City, Japan) will be used on the points of the unilateral auricle and removed by the acupuncturists in the subsequent treatment. Auricular acupuncture will be conducted on one side of the auricle for one treatment and the other side for the next treatment, alternating between the two sides of auricles. For sham conditions, fake auricular needles will be used and not placed (**Figures 7B,C**).

The participants in the combined treatment and the iAM-only groups will receive the iAM treatment, while in sertraline-only and placebo group will receive the sham iAM treatment.

Treatment Vs. Placebo

The participants in the combined treatment and sertraline-only groups will be given a bottle filled with the antidepressant

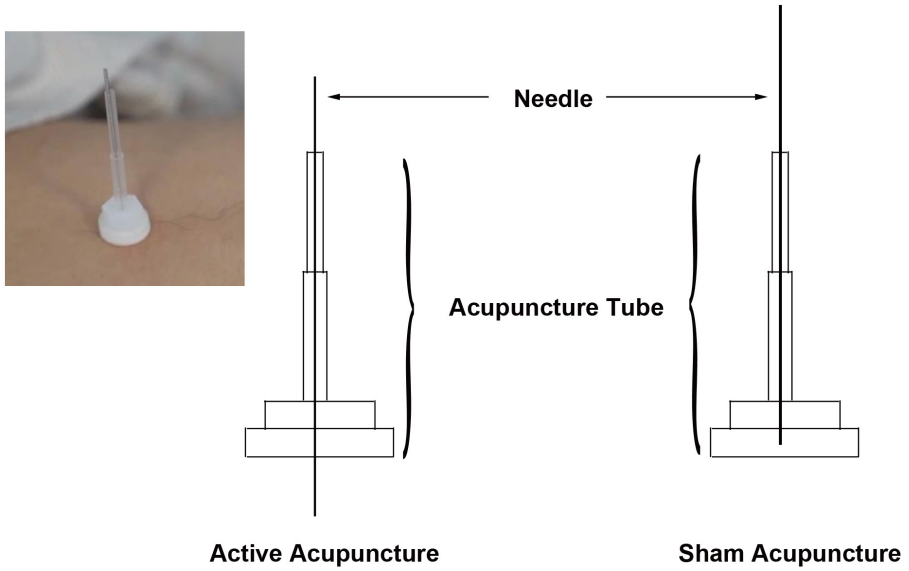

FIGURE 4 | Active vs. sham acupuncture. Shaft needles are punctured subcutaneously in active acupuncture, while blunt needles are not punctured subcutaneously in sham acupuncture.

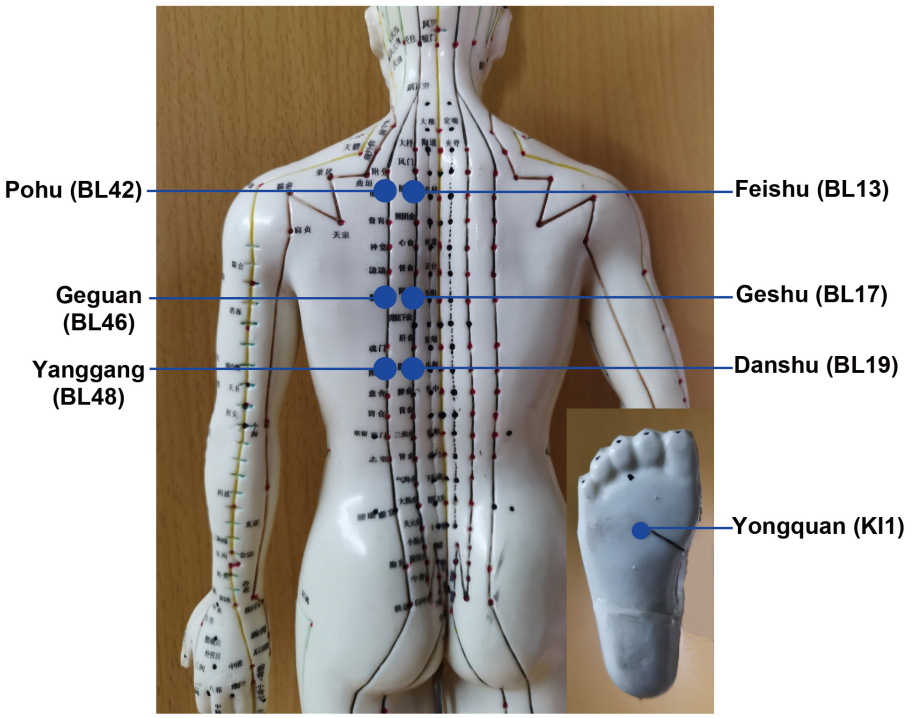

FIGURE 5 | Moxibustion acupoints and locations.

medication sertraline (Approval No. h10980141, Pfizer Co., Ltd.), and instructed to take 50 mg orally after breakfast, once a day, continuously for 12 weeks. The participants in the iAM-only and placebo groups will receive a bottle filled with placebo pills made of starch instead. The placebo pills do not differ from the antidepressants in appearance.

Due to changes in the participants' illness, there will be no prohibitions on medication changes during the trial. However, any changes will be checked by psychiatrists and recorded.

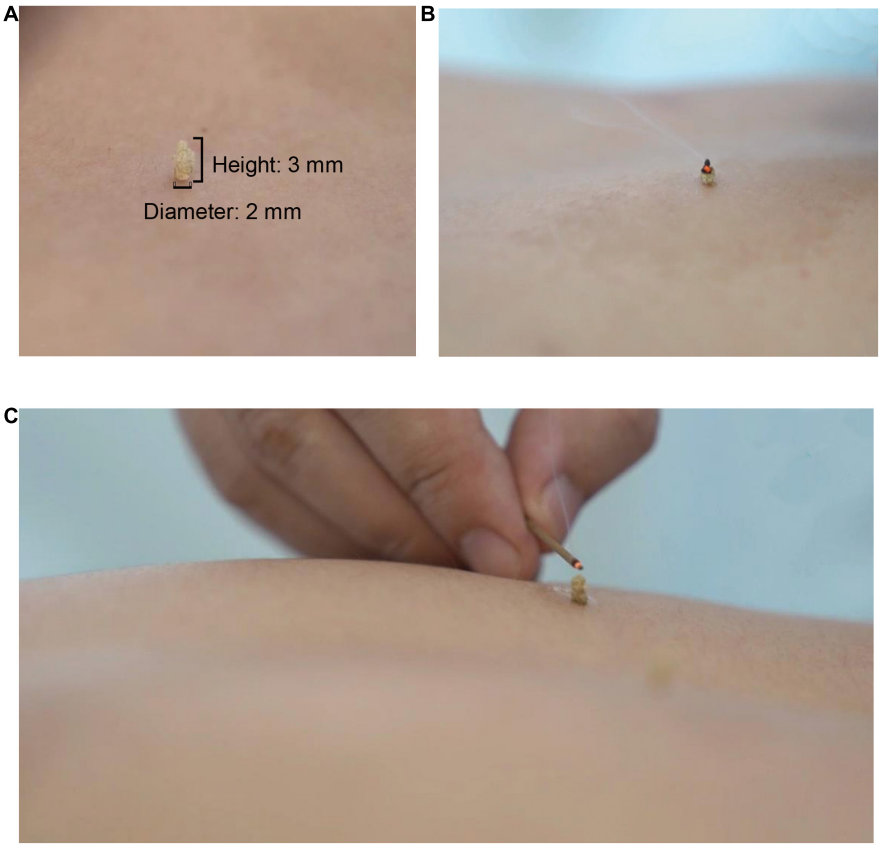

FIGURE 6 | Moxibustion intervention. **(A)** Presents the size of a moxa cone; **(B)** and **(C)** present the demonstration of lighting a moxa cone.

Study Visits and Measures

Five visits will be performed in this study. The study schedules and measures are displayed in **Table 2**. During the screening visit, patients will undergo psychological interviews to screen for eligibility according to the inclusion and exclusion criteria. Diagnosis of TCM syndrome types, demographic information, and medical history will be obtained. Eligible patients will then undergo assessment and baseline data collection. Three visits will take place during the treatment period (weeks 4, 8, and 12), with one follow-up visit at week 20. The study assessments will include HAMD-17, the Patient Health Questionnaire-9 (PHQ-9), the Hamilton Anxiety Scale (HAMA), the Pittsburgh Sleep Quality Index (PSQI), sleep actigraphy, and adverse event reporting. Liver and kidney function tests will also be performed to assess safety, and blood samples will be collected only at baseline (week 0) and after treatment (week 12). Data collection and analysis will be conducted by three research assistants blinded to the treatment allocations.

Outcome Measures

Primary Outcomes

The HAMD-17 is an established clinician-rated assessment of depressive symptom severity and encompasses psychological and somatic symptoms (34). It assesses the severity of 17 depression symptom items over the previous week and has five factors: anxiety/somatization, cognitive impairment, "psychomotor retardation," weight, and sleep changes. A higher HAMD-17 score implies more severe depressive symptoms. Severity is classified as no depression (score ≤ 7), mild (score 8–16), moderate (score 17–24), and severe (score ≥ 25). A score of <8 will be interpreted as remission, and a decrease greater than 50% as an effective treatment response. This measure has been shown to be reliable (35, 36). And in this study, the HAMD-17 scores measured in week 12 after treatment are defined as the primary outcome.

Secondary Outcomes

The PHQ-9 is a self-administered version of the Primary Care Evaluation of Mental Disorders diagnostic instrument for depression disorders (37). Its score for each criterion ranges from 0 ("not at all") to 3 ("nearly every day") (37). Thus, the higher the score, the more severe the patient's depression. It is sensitive to changes and is used as a brief depression severity measurement (38, 39). The 14-item version of the HAMA is designed to standardize clinical assessments of anxiety, where the scores for each item are rated on a 5-point scale from 0 ("not at all") to 4 ("nearly every day"). HAMA-14 classifies anxiety factors into physical and mental symptoms. Items 1–6 and 14 reflect mental symptoms, while items 7–13 reflect somatic anxiety symptoms.

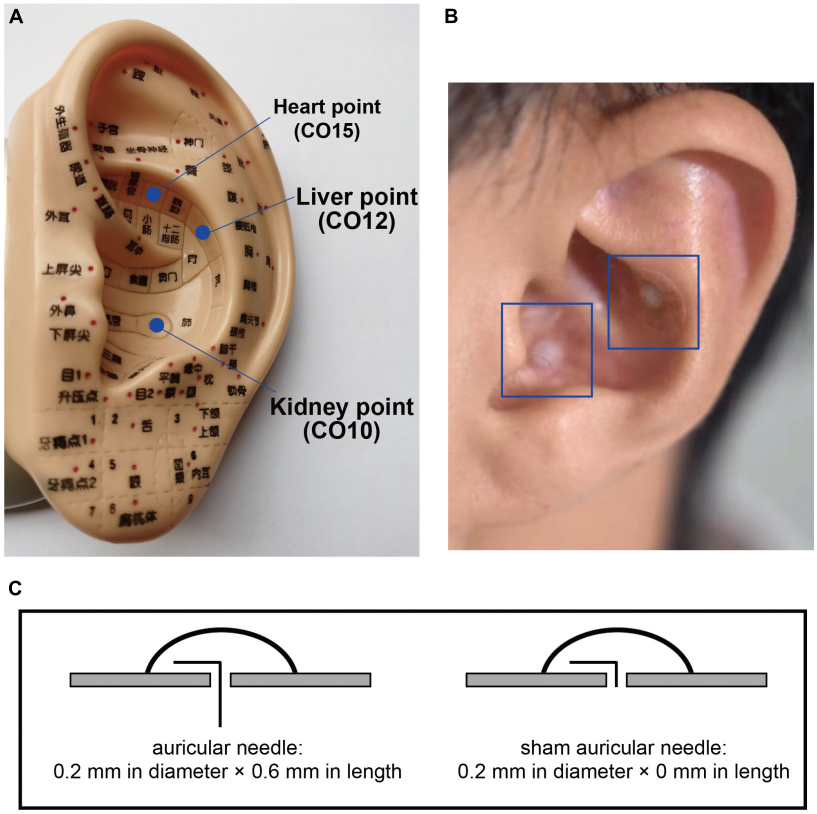

FIGURE 7 | Auricular points and needles. The locations of selected auricular points are shown in **(A)**; the intervention methods of the auricular needles and sham needles are shown in **(B)** and **(C)**.

The PSQI is a self-rated questionnaire that assesses sleep quality and disturbances for up to a 1-month period. It is a recommended measure for treatment effectiveness studies of global sleep quality. From the 19 included items, seven sub-scores will be calculated, including sleep quality, latency, habitual efficiency, duration, disturbances, medication use, and daytime dysfunction (40). The sum of these components yields a global score of 21, with a higher score indicating poorer sleep quality. A PSQI score > 5 implies poor sleep, and a three-point change suggests a clinical effect. In addition, actigraphy, which used a bracelet-like monitor, will be used as a supplemental tool to assess sleep quality (YWK-P9; BOZHILUN, Inc., Shenzhen, Guangdong, China). It can monitor data regarding real-time sleep quality, blood oxygen saturation, blood pressure, heart rate, exercise situation, and other indicators, which can be used to assess the physical condition. In our study, we will use this method to collect quantitative sleep data and supplement sleep quantity information. Each participant will have a smartphone application connected by Bluetooth to the actigraphy to collect data.

Safety Assessment

Venous blood samples will be collected through venipuncture to assess liver and kidney function at baseline and posttreatment. Adverse event records will be monitored and recorded at every visit to assess the safety of the intervention.

Quality Control

Before the start of the recruitment period, a training workshop will be organized for the entire research team, including acupuncturists and research assistants. Acupuncturists who will provide the treatment are licensed by the Ministry of Health of the People's Republic of China and will have received training in the application of integrative acupuncture and moxibustion treatment. These requirements will be ensured before the trial to ensure strict adherence to the study protocol and familiarity with the administration process. The research team will also be supplied with a written protocol and standard operating procedure documents.

Data will be collected from information recorded in the case report forms and an online data management platform (eMedInform, LinkerMedTechCo., Ltd., Beijing, China), a data recording and control system for clinical scientific research. All modifications will be marked on the case report forms and system. Data quality will be checked regularly by research assistants, which will be overseen by monitors. Audits will be performed regularly by the GPHCM Department of Science Research. Data monitoring will be conducted regularly, adhering to standard operating procedures by the Guangdong

TABLE 2 | Schedule of study visits and assessments during the enrolment, treatment, and follow-up periods.

		Treatment period			Follow-up
	Screening and baseline	Week 4	Week 8	Week 12	Week 20
Inclusion/exclusion criteria	×				
Consent	×				
Allocation	×				
Active/Sham iAM intervention		×	×	×	
Sertraline/placebo drug intervention		×	×	×	
Demographic information	×				
Medical history taking	×				
TCM Syndrome Types	×				
PHQ-9	×	×	×	×	×
HAMD-17	×	×	×	×	×
HAMA	×	×	×	×	×
PSQI	×	×	×	×	×
Sleep Actigraphy measures	×	×	×	×	×
Blood sample for liver and kidney function tests	×				×
Adverse events report		×	×	×	×
Adherence scale		×	×	×	×
Medication record		×	×	×	×

iAM, integrative acupuncture and moxibustion; TCM, Traditional Chinese Medicine; PHQ-9, Patient Health Questionnaire; HAMD-17, Hamilton Depression Scale-17; HAMA, Hamilton Anxiety Scale; PSQI, Pittsburgh Sleep Quality Index; ×, required item.

International Clinical Research Center of Chinese Medicine (Guangzhou, China). In case of participant withdrawal during the treatment period or the follow-up phase, the reasons shall be clarified only if the participant is willing and the rate statistically analyzed.

Statistical Analysis

Efficacy analyses will be performed based on the intent-to-treat (ITT) principle unless stated otherwise. We have defined two study populations: full analysis set (FAS) and per-protocol set (PPS). The FAS is defined as all treated patients with a baseline HAMD-17 score and at least one HAMD-17 score on treatment. PPS includes only those patients who all completed the allocated treatment. Demographic and other basic characteristics will be summarized according to the assigned groups. Categorical variables will be summarized as frequencies and percentages, and continuous variables as mean ± SD for normally distributed and median for non-normally distributed data.

Two separate analyses of the primary outcomes will be performed. First, for the long-term effect, an analysis of variance model with a 95% confidence interval will be used to compare the mean change in the HAMD score among the groups without adjusting for any covariates. Second, a linear mixed model will be performed. For the secondary outcomes, the chi-square test will be used to assess the effect of the intervention on binary outcomes. Logistic regression analysis or generalized estimating equations will be used to analyze the categorical outcomes. The reliability, validity, and item responses of the questionnaires will also be included in our analyses. Safety analyses will be performed on safety set (SS), i.e., all randomized participants administered at least one treatment. All statistical analyses will be performed using SAS 9.3 or SPSS 18.0. All tests of significance will be two-sided, with a maximal type I error risk of 5%.

Sample Size Calculation

The sample size was calculated based on the mean and SD of the HAMD-17. According to our previous pilot study, the data measured after treatment for 8 weeks can be considered clinically significant (10.47 ± 6.24 in the combined treatment, 12.03 ± 5.78 in the iAM-only, 12.17 ± 5.78 in the sertraline-only, and 19.11 ± 3.92 in the placebo groups). PASS 11.0 was applied to calculate sample size. We applied a 2:2:2:1 ratio to the four groups. For a statistical power of 90%, with a two-tailed significance level of 5%, the sample size was 72 for the placebo group and 144 for each of the other three groups (alpha = 0.05, beta = 0.10, n = 503). Assuming a dropout rate of approximately 15%, we required a sample size of 85 for the placebo group and 169 for each of the remaining three groups.

DISCUSSION

MDD is a prevalent disorder with a low treatment rate and significant burden (12). Pharmaceutical therapies are the guideline-recommended front-line treatments (15, 41), but 30–50% of patients with MDD do not respond to antidepressant medications (4). Moreover, the side effects of these treatments often result in low adherence rates and limited efficacy. As traditional Chinese therapies, acupuncture and moxibustion are popular and widely accepted. An increasing number of studies in the literature have demonstrated that acupuncture and related treatments may be safe and effective for depressive disorders (42, 43).

Although the antidepressant qualities of acupuncture are widely recognized, a single-acupuncture technique can not completely treat the clinical issue due to the complexity and diversity of depressive symptoms. We formulated iAM treatments to improve the efficiency of single-acupuncture therapy. The acupoint selection protocol was based on the TCM theory of regulating "Liver Qi" and spirit, which, according to this theory, play essential roles in managing mood and energy. If "Liver Qi" stagnates, it may negatively affect mood and result in apathy and low self-esteem (29). The integrative treatment process is formulated in three steps; the first step applies acupuncture to regulate "Liver Qi" and spirit, the second applies moxibustion to augment the effect, and the third applies auricular acupuncture to consolidate and extend the curative effect.

To our knowledge, this study is the first multicenter, large-sample, 2 × 2 factorial randomized trial to date assessing the efficacy and safety of iAM treatment in patients with MDD.

In this regard, this will be the most comprehensive nationwide trial in terms of the participants, as they will be recruited from nine clinical centers covering seven provinces and eight cities in China. In this four-armed clinical trial, we will focus on the different efficiencies among active iAM plus antidepressants, active iAM plus placebo, sham iAM plus antidepressants, and sham iAM plus placebo to treat MDD. We will then investigate whether iAM positively affects MDD, as well as any augmentative effects and safety concerns when combined with sertraline. Given these outcomes, we can assume that this treatment program could play an essential role in preventing depression. In addition, this study is also expected to form a standardized, applicable integrative regimen to promote extensive implementation in improving depressive disorders.

This study has two limitations. Due to the nature of intervention in clinical trials, the study cannot be conducted in an ideal blind-controlled setting. Moreover, the outcomes of our study cannot be interpreted as effective for any one of these three therapies in iAM treatments when used separately.

CONCLUSION

This clinical trial is designed to prove the hypothesis that integrative acupuncture and moxibustion treatment can augment the effectiveness and safety of conventional antidepressant therapy. The results will also provide more evidence regarding the effectiveness and safety of integrative acupuncture and moxibustion, potentially leading to alterations or supplementations to the routine treatment strategy.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of The Second Affiliated Hospital of Guangzhou Chinese Medical University (No. BF2020-186). The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

YZ, BN, LS, and WF designed the study protocol. YL, LY, LW, and DZ assisted in the development and operationalization of the study methods. YZ, YL, BN, LY, LW, and DZ wrote the grant proposal. YZ and YL drafted the manuscript. YZ, YL, BN, LY, and LS participated in coordinating the study. CL and DZ assisted with data collection activities. WO generated the random allocation sequences for randomization and was the trial statistician. SS, SJ, GZ, JX, ZW, ZZ, DZ, and SC developed and delivered study implementation strategies in collaboration with YZ, YL, and LW. All authors critically reviewed and approved the final manuscript.

FUNDING

This study was supported by grants from the Special Project for Research and Development in Key areas of Guangdong Province (No: 2020B1111100007), Guangdong Provincial Hospital of Traditional Chinese Medicine for Scientific and Technological Research of TCM in the Special Funding (YN2019ML13), Shenzhen Bao'an Research Center for Acupuncture and Moxibustion (BAZJ2018239), Sanming Project of Medicine in Shenzhen (SZSM201806077), and the Project of "Double First-Class" and "High-level" University Discipline Collaborative Innovation Team of Guangzhou University of Chinese Medicine (No: 2021xk22).

ACKNOWLEDGMENTS

We would like to thank all the researchers and staff at participating clinical institutions for their valuable contribution to ensure the implementation of this study protocol. We also would like to thank Editage (www.editage.com) for English language editing. This study protocol was posted as a preprint on Research Square (44).

REFERENCES

1. World Health Organization. *Depression and Other Common Mental Disorders: Global Health Estimates*. Geneva: World Health Organization (2017).
2. World Health Organization. *The Global Burden of Disease: 2004 Update*. Geneva: World Health Organization (2008).
3. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med.* (2006) 3:11. doi: 10.1371/journal
4. Ng CH, Kato T, Han C, Wang G, Trivedi M, Ramesh V, et al. Definition of treatment-resistant depression – Asia Pacific perspectives. *J Affect Disord.* (2019) 245:626–36. doi: 10.1016/j.jad.2018.11.038
5. Kennedy SH, Lam RW, Morris B, Grp CDW. Clinical guidelines for depressive disorders – summary of recommendations relevant to family physicians. *Can Fam Physician.* (2003) 49:489–91.
6. Demyttenaere K, Bruffaerts R, Posada-Villa J, Gasquet I, Kovess V, Lepine JP, et al. Prevalence, severity, and unmet need for treatment of mental disorders in the World Health Organization world mental health surveys. *JAMA.* (2004) 291:2581–90. doi: 10.1001/jama.291.21.2581
7. Ter Meulen WG, Draisma S, van Hemert AM, Schoevers RA, Kupka RW, Beekman ATF, et al. Depressive and anxiety disorders in concert-A synthesis of findings on comorbidity in the NESDA study. *J Affect Disord.* (2021) 284:85–97. doi: 10.1016/j.jad.2021.02.004
8. Luca A, Luca M, Calandra C. Sleep disorders and depression: brief review of the literature, case report, and nonpharmacologic interventions for depression. *Clin Interv Aging.* (2013) 8:1033–8. doi: 10.2147/cia.s47230
9. Richardson L, Adams S. Cognitive deficits in patients with depression. *J Nurse Pract.* (2018) 14:437–43. doi: 10.1016/j.nurpra.2018.03.006
10. Gelenberg AJ. Using assessment tools to screen for, diagnose, and treat major depressive disorder in clinical practice. *J Clin Psychiatry.* (2010) 71(Suppl. E1):e01. doi: 10.4088/JCP.9058se1c.01gry
11. Wang PS, Aguilar-Gaxiola S, Alonso J, Angermeyer MC, Borges G, Bromet EJ, et al. Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. *Lancet.* (2007) 370:841–50. doi: 10.1016/s0140-6736(07)61414-7
12. Phillips MR, Zhang J, Shi Q, Song Z, Ding Z, Pang S, et al. Prevalence, treatment, and associated disability of mental disorders in four provinces in

China during 2001-05: an epidemiological survey. *Lancet.* (2009) 373:2041–53. doi: 10.1016/s0140-6736(09)60660-7
13. Cipriani A, Furukawa TA, Salanti G, Chaimani A, Atkinson LZ, Ogawa Y, et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *Lancet.* (2018) 391:1357–66. doi: 10.1016/S0140-6736(17)32802-7
14. Gelenberg AJ, Freeman MP, Markowitz JC, Rosenbaum JF, Thase ME, Trivedi MH, et al. *Practice Guideline for the Treatment of Patients With Major Depressive Disorder*. Washington, DC: American Psychiatric Association (2010).
15. Qaseem A, Barry MJ, Kansagara D, Amer Coll P. Nonpharmacologic versus pharmacologic treatment of adult patients with major depressive disorder: a clinical practice guideline from the American college of physicians. *Ann Intern Med.* (2016) 164:350–9. doi: 10.7326/m15-2570
16. Mnie-Filali O, Faure C, Lambás-Señas L, El Mansari M, Belblidia H, Gondard E, et al. Pharmacological blockade of 5-HT7 receptors as a putative fast acting antidepressant strategy. *Neuropsychopharmacology.* (2011) 36:1275–88. doi: 10.1038/npp.2011.13
17. Gao N, Tiliwaerde M, Zheng W, Xiong J, Li X, Jin Z. Neuropharmacological and antidepressant-like effects of ZY-1408: A novel serotonin/norepinephrine reuptake inhibitor and serotonin receptor 2C antagonist. *Neuropharmacology.* (2021) 182:108376. doi: 10.1016/j.neuropharm.2020.108376
18. Fava M. Diagnosis and definition of treatment-resistant depression. *Biol Psychiatry.* (2003) 53:649–59. doi: 10.1016/s0006-3223(03)00231-2
19. Rossom R, Shortreed S, Coleman K, Beck A, Waitzfelder B, Stewart C, et al. Antidepressant adherence across diverse populations and healthcare settings. *Depress Anxiety.* (2016) 33:765–74. doi: 10.1002/da.22532
20. Jorm AF, Medway J, Christensen H, Korten AE, Jacomb PA, Rodgers B. Public beliefs about the helpfulness of interventions for depression: effects on actions taken when experiencing anxiety and depression symptoms. *Aust N Z J Psychiatry.* (2000) 34:619–26. doi: 10.1080/j.1440-1614.2000.00761.x
21. Robinson N, Lorenc A, Ding W, Jia J, Bovey M, Wang X. Exploring practice characteristics and research priorities of practitioners of traditional acupuncture in China and the EU-A survey. *J Ethnopharmacol.* (2012) 140:604–13. doi: 10.1016/j.jep.2012.01.052
22. Hopton AK, Curnoe S, Kanaan M, MacPherson H. Acupuncture in practice: mapping the providers, the patients and the settings in a national cross-sectional survey. *BMJ Open.* (2012) 2:e000456. doi: 10.1136/bmjopen-2011-000456
23. MacPherson H, Richmond S, Bland M, Brealey S, Gabe R, Hopton A, et al. Acupuncture and counselling for depression in primary care: a randomised controlled trial. *PLoS Med.* (2013) 10:e1001518. doi: 10.1371/journal.pmed.1001518
24. Sherman KJ, Hogeboom CJ, Cherkin DC. How traditional Chinese medicine acupuncturists would diagnose and treat chronic low back pain: results of a survey of licensed acupuncturists in Washington State. *Complement Ther Med.* (2001) 9:146–53. doi: 10.1054/ctim.2001.0458
25. Tian C, Fan Y, Xu J, Huang Y, Wang W, Wang S, et al. The efficacy and safety of acupuncture and moxibustion combined with western medicine for obsessive-compulsive disorder: a protocol for systematic review and meta-analysis. *Medicine.* (2020) 99:e21395. doi: 10.1097/md.0000000000021395
26. Li M, Niu J, Yan P, Yao L, He W, Wang M, et al. The effectiveness and safety of acupuncture for depression: an overview of meta-analyses. *Complement Ther Med.* (2020) 50:102202. doi: 10.1016/j.ctim.2019.102202
27. Fan L, Gong J, Fu W, Chen Z, Xu N, Liu J, et al. Gender-related differences in outcomes on acupuncture and moxibustion treatment among depression patients. *J Altern Complement Med.* (2015) 21:673–80. doi: 10.1089/acm.2015.0068
28. Fan L, Fu W, Chen Z, Xu N, Liu J, Lü A, et al. Curative effect of acupuncture on quality of life in patient with depression: a clinical randomized single-blind placebo-controlled study. *J Tradit Chin Med.* (2016) 36:151–9. doi: 10.1016/s0254-6272(16)30021-8
29. Wen X, Wu Q, Liu J, Xu Z, Fan L, Chen X, et al. Randomized single-blind multicenter trial comparing the effects of standard and augmented acupuncture protocols on sleep quality and depressive symptoms in patients with depression. *Psychol Health Med.* (2018) 23:375–90.
30. Moher D, Hopewell S, Schulz K, Montori V, Gøtzsche P, Devereaux P, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ.* (2010) 340:c869. doi: 10.1136/bmj.c869
31. Chan A, Tetzlaff J, Altman D, Laupacis A, Gøtzsche P, Krleža-Jerić K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Ann Intern Med.* (2013) 158:200–7. doi: 10.7326/0003-4819-158-3-201302050-00583
32. Lai X, Zhang G, Huang Y, Tang C, Yang J, Wang S, et al. A cerebral functional imaging study by positron emission tomography in healthy volunteers receiving true or sham acupuncture needling. *Neurosci Lett.* (2009) 452:194–9. doi: 10.1016/j.neulet.2009.01.052
33. Margolin A, Kleber HD, Avants SK, Konefal J, Gawin F, Stark E, et al. Acupuncture for the treatment of cocaine addiction: a randomized controlled trial. *JAMA.* (2002) 287:55–63. doi: 10.1001/jama.287.1.55
34. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry.* (1960) 23:56–62. doi: 10.1136/jnnp.23.1.56
35. Trajkovic G, Starcevic V, Latas M, Lestarevic M, Ille T, Bukumiric Z, et al. Reliability of the Hamilton rating scale for depression: a meta-analysis over a period of 49 years. *Psychiatry Res.* (2011) 189:1–9. doi: 10.1016/j.psychres.2010.12.007
36. Schwab JJ, Bialow MR, Holzer CE. A comparison of two rating scales for depression. *J Clin Psychol.* (1967) 23:94–6. doi: 10.1002/1097-4679(196701)23:13.0.co;2-k
37. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9–validity of a brief depression severity measure. *J Gen Intern Med.* (2001) 16:606–13. doi: 10.1046/j.1525-1497.2001.016009606.x
38. Lowe B, Schenkel I, Carney-Doebbeling C, Gobel C. Responsiveness of the PHQ-9 to psychopharmacological depression treatment. *Psychosomatics.* (2006) 47:62–7. doi: 10.1176/appi.psy.47.1.62
39. McMillan D, Gilbody S, Richards D. Defining successful treatment outcome in depression using the PHQ-9: a comparison of methods. *J Affect Disord.* (2010) 127:122–9. doi: 10.1016/j.jad.2010.04.030
40. Buysse DJ, Reynolds CF III, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry Res.* (1989) 28:193–213. doi: 10.1016/0165-1781(89)90047-4
41. van der Lem R, van der Wee NJA, van Veen T, Zitman FG. The generalizability of antidepressant efficacy trials to routine psychiatric out-patient practice. *Psychol Med.* (2011) 41:1353–63. doi: 10.1017/s0033291710002175
42. Zhao B, Li Z, Wang Y, Ma X, Wang X, Wang X, et al. Can acupuncture combined with SSRIs improve clinical symptoms and quality of life in patients with depression? Secondary outcomes of a pragmatic randomized controlled trial. *Complement Ther Med.* (2019) 45:295–302. doi: 10.1016/j.ctim.2019.03.015
43. Kim M, Choi E, Kwon O, Park H, Kim A, Seo B, et al. Electroacupuncture plus moxibustion for major depressive disorder: a randomized, sham-controlled, pilot clinical trial. *Integr Med Res.* (2021) 10:100727. doi: 10.1016/j.imr.2021.100727
44. Zhang Y, Liu Y, Ning B, Yan L, Wu L, Zhang D, et al. Implementation and evaluation of the integrated acupuncture and moxibustion treatment (acupuncture, moxibustion and auricular acupuncture) for major depressive disorder: a trial protocol for a randomized clinical trial. *Res Square.* [Preprint]. (2021). doi: 10.21203/rs.3.rs-801003/v1

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Zhang, Liu, Ning, Yan, Wu, Zhang, Li, Ouyang, Su, Jiang, Zhang, Xu, Wang, Zheng, Zheng, Chen, Sun and Fu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Public Health

SYSTEMATIC REVIEW
published: 06 June 2022
doi: 10.3389/fpubh.2022.780538

Family Socioeconomic Position and Lung Cancer Risk: A Meta-Analysis and a Mendelian Randomization Study

Xusen Zou[1†], Runchen Wang[2,3†], Zhao Yang[4†], Qixia Wang[2,3†], Wenhai Fu[2,5†], Zhenyu Huo[2,3], Fan Ge[2,5], Ran Zhong[2], Yu Jiang[2,3], Jiangfu Li[2], Shan Xiong[2], Wen Hong[1] and Wenhua Liang[2*]*

[1] South China University of Technology, School of Public Administration, Guangzhou, China, [2] Department of Thoracic Oncology and Surgery, China State Key Laboratory of Respiratory Disease and National Clinical Research Center for Respiratory Disease, the First Affiliated Hospital of Guangzhou Medical University, Guangzhou, China, [3] Nanshan School, Guangzhou Medical University, Guangzhou, China, [4] Peking University First Hospital, Beijing, China, [5] First Clinical School, Guangzhou Medical University, Guangzhou, China

OPEN ACCESS

Edited by:
Yiqiang Zhan,
Helmholtz Association of German Research Centers (HZ), Germany

Reviewed by:
Laercio Lopes DaSilva,
MetroWest Medical Center,
United States
Cheng Wang,
Nanjing University, China

***Correspondence:**
Wenhua Liang
liangwh1987@163.com
Wen Hong
wenhong@scut.edu.cn

[†] These authors share first authorship

Specialty section:
This article was submitted to Public Health Policy, a section of the journal Frontiers in Public Health

Received: 21 September 2021
Accepted: 11 April 2022
Published: 06 June 2022

Citation:
Zou X, Wang R, Yang Z, Wang Q, Fu W, Huo Z, Ge F, Zhong R, Jiang Y, Li J, Xiong S, Hong W and Liang W (2022) Family Socioeconomic Position and Lung Cancer Risk: A Meta-Analysis and a Mendelian Randomization Study. Front. Public Health 10:780538. doi: 10.3389/fpubh.2022.780538

Background: Family socioeconomic position (SEP) in childhood is an important factor to predict some chronic diseases. However, the association between family SEP in childhood and the risk of lung cancer is not clear.

Methods: A systematic search was performed to explore their relationship. We selected education level, socioeconomic positions of parents and childhood housing conditions to represent an individual family SEP. Hazard ratios (HRs) of lung cancer specific-mortality were synthesized using a random effects model. Two-sample Mendelian randomization (MR) was carried out with summary data from published genome-wide association studies of SEP to assess the possible causal relationship of SEP and risk of lung cancer.

Results: Through meta-analysis of 13 studies, we observed that to compared with the better SEP, the poorer SEP in the childhood was associated with the increased lung cancer risk in the adulthood (HR: 1.25, 95% CI: 1.10 to 1.43). In addition, the dose-response analysis revealed a positive correlation between the poorer SEP and increased lung cancer risk. Same conclusion was reached in MR [(education level) OR 0.50, 95% CI: 0.39 to 0.63; $P < 0.001$].

Conclusion: This study indicates that poor family socioeconomic position in childhood is causally correlated with lung cancer risk in adulthood.

Systematic Review Registration: identifier: 159082.

Keywords: lung cancer risk, socioeconomic position (SEP), meta-analysis, dose response, Mendelian randomization

BACKGROUND

The relationship between everyday-life situations in childhood and chronic diseases has received increased attention across a number of disciplines in recent years (1–4). In addition to the genetic effects, the effects of social environment in early life can also impact one's later life (5, 6). Early prediction of high incidence of lung cancer can play an important role in preventing lung

cancer-specific death. Early detection and treatment of lung cancer is a promising strategy to reduce lung cancer mortality (7).

SEP is one of the interfering factors of health of family members (8, 9). Family SEP in childhood consists of many conditions, including years of schooling, parents' social status, childhood housing conditions and so on (10, 11). So far, family SEP in childhood is an increasingly important factor to predict chronic diseases. Multiple studies have indicated that there is an association between SEP and the risk of chronic disease (12–15). Some cohort studies have also demonstrated the impact of family SEP in childhood on adult risk of mortality and cancer incidence (16–29). Overall, poor SEP in early years has an impact on adult morbidity and mortality, especially in cardiovascular and gastrointestinal disorders (30, 31). However, there has been little reliable evidence that lung cancer in adulthood may be related to family SEP in childhood (32, 33).

In this study, we sought to investigate the association between SEP in childhood and lung cancer risk in adulthood. A meta-analysis with dose response and a two-sample Mendelian Randomization (MR) were performed using single or multiple single nucleotide polymorphisms (SNPs) as an instrument in instrumental variable analyses by studying known genetic determinants of the exposure variable of interest.

MR uses genetic variants as instrumental variables for assessing causal relationships from observational data (34) and is an established method for probing questions of causality in observational epidemiology (35). MR technique is being extensively applied to estimate the long-term causal effects of various exposures on clinical and epidemiological outcomes using observational data (36). MR is an established approach to evaluate the effect of an exposure on an outcome. Most importantly, unlike main effect MR studies, gene x environment interaction studies are susceptible to confounding factors (37).

METHODS

Search Strategy and Inclusion Criteria

Two investigators (W.R., L.C.) independently searched, PubMed, Web of Science from their inceptions (1966 and 1947, respectively) to June 1, 2020, for cohort studies related to family SEP in childhood and lung cancer in adulthood, without language restrictions. Keywords used were "socioeconomic position (SEP)", "cohort", "lung cancer", "risk", as well as their Medical Subject Headings (MeSH) terms. Only papers published in English were included. Details of the search terms and inclusion/exclusion criteria are shown in **Figure 1**.

The inclusion criteria were: (i) cohort study design; (ii) all cases were lung cancer patients, and data were classified according to different SEP, regardless of subtypes of lung cancer; (iii) In addition, the exclusion criteria were: (i) small sample sizes; (ii) non-English publications; (iii) lack of data and necessary information for meta analysis or dose response analysis.

All the included studies were divided into two groups: (I) studies reported adjusted estimates of the hazard ratios (HR) (including relative risk and odds ratios) and 95% confidence interval (95% CI); (II) studies missed data mentioned above. We also included studies that graded data based on parents' financial status and family life in childhood, and overall impact. The quintile or quartile method was incorporated and the first level was used as the reference.

Data Acquisition Quality Assessment

Data were extracted by two investigators (W.R., H.Z.) independently, and any disagreements came to consensus after discussion. Basic data were recorded from all eligible studies, including the first author's name, publication year, country, study period, period of follow-up, number of lung cancer patients and HR with their 95% Cl. The results were reviewed by two senior investigators (L.W., H.J.).

Quality Assessment

The quality of the studies was evaluated using a score system that was designed with reference to the Newcastle-Ottawa Scale (NOS) tool (38–40). The system is based on a 0-9 points, with 9 reflecting the highest quality and 0 the lowest. Each point was allocated for the following: (I) representativeness of exposure arm(s), (II) selection of the comparative arm(s), (III) origin of exposure source, (IV) demonstration that outcome of interest was not present at start of study, (V) studies controlling the most important factors, (VI) studies controlling the other main factors, (VII) assessment of outcome with independency, (VIII) adequacy of follow-up length (to assess outcome), (IX) lost to follow-up acceptable (< 10% and reported).

Two researchers (W.R. and G.F.) independently evaluated the methodological quality of each included published cohort study. The results of the quality assessment were used for descriptive purposes, to provide an overall assessment of the quality of the included studies (38, 41–44).

A meta-analysis was performed following the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines and the study protocol of this study is 1,59,082 (45, 46). We collected study-specific HR with 95% Cl for lung cancer to combine the data. Through Cochran's I^2 statistic, we examined the heterogeneity across studies and statistical heterogeneity was considered if an I^2 statistic $\geq$ 50% (47). A random effect model was employed to synthesize hazard ratios (HRs) for lung cancer-specific mortality if high heterogeneity existed ($p < 0.5$, $I^2 > 50\%$), otherwise a fixed effect model was conducted.

We utilized Funnel plot tests, Begg's test and Egger's test to evaluate the publication bias (48–50). In addition, subgroup analysis was performed according to gender. Sensitivity analysis was conducted by excluding each study in turn to access the stability of results and potential sources of heterogeneity. All statistical manipulation was employed by Stata software (version 12, StataCorp, TX, USA). All *P*-values were 2-tailed; statistical significance was set at *P*-value < 0.05 (51).

Abbreviations: CI, Confidence interval; GWASs, Genome-wide association studies; HR, Hazard ratio; ILCCO, International Lung Cancer Consortium; IVW, Inverse Variance-Weighted; MeSH, Medical Subject Headings; MR, Mendelian randomization; OR, Odds ratios; RR, Relative risk; SCLC, Small cell lung cancer; SNP, Single-nucleotide polymorphism; TSMR, Two-sample Mendelian randomization; BMDS, Benchmark Dose Software.

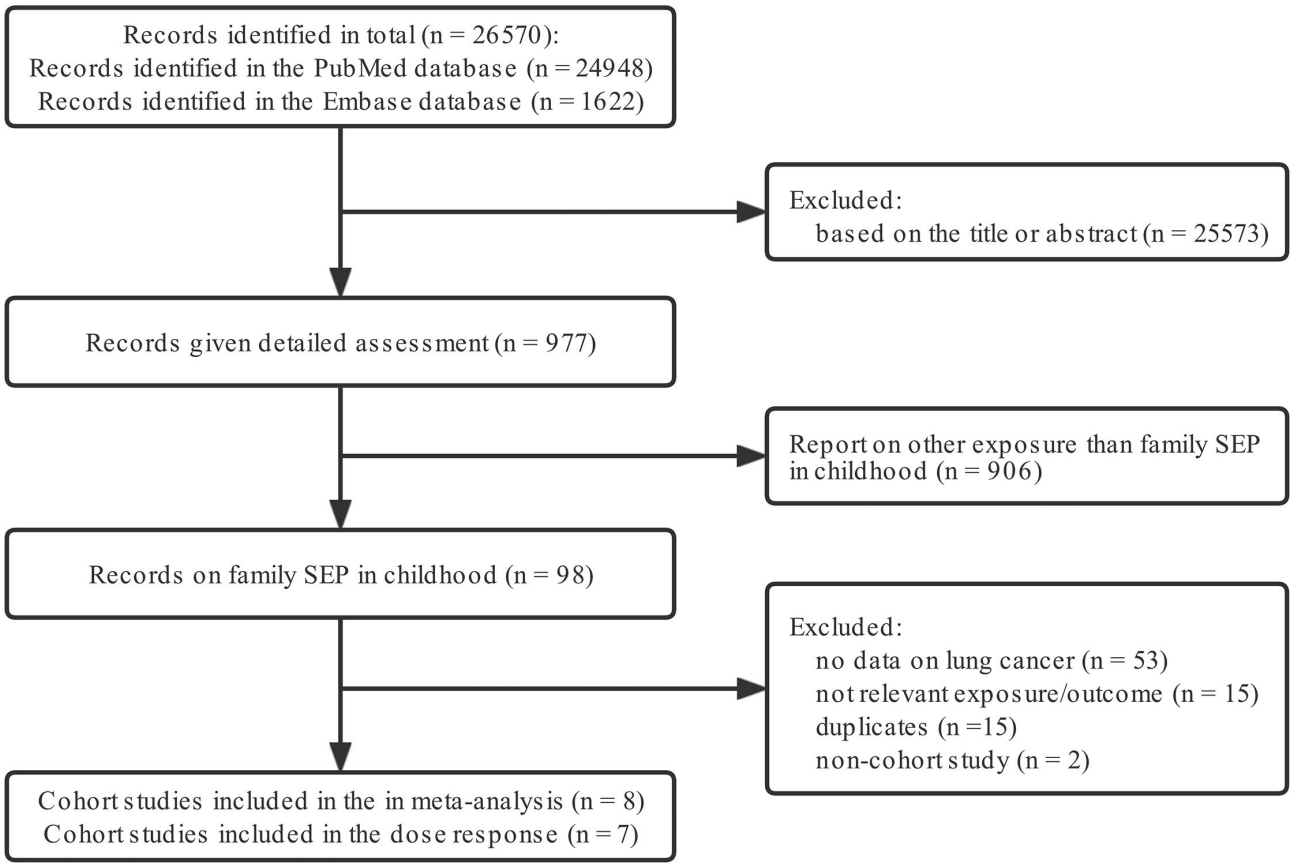

FIGURE 1 | Flow diagram detailing the search strategy and identification of studies used in meta-analysis.

The dose-response study was analyzed by Benchmark Dose Software 3.1.2(BMDS 3.1.2) (52). All steps were established under the BMDS guidelines (53–56). Multiple models were selected for analysis, including extra risk assumptions for background and a benchmark response of 5% (39, 40). According to the included studies, the dose response was divided into two categories: family economic status in childhood and childhood housing conditions. At the same time, according to the original data with or without adjustment, it was divided into adjusted and unadjusted group for subgroup analysis and publication bias analysis. In all dose-response models, we used Exponential models and Hill models as the main dose-response observation results, (57) while using other results for verification. Duplicate studies were excluded.

MR Analysis Using Summary Statistics

The MR method was based on the following three assumptions: (i) the instrumental variables are strongly associated with the family SEP in childhood; (ii) the instrumental variables affect cancer only through their effect on family SEP in childhood and not through any alternative causal pathway; and (iii) the instrumental variables are independent of any confounders (58, 59). To avoid the potential violation of the first assumption, we selected SNPs that meet the threshold of genome-wide significance ($P < 5 \times 10^{-8}$). Besides, we applied some methods including weighted median and MR–Egger to test the second assumption. Then we evaluated the directional pleiotropy based on the intercept obtained from the MR Egger analysis to satisfy the third assumption (60). We also performed a leave-one-out analysis in which we sequentially omitted one SNP at a time, to evaluate whether the MR estimate was driven or biased by a single SNP.

The analysis was conducted to estimate the effect of family SEP in childhood (X) on the risk of lung cancer (Y) using genetic variants (g), and the causal estimate is equal to Yg/Xg (61). For the association between genetic variants and family SEP in childhood (Xg), summary data were utilized from published Genome-Wide Association Studies (GWASs), (62–77) including Social Science Genetic Association Consortium (SSGAC) (1,060,068 individuals), MRC Integrative Epidemiology Unit (MRC-IEU) (2,49,790 individuals), UK Biobank (75,244 individuals) and Neale Lab (4,55,571 individuals) (**Supplementary Table 1**) (78–80). Summary statistics for the association between genetic variants and lung cancer (Yg) are from the International Lung Cancer Consortium (ILCCO) (27,209 individuals) (**Supplementary Table 2**) (69).

TABLE 1 | Pooled estimates from random effects meta-analysis, expressing the HR of risk of lung cancer.

Outcome	Adjusted group			Unadjusted group		
	HR (95%CI)	*P*-value	I^2	HR (95%CI)	*P*-value	I^2
Overall	1.25 (1.10–1.43)	0.04	48.90%	1.42 (1.21–1.66)	<0.001	73.40%
Male	1.10 (0.96–1.26)	0.78	0.00%	1.52 (1.13–2.06)	0.01	87.10%
Female	1.09 (0.88–1.35)	0.57	0.00%	1.28 (1.07–1.54)	0.01	19.40%
Both sexes	1.40 (1.20–1.62)	0.04	48.80%	1.63 (1.15–2.32)	0.01	58.70%

HR, Hazard Ratio.

We selected uncorrelated variants to construct the instrumental variables in the two-sample MR analysis. Using both summary statistics for *Yg* and *Xg*, an Inverse Variance-Weighted (IVW) meta-analysis was employed to estimate the effect of genetically determined family SEP in childhood on the risk of lung cancer using the method of Burgess et al. (81):

$$\hat{\beta}_{IVW} = \frac{\sum_{i=1}^{g} X_g Y_g \sigma_{Yg}{}^{-2}}{\sum_{i=1}^{g} X_g{}^2 \sigma_{Yg}{}^{-2}},\ se\left(\hat{\beta}_{IVW}\right) = \sqrt{\frac{1}{\sum_{i=1}^{g} X_g{}^2 \sigma_{Yg}{}^{-2}}}$$

where *Xg* is the beta estimate for the association between the SNP and family SEP in childhood, *Yg* is the beta estimate for the association between the SNP and lung cancer, and σYg is the standard error for *Yg*. Corresponding HR and 95% CIs were calculated using β^{IVW} and se(β^{IVW}).

In addition, MR-Egger and Weighted Median were also conducted to identify the causality. Ward ratio would only be observed when the first three models cannot be used due to lack of SNPs (82). Leave-one-out analysis was conducted to estimate whether the result was driven by a single SNP. MR-Egger regression was also performed to access the pleiotropy.

RESULTS

Family Socioeconomic Position in Childhood and the Risk of Lung Cancer in Adulthood Observational Analyses

A total of eight cohort studies consisting of 2,779,242 cases were included in the final meta-analysis, (16, 17, 20, 23–28) and seven studies were included in the final dose-response analysis (16–22). The age of the included participants ranged from 14 to 74, the majority of the respondents were between the age of 30 to 64 with 2,413 respondents recorded to have died from lung cancer, details about patients characteristics of these studies can be seen in **Supplementary Table 3**. **Supplementary Table 4** presented the result of the quality assessment. Results and study characteristics including name of first author, publication year, country or region, follow-up period, sample size and number of cases or deaths, type of outcome, gender, age range, HRs and 95% CIs were extracted and can be seen in **Supplementary Table 5**.

The results of quality assessment did not suggest any evidence of publication bias (**Supplementary Figures 1**, **2**) [Unadjusted group: (Begg: 0.673; Egger: 0.182)]; [Adjusted group: (Begg:1; Egger:0.228)]. And sensitivity analysis was performed by sequentially excluding individual studies (**Supplementary Figures 3**, **4**).

In unadjusted group, the pooled HR for family economy conditions with best family SEP in childhood group was 1.42 (95% CI, 1.21, 1.66) (I2 = 73.4%, $P < 0.001$) (**Table 1**, **Figure 2**). The risk was slightly lower than adjusted group: 1.25 (95% CI, 1.10, 1.42) (I2 = 48.9%, $P = 0.04$) (**Table 1**, **Figure 3**). Subgroup analysis based on sex using the provided adjusted original data resulted in a pooled HR of 1.09 (95% CI 0.95, 1.26) (I2 = 0.0%, $p = 0.78$) for male, a pooled HR of 1.10 (95%CI 0.88, 1.35) (I2 = 0.0%, $p = 0.57$) for female, and a pooled HR of 1.39 (95%CI 1.20, 1.62) (I2 = 48.9%, $p = 0.04$) for both sexes. Subgroup analysis based on sex that provided unadjusted original data resulted in a pooled HR of 1.52 (95% CI 1.13, 2.06) (I2 = 87.1%, $p = 0.006$) for male and a pooled HR of 1.28 (95%CI 1.07, 1.54) (I2 = 19.4%, $p = 0.007$) for female and a pooled HR of 1.63 (95%CI 1.15, 2.32) (I2 = 58.7%, $p = 0.006$) for both sexes. The above results was consistent in dose response used the model summary with BMR of 1 Std. Dev. including the family economic conditions of parents and childhood housing conditions (**Supplementary Figure 5**).

Genetically Determined Family SEP in Childhood Was Associated With Risk of Lung Cancer in Adulthood

Genetically predicted lower family SEP in childhood was associated with significantly higher odds of lung cancer (**Table 2**). The causal relationship between length of education and lung cancer was verified first. Using conventional MR analysis, one SD longer education was associated with a 50% lower risk of lung cancer (SSGAC 1: OR 0.50, 95% CI: 0.39, 0.63; *P*: 0.001). Another gene pool with the same exposure factor was also included for verification in our study, showing one SD longer education (genetic predisposition by 4 SNPs) was associated with a 60% lower risk of lung cancer (SSGAC 2: OR 0.40, 95% CI 0.32, 0.50; *p*: 0.001). With a sample size of 27 209, our sample provided sufficient statistical power (> 80%) to detect a causal effect of educational attainment on lung cancer. Power calculations on MR analysis were performed according to Brion et al. (83).

Meanwhile, associations were consistent in sensitivity analyses using weighted median (OR 0.57, 95% CI 0.45, 0.71; *P*: 0.001;

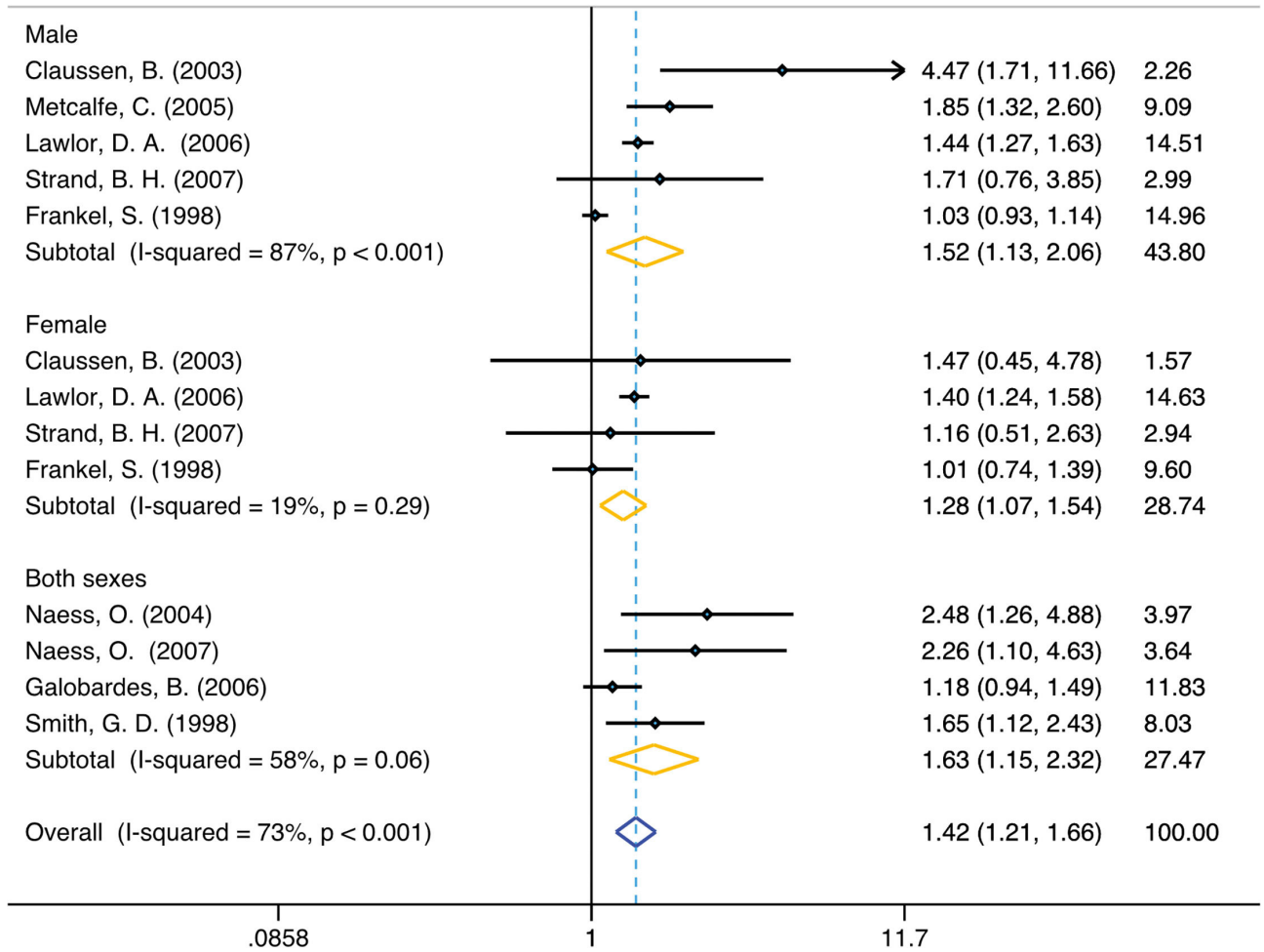

FIGURE 2 | Forest plot of population-based cohort studies of family SEP and lung cancer, stratified by sex (unadjusted group).

Okbay *et al.*) (OR 0.58, 95% CI 0.48, 0.70; $P < 0.001$; Lee *et al.*) and MR-Egger method (OR 0.56, 95% CI 0.17, 1.88; *P*: 0.35; Okbay *et al.*) (OR 0.75, 95% CI 0.60, 0.94; *P*: 0.01; Lee *et al.*), but provided less precise estimates than with conventional MR (IVW method). Nonetheless, their causal estimates were similar in terms of direction and magnitude, and they were unlikely to occur by chance alone. These results were consistent with the hypothesis that genetic pleiotropy does not drive the result.

We also analyzed and observed the relationship between other exposure factors related to family SEP in childhood and lung cancer, and we found that the *P*-values for the intercept were large and the estimates adjusted for pleiotropy suggested null effects in four exposure factors: number of children fathered (Neale Lab): (OR 0.92, 95% CI: 0.29, 2.84; *p*: 0.88); mother's age at death (Neale Lab): (OR 1.00, 95% CI: 0.19, 5.26; *p*: 1.00); and average total household income before tax (MRC-IEU): (OR 0.90, 95% CI: 0.77, 1.04; *p*: 0.156). Nine out of the 15 independent exposure factor gene pools were unable to be analyzed due to the lack of SNPs corresponding to lung cancer.

DISCUSSION

Based on data from 2,779,242 respondents, we observed that poorer family SEP in childhood means higher lung cancer risk in adulthood compared to controls. The overall dose-response of multiple models also demonstrated this trend. This phenomenon was partially verified by Mendelian randomization of two samples, but the causal relationship between some genes was still unclear. Through meta-analysis, we analyzed both the adjusted (age/adult social class) and unadjusted data, and strengthened the credibility by mutual authentication. The results showed a link between family SEP in childhood and the risk of lung cancer in adulthood in the unadjusted group.

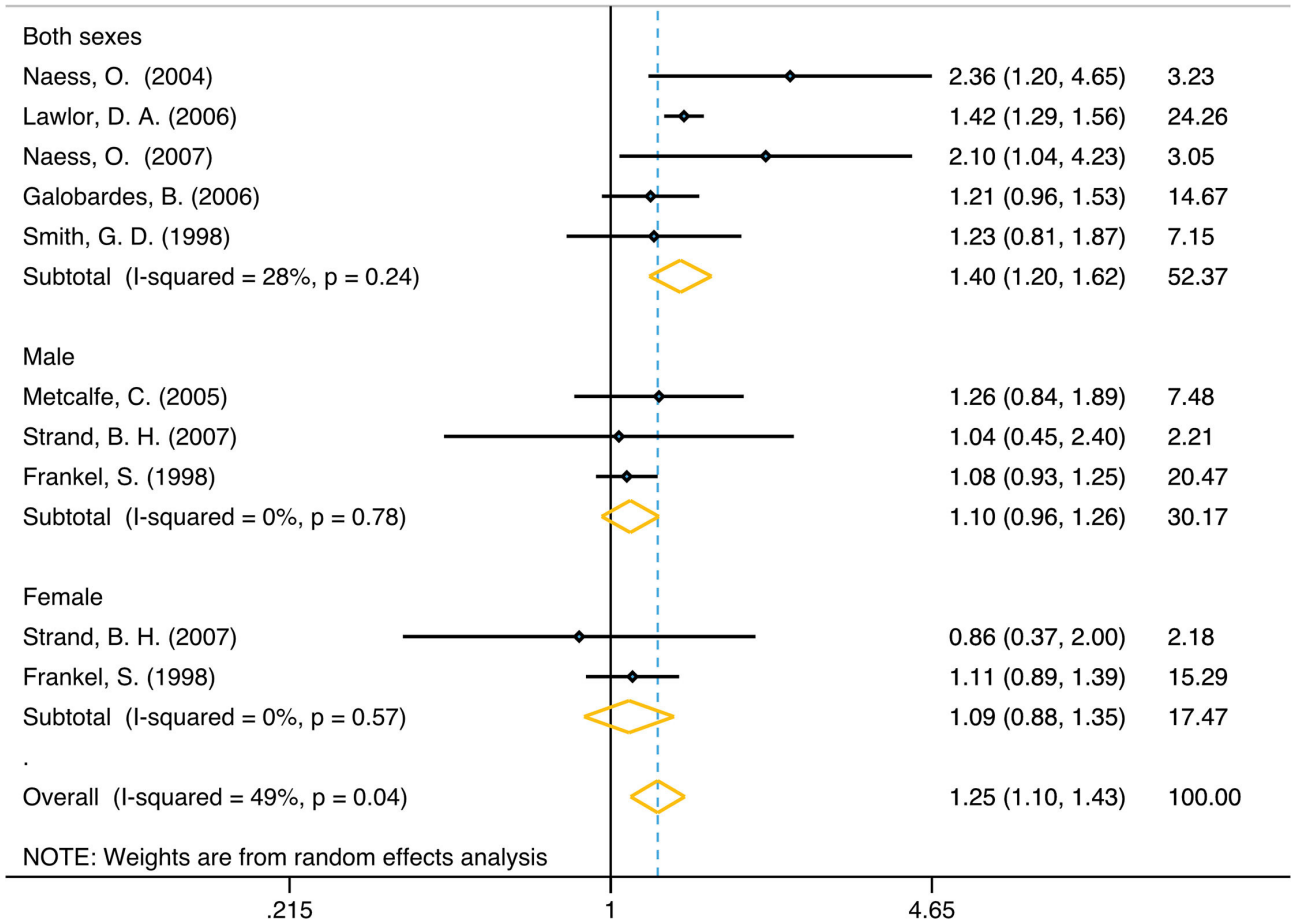

FIGURE 3 | Forest plot of population-based cohort studies of family SEP and lung cancer, stratified by sex (adjusted group).

People in poor family economic conditions were more likely to smoke, leading to a higher risk of lung cancer (84). Furthermore, the poor face a variety of problems that may pose other risks including malnutrition, violence, AIDS, and other infectious diseases, which also correlate with lung cancer risk (85, 86). The dose-response curve of family SEP in childhood also contribute to risk. Specifically, the lower the family economic conditions in childhood and the worse the childhood housing conditions, the more likely a person is to develop lung cancer in adulthood. This result has been verified in both the adjusted and unadjusted groups. We selected factors as comprehensive as possible to explore the relationship between family SEP in childhood and the risk of lung cancer in adulthood using two sample MR. Only the data from the official database were used, which made our data more reliable and useful. The results show that there is a clear causal relationship between the length of education and the occurrence of lung cancer. People with low education during childhood were more likely to smoke or be exposed to undesirable living environments, and lack relevant basic medical knowledge. Moreover, Children of parents who smoke perceived casual smoking to be safer and then reported wanting to smoke in response to smoking-related cues more than children of non-smoking parents (87). It should be noted that further research is needed to prove the relationship between father's age of death and lung cancer. On the one hand, although our results showed a strong correlation between SEP and the risk of lung cancer, only one SNP was included. This is too small to calculation of explained variance. This limited the representativeness of the result of father's death of age and lung cancer. On the other hand, early parental death, when a child experiences the loss of a parent before age 18, is one of the most severe stressors and a potential risk factor for adult psychopathology (88). It may cause an increased likelihood of poor habits, for example, smoking, which in turn lead to lung cancer. Worth mentioning, among the results related to education duration, the results of incorporating the genome determined by Okbay et al. were slightly different from the previous results of Zhou et al. (89).

TABLE 2 | Mendelian randomization estimates of the associations between family SEP in childhood and risk of lung cancer.

Outcome	Inverse variance weighted		Weighted median		MR Egger		Ward ratio	
	OR (95%CI)	*P*-value	OR (95%CI)	*P*-value	OR (95%CI)	*P*-value	OR (95%CI)	*P*-value
The associations between education attainment and risk of lung cancer overall and histological types (Okbay et al.)								
Lung cancer overall	0.49 (0.35–0.69)	<0.01	0.60 (0.39–0.91)	0.02	0.65 (0.11–3.81)	0.63	N/A	N/A
Lung adenocarcinoma	0.64 (0.41–1.02)	0.06	0.66 (0.36–1.20)	0.17	1.35 (0.12–14.80)	0.80	N/A	N/A
Squamous cell lung cancer	0.40 (0.26–0.62)	0.00	0.53 (0.29–0.97)	0.04	0.20 (0.02–1.89)	0.16	N/A	N/A
The associations between education attainment and risk of lung cancer overall and histological types (Lee et al.)								
Lung cancer overall	0.42 (0.28–0.62)	<0.05	0.66 (0.51–0.76)	0.016	0.82 (0.54–1.25)	0.63	N/A	N/A
Lung adenocarcinoma	0.23 (0.08–0.67)	0.06	0.38 (0.19–0.76)	0.17	0.58 (0.20–1.65)	0.80	N/A	N/A
Squamous cell lung cancer	0.40 (0.31–0.52)	<0.01	0.55 (0.46–0.65)	0.04	0.74 (0.57–0.96)	0.16	N/A	N/A
The associations between father's age at death and risk of lung cancer overall and histological types								
Lung cancer overall	N/A	N/A	N/A	N/A	N/A	N/A	2.15E−05 (1.42E−06–3.26E−04)	<0.01
Lung adenocarcinoma	N/A	N/A	N/A	N/A	N/A	N/A	1.77E−05 (3.00E−06–1.04E−04)	<0.01
Squamous cell lung cancer	N/A	N/A	N/A	N/A	N/A	N/A	3.39E−05 (2.33E−06–4.93E−04)	<0.01

OR, Odds Ratio; N/A, Not available.

This may be explained by these observations: the statistical significance was set as *P*-value < 0.001 instead of 0.01 in the R package used for two-sample MR (34, 90). Because of this, the positive results of our study are more reliable. For the negative result in two-sample Mendelian randomization, we believe that further researcher is required to verify the result. Because the data of the SNPs in the existing studies are still limited, the final results of these studies have not yet been proven.

Three strengths of our study should be highlighted. First, to our knowledge, this is the most comprehensive meta-analysis estimating the relationship between family economic position in childhood and the risk of lung cancer in adulthood. The collection of such large sample data to assess the cancer risk for different family SEP in childhood around the world has not previously been performed. Second, we included adjusted and non-adjusted data for all-factors to mutually verify the dose-response of parents' socioeconomic status and family congestion in childhood. Third, this is the first study to explore the causal relationship between family SEP in childhood and the risk in adulthood by two sample MR. This research offered evidence for researcher focus on epidemiological characters in early detection of lung cancer and clinical staff of early diagnosis of lung cancer.

There are several limitations in our study. First, the studies included for meta-analysis and dose-response studies originated from the UK, Norway, and the Netherlands, there is the risk of introducing potentially heterogeneity. Second, because of the long-term follow-up needs of this study, included studies do not have data from recent years. Third, because the included cohort studies were all-factor epidemiological studies, subgroups of different lung cancer subtypes or specific age could not be established in the meta-analysis. Fourth, in terms of meta-analysis, due to the lack of subgroup data, the results obtained by the adjusted male and female subgroups were not statistically significant. Another potential limitation is that we only included common polymorphisms. Finally, although we have included many different exposure genes to build connections with lung cancer gene pool, some Mendelian randomization results are negative. However, limited SNPs of SEP can lead to false negative result. Further verification is needed to make a precise conclusion for the reason given above.

This research needs to be improved. Cohort results from other countries and expanded observations from lung cancer to all-factor risk are necessary for future studies. Genetic characteristics of continents other than Europe should also be included. Overall, our research confirms our conjecture: the family SEP in childhood is inversely proportional to the risk of lung cancer, which has a positive effect on the early diagnosis and intervention of lung cancer.

CONCLUSION

This study indicates that poor family SEP in childhood is a causal risk factor for lung cancer, and thus lung

cancer screening should be more heavily considered for these populations. More research is needed to cross-validate these findings.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

XZ and RW: conceptualization and project administration. ZY, XZ, and RW: investigation. ZH, ZY, XZ, YJ, and RW: data curation. ZH: visualization and writing—review and editing. FG, XZ, and RW: formal analysis. FG, ZY, XZ, and RW: writing—original draft. RZ: methodology and software. JL: definition. JL and SX: validation. SX: supervision and funding acquisition. WH: resources. WH and ZH: writing—review and editing. All authors were involved in the conception and design of the study, critically reviewed, and approved the final manuscript.

FUNDING

This work was supported by the China National Science Foundation (Grant No. 81871893) and Key Project of Guangzhou Scientific Research Project (Grant No. 201804020030).

ACKNOWLEDGMENTS

We sincerely thank all the medical staff who are fighting against this public health crisis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2022.780538/full#supplementary-material

Supplementary Figure 1 | Begg funnel plots of population-based cohort studies of family SEP and lung cancer (unadjusted group).

Supplementary Figure 2 | Begg funnel plots of population-based cohort studies of family SEP and lung cancer (adjusted group).

Supplementary Figure 3 | Sensitivity analysis of population-based cohort studies of family SEP and lung cancer (unadjusted group).

Supplementary Figure 4 | Sensitivity analysis of population-based cohort studies of family SEP and lung cancer (adjusted group).

Supplementary Figure 5 | Dose response of childhood housing conditions and lung cancer (adjusted group) (top left), family economic conditions and lung cancer (adjusted group) (top right), childhood housing conditions and lung cancer (unadjusted group) (bottom left), family economic conditions and lung cancer (unadjusted group) (bottom right).

Supplementary Table 1 | Summary of identified MR gene bank of exposure.

Supplementary Table 2 | Summary of identified MR gene bank of outcome.

Supplementary Table 3 | Characteristic of the included studies in the meta-analysis.

Supplementary Table 4 | The Newcastle-Ottawa quality assessment scale of including studies.

Supplementary Table 5 | Data from included cohort studies.

REFERENCES

1. Hertzman C. The biological embedding of early experience and its effects on health in adulthood. *Ann N Y Acad Sci.* (1999) 896:85–95. doi: 10.1111/j.1749-6632.1999.tb08107.x
2. Carey G, Crammond B. Systems change for the social determinants of health. *BMC Public Health.* (2015) 15:662. doi: 10.1186/s12889-015-1979-8
3. Blackwell DL, Hayward MD, Crimmins EM. Does childhood health affect chronic morbidity in later life? *Soc Sci Med (1982).* (2001). 52:1269–84. doi: 10.1016/S0277-9536(00)00230-6
4. Nordestgaard AT, Thomsen M, Nordestgaard BG. Coffee intake and risk of obesity, metabolic syndrome and type 2 diabetes. a Mendelian randomization study. *Int J Epidemiol.* (2015) 44:551–65. doi: 10.1093/ije/dyv083
5. Mackenbach JP, Kulhánová I, Bopp M, Deboosere P, Eikemo TA, Hoffmann R, et al. Variations in the relation between education and cause–specific mortality in 19 European populations a test of the "fundamental causes" theory of social inequalities in health. *Soc Sci Med.* (2015) 127:51–62. doi: 10.1016/j.socscimed.2014.05.021
6. Oates GR, Schechter MS. Socioeconomic status and health outcomes. cystic fibrosis as a model. *Expert Rev Respir Med.* (2016) 10:967–77. doi: 10.1080/17476348.2016.1196140
7. Yee J, Sadar MD, Sin DD, Kuzyk M, Xing L, Kondra J, et al. Connective tissue-activating peptide III. a novel blood biomarker for early lung cancer detection. *J Clin Oncol.* (2009) 27:2787–92. doi: 10.1200/JCO.2008.19.4233
8. Poulain T, Vogel M, Kiess W. Review on the role of socioeconomic status in child health and development. *Curr Opin Pediatr.* (2020) 32:308–14. doi: 10.1097/MOP.0000000000000876
9. Komulainen K, Mittleman MA, Jokela M, Laitinen TT, Pahkala K, Elovainio M, et al. Socioeconomic position and intergenerational associations of ideal health behaviors. *Eur J Prev Cardiol.* (2019) 26:1605–12. doi: 10.1177/2047487319850959
10. Moor I, Kuipers MAG, Lorant V. Pförtner T-K, Kinnunen JM, Rathmann K, Perelman J, Alves J, Robert P-O, Rimpelä A et al. Inequalities in adolescent self–rated health and smoking in Europe comparing different indicators of socioeconomic status. *J Epidemiol Community Health.* (2019) 73:963–70. doi: 10.1136/jech-2018-211794
11. Adler NE, Boyce T, Chesney MA, Cohen S, Folkman S, Kahn RL, et al. Socioeconomic status and health. the challenge of the gradient. *Am Psychol.* (1994) 49:15–24. doi: 10.1037/0003-066X.49.1.15
12. Disano J, Goulet J, Muhajarine N, Neudorf C, Harvey J. Social-economic status and rates of hospital admission for chronic disease in urban Canada. *Can Nurse.* (2010) 106:24–9.
13. Kalyesubula R, Nankabirwa JI, Ssinabulya I, Siddharthan T, Kayima J, Nakibuuka J, et al. kidney disease in Uganda. a community based study. *BMC Nephrol.* (2017) 18:116. doi: 10.1186/s12882-017-0521-x
14. Forootan M, Bagheri N, Darvishi M. Chronic constipation. a review of literature. *Medicine (Baltimore).* (2018) 97:e10631. doi: 10.1097/MD.0000000000010631
15. Yankson R, Anto EA, Chipeta MG. Geostatistical analysis and mapping of malaria risk in children under 5 using point-referenced prevalence data in Ghana. *Malar J.* (2019) 18:67–67. doi: 10.1186/s12936-019-2709-y
16. Frankel S, Gunnell DJ, Peters TJ, Maynard M, Davey Smith G. Childhood energy intake and adult mortality from cancer. the boyd orr cohort study. *Bmj.* (1998) 316:499–504. doi: 10.1136/bmj.316.7130.499

17. Smith GD, Hart C, Blane D, Hole D. Adverse socioeconomic conditions in childhood and cause specific adult mortality. prospective observational study. *Bmj*. (1998) 316:1631–5. doi: 10.1136/bmj.316.7145.1631
18. Hart CL, Davey Smith G. Relation between number of siblings and adult mortality and stroke risk. 25 year follow up of men in the collaborative study. *J Epidemiol Community Health*. (2003) 57:385–91. doi: 10.1136/jech.57.5.385
19. Power C, Hypponen E, Smith GD. Socioeconomic position in childhood and early adult life and risk of mortality. a prospective study of the mothers of the (1958). British birth cohort. *Am J Public Health*. (2005) 95:1396–02. doi: 10.2105/AJPH.2004.047340
20. Galobardes B, Davey Smith G, Jeffreys M, McCarron P. Childhood socioeconomic circumstances predict specific causes of death in adulthood. the glasgow student cohort study. *J Epidemiol Community Health*. (2006) 60:527–9. doi: 10.1136/jech.2005.044727
21. de Kok IM, van Lenthe FJ, Avendano M, Louwman M, Coebergh JW, Mackenbach JP. Childhood social class and cancer incidence. results of the globe study. *Soc Sci Med*. (2008) 66:1131–9. doi: 10.1016/j.socscimed.2007.11.035
22. Dedman DJ, Gunnell D, Davey Smith G, Frankel S. Childhood housing conditions and later mortality in the boyd orr cohort. *J Epidemiol Community Health*. (2001) 55:10–5. doi: 10.1136/jech.55.1.10
23. Claussen B, Davey Smith G, Thelle D. Impact of childhood and adulthood socioeconomic position on cause specific mortality. the oslo mortality study. *J Epidemiol Community Health*. (2003) 57:40–5. doi: 10.1136/jech.57.1.40
24. Naess O, Claussen B, Davey Smith G. Relative impact of childhood and adulthood socioeconomic conditions on cause specific mortality in men. *J Epidemiol Community Health*. (2004) 58:597–8. doi: 10.1136/jech.2003.012229
25. Metcalfe C, Davey Smith G, Sterne JA, Heslop P, Macleod J, Hart CL. Cause-specific hospital admission and mortality among working men. association with socioeconomic circumstances in childhood and adult life, and the mediating role of daily stress. *Eur J Public Health*. (2005) 15:238–44. doi: 10.1093/eurpub/cki063
26. Lawlor DA, Sterne JA, Tynelius P, Davey Smith G, Rasmussen F. Association of childhood socioeconomic position with cause-specific mortality in a prospective record linkage study of 1,839,384 individuals. *Am J Epidemiol*. (2006) 164:907–15. doi: 10.1093/aje/kwj319
27. Naess O, Claussen B, Smith GD. Housing conditions in childhood and cause-specific adult mortality. the effect of sanitary conditions and economic deprivation on 55,761 men in Oslo Scand. *J Public Health*. (2007) 35:570–6. doi: 10.1080/14034940701320846
28. Strand BH, Kunst A. Childhood socioeconomic position and cause-specific mortality in early adulthood. *Am J Epidemiol*. (2007) 165:85–93. doi: 10.1093/aje/kwj352
29. Melchior M, Berkman LF, Kawachi I, Krieger N, Zins M, Bonenfant S, et al. Lifelong socioeconomic trajectory and premature mortality (35–65 years) in France. findings from the GAZEL cohort study. *J Epidemiol Community Health*. (2006) 60:937–44. doi: 10.1136/jech.2005.042440
30. Galobardes B, Smith GD, Lynch JW. Systematic review of the influence of childhood socioeconomic circumstances on risk for cardiovascular disease in adulthood. *Ann Epidemiol*. (2006). 16:2. doi: 10.1016/j.annepidem.2005.06.053
31. Vohra J, Marmot MG, Bauld L, Hiatt RA. Socioeconomic position in childhood and cancer in adulthood. *a rapid-review. J. Epidemiol. Community Health*. (2016) 70:629–34. doi: 10.1136/jech-2015-206274
32. Hrubá F, Fabiáová E, Bencko V, Cassidy A, Lissowska J, Mates D, et al. Socioeconomic indicators and risk of lung cancer in central and eastern Europe cen. *Eur J Public Health*. (2009) 17:115–21. doi: 10.21101/cejph.a3516
33. Sharpe KH, McMahon AD, McClements P, Watling C, Brewster DH, Conway DI. Socioeconomic inequalities in incidence of lung and upper aero-digestive tract cancer by age, tumour subtype and sex. a population-based study in Scotland (2000–2007). *Cancer Epidemiol*. (2012) 36:e164–e170. doi: 10.1016/j.canep.2012.01.007
34. Yavorska OO, Burgess S. Mendelian randomization. an R package for performing Mendelian randomization analyses using summarized data. *Int J Epidemiol*. (2017) 46:1734–9. doi: 10.1093/ije/dyx034
35. Bowden J, Del Greco M F, Minelli C, Davey Smith G, Sheehan NA, Thompson JR. Assessing the suitability of summary data for two-sample mendelian randomization analyses using MR-egger regression. the role of the I2 statistic. *Int J Epidemiol*. (2016) 45:1961–74. doi: 10.1093/ije/dyw220
36. Burgess S, Daniel RM, Butterworth AS, Thompson SG. Network mendelian randomization. using genetic variants as instrumental variables to investigate mediation in causal pathways. *Int J Epidemiol*. (2015) 44:484–95. doi: 10.1093/ije/dyu176
37. Tyrrell J, Wood AR, Ames RM, Yaghootkar H, Beaumont RN, Jones SE, et al. Gene-obesogenic environment interactions in the UK Biobank study. *Int J Epidemiol*. (2017) 46:559–75. doi: 10.1093/ije/dyw337
38. Stang A. Critical evaluation of the newcastle-ottawa scale for the assessment of the quality of nonrandomized studies in meta–analyses. *Eur J Epidemiol*. (2010) 25:603–5. doi: 10.1007/s10654-010-9491-z
39. Bogen KT. Generic hockey-stick model for estimating benchmark dose and potency. performance relative to BMDS and application to anthraquinone. *Dose Response*. (2011) 9:182–208. doi: 10.2203/dose-response.10-018.Bogen
40. Haber LT, Dourson ML, Allen BC, Hertzberg RC, Parker A, Vincent MJ, et al. Benchmark dose (BMD) modeling. current practice, issues, and challenges. *Crit Rev Toxicol*. (2018) 48:387–415. doi: 10.1080/10408444.2018.1430121
41. Friedemann Smith C, Tompson A, Holtman GA, Bankhead C, Gleeson F, Lasserson D, et al. General practitioner referrals to one–stop clinics for symptoms that could be indicative of cancer. a systematic review of use and clinical outcomes. *Fam Pract*. (2019) 36:255–61. doi: 10.1093/fampra/cmy069
42. Lo CK-L, Mertz D, Loeb M. Newcastle-ottawa Scale. comparing reviewers' to authors' assessments. *BMC Med Res Methodol*. (2014). 14:45. doi: 10.1186/1471-2288-14-45
43. Wang F-G, Yan W-M, Yan M, Song M-M. Outcomes of mini vs roux–en–Y gastric bypass. a meta–analysis and systematic review. *Int J Surg*. (2018). 56:7–14. doi: 10.1016/j.ijsu.2018.05.009
44. Suriyong P, Ruengorn C, Shayakul C, Anantachoti P, Kanjanarat P. Prevalence of chronic kidney disease stages 3-5 in low- and middle-income countries in Asia: A systematic review and meta-analysis. *PLoS One*. (2022). 17:e0264393. doi: 10.1371/journal.pone.0264393
45. Bernardo WM, PRISMA. statement and PROSPERO. *Int Braz J Urol*. (2017) 43:383–4. doi: 10.1590/s1677-5538.ibju.2017.03.02
46. Selçuk AA, A. Guide for systematic reviews. *PRISMA Turk Arch Otorhinolaryngol*. (2019) 57:57–8. doi: 10.5152/tao.2019.4058
47. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta–analyses. *BMJ*. (2003) 327:557–60. doi: 10.1136/bmj.327.7414.557
48. Fan D, Liu L, Xia Q, Wang W, Wu S, Tian G, et al. Female alcohol consumption and fecundability a systematic review and dose–response meta–analysis. *Sci rep*. (2017) 7:13815. doi: 10.1038/s41598-017-14261-8
49. Lin L, Chu H. Quantifying publication bias in meta-analysis. *Biometrics*. (2018) 74:785–94. doi: 10.1111/biom.12817
50. Shi L, Lin L. The trim-and-fill method for publication bias. practical guidelines and recommendations based on a large database of meta-analyses. *Medicine (Baltimore)*. (2019) 98:e15987. doi: 10.1097/MD.0000000000015987
51. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta–analysis of observational studies in epidemiology. a proposal for reporting meta-analysis of observational studies in epidemiology (MOOSE) group. *JAMA*. (2000) 283:2008–12. doi: 10.1001/jama.283.15.2008
52. Jensen SM, Kluxen FM, Ritz C, A. Review of recent advances in benchmark dose methodology. *Risk Anal*. (2019) 39:2295–315. doi: 10.1111/risa.13324
53. Shao K, Shapiro AJ, A. Web–based system for bayesian benchmark dose estimation. *Environ Health Perspect*. (2018) 126:017002. doi: 10.1289/EHP1289
54. Filipsson AF, Sand S, Nilsson J, Victorin K. The benchmark dose method—-review of available models, and recommendations for application in health risk assessment. *Crit Rev Toxicol*. (2003) 33:505–42. doi: 10.1080/748638748
55. Davis JA, Gift JS. Zhao QJ. Introduction to benchmark dose methods and US EPA's benchmark dose software (BMDS) version 211. *Toxicol Appl Pharmacol*. (2011) 254:181–91. doi: 10.1016/j.taap.2010.10.016
56. Agency, USEP. *BMDS Wizard User's Guide*. (2015). Available online at: https://www.epa.gov/sites/default/files/2015-11/documents/bmds_manual.pdf
57. Ma J. Estimating epidemic exponential growth rate and basic reproduction number. *Infect Dis Model*. (2020) 5:129–41. doi: 10.1016/j.idm.2019.12.009

58. Boef AGC, Dekkers OM. le Cessie S. Mendelian randomization studies a review of the approaches used and the quality of reporting. *Int J Epidemiol.* (2015) 44:496–511. doi: 10.1093/ije/dyv071
59. Burgess S, Dudbridge F, Thompson SG. Combining information on multiple instrumental variables in mendelian randomization. comparison of allele score and summarized data methods. *Stat Med.* (2016) 35:1880–906. doi: 10.1002/sim.6835
60. Burgess S, Thompson SG. Interpreting findings from mendelian randomization using the MR-egger method. *Eur J Epidemiol.* (2017) 32:377–89. doi: 10.1007/s10654-017-0255-x
61. Ordóñez–Mena JM, Schöttker B, Mons U, Jenab M, Freisling H, Bueno-de-Mesquita B, et al. Quantification of the smoking–associated cancer risk with rate advancement periods. meta–analysis of individual participant data from cohorts of the CHANCES consortium. *BMC Med.* (2016). 14:62. doi: 10.1186/s12916-016-0607-5
62. Tang H, Wei P, Chang P, Li Y, Yan D, Liu C, et al. Genetic polymorphisms associated with pancreatic cancer survival. a genome–wide association study. *Int J Cancer.* (2017) 141:678–86. doi: 10.1002/ijc.30762
63. Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, et al. The allelic landscape of human blood cell trait variation and links to common complex disease. *Cell.* (2016). 167:5. doi: 10.1016/j.cell.2016.10.042
64. Lutz SM, Cho MH, Young K, Hersh CP, Castaldi PJ. McDonald M-L, et al. A genome-wide association study identifies risk loci for spirometric measures among smokers of European and African ancestry. *BMC Genet.* (2015) 16:138. doi: 10.1186/s12863-015-0299-4
65. Wheeler HE, Gamazon ER, Frisina RD. Perez–Cervantes C, El Charif O, Mapes B, et al. Variants in and other mendelian deafness genes are associated with cisplatin–associated ototoxicity. *Clin Cancer Res.* (2017) 23:3325–33. doi: 10.1158/1078-0432.CCR-16-2809
66. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature.* (2015) 518:197–206. doi: 10.1038/nature14177
67. Al Olama AA. Kote— Z, Berndt SI, Conti DV, Schumacher F, Han Y, et al. A meta–analysis of 87,040 individuals identifies 23 new susceptibility loci for prostate cancer. *Nat Genet.* (2014) 46:1103–9.
68. Broer L, Buchman AS, Deelen J, Evans DS, Faul JD, Lunetta KL, et al. GWAS of longevity in CHARGE consortium confirms APOE and FOXO3 candidacy. *J Gerontol A Biol Sci Med Sci.* (2015) 70:110–8. doi: 10.1093/gerona/glu166
69. Wang Y, McKay JD, Rafnar T, Wang Z, Timofeeva MN, Broderick P, et al. Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer. *Nat Genet.* (2014) 46:736–41.
70. Mayerle J, den Hoed CM, Schurmann C, Stolk L, Homuth G, Peters MJ, et al. Identification of genetic loci associated with Helicobacter pylori serologic status. *JAMA.* (2013) 309:1912–20. doi: 10.1001/jama.2013.4350
71. Lauc G, Huffman JE, Pučić M, Zgaga L, Adamczyk B, MuŽinić A, et al. Loci associated with N–glycosylation of human immunoglobulin G show pleiotropy with autoimmune diseases and haematological cancers. *PLoS Genet.* (2013) 9:e1003225. doi: 10.1371/journal.pgen.1003225
72. Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature.* (2010) 467:832–8. doi: 10.1038/nature09410
73. Genome–wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature.* (2007). 447:661–78. doi: 10.1038/nature05911
74. Wu C, Kraft P. Stolzenberg-Solomon R, Steplowski E, Brotzman M, Xu M, et al. Genome–wide association study of survival in patients with pancreatic adenocarcinoma. *Gut.* (2014) 63:152–60. doi: 10.1136/gutjnl-2012-303477
75. Radstake TRDJ, Gorlova O, Rueda B. Martin J-E, Alizadeh BZ, Palomino-Morales R, et al. Genome–wide association study of systemic sclerosis identifies CD247 as a new susceptibility locus. *Nat Genet.* (2010) 42:426–9. doi: 10.1038/ng.565
76. Allanore Y, Saad M, Dieudé P, Avouac J, Distler JHW, Amouyel P. et al. Genome–wide scan identifies TNIP1, PSORS1C1, and RHOB as novel risk loci for systemic sclerosis. *PLoS Genet.* (2011) 7:e1002091. doi: 10.1371/journal.pgen.1002091
77. Gorlova O. Martin J–E, Rueda B, Koeleman BPC, Ying J, Teruel M, et al. Identification of novel genetic markers associated with clinical phenotypes of systemic sclerosis through a genome–wide association strategy. *PLoS Genet.* (2011) 7:e1002178.
78. Benyamin B, Pourcain B, Davis OS, Davies G, Hansell NK, Brion MJA, et al. Childhood intelligence is heritable, highly polygenic and associated with FNBP1L. *Mol Psychiatry.* (2014) 19:253–8. doi: 10.1038/mp.2012.184
79. Pilling LC, Atkins JL, Bowman K, Jones SE, Tyrrell J, Beaumont RN, et al. Human longevity is influenced by many genetic variants evidence from 75,000 UK Biobank participants. *Aging (Albany NY).* (2016) 8:547–60. doi: 10.18632/aging.100930
80. Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, et al. Gene discovery and polygenic prediction from a genome–wide association study of educational attainment in 11 million individuals. *Nat Genet.* (2018) 50:1112–21. doi: 10.1038/s41588-018-0147-3
81. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol.* (2013) 37:658–65. doi: 10.1002/gepi.21758
82. Pierce BL, Burgess S. Efficient design for mendelian randomization studies. subsample and 2–sample instrumental variable estimators. *Am J Epidemiol.* (2013) 178:1177–84. doi: 10.1093/aje/kwt084
83. Brion M–JA, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. *Int J Epidemiol.* (2013). 42:1497–501. doi: 10.1093/ije/dyt179
84. Chae DH, Gavin AR, Takeuchi DT. Smoking prevalence among asian americans. findings from the national latino and Asian American study (NLAAS). *Public Health Rep.* (2006) 121:755–63. doi: 10.1177/003335490612100616
85. Pampel FC. Patterns of tobacco use in the early epidemic stages. Malawi and Zambia, 2000–2002. *Am J Public Health.* (2005). 95:1009–15. doi: 10.2105/AJPH.2004.056895
86. Stubbs B, Veronese N, Vancampfort D, Prina AM. Lin P-Y, Tseng P-T, et al. Perceived stress and smoking across 41 countries a global perspective across Europe, Africa, Asia and the Americas. *Sci rep.* (2017) 7:7597. doi: 10.1038/s41598-017-07579-w
87. Faletau J, Glover M, Nosa V. Pienaar F. Looks like smoking, is it smoking? children's perceptions of cigarette–like nicotine delivery systems, smoking and cessation. *Harm Reduct J.* (2013) 10:30–30. doi: 10.1186/1477-7517-10-30
88. Misra S, Gelaye B, Koenen KC, Williams DR, Borba CPC, Quattrone D, et al. Early parental death and risk of psychosis in offspring a six-country case-control study. *J Clin Med.* (2019) 8:1081. doi: 10.3390/jcm8071081
89. Zhou H, Zhang Y, Liu J, Yang Y, Fang W, Hong S, et al. Education and lung cancer a Mendelian randomization study. *Int J Epidemiol.* (2019) 48:743–50. doi: 10.1093/ije/dyz121
90. Burgess S. Sample size and power calculations in mendelian randomization with a single instrumental variable and a binary outcome. *Int J Epidemiol.* (2014) 43:922–9. doi: 10.1093/ije/dyu005

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Zou, Wang, Yang, Wang, Fu, Huo, Ge, Zhong, Jiang, Li, Xiong, Hong and Liang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Pharmacology

SYSTEMATIC REVIEW
published: 06 June 2022
doi: 10.3389/fphar.2022.859296

Clinical Effectiveness of Conservative Treatments on Lumbar Spinal Stenosis: A Network Meta-Analysis

Xuanwei Chen[1], *Zhizhan Zheng*[2] *and Jianhua Lin*[1]*

[1]*The First Affiliated Hospital of Fujian Medical University, Fujian, China,* [2]*Xianyou Country General Hospital, Fujian, China*

Objective: To systematically evaluate the clinical effectiveness of conservative treatments including pharmacological treatments and nonpharmacological treatments on patients with lumbar spinal stenosis.

Methods: We searched six electronic databases systematically for randomized clinical trials published between January 2000 and July 2021, including the China National Knowledge Infrastructure, WanFang Data, PubMed, MEDLINE, Embase, and the Cochrane library. The studies focused on the therapeutic effects of pharmacological treatments including calcitonin, antiepileptics, neurotrophic drugs, nonsteroidal anti-inflammatory drugs, Chinese Traditional Medicine, limaprost, and nonpharmacological treatments like physiotherapy for treating lumbar spinal stenosis were included. The outcome was measured using the visual analog scale, Oswestry Dysfunction Index, Japanese Orthopaedic Association Score, and EuroQol Five Dimensions Questionnaire. The quality of eligible studies was assessed by using the Cochrane recommended bias risk assessment tool. Stata was used to conduct the network meta-analysis.

Results: A total of 12 randomized control trials with 1,194 patients were included. The network meta-analysis showed that for the visual analog scale, a better therapeutic effect was noted while using Chinese Traditional Medicine and physiotherapy, followed by analgesics drugs and limaprost. Limaprost and calcitonin were better in decreasing the Oswestry Dysfunction Index. In terms of the Japanese Orthopaedic Association Score, the use of traditional Chinese Medicine and limaprost were associated with a better improvement than other treatments. Meanwhile, limaprost combined with analgesics drugs was found to be effective to improve the EuroQol Five Dimensions Questionnaire.

Conclusion: Among the commonly used conservative treatments for the treatment of lumbar spinal stenosis, limaprost may have better efficacy in improving the Japanese Orthopaedic Association Score and decreasing the Oswestry Dysfunction Index, with a beneficial effect on decreasing the visual analog scale and improving the EuroQol Five Dimensions Questionnaire.

Systematic Review Registration: website, identifier registration number.

Keywords: visual analog scale, Japanese Orthopaedic Association Score, Oswestry Dysfunction Index, Euroqol Five Dimensions Questionnaire, limaprost

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Francesco Agostini,
Sapienza University of Rome, Italy
Mohammad Sami Al-Wardat,
Aqaba University of Technology,
Jordan

***Correspondence:**
Jianhua Lin
jianhual@126.com

Specialty section:
This article was submitted to
Drugs Outcomes Research and
Policies,
a section of the journal
Frontiers in Pharmacology

Received: *21 January 2022*
Accepted: *29 March 2022*
Published: *06 June 2022*

Citation:
Chen X, Zheng Z and Lin J (2022)
Clinical Effectiveness of Conservative
Treatments on Lumbar Spinal
Stenosis: A Network Meta-Analysis.
Front. Pharmacol. 13:859296.
doi: 10.3389/fphar.2022.859296

INTRODUCTION

Lumbar spinal stenosis (LSS) is a degenerative disease that commonly occurs in the elderly population (Verbiest, 1980), with a high prevalence rate worldwide (LaBan and Imas, 2003). Studies in Japan have shown that the prevalence of symptomatic LSS is 10.0% in the general population (Ishimoto et al., 2012) and approximately 22.5% in the general population of America (Kalichman et al., 2009). In China, the prevalence rate of LSS is approximately 3.9%–11.0% (Letao, 2021), and the rate is expected to increase with the aging of the Chinese population.

Degenerative changes in the intervertebral discs, ligamentum, and synovial joints lead to normal narrowing of the spinal canal and reduction of the internal diameter of the nerve root canal, which irritates or compresses the neurovascular vessels and causes a series of clinical symptoms (Rui, 2020). The typical symptoms of patients with LSS include back pain or intermittent claudication, resulting in the reduction of the patients' quality of life (QoL) and mobility (Mangone et al., 2020; Rui, 2020).

The major objective of LSS treatment is to reduce pain, improve numbness symptoms, and increase the patient's mobility. The guideline of LSS (Rui, 2020) recommends surgery as the best treatment for patients with severe LSS, and conservative treatments including pharmacological treatments like medications, nonpharmacological treatments like physiotherapy, lifestyle changes, and others for patients with mild-to-moderate LSS. The commonly used drugs include calcitonin, mecobalamin, gabapentin, celecoxib, loxoprofen, indomethacin, Chinese Traditional Medicine (CTM), and limaprost (Rui, 2020). Although these interventions are commonly used among patients with LSS in the real world, the efficacy of the interventions still remains unknown for related studies are limited. Therefore, our study aims to systematically evaluate the clinical effectiveness of conservative treatments on patients with LSS by network meta-analysis, thus providing evidence for clinical practice.

METHODS

Searching Strategy

A systematic search was conducted in six electronic databases, including the China National Knowledge Infrastructure (CNKI), WanFang Data, PubMed, MEDLINE, Embase, and Cochrane library from January 2000 to July 2021. The terms used for searching included LSS, the visual analog scale (VAS), the Japanese Orthopaedic Association (JOA) scores, the Oswestry Disability Index (ODI), and 6 min walking test. The detailed searching strategy is shown in the **Supplementary Material**.

Inclusion and Exclusion Criteria

The inclusion and exclusion criteria were followed by Patient-Intervention-Control-Outcome methods (da Costa Santos et al., 2007).

The main inclusion criteria were as follows: 1) individuals with LSS were included, excluding those with combined orthopedic diseases such as lumbar disc herniation and fracture, 2) randomized controlled studies (RCTs) focusing on the effectiveness comparison of conservative treatments were eligible to be included in this study, and 3) various outcomes related to the disease were considered to be included, but inclusion was not limited to pain (including leg pain and back pain), disability, physical function, the QoL, adverse events, the satisfaction rate of patients, and recovery. Moreover, the studies focusing on the effects of limaprost in Japanese were included for supplementary, as limaprost was widely used in Japan and generated a lot of clinical evidence.

The studies meeting one of the following criteria would be excluded: 1) publications that are duplicated or studies without full text and 2) studies with a study period of more than 3 months were excluded to maintain consistency.

Conservative treatments in clinical practice were far more complex than the guideline recommendation, so treatment classification was conducted. The pharmacological treatments were categorized as calcitonin, antiepileptics, neurotrophy drugs, analgesics, CTM, and limaprost according to their mechanism and effects. Since nonpharmacological treatments are messy in the real world and the related studies are limited, it is difficult to make further categories of nonpharmacological treatments; nonpharmacological treatments were defined as physiotherapy in this study. Limaprost was the only drug with a specific indication for the treatment of LSS, and its mechanism is to improve the microcirculation of patients with LSS which is different from other drugs (Kurokawa et al., 2011; Kim et al., 2016), making it a separate category. The treatment included in each category were as follows: antiepileptics included gabapentin; neurotrophy drugs included mecobalamin, vitamin B6, etc.; physiotherapy included acupuncture, massage, cupping, etc.; analgesics included nonsteroidal anti-inflammatory drugs (NSAIDs) such as celecoxib and etodolac, pregabalin, etc.; CTM included Biqi granules, Chinese herbal soup, etc. Drug combined with physiotherapy was defined as drug medication. As for the combined treatment of analgesics and neurotrophic drugs, it was classified as analgesics in this study. Moreover, abbreviations were used in tables/figures to be clearer. Lim represents limaprost, Ana represents analgesics drugs, Phy represents physiotherapy, Cal represents calcitonin, Antiep represents antiepileptic drugs, and Neu represents neurotrophic drugs.

As these clinical outcomes are assessed by different scales and methods, a wide range of indicators will be considered for inclusion in this study. Pain was measured using different scales, such as the VAS, Numeric Rating Scale (NRS), and Verbal Rating Scale (VRS). Disability was measured using scales such as the ODI and the Roland Morris Disability Questionnaire (RMDQ). Physical function was measured using the JOA score, max walking distance, pain-free walking distance, and 6-min walking distance. As for QoL, scales like the EuroQol Five Dimensions Questionnaire (EQ-5D) and the short-form health survey questionnaire were commonly used. Recovery was measured by the effectiveness rate and improvement rate. Adverse events were defined as any adverse events reported in the literature, with no restriction on the level of adverse events.

Study Screening

Following the search, all references of identified studies were imported into the reference management software, NoteExpress. After removing duplicates, two reviewers evaluated the titles and

abstracts of the identified studies according to the inclusion and exclusion criteria. Reasons for the exclusion were recorded and reported. For studies included by the title and abstract screening, the full text was reviewed carefully to determine the final inclusion decision. Any disagreement between the two reviewers regarding inclusion and exclusion at each stage of the selection process was resolved through discussions, or by a third reviewer.

Data Extraction and Quality Evaluation

Two reviewers independently collected data from the selected articles. The corresponding authors were contacted through email to collect detailed information, in case of any uncertainty. The extracted information from the included studies was as follows: 1) the basic information of the study, including the title, author, and year; 2) baseline characteristics of the study participants and the treatment; 3) outcomes and the results; and 4) the risk assessment of the study.

The quality of included studies was assessed by the bias risk assessment tool, which is recommended by the Cochrane Handbook version 6.3 for Systematic Reviews of Interventions (Higgins et al., 2022). This tool includes seven items, including random methods, allocation concealment, blinding researchers and participants, blinding research results, completeness of data, selective reporting of research results, and other biases. These items were classified into low risk, high risk, and unclear based on the specific research content.

Statistical Methods

A network meta-analysis was conducted by using Stata. For each clinical indicator, the network plot was drawn, respectively, to demonstrate the comparison of interventions. In the network plot, the lines represented the number of included studies and the node represented the number of patients included in the intervention. The thicker the node, the more studies included; the larger the node, the more patients included. As for the effectiveness, it was measured by the ranking of the area under the curve. The larger the area under the curve, the better the effect of the intervention. The node-splitting method was performed to test consistency, and $p < 0.05$ was used as the level of statistical significance. The sensitivity analysis was conducted by using the inverted funnel chart.

RESULTS

There were 7,302 articles with 4 **Supplementary Material** in total according to the searching strategy. After exclusion, 12 studies involving 1,194 patients with LSS were included. Among the included studies, eight studies were from China (750 patients), two studies were from Japan (172 patients), one study was from Iran (90 patients), and one study was from Korea (182 patients). The study selection flowchart is shown in **Figure 1**. The basic information of all included studies is shown in **Table 1**, and the results of the bias risk assessment are shown in **Table 2**.

Among all outcomes, only the VAS, ODI, JOA, and EuroQol Five Dimensions Questionnaire (EQ-5D) were available for analysis, while other outcomes could not be analyzed due to various reasons. Among the included articles, there was no article reporting any results of the VRS, SF-8, or 6-min walking distance. The NRS, the Roland Morris Disability Questionnaire (RMDQ), the MOS item short from the health survey (SF-36), and adverse effects could not be analyzed due to the limited data and unformed network. Moreover, there were various definitions of walking distance, improvement rate, and effectiveness rate in studies and therefore these outcomes were not analyzed.

Visual Analog Scale

There were 8 studies with 821 patients reporting the VAS. The evaluated treatments included the treatment with limaprost, analgesics drugs, physiotherapy, TCM, and limaprost combined with analgesics drugs (see **Figure 2**). As illustrated in the figure, a larger number of studies were noted in the intervention of TCM and physiotherapy, and more patients were reported in the intervention of analgesics drugs and physiotherapy.

The consistency model used for the analysis of the inconsistency test results was 0.07 > 0.05. The differences in interventions for reducing the VAS are shown in **Figure 3**. The area under the curve and the mean rank are shown in **Table 3**. Therefore, in terms of reducing the VAS, the ranking of efficacy was CTM > physiotherapy > analgesics drugs > limaprost > limaprost combined with analgesics drugs.

Oswestry Disability Index

There were five studies with 610 patients reporting the ODI. The evaluated treatments included the treatment with calcitonin, analgesics drugs, antiepileptic drugs, limaprost, limaprost combined with analgesics drugs, physiotherapy, and CTM (see **Figure 4**). As illustrated in the figure, more patients were noted in the intervention of physiotherapy.

The consistency model used for the analysis of the inconsistency test results was 0.81 > 0.05. The differences in interventions for reducing the ODI are shown in **Figure 5**, and the area under the curve and the mean rank are shown in **Table 4**. The results indicate that in terms of reducing the ODI, the therapeutic efficacy was ranked as limaprost > calcitonin > CTM > limaprost combined with analgesics drugs > analgesics drugs > antiepileptic drugs > physiotherapy.

Japanese Orthopaedic Association Scores

There were five studies with 445 patients reporting the JOA. The evaluated treatments included the treatment with analgesics drugs, limaprost, physiotherapy, neurotrophic drugs, and CTM (see **Figure 6**). As shown in the figure, more patients were noted in interventions of CTM and physiotherapy.

The inconsistency model used for the analysis of the inconsistency test results was 0.0008 < 0.05. The differences in interventions on increasing the JOA are shown in **Figure 7**, and the area under the curve and the mean rank are shown in **Table 5**. Based on these results, the efficacy ranking was CTM > limaprost > physiotherapy > neurotrophic drugs > analgesics drugs.

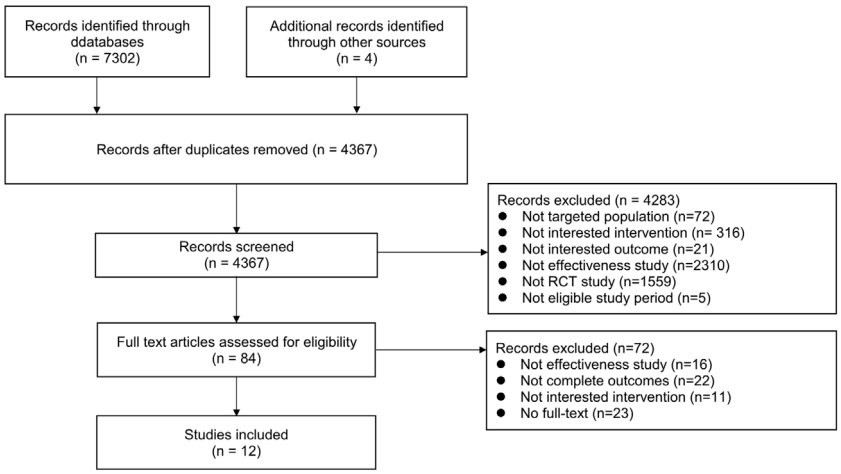

FIGURE 1 | Flowchart for selection of studies.

TABLE 1 | Characteristics of included studies.

Study	Number of Cases	Intervention Measure		Duration/Month	Outcome
		Intervention Group	Control Group		
Haddadi et al., 2016	30/30/30	Calcitonin	Antiepileptic drugs; physiotherapy	3	②
Kim et al., 2016	60/61/61	Limaprost	Analgesics; limaprost + Analgesics	2	①②④
Wei, 2013a	35/34	Neurotrophic drugs	CTM	1	③
Wei, 2013b	60/60	Analgesics	Physiotherapy	1.25	①②
Zhenqiu, 2019	40/40	CTM	Physiotherapy	2	①③
Jian, 2013	66/69	Analgesics	Physiotherapy	1	③
Yang, 2019	30/30	Analgesics	Physiotherapy	1	①
Haishan, 2020	58/58	CTM	Physiotherapy	3	①②
Xiaoping, 2020	34/34	CTM	Physiotherapy	1	①③
Xiaoxia, 2020	51/51	CTM	Physiotherapy	1	①②
Matsuyama, 2005	67/26	Limaprost	Physiotherapy	2	①③
Takahashi et al., 2013	56/23	Limaprost	Analgesics	2	④

Note: ①,VAS; ②, ODI; ③, JOA; ④, EQ-5D.

TABLE 2 | Bias assessment of included studies.

Studies	Method for randomization	Allocation concealment	Blinding		Completeness of outcome data	Selective outcome reporting	Other Bias
			Patients and researchers	Personnel for outcome measurement			
Haddadi et al., 2016	Low risk	Low risk	Unclear	Unclear	Low risk	Low risk	Low risk
Kim et al., 2016	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
Wei, 2013a	Unclear	Unclear	Unclear	Unclear	Low risk	Low risk	Low risk
Wei, 2013b	Unclear	Unclear	Unclear	Unclear	Low risk	Low risk	Low risk
Zhenqiu, 2019	Low risk	Unclear	Unclear	Unclear	Low risk	Low risk	Low risk
Jian, 2013	Low risk	Unclear	Unclear	Unclear	Low risk	Low risk	Low risk
Yang, 2019	Low risk	Unclear	Unclear	Unclear	Low risk	Low risk	Low risk
Haishan, 2020	Low risk	Unclear	Unclear	Unclear	Low risk	Low risk	Low risk
Xiaoping, 2020	Unclear	Unclear	Unclear	Unclear	Low risk	Low risk	Low risk
Xiaoxia, 2020	Unclear	Unclear	Unclear	Unclear	Low risk	Low risk	Low risk
Matsuyama, 2005	Low risk	Unclear	Unclear	Unclear	Low risk	Low risk	Low risk
Takahashi et al., 2013	Unclear	Unclear	Unclear	Unclear	Low risk	Low risk	Low risk

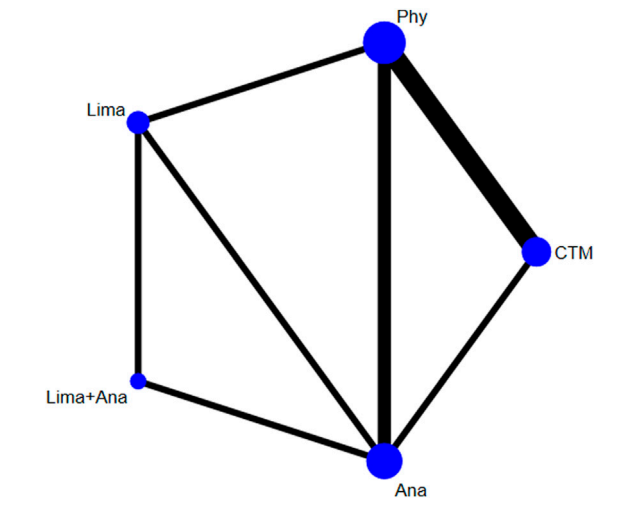

FIGURE 2 | Network plot for included studies on VAS.

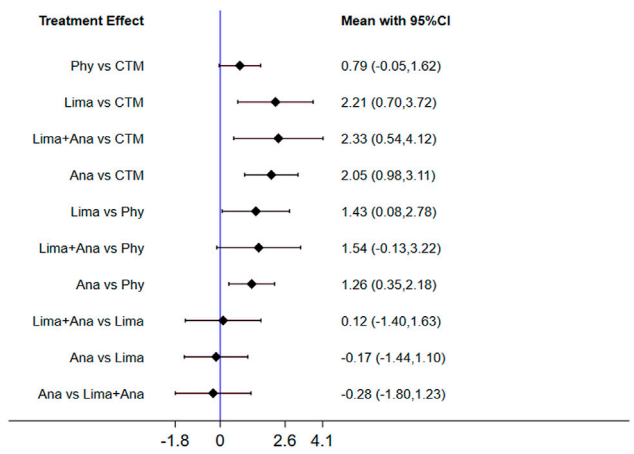

FIGURE 3 | Mean difference for VAS between treatments.

TABLE 3 | Area under the curve and the mean rank for VAS.

Treatments	Area under the Curve (%)	Mean Rank
CTM	99.0	1.0
Phy	74.4	2.0
Ana	31.0	3.8
Lima	24.5	4.0
Lima + Ana	21.1	4.2

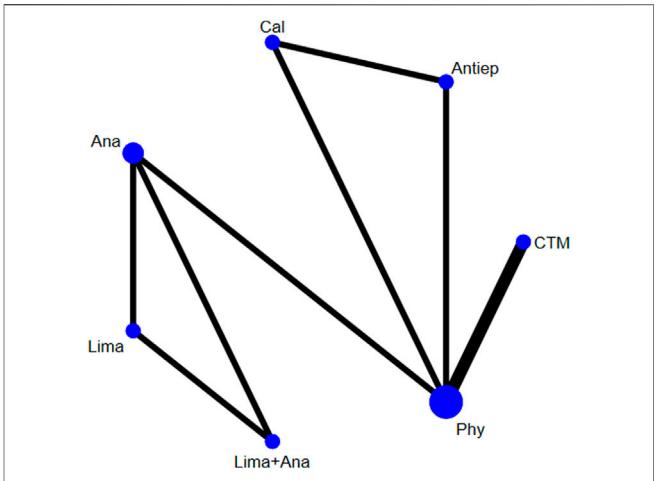

FIGURE 4 | Network plot for included studies on ODI.

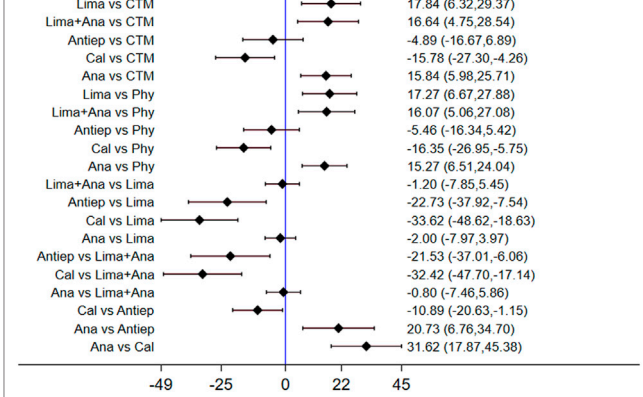

FIGURE 5 | Mean difference for ODI between treatments.

TABLE 4 | Area under the curve and the mean rank for ODI.

Treatments	Area Under the Curve (%)	Mean Rank
Lima	75.7	2.5
Cal	66.3	3.0
CTM	63.2	3.2
Lima + Ana	61.1	3.3
Ana	46.5	4.2
Antiep	44.3	4.3
Phy	26.1	5.4

EuroQol Five Dimensions Questionnaire

There were 2 studies with 261 patients reporting the EQ-5D. The evaluated treatments included the treatment with limaprost, analgesics drugs, and limaprost combined with analgesics drugs (see **Figure 8**).

The consistency model fitted for the analysis of the inconsistency test results was 0.90 > 0.05. The differences in interventions on increasing the EQ-5D are shown in **Figure 9**, and the area under the curve and the mean rank are shown in **Table 6**. These results demonstrate that the efficacy ranking was

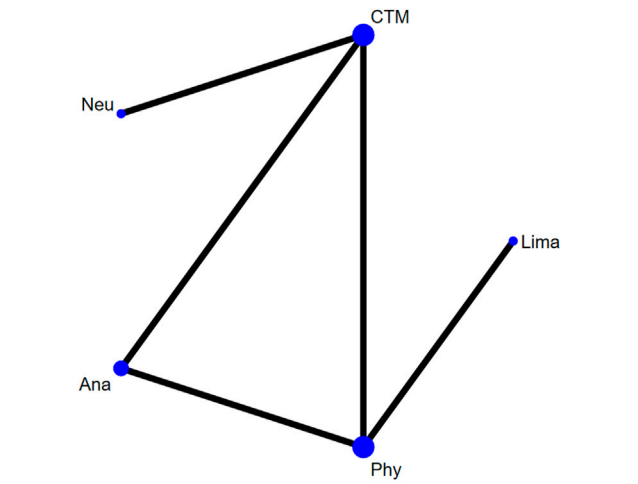

FIGURE 6 | Network plot for included studies on JOA.

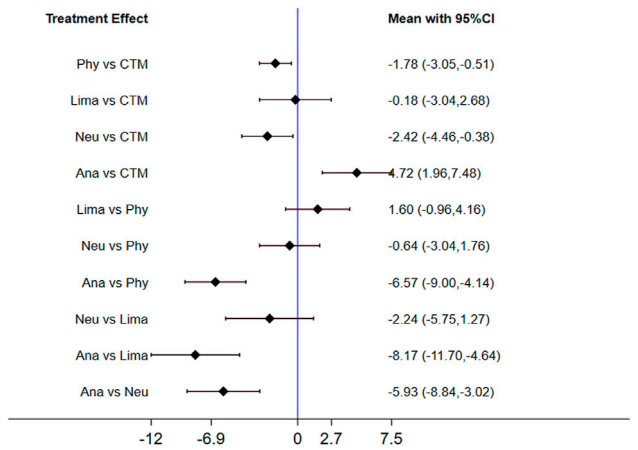

FIGURE 7 | Mean difference for JOA between treatments.

limaprost combined with analgesics drugs > analgesics drugs > limaprost.

Sensitivity Analysis

The inverted funnel charts of the VAS, ODI, and JOA were drawn to assess the publication bias, as shown in **Figure 10**. There was

TABLE 5 | Area under the curve and the mean rank for JOA.

Treatments	Area under the curve (%)	Mean Rank
CTM	88.3	1.5
Lima	81.0	1.8
Phy	45.8	3.2
Neu	34.9	3.6
Ana	0.0	5.0

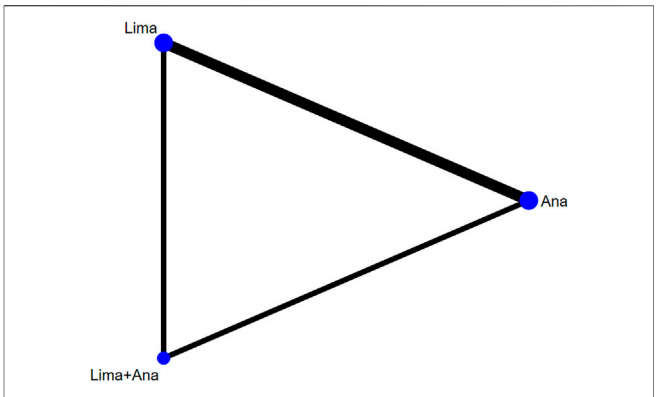

FIGURE 8 | Network plot for included studies on EQ-5D.

no further analysis on the EQ-5D due to the comparatively limited number of included studies using the EQ-5D as an outcome. These figures were basically symmetrically distributed, indicating the minimized publication bias among included studies. As for the clinical similarity and methodological similarity, an assessment was conducted by two reviews and the result showed that the similarity was robust and valid.

DISCUSSION

In the real world, the conservative treatment for patients with LSS includes pharmacological treatments like various medications and nonpharmacological treatments like exercising, massage, and so on. This study aimed to evaluate the effectiveness of the commonly used conservative treatment for patients with LSS, hoping to provide evidence for the clinical practice.

For reducing the VAS, CTM and physiotherapy were found to be the most effective treatments, which was consistent with the Chinese guideline (Group, 2014) and other studies. As mentioned in the Chinese expert consensus in 2014 (Group, 2014), physiotherapy including massage, acupuncture, electrotherapy, hyperthermia, and other treatments has relatively positive short-term efficacy than others. However, the long-term efficacy is not clear, which needs further study. A systematic review by Kim et al. (2013) evaluated the effects and safety of acupuncture, and the results showed that the acupuncture group had significantly favorable improvement over the control group on pain intensity, overall symptoms, and function outcomes related to LSS and the QoL. In this study, half of the physiotherapy was acupuncture, and all the CTM treatments were treatments combining physiotherapy and related drugs, which makes the ranking results more reasonable and comprehensible.

In terms of reducing the ODI and improving the JOA, the use of limaprost also demonstrated good efficacy than other commonly used drugs, with second ranking in the ODI and first ranking in the JOA, respectively. Up to now, limaprost is the only chemistry drug that is reported to have the specific indication of LSS, making it an important and preferred treatment in clinical practice (Matsudaira et al., 2009; Wang, 2019). As for other drugs, they are only able to relieve the

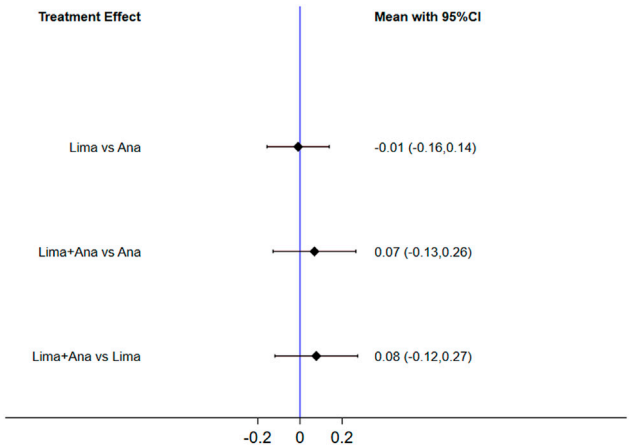

FIGURE 9 | Mean difference for EQ-5D between treatments.

TABLE 6 | Area under the curve and the mean rank for EQ-5D.

Treatments	Area under the curve (%)	Mean Rank
Lima + Ana	77.4	1.5
Ana	39.4	2.2
Lima	33.2	2.3

symptoms, to some extent. And this superior efficacy may be due to the unique mechanism of limaprost, to reduce pain by improving the blood flow and antiplatelet aggregation in the cauda equina or nerve roots (Matsudaira et al., 2009; Wang, 2019).In addition, evidence has been accumulated for limaprost, which has been preliminarily used in Japan since the 1980s. RCT conducted by Matsudaira et al. (2009) showed that limaprost resulted in considerably greater improvements in the SF-36 subscales, and was also significantly better than etodolac for leg numbness, neurogenic intermittent claudication distance, and subjective improvement and satisfaction. Kim et al. (2016) compared the effects of limaprost and pregabalin individually and in combination for the treatment of LSS. The results showed that the efficacy of limaprost for patients with LSS was not inferior compared to that of pregabalin or the combination of limaprost and pregabalin in terms of disability.

In addition, limaprost also showed better efficacy in improving patients' QoL. Kim et al. (2016) compared the EQ-5D scores of limaprost, pregabalin, and limaprost combined with pregabalin. The results showed that the baseline-adjusted EQ-5D scores improved significantly over time in all three groups after treatment. Takahashi et al. (2013) performed a survey to assess the relationships between the QoL and the therapy with limaprost or etodolac (an NSAID). The mean EQ-5D utility value for patients was 0.59 ± 0.12, which is much lower than the reported utility for patients with type 2 diabetes and stroke (0.86 and 0.84, respectively), suggesting that patients with LSS had poorer QoL. The EQ-5D utility value was increased by ≥ 0.1 points in significantly more patients in the limaprost group, whereas no significant changes occurred in the etodolac group.

This is the first study to evaluate the therapeutic efficacy of commonly used conservative treatments for LSS by network meta-analysis. However, there are some limitations to our study. First, the number of studies included in specific endpoints and the sample size of some studies are small, which may lead to the uncertainty of the results. Therefore, it is necessary to use further high-quality RCT studies to verify the findings of our study. Second, the clinical outcomes in this study are limited too. These results also indicate the lack of the current study in the LSS field, making it difficult for clinical practice. Further studies with long-term observation and a large sample shall be called on for better guidance and evidence support in clinics.

CONCLUSION

This network meta-analysis systematically assessed the clinical effectiveness of conservative treatments in patients with LSS worldwide. Our findings demonstrated that limaprost has better efficacy in reducing the ODI, improving the JOA, and improving the QoL, while physiotherapy and CTM are found most effective in reducing the VAS in the short term after treatment. Our study provides a ranking of the therapeutic efficacy of treatments, and thus, provides the appropriate evidence to support clinical practice. However, our study still has some limitations, including the lack of studies from western countries and the small population size in included

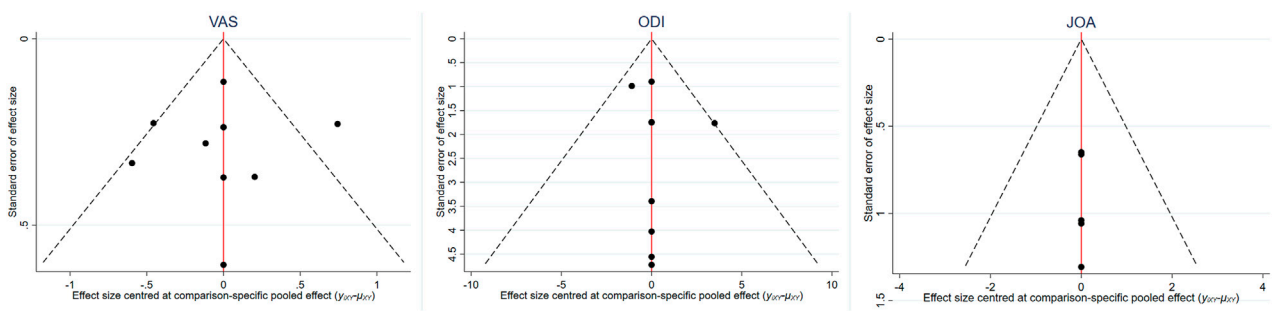

FIGURE 10 | Inverted funnel chart on VAS, ODI, and JOA.

studies. Thus, further studies with a large sample and high quality are preferred to generate more valid evidence and to make up for the limitations of this study.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

JL conceived and designed research; XC and ZZ collected data, conducted research, and analyzed data; XC wrote the initial paper; JL revised the paper; JL is primarily responsible for the final content. All authors reviewed the results and approved the final version.

ACKNOWLEDGMENTS

We thank Jia Fang of Tianjin Happy Life Tech Co., Ltd. for her contributions to the statistical analyses.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2022.859296/full#supplementary-material

REFERENCES

da Costa Santos, C. M., de Mattos Pimenta, C. A., and Nobre, M. R. C. (2007). The PICO Strategy for the Research Question Construction and Evidence Search. *Rev. Lat Am. Enfermagem* 15 (3), 508–511. doi:10.1590/s0104-11692007000300023

Group, C. E. C. (2014). Chinese Expert Consensus on the Standard of Surgical Treatment for Lumbar Spinal stenosis(2014)[J]. *Natl. Med. J. China* 94 (35), 2724–2725. doi:10.3760/cma.j.issn.0376-2491.2014.35.002

Haddadi, K., Asadian, L., and Isazade, A. (2016). Effects of Nasal Calcitonin vs. Oral Gabapentin on Pain and Symptoms of Lumbar Spinal Stenosis; A Clinical Trial Study[J]. *Clin. Med. Insights : Arthritis Musculoskelet. Disord.* 9, 133–138. doi:10.4137/CMAMD.S39938

Haishan, S. H. T. S. T. (2020). Clinical Observation of Treating Lumbar Spinal Stenosis Using Modified Buyan Huanwu Decoction Combined with Core Stability Exercises[J]. *Chin. J Trad Med Traum Orthop* 28 (07), 21–23.

Ishimoto, Y., Yoshimura, N., Muraki, S., Yamada, H., Nagata, K., Hashizume, H., et al. (2012). Prevalence of Symptomatic Lumbar Spinal Stenosis and its Association with Physical Performance in a Population-Based Cohort in Japan: The Wakayama Spine Study. *Osteoarthritis Cartilage* 20 (10), 1103–1108. doi:10.1016/j.joca.2012.06.018

Jian, Q. Z. C. G. W. (2013). Clinical Research on Acupuncture Plus Cupping with Retaining of Needle in the Treatment of Treating Lumbar Spinal Stenosis[J]. *Chinese Medical Herald* 10 (8), 3. doi:10.3969/j.issn.1673-7210.2013.08.048

Higgins, J. P. T., Savović, J., Page, M. J., Elbers, R. G., and Sterne, J. A. C. (2022). Chapter 8: Assessing Risk of Bias in a Randomized trial[EB/OL]. [2022-03-22]. Available at: https://training.cochrane.org/handbook/current/chapter-08.

Kalichman, L., Cole, R., Kim, D. H., Li, L., Suri, P., Guermazi, A., et al. (2009). Spinal Stenosis Prevalence and Association with Symptoms: The Framingham Study. *Spine J.* 9 (7), 545–550. doi:10.1016/j.spinee.2009.03.005

Kim, H. J., Kim, J. H., Park, Y. S., Suk, K. S., Lee, J. H., Park, M. S., et al. (2016). Comparative Study of the Efficacy of Limaprost and Pregabalin as Single Agents and in Combination for the Treatment of Lumbar Spinal Stenosis: A Prospective, Double-Blind, Randomized Controlled Non-inferiority Trial. *Spine J.* 16 (6), 756–763. doi:10.1016/j.spinee.2016.02.049

Kim, K. H., Kim, T. H., Lee, B. R., Kim, J. K., Son, D. W., Lee, S. W., et al. (2013). Acupuncture for Lumbar Spinal Stenosis: A Systematic Review and Meta-Analysis. *Complement. Ther. Med.* 21 (5), 535–556. doi:10.1016/j.ctim.2013.08.007

Kurokawa, R., Nagayama, E., Murata, H., and Kim, P. (2011). Limaprost Alfadex, a Prostaglandin E1 Derivative, Prevents Deterioration of Forced Exercise Capability in Rats with Chronic Compression of the Spinal Cord. *Spine (Phila Pa 1976)* 36 (11), 865–869. doi:10.1097/BRS.0b013e3181e878a1

LaBan, M. M., and Imas, A. (2003). "Young" Lumbar Spinal Stenotic: Review of 268 Patients Younger Than 51 Years. *Am. J. Phys. Med. Rehabil.* 82 (1), 69–71. doi:10.1097/01.PHM.0000034918.76195.5E

Letao, D. J. J. (2021). Advances in the Treatment of Lumbar Spinal Stenosis [J]. *J. Reg. Anat. Oper. Surg.* 30 (11), 1012–1017. doi:10.11659/jjssx.03E021036

Mangone, M., Paoloni, M., Procopio, S., Venditto, T., Zucchi, B., Santilli, V., et al. (2020). Sagittal Spinal Alignment in Patients with Ankylosing Spondylitis by Rasterstereographic Back Shape Analysis: An Observational Retrospective Study. *Eur. J. Phys. Rehabil. Med.* 56 (2), 191–196. doi:10.23736/S1973-9087.20.05993-6

Matsudaira, K., Seichi, A., Kunogi, J., Yamazaki, T., Kobayashi, A., Anamizu, Y., et al. (2009). The Efficacy of Prostaglandin E1 Derivative in Patients with Lumbar Spinal Stenosis. *Spine (Phila Pa 1976)* 34 (2), 115–120. doi:10.1097/BRS.0b013e31818f924d

Matsuyama, H. Y. E. A. Y. (2005) Conservative Treatment Using PGE1 for Lumbar Canal Stenosis. Multicenter Comparative Prospective Clinical Study[J]. *Clinical Surgery* 6: 673–680.

Rui, Y. J. X. (2020). Surgical and Non-surgical Treatment of Patients with Lumbar Spinal Stenosis: Interpretation of Danish National Clinical Guidelines[J]. *Tianjin J. Traditional Chin. Med.* 37 (01), 33–37. doi:10.11656/j.issn.1672-1519.2020.01.08

Takahashi, J., Kobayashi, H., Wakabayashi, S., Deguchi, M., Ito, H., Mogami, Y., et al. (2013). The Effect of a Prostaglandin E1 Derivative on the Symptoms and Quality of Life of Patients with Lumbar Spinal Stenosis. *J. Orthop. Sci.* 18 (2), 208–215. doi:10.1007/s00776-012-0342-5

Verbiest, H. (1980). Stenosis of the Lumbar Vertebral Canal and Sciatica. *Neurosurg. Rev.* 3 (1), 75–89. doi:10.1007/BF01644422

Wang, K. M. W. (2019). Progress in the Treatment of Lumbar Spinal Stenosis [J]. *Orthopaedics* 10 (03), 248–252. doi:10.3969/j.issn.1674-8573.2019.03.018

Wei, C. J. S. B. J. (2013). Clinical Trails of Shi's Zhilong Juanbi Pian Treating Lumbar Spinal Stenosis[J]. *Chin. J Trad Med Traum Orthop* 21 (9), 3.

Wei, H. J. X. H. L. (2013). A Randomized Controlled Study for Non-surgical Comprehensive Treatment for Degenerative Lumbar Spinal Stenosis[J]. *Chin. J Trad Med Traum Orthop* 21 (12), 17–20.

Xiaoping, Y. Y. S. R. L. (2020). Clinical Efficacy of Duhuo Jisheng Decoction Combined with Wentong Acupuncture Therapy on Degenerative Lumbar Spinal Stenosis[J]. *World Chinese Medicine* 15 (21), 5. doi:10.3969/j.issn.1673-7202.2020.21.021

Xiaoxia, Y. H. W. Y. Y. (2020). Observation on the Curative Effect of Modified Buyang Huanwu Decoction Combined with Manipulation in the Treatment of

Lumbar Spinal Stenosis[J]. *Chinese Community Doctors* 36 (36), 2. doi:10.3969/j.issn.1007-614x.2020.36.048

Yang, B. G. B. (2019). Clinical Observation of Minimally Invasive Treatment of Lumbar Spinal Stenosis by Traditional Chinese Medicine[J]. *Diabetes world* 16 (11), 13.

Zhenqiu, Z. M. X. (2019). Clinical Effect of Yishen Huoxue Decoction Combined with Electricity on Patients with Degenerative Lumbar Spinal Stenosis[J]. *Chin. Traditional Patent Med.* 41 (11), 4. doi:10.3969/j.issn.1001-1528.2019.11.052

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Chen, Zheng and Lin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Public Health

POLICY AND PRACTICE REVIEWS
published: 17 June 2022
doi: 10.3389/fpubh.2022.861067

Influencing Factors of Health Technology Assessment to Orphan Drugs: Empirical Evidence in England, Scotland, Canada, and Australia

Na Zhou[1,2†], Hong Ji[3†], Zheng Li[1], Jun Hu[4], Jia-Hua Xie[1], Yu-Heng Feng[1] and Ni Yuan[1]*

[1] Department of Social Medicine, School of Public Health, Dalian Medical University, Dalian, China, [2] Department of Health Policy and Management, School of Public Health, Peking University, Beijing, China, [3] Department of Pediatrics, The First Affiliated Hospital of Dalian Medical University, Dalian, China, [4] Shanghai Medical Products Administration, Shanghai, China

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Pei Wang,
Fudan University, China
Ana V. Pejcic,
University of Kragujevac, Serbia
Haiyin Wang,
Shanghai Health Development
Research Center, China

***Correspondence:**
Ni Yuan
nier1209@163.com

[†] These authors share first authorship

Specialty section:
This article was submitted to
Public Health Policy,
a section of the journal
Frontiers in Public Health

Received: 24 January 2022
Accepted: 23 May 2022
Published: 17 June 2022

Citation:
Zhou N, Ji H, Li Z, Hu J, Xie J-H,
Feng Y-H and Yuan N (2022)
Influencing Factors of Health
Technology Assessment to Orphan
Drugs: Empirical Evidence in England,
Scotland, Canada, and Australia.
Front. Public Health 10:861067.
doi: 10.3389/fpubh.2022.861067

This study summarizes the intrinsic criteria for the recommendation of orphan drugs in England, Scotland, Canada, and Australia with the aim of understanding the rationale for the variability in decision-making and to provide a reference for the establishment of criteria in the process of access to health insurance for orphan drugs in different countries and the construction of national uniform criteria. A comparative analysis of 60 health technology assessment (HTA) guidelines of 15 drug-indication pairs appraised by four countries (England, Scotland, Canada, and Australia) from 2017 to 2018 was done, including an in-depth analysis of a case study. Agreement levels were measured using kappa scores. Associations were explored through correspondence analysis. The four countries possess some homogeneity in the assessment, but each has its own preferences. Poor agreement exists between England, Scotland, and Canada ($-0.41 <$ kappa score < 0.192). In the correspondence analysis, England placed more emphasis on treatment methods in terms of control type when making recommendations. Canada and Scotland focused more on trial type with Canada placing more emphasis on phase III and open-label trials and on cost-utility analysis, while Australia was less studied in terms of economic models. Different countries have different goals when establishing HTA decisions for orphan drugs due to their different degrees of orphan drug coverage. Different countries should not only combine their unique values of clinical benefit and cost-effectiveness in the assessment of orphan drugs but also give different weights during the HTA process, after considering account the development of the country itself.

Keywords: orphan drugs, health technology assessment (HTA), consistency analysis, correspondence analysis, health insurance reimbursement

INTRODUCTION

The incidence of rare disease is low; however, rare diseases, which are defined differently in different countries and regions, often endanger health and life. The general principle of the definition of a rare disease is based on the prevalence or number of people with the disease. In the United States (US), a rare disease is defined as a condition that affects fewer than 200,000 people in the country (1).

In the European Union (EU), a rare disease is one that affects no more than 1 person in 2,000 people (2). In Australia, a disease is considered rare if it affects <5 in 10,000 people (3).

Health technology assessment (HTA) is a systematic evaluation of the safety, effectiveness, and economic and social aspects of medical technologies *via* an analysis of knowledge about health technologies based on research and practice using the principles and methods of evidence-based medicine (4). In terms of safety, the number of patients with rare diseases is small compared to common diseases. This factor makes it difficult to conduct traditional randomized controlled double-blind trials of orphan drugs to test clinical efficacy, resulting in higher difficulties in the design and implementation of clinical trials, often allowing only single-arm trials, with historical treatment as the default control, or real-world studies to observe the actual clinical effects of orphan drugs after they are launched (5). In terms of economics, the small population of indications for orphan drugs and the fact the cost of diagnosis and treatment of orphan drugs is often high make it difficult to meet the requirements of the traditional incremental cost-effectiveness ratio (ICER) or quality-of-life (QoL) ratio thresholds for evaluation (6). Novartis' Zolgensma, a one-time gene therapy for spinal muscular atrophy tops the list with a price tag of $2.125 million in 2021 (7). Some articles point out that by 2024, orphan drugs are expected to reach $242 billion, capturing one-fifth of worldwide prescription sales, and orphan drugs sale in the world are expected to grow at a compound annual growth rate of 12.3% from 2019 to 2024, which is approximately double the rate foreseen for non-orphan drugs market (8). In addition, other reasons for these high costs, such as the high heterogeneity and low recognition rate of rare diseases that make the diagnosis and treatment difficult and the cost of orphan drugs higher, have been reported. Due to these specificities of HTA for rare diseases, traditional methods, such as randomized controlled trials (RCTs) and cost-effectiveness analyses, may not be applicable to many rare diseases, and orphan drugs are often more prone to uncertainty in cost-effectiveness assessments than common diseases, which makes the process of HTA and health decision making more difficult (9).

It has been shown that an increasing number of countries use HTA to provide relevant information for health technology reimbursement (10). In HTA, cost-effectiveness analysis is one of the commonly used methods for which ICER, especially ICER based on quality-adjusted life years (QALY), has a wide range of applications (11). In the reimbursement of orphan drugs, quality adjusted life year (QALY) is the main QoL indicator for reimbursement decisions, while the treatment of rare diseases can meet the cost effectiveness criteria required for the use of ICER. Relatively speaking, ICER and QALY are not the only factors influencing the reimbursement decision as other social factors that influence the reimbursement decision can be found (12). The National Institute for Health and Clinical Excellence (NICE) gives a relatively generous margin to rare diseases when applying QALY as the threshold for evaluation indicators (13).

At the same time, reimbursement for orphan drugs has encountered challenges, both because clinical evidence may be difficult to collect due to the small number of patients and because their treatments do not meet the cost effectiveness criteria required to use ICERs and are associated with the high price and uncertainty of orphan drugs. More flexible criteria for assessing the clinical value of orphan drugs along with more high-quality clinical evidence for the use of orphan drugs may be needed to increase the likelihood of orphan drug reimbursements by enabling regulators and payers to consistently assess the risks and benefits of orphan drugs in a more efficient manner (12).

Drug reimbursement methods vary across countries ranging from universal health insurance for reimbursement to third-party payment reimbursement by insurance companies, which are separately funded and administered with different reimbursement rates. However, although a proportion of orphan drugs are reimbursed, the proportion of orphan drugs that are reimbursed is still only a small fraction of the total number of orphan drugs (14, 15). In terms of orphan drugs, countries have successively conducted the experience of exploring HTA for rare diseases; however, a need to actively explore a more normative HTA program still exists (9).

After considering well-established assessment systems, similar decision-making criteria, different technical approaches, publicly available HTA reports, and language, four countries were selected for this study: (1) England, (2) Scotland, (3) Australia, and (4) Canada, with the National Institute for Health and Care Excellence (NICE) (16), Scottish Medicines Consortium (SMC) (17), Pharmaceutical Benefits Advisory Committee (PBAC) (18, 19), and Canadian Agency for Drugs and Technologies in Health (CADTH) (20). The purpose of this study was to systematically compare health technology assessment procedures of orphan drugs in these four countries to summarize the intrinsic criteria of orphan drugs recommended in these four countries. The reasons for the variability in access decisions were explained, and countermeasures were proposed to reduce the variability and provide a reference for the establishment of criteria in the process of access to health insurance for orphan drugs. This information will provide a reference for the establishment of criteria for orphan drugs in the health insurance access process.

OBJECTS AND METHODS

Research Subjects

The report was selected from the official websites of the health technology assessment agencies of the four countries, namely NICE, SMC, PBAC, CADTH. NICE produces health technology appraisal guidance to ensure that all National Health Service (NHS) patients have equal access to the most clinically effective and cost-efficient treatments (21). The SMC evaluates drugs to make decisions to recommend, not recommend, or recommend use with restrictions, and drugs approved by the SMC are automatically entered into the Area Drug and Therapeutics Committees (ADTC) reimbursement list. Drugs that are not approved are not entered on the ADTC's reimbursement list, and use of such drugs will be restricted (22). Therapeutic Goods Administration (TGA) is responsible for the marketing approval and orphan drug status of pharmaceutical products in Australia. The Life Saving Drugs Program (LSDP), a free orphan drug program for patients with rare diseases, is available from the

LSDP. Before a drug can be offered on the LSDP, it must be evaluated by PBAC as clinically necessary and effective, but it is not recommended for inclusion in the Pharmaceutical Benefits Scheme (PBS) due to unacceptable cost effectiveness (23, 24). In Australia, the reimbursement evaluation of orphan drugs is independent of their cost-effectiveness as long as sufficient evidence of their clinical benefits has been demonstrated (25). The CADTH's recommendation for inclusion on the provincial reimbursement list can be classified as "recommended for inclusion," "recommended for conditional inclusion," "not recommended for inclusion at current submission prices," and "not recommended for inclusion," based on a combination of clinical effectiveness and economics (26–28).

This study used the orphan drugs queried by NICE in 2017–2018 as a benchmark and queried the corresponding orphan drugs reports for the same drug in three other countries, including a total of 65 orphan drugs. Since this study was expected to explore the factors that influence the relevant authorities in the four countries when conducting orphan drug assessments by having different recommendations for the same drug in different country assessments, orphan drugs that did not have identical assessment results in the four countries were included in this study; thus, 15 drugs were finally included (**Table 1**).

Research Methods

Nvivo Qualitative Analysis

This study used the Nvivo software to code the reports and systematically present trends in the focus of clinical evidence and economic models found in different country reports. NVivo is one of the computer-assisted qualitative data analysis software systems (CAQDAS) developed by QSR International (Melbourne, Australia), the world's largest qualitative research software developer. This software allows for qualitative inquiry beyond coding, sorting, and retrieval of data. It was also designed to integrate coding with qualitative linking, shaping, and modeling (29). By importing the queried assessment reports into Nvivo 12.0, the reports in the file were coded correspondingly (**Appendix Table 1**) after the creation of the node system (**Appendix Figures 1A–C**), and the criteria present in the access decisions in the four countries in terms of clinical evidence were explored through the case node classification and node matrix.

SPSS Consistency Analysis and Correspondence Analysis

The coded nodes were exported to an Excel spreadsheet, and the data were entered with the help of SPSS 20.0. A kappa consistency and correspondence analyses were performed for the HTA of orphan drugs in four countries. In this paper, the secondary nodes "clinical trials," "clinical endpoint," and "economic models" under the primary node "clinical evidence" were selected. By analyzing the correspondence between the country or recommendation situation and the assessment evidence, a scatter plot distribution between the country or recommendation situation and each node element is presented. This analysis was done to analyze the preference of countries by vector distance, specifically by connecting a line from the center (0.0) to any point and extending it in the reverse direction, and then drawing the other points to this line and its extension to make a vertical line. The closer the vertical point is to the positive vector direction, indicates the closer the relationship is, and the preference can be ranked.

RESULTS

Consistency Analysis of the Four Countries' Assessments

The values in the **Table 2** are Kappa scores with progressively higher concordance ranging from −1 to 1 during the consistency testing of these 15 drugs. This time, only three other national technology assessments were tested for consistency with each other, and the results are shown in **Table 2**. The data in the table show a poor consistency between England, Scotland, and Canada ($-0.410 <$ kappa score < 0.192), but in comparison, the consistency between England and Scotland is stronger.

Analysis of the Main Influencing Factors in Each Country

Correspondence Analysis Between Country and Clinical Evidence

After performing the corresponding analysis of national and clinical evidence, the SPSS result indicated that the first and second dimensions carried 89% of the information with a chi-square value of 10.436, $p = 0.316 > 0.05$, meaning that the association was not significant (**Figure 1**).

According to the vector analysis method, using the "trial type" in **Figure 1** as an example, a vector from the center to the "trial type" was drawn and then a vertical line (blue line) from England, Scotland, Australia, and Canada to this line and the extension line (red line) was drawn. Scotland's distance to the vertical line was closer to the positive vector direction, indicating that Scotland is more concerned about the "trial type". The other results are analyzed in the same way to roughly derive the main concerns of each country.

Each of the four countries has its own focus on the process of health technology assessment with England being strongly associated with the control type, indicating that England focuses more on the control type. Australia focuses more on the effect type followed by the assessment of the main effect evidence. Canada focuses more on the trial type followed by Scotland. In terms of the control type, the emphasis extended from highest to lowest in England, Scotland, Australia, and Canada, while the four countries with a higher emphasis on the effect type were Australia and England.

Correspondence Analysis Between Country and Economic Model

After the corresponding analysis of the country and economic model, the SPSS result indicates that the first and second dimensions carry 97.4% of the information, and the chi-square value was 25.038, $p = 0.124 > 0.05$, showing that the association was not significant (**Figure 2**).

According to the vector analysis method, "cost-effectiveness analysis" in **Figure 2** was used as an example, and a vector from

TABLE 1 | HTA assessment for orphan drugs.

No.	Generic name (trade name)	Manufacturers	Year	Whether to review (yes, recommended and type of recommendation/yes, not recommended/no)			
				NICE (England)	SMC (Scotland)	PBAC (Australia)	CADTH (Canada)
1	Ibrutinib (Imbruvica)	Janssen	2017	Recommended	Not recommended	–	Not recommended
2	Everolimus (Afinitor)	Novartis Pharmaceuticals	2017	Recommended	Recommended	–	Not recommended
3	Daratumumab (Darzalex)	Janssen	2017	Recommended	Restricted use	–	NOT recommended
4	Pegylated liposomal irinotecan (Onivyde)	Shire	2017	Not recommended	NOT recommended	Recommended	–
5	Afatinib (Giotrif)	Boehringer Ingelheim Pty Ltd	2017	Not recommended	Not recommended	Recommended	Recommended
6	Vandetanib (Caprelsa)	Sanofi	2018	Not recommended	Not recommended	Recommended	Recommended
7	Romiplostim (Nplate)	Amgen	2018	Recommended	Restricted use	–	Not recommended
8	Eltrombopag (Revolade)	GlaxoSmithKline	2018	Recommended	Restricted use	Recommended	Not recommended
9	Cabozantinib (Cabometyx)	Ipsen	2018	Recommended	Not recommended	–	–
10	Cenegermin (Oxervate)	Dompé	2018	Not recommended	Not recommended	Recommended	–
11	Cabozantinib (Cometriq)	–	2018	Recommended	Not recommended	–	–
12	Obinutuzumab (Gazyvaro)	Roche	2018	Recommended	Not recommended	Recommended	Recommended
13	Ixazomib with lenalidomide and dexamethasone (Ninlaro)	Takeda	2018	Recommended	Not recommended	Recommended	Not recommended
14	Pirfenidone (Esbriet)	Roche	2018	Recommended	Restricted use	Recommended	Not recommended
15	Lenvatinib with everolimus (Kisplyx)	Eisai	2018	Recommended	Recommended	Recommended	Not recommended

"-" Reports on health technology assessment opinions on pharmaceuticals are not available on the official website of health technology assessment agencies in Canada or Australia.

the center to the "cost-effectiveness analysis" was constructed after which a vertical line (blue line) from England, Scotland, Australia, and Canada to this line and its extension (red line) was drawn. The vertical line corresponding to Australia was closer to the positive vector direction, indicating that Australia is more concerned with "cost-effectiveness analysis"

After performing other vector distance analyses, it can be seen that in the analysis of health technology assessment reports in the four countries, Canada focuses most on the assessment of cost-utility and cost-minimization analyses. Scotland is closely related to many elements in the assessment process, including cost-minimization analysis, series of treatment models, and cost-utility analysis. England and Australia focus on cost-effectiveness and for cost-utility analyses, and the emphasis is on the order of Canada, Scotland, England, and Australia.

Correspondence Analysis Between Recommendation and Clinical Evidence

After obtaining the correspondence analysis of the recommendation situation and clinical evidence, the SPSS result indicated that the first and second dimensions carried nearly 100% of the information, and the chi-square value was 6.621, $p = 0.357 > 0.05$, indicating that the association was not significant (**Figure 3**).

According to the vector analysis method, taking the "recommendation" in **Figure 3** as an example, the vector was

TABLE 2 | Kappa consistency test scores for the four country assessments.

	England	Scotland	Canada	Australia
England	—	0.192	−0.410	*
Scotland		—	−0.320	*
Canada			—	*
Australia				—

The column "" is not in line with the Kappa test conditions.*

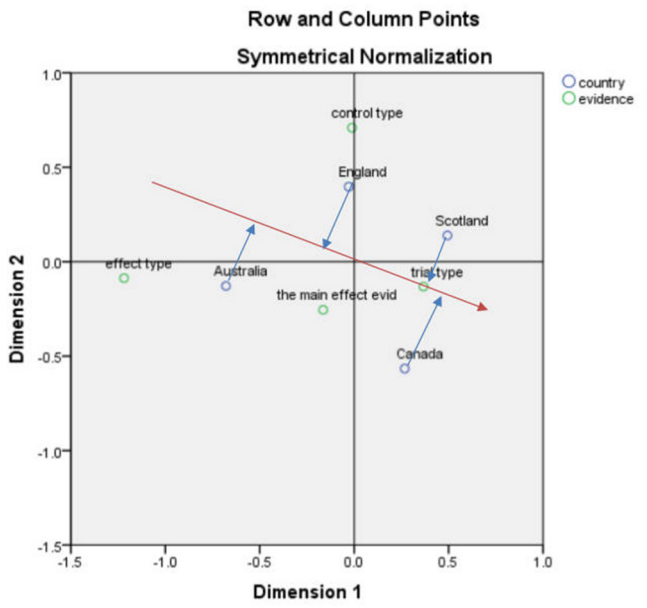

FIGURE 1 | Correspondence analysis biplot (country and clinical evidence).

constructed from the center to the "recommendation", and then the "effect type", "main effect evidence" and "control type" were made to this straight line and the extension line (red line) to make a vertical line (blue line). The vertical line made by the "effect type" was closer to the vector positive direction, which indicated the "effect type" was closer to the "recommendation".

Among the three types of recommendation, the recommendation focused on the effect type, evidence of main effects, and control type in order, which were optimized to facilitate good results in the HTA evaluation of drugs, The non-recommendation and restriction types focused more on the evaluation of the trial type, which may indicate that the trial type influences the national recommendation of a drug when it is evaluated.

Correspondence Analysis Between Recommendation and Economic Model

After the correspondence analysis of the recommended and economic models, the SPSS result indicated that the first and second dimensions carried nearly 100% of the information with a chi-square value of 17.555, $p = 0.130 > 0.05$, showing that the association was not significant (**Figure 4**).

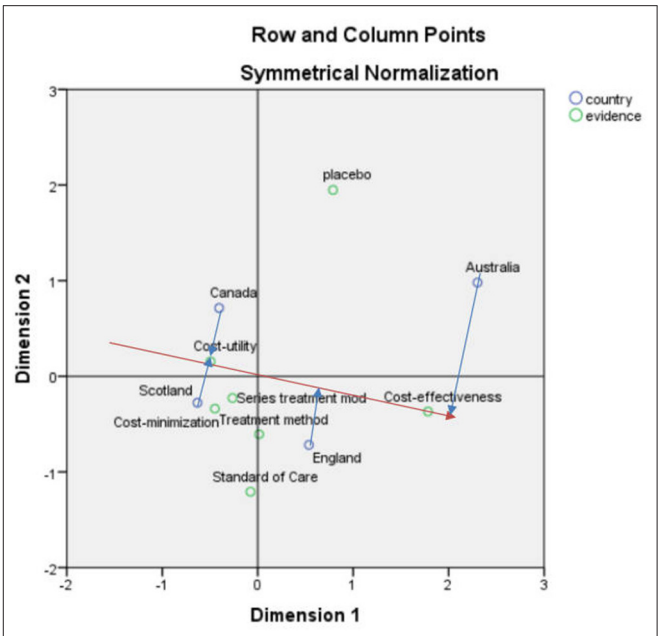

FIGURE 2 | Scatterplot of country and economic model correspondence analysis.

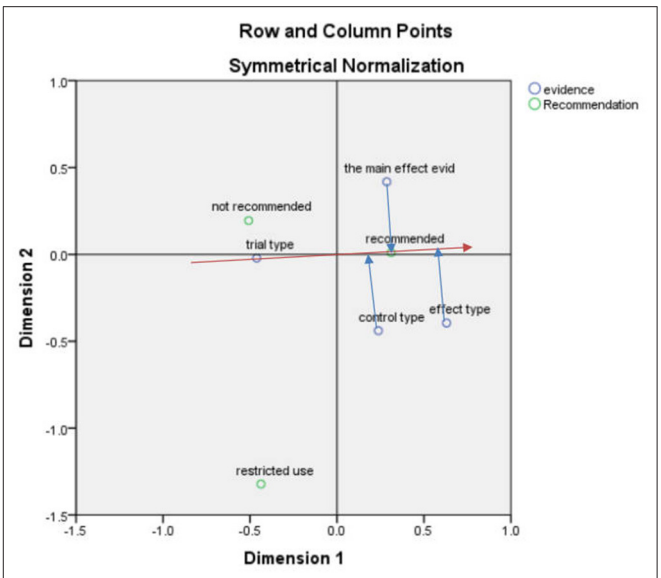

FIGURE 3 | Correspondence analysis biplot (recommendation and clinical evidence).

According to the vector analysis method, taking the "recommendation" in **Figure 4** as an example, the vector was constructed from the center to the "recommendation" after which the "placebo," "series of treatment models," "cost-minimization analysis" and "cost-effectiveness analysis" were drawn to this straight line and the extension line (red line) to the vertical line (blue line). The vertical line formed by the "cost-effectiveness analysis" was closer to the positive vector direction,

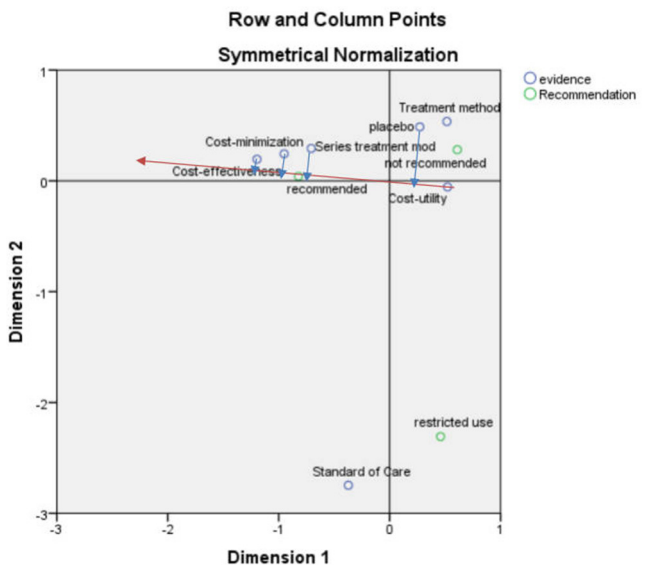

FIGURE 4 | Correspondence analysis biplot (recommended scenarios and economic model).

indicating that the "cost-effectiveness analysis" was closer to the "recommendation".

According to the scatter plot (**Figure 4**), the recommendation focused on the elements of cost-effectiveness and cost-minimization analyses, serial of treatment models, standard of care, and placebo in that order, which can be considered to have some influence on the recommendation, but most focus was on cost-effectiveness analysis. The non-recommendation focused more on treatment method, cost-utility analysis, and placebo, and the restricted recommendation focused more on standard of care.

Case Analysis

To present a more intuitive picture of the process of forming opinions on drug reviews by national health technology assessment bodies, the drug assessment process can be outlined by means of a case study, which can then be combined with the above conclusions for the purpose of promoting understanding or mutual argumentation. The subject of this case study was Obinutuzumab.

Introduction to Obinutuzumab

Rituximab was first approved by the United States Food and Drug Administration (FDA) on November 26, 1997 for the treatment of relapsed or refractory non-Hodgkin's lymphoma (NHL). As the concept of rituximab resistance emerged, in this context, scientists developed Obinutuzumab as a way to compensate for the lack of rituximab. Obinutuzumab is a novel, humanized, type II glycoengineered monoclonal antibody (mAb) directed against the CD20 antigen found on the surface of most malignant and benign B-cell-derived cells. The glycoengineering process used in the development of this agent increases its anti-lymphoma activity by enhancing binding affinity to the FcγRIII receptor on immune effector cells (30).

Obinutuzumab Recommendations by Four Countries

Roche Pharmaceuticals submitted applications for evaluation of Obinutuzumab to four countries in 2016–2017 and the HTAs and agencies in each of the four countries gave different evaluations. Overall, the four countries analyzed the value of the drug in terms of clinical effectiveness, safety, and cost-effectiveness to reach a conclusion on whether to recommend the drug.

Scotland explicitly does not recommend the drug, whereas Australia and Canada both recommend it. England also recommended the drug although with two conditions: (1) a Follicular Lymphoma International Prognostic Index (FLIPI) score >2 in patients with follicular lymphoma and (2) the company offered Obinutuzumab at the same discount as agreed in the patient access program.

Analysis of the Four-Country Assessment Process

Throughout the health assessment reports from the four countries, the first common point of assessment that was addressed was drug effectiveness. The clinical evidence from Australia, Scotland, and the UK all came from the same trial, GALLIUM, while the clinical evidence from Canada came primarily from another clinical trial, GADOLIN. The clinical evidence from the GALLIUM trial is representative of the three countries and meets the requirements for testing the clinical effectiveness of Obinutuzumab. The Australian evaluators concluded that Obinutuzumab was significantly better than rituximab in terms of progression-free survival (PFS). The Canadian evaluators based clinically meaningful improvements in PFS and overall survival (OS) while under treatment with Obinutuzumab and the net clinical benefit of the drug in combination with bendamustine. Scotland verified obinutuzumab's good clinical performance in terms of PFS while England concluded that Obinutuzumab delayed disease progression in the short-term, but its long-term impact on PFS is uncertain.

Similar to clinical outcomes were drug safety studies, which were given more attention in all four countries. The Australian evaluation found more patients with renal function-related adverse events and toxic effects (AEs) while taking Obinutuzumab. The Canadian evaluation noted a higher frequency of serious AEs in the Obinutuzumab combination group. The Scottish safety evaluation concluded that because Obinutuzumab was used concurrently with chemotherapy during the induction phase, it was difficult to fully determine the overall impact of AEs due solely to Obinutuzumab. England had similar findings to the first three with obinutuzumab-based therapy having a higher incidence of AEs when compared with rituximab-based therapy.

In addition to the aforementioned efficacy and safety, the cost-effectiveness of the drug was also highlighted in the four countries' evaluations. While Canada, Scotland, and England all reported extensively on the cost-effectiveness of Obinutuzumab, Australia did not mention it throughout the report, consistent with the findings from the corresponding analysis above. Canada did not consider the uncertainty of the incremental effect of Obinutuzumab in the economic model but rather placed great importance on the assessment model, the selection of indicators,

and the reasonableness of the data in the cost-effectiveness assessment. A key factor in Scotland's refusal to recommend the drug for use in the country was the lack of stability in the cost- effectiveness analysis model submitted by the applicant company. England, while also acknowledging the immaturity of the duration of treatment effect and clinical data in the model analysis, also felt that significant uncertainty in the evidence base and incremental cost effectiveness ratio existed. However, the Committee concluded that Obinutuzumab was a cost-effective approach to the use of NHS resources in England for patients with untreated follicular lymphoma (FLIPI score >2).

The four countries possess some homogeneity among their assessments, but some particularities in certain countries could be noted. The most notable feature of the Australian assessment body is the two rounds of risk-benefit assessment that it conducts, which aims to achieve optimal treatment outcomes while controlling treatment risks. In addition, Canada has a pharmacovigilance in addition to a risk management program designed to identify identifiable and potential risks and missing information to ensure safe use of medications. One element of the Scottish assessment process is novel in that it sought the views of physicians and patients who use the drug in order to consider the added value of Obinutuzumab as an orphan drug in the context of existing treatment in the NHS in Scotland. England's assessment also considered the innovative nature of the drug with the committee hearing from clinical experts that Obinutuzumab has a similar mechanism to rituximab and that Obinutuzumab is not an innovative drug.

DISCUSSION

Of the 15 drugs studied above, Australia was found to have a recommended status for the nine drugs for which assessments were available with a single assessment that did not qualify for Kappa analysis and therefore could not be tested for consistency between Australia and the other countries. Canada and England had only recommendation and non-recommendation statuses, while Scotland also had a restricted status, which allowed for better application of drugs to achieve better symptomatic use and reduce adverse effects of drugs. The four countries differed in their preferences for assessing clinical evidence and the economic model for orphan drugs. England placed more emphasis on treatments in terms of control type than did the other countries, Canada had a greater proportion of Phase III and open-label trial types and focused more on cost-utility analysis, while Australia was less studied in terms of economic model, probably because of the different review processes in these four countries. In Australia, the PBAC's primary role is to recommend which medicines should be subsidized, and the Australian government determines the final listing decision (31). The NHS in England must comply with positive recommendations by NICE, a process that resembles the Scottish NHS boards. These differences may explain why less attention has been paid to the economic modeling aspect in Australia.

For national clinical trials, England and Scotland focus more on the assessment of clinical trial directions, while Canada and Australia focus on the assessment of clinical endpoints. PBAC has a clear preference for head-to-head RCT (if available) and indirect comparisons (32), while placebo comparisons are listed as acceptable evidence for NICE and SMC (5), the strength of evidence required for clinical trials vary to different countries. In the analysis of clinical uncertainty, it can also be seen that uncertainty in evidence and study design occupies most of the English and Scottish evaluation reports followed by clinical effectiveness, indicating that these two countries are likely to focus on these two aspects of uncertainty when conducting their evaluations. Their processes precisely indicate the focus and difficulty of addressing orphan drugs in health evaluation and needs more attention.

The case study demonstrates that although Scotland affirmed the good clinical performance of Obinutuzumab in PFS, it still refused to recommend the drug for reimbursement in this country due to the lack of stability of the submitted cost-effectiveness analysis model. In contrast, England recommended the drug for reimbursement by limiting its indications and taking advantage of discounts agreed upon the patient access scheme. The decision making in Canada and Australia was focused on the clinical benefit of the drug and neglected to assess the uncertainty of the economic benefit. This finding also matched the results of the analysis in the article by Nicod (5) in which the SMC seems less likely to approve drugs with high and uncertain clinical cost-effectiveness and the CDR seems to place more emphasis on the relative effectiveness of new treatments among others.

The poor consistency of the four national assessments may be due to the fact that the focus of the assessment of orphan drugs in different countries is based on the prevalence of rare diseases in their countries, and that the efficacy requirements for orphan drugs may differ from country to country, with each country focusing on the criteria of the actual development in their country when conducting the assessment. Selection of study subjects is a component of health technology assessment report with variability in the review opinions of the four countries. The differences may also be due to the existence of manual coding errors and the limitations of the information obtained from the 15 orphan drug reports that were selected. In future studies, the sample size should be enlarged, and more representative countries should be selected to extrapolate the findings.

ACTIONABLE RECOMMENDATIONS

Due to the specificity of rare diseases, clinical trials have a small subject population, and the efficacy is often uncertain while the effectiveness and efficacy of drugs are factors of great concern to countries with respect to the HTAs of orphan drugs. The uncertainty of efficacy will lead to an increase in the instability of pharmacoeconomic models, thus making the results of cost effectiveness of drugs unstable and often not cost effective due to the high price of orphan drugs, which eventually leads to orphan drugs not being recommended to the health insurance list of each country or strictly limited for certain indications.

Countries have different levels of economic development and social conditions, such as population, and large uncertainties in their clinical evidence and economic evidence exists, which brings challenges to the HTA review process in each country. Currently, no uniform standards in the review of orphan drugs exist, so countries show differences in their attention to different types of evidence when making decisions. In the process of HTA of orphan drugs, the analysis should not be limited to cost effectiveness but also evaluate various factors, such as the innovativeness of orphan drugs, the increase in the utility of family members, and the improvement of life satisfaction when treating them on the basis of the sustainability of health insurance funds and through the collection of real-world data, in order to enhance the accessibility of drugs for patients. In addition, the uncertainty of the clinical use of orphan drugs through real-world data collection can be reduced and the economics of orphan drugs not only through NHS reimbursement but also through multi-channel economic sharing mechanisms could be improved, so that more patients with rare diseases can be treated and their economic burden can be effectively alleviated through price agreements or joint multi-level protection.

For example, in England, ICERs for common drugs must meet the threshold of £20,000–30,000 per QALY to enter the NHS system, but the majority of ICERs for orphan drugs are higher than the upper threshold of willingness to pay set by England, and NICE has introduced social value judgments to evaluate orphan drugs, namely combining the social benefits of orphan drugs, innovation, the number of people with the rare disease, the specificity of the disease condition, and others to decide whether to include them in the NHS system (33). In addition, NICE conducted a highly specialized technology evaluation (HSTE) for high-value drugs (ultra-rare drugs) with a willingness-to-pay threshold adjusted to cost £100,000 per QALY obtained. It is also clear from the case studies that in response to the uncertainty of the long-term effects of orphan drug efficacy and the instability of economic models, England has saved some patients first by entering into price agreements and then collecting evidence and revising the results in subsequent evaluations.

Currently, the focus of HTA development is very different among different countries, and different countries have different degrees of orphan drug coverage and different goals when making HTA decisions for orphan drugs. A need to fix the HTA shortcomings, build a more mature assessment mechanism, and explore a more complete and appropriate reimbursement pathway in the field of rare diseases exists with a view to providing a basis for health decision-making. Concurrently, exploring the use of real-world data for orphan drug access and security and making good use of HTA to promote more marketing and inclusion of orphan drugs in health insurance combined with the country's own development, can guarantee the safety and accessibility of orphan drugs, and actively safeguard the rights of patients with rare diseases.

AUTHOR CONTRIBUTIONS

NZ carried out the search and selection of the articles and drafted the manuscript. HJ was involved in the design of the study and critically reviewed the manuscript. ZL participated in data analysis and literature review. JH gave guidance on the logic and key problem of the article. J-HX participated in data analysis. Y-HF performed the literature searches. NY was involved in the design of the study, conceived of the study, and critically reviewed the manuscript. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2022.861067/full#supplementary-material

REFERENCES

1. *FAQs about Rare Diseases | Genetic and Rate Diseases Information Center(GARD)- an NCATS Program.* Available online at: https://rarediseases.info.nih.gov/diseases/pages/31/faqs-about-rare-diseases (accessed March 20, 2021).
2. *Rare Diseases | European Commission.* Available online at: https://ec.europa.eu/info/research-and-innovation/research-area/health-research-and-innovation/rare-diseases_en#:$\sim$:text=In%20the%20EU%2C%20a%20rare,million%20people%20in%20the%20EU (accessed March 22, 2021).
3. *What We're Doing About Rare Diseases | Australian Government Department of Health.* Available online at: https://www.health.gov.au/health-topics/chronic-conditions/what-were-doing-about-chronic-conditions/what-were-doing-about-rare-diseases#what-are-rare-diseases (accessed April 25, 2022).
4. Zhang H, Xia HL, Gao GX, Yang T, Xia T. The enlightenment of the development of European HTA for China. *J China Medical Equipment.* (2019) 16:182–6. doi: 10.3969/J.ISSN.1672-8270.2019.06.046
5. Nicod E, Kanavos P. Commonalities and differences in HTA outcomes: a comparative analysis of five countries and implications for coverage decisions. *Health Policy.* (2012) 108:167–77. doi: 10.1016/j.healthpol.2012.09.012
6. Hu SL. Characteristics of drug policy and pharmacoeconomics study on rare diseases. *J Int Pharm Res.* (2019) 46:652–8. doi: 10.13220/j.cnki.jipr.2019.09.002
7. *Top 10 Most Expensive Drugs on the Market.* BioSpace. Available online at: https://www.biospace.com/article/gene-therapy-zolgensma-tops-goodrx-list-of-10-most-expensive-drugs/ (accessed January 2, 2022).
8. Villa F, Filippo AD, Pierantozzi A, Genazzani AA, Addis A, Trifirò G, et al. Orphan drug prices and epidemiology of rare diseases: a cross-sectional study in italy in the years 2014–2019. *Front Med.* (2022) 9:2296-858X. doi: 10.3389/fmed.2022.820757
9. Xu HP, Shen Y, Zhu C, Gong MC, Chen L, Zhao K, et al. Challenges and solution to health technology assessment of rare diseases. *J Int Pharm Res.* (2019) 46:645–51. doi: 10.13220/j.cnki.jipr.2019.09.001
10. Hutton J, Trueman P, Facey K. Harmonization of evidence requirements for health technology assessment in reimbursement decision making. *Int J Technol Assess Health Care.* (2008) 24:511–7. doi: 10.1017/S0266462308080677
11. Zhang N, Shi XF, Wu J. The application of incremental cost-effectiveness ratio to health technology assessment. *Chin J Health Policy.* (2012) 5:64–8. doi: 10.3969/j.issn.1674-2982.2012.02.012
12. Iskrov G, Stefanov R. Post-marketing access to orphan drugs: a critical analysis of health technology assessment and reimbursement

decision-making considerations. *Orphan Drugs Res Rev*. (2014) 4:1–9. doi: 10.2147/ODRR.S43409
13. Li H, Wang XY, Yue XM, Wu JH. Considerations on the health technology assessment for rare diseases and orphan drugs in China. *J Int Pharm Res*. (2019) 46:666–72. doi: 10.13220/j.cnki.jipr.2019.09.004
14. Zhang ZZ, Chen C, Liu X, Li XY. Survey of application status of health technology assessment in China. *J Med Inf*. (2019) 32:22—5. doi: 10.3969/j.issn.1006-1959.2019.23.006
15. Hu SL. The role of health technology assessment in the adjustment of national drugs reimbursement list. *World Clin Drugs*. (2020) 41:915–6. doi: 10.13683/j.wph.2020.12.001
16. *Published Guidance, NICE Advice and Quality Standards | Guidance*. NICE. Available online at: https://www.nice.org.uk/guidance/published?type=ta&fromdate=January%202018&todate=January%202019 (accessed May 22, 2019).
17. *Medicines Advice*. Available online at: https://www.scottishmedicines.org.uk/medicines-advice/ (accessed June 22, 2019).
18. *AusPAR Search | Therapeutic Goods Administration (TGA)*. Available online at: https://www.tga.gov.au/ws-auspar-index?search_api_views_fulltext=Venetoclax&sort_by=field_date&sort_order=DESC&items_per_page=10 (accessed May 28, 2019).
19. *Pharmaceutical Benefits Scheme (PBS)*. Available online at: https://www.pbs.gov.au/info/industry/listing/elements/pbac (accessed May 5, 2019).
20. *Search*. CADTH. Available online at: https://cadth.ca/pcodr (accessed July 12, 2019).
21. Zhang HJ, Liang Y, Yuan N. Introduction of the health technology assessment process on orphan drugs by NICE and its implications. *Chin J New Drugs*. (2018) 27:1587–94.
22. Liu F, Huang SJ, Shi YY, Zhu WT. Benchmarking study of abroad pharmacoeconomic evaluation in the policy decisions. *Chin J Drug Eval*. (2014) 31:184–8. doi: 10.3969/j.issn.2095-3593.2014.03.015
23. *OECD*. Available online at: https://www.oecd.org/health/value-in-pharmaceutical-Pricing-Australia (accessed June 6, 2021).
24. Guan YR, Xiang W, Zhang F. The Australian life saving drugs and its implications for medical assistance system for the rare diseases in China. *Chin J Health Policy*. (2015) 8:51–5. doi: 10.3969/j.issn.1674-2982.2015.08.010
25. Hu JJ, Chen X, Gong SW. Selection principle of orphan drug under UK, Canada and Australia National Health Service Security. *Med Soc*. (2016) 29:10–12+16. doi: 10.13723/j.yxysh.2016.01.004
26. McCormick JI, Berescu LD, Tadros N. Common drug review recommendations for orphan drugs in Canada: basis of recommendations and comparison with similar reviews in Quebec, Australia, Scotland and New Zealand. *J Orphanet J Rare Dis*. (2018) 13:1–12. doi: 10.1186/s13023-018-0759-9
27. Chen T, Jiang R, Shao R. Research on the Canadian Pharmaceutical Rebate Agreements and its enlightenment to China. *Chin J Health Policy*. (2021) 14:68–73. doi: 10.3969/j.issn.1674-2982.2021.04.010
28. Xu WY, Xue D. Health technology assessment in United States, Canada and Australia. *Chin Health Q Manag*. (2011) 18:8–10. doi: 10.13912/j.cnki.chqm.2011.01.007
29. Wong L. Data analysis in qualitative research: a brief guide to using nvivo. *Malaysian Fam Phys*. (2008) 3:14–20.
30. Freeman CL, Laurie HS, A. tale of two antibodies: obinutuzumab versus rituximab. *Br J Haematol*. (2018) 182:29–45. doi: 10.1111/bjh.15232
31. Allen N, Walker SR, Liberti L, Salek S. Health Technology Assessment (HTA) Case Studies: Factors Influencing Divergent HTA Reimbursement Recommendations in Australia, Canada, England, and Scotland. *Value Health*. (2017) 20:320–8. doi: 10.1016/j.jval.2016.10.014
32. *Pharmaceutical Benefits Advisory Committee (PBAC)*. Available online at: http://www.health.gov.au/internet/main/publishing.nsf/content/pbacguidelines-index (accessed December 11, 2020).
33. Ceri JP, Dyfrig AH. HTAs and rare diseases: how to assess their cost-effectiveness. *Pharm Policy and Law*. (2011) 13:161—5. doi: 10.3233/PPL-2011-0321

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Zhou, Ji, Li, Hu, Xie, Feng and Yuan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Public Health

SYSTEMATIC REVIEW
published: 20 June 2022
doi: 10.3389/fpubh.2022.832881

Chronic Noise Exposure and Risk of Dementia: A Systematic Review and Dose-Response Meta-Analysis

Linghao Meng[1†], Yang Zhang[2,3†], Shushan Zhang[4], Fugui Jiang[5], Leihao Sha[6], Yajia Lan[7] and Lei Huang[7,8]*

[1] Department of Urology, West China School of Medicine/West China Hospital, Sichuan University, Chengdu, China, [2] Cochrane China Center, Chinese Evidence-Based Medicine Center, West China Hospital, Sichuan University, Chengdu, China, [3] Department of Periodical Press and National Clinical Research Center for Geriatrics, West China Hospital, Sichuan University, Chengdu, China, [4] Department of Neurology, Affiliated Hospital of North Sichuan Medical College, Nanchong, China, [5] Sichuan Provincial Center for Mental Health, Sichuan Academy of Medical Sciences & Sichuan Provincial People's Hospital, Chengdu, China, [6] Department of Neurology, West China Hospital, Sichuan University, Chengdu, China, [7] Department of Environmental Health and Occupational Medicine, West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, China, [8] Department of Occupational Hazard Assessment, West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, China

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Jie Kuang,
Nanchang University, China
Agnieszka Micek,
Jagiellonian University, Poland

***Correspondence:**
Lei Huang
hlst0610@163.com

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Aging and Public Health, a section of the journal Frontiers in Public Health

Received: 10 December 2021
Accepted: 23 May 2022
Published: 20 June 2022

Citation:
Meng L, Zhang Y, Zhang S, Jiang F, Sha L, Lan Y and Huang L (2022) Chronic Noise Exposure and Risk of Dementia: A Systematic Review and Dose-Response Meta-Analysis. Front. Public Health 10:832881. doi: 10.3389/fpubh.2022.832881

Objective: Evidence is scarce about the effect of noise exposure on the risk of dementia. We conducted a systematic review and dose-response meta-analysis, aiming to explore the association between noise exposure and the risk of dementia.

Methods: We searched PubMed, EMBASE and the Cochrane Library to collect studies on chronic noise exposure and the risk of dementia from database inception to September 18, 2021 without language limitations. Two authors independently screened the literature, extracted data and assessed the risk of bias of the included studies. A dose-response meta-analysis and subgroup analysis were then conducted to detect the association between noise exposure and the risk of dementia by using Stata 14.0 software. This study is registered on PROSPERO (CRD42021249243).

Results: A total of 11 studies were eligible for qualitative synthesis, and nine were eligible for quantitative data synthesis. All of them showed moderate to high quality scores in the assessment of risk of bias. We found a positive linear association between the noise increment and dementia risk (R^2 = 0.58). When noise exposure increased 57 dB, the RR of dementia was 1.47 (95% CI: 1.21–1.78). From the outcome subgroup of AD, AD and dementia, VaD and NAD, we also found a positive association (R^2 = 0.68, 0.68, 0.58, respectively). When noise exposure increased by 25 dB, the RRs were 1.18 (95% CI: 1.14–1.23), 1.19 (95% CI: 1.14–1.23) and 1.17 (95% CI: 1.06–1.30), respectively. We found a nonlinear association between the noise increment and dementia risk when only cohort studies were included (R^2 = 0.58). When noise exposure increased by 25 dB, the RR of dementia was 1.16 (95% CI: 1.12–1.20). From the subgroup of AD, AD and dementia, VaD and NAD of cohort studies, the regression curve showed a nonlinear positive association (R^2 = 0.74, 0.71, 0.43, respectively). When noise exposure increased by

25 dB, the RRs were 1.17 (95% CI: 1.12–1.21), 1.17 (95% CI: 1.12–1.22) and 1.13 (95% CI: 0.99–1.28), respectively.

Conclusion: Based on the current evidence, exposure to noise may be a specific risk factor for dementia. To better prevent dementia, more rigorously designed studies are needed to explore the etiological mechanism of noise and dementia.

Keywords: noise exposure, dementia, mild cognitive impairment, Alzheimer's disease, dose-response meta-analysis

INTRODUCTION

Dementia is a kind of disease characterized by severe cognitive impairment, including several subtypes, such as Alzheimer's disease (AD), cerebrovascular disease, Lewy body dementia, and Huntington's disease (1). Alzheimer's disease contributes the majority of patients with dementia, accounting for approximately 60∼80% of cases (2). Studies have shown that there were approximately 46 million people diagnosed with AD globally in 2015, and the number may reach 115.4 million by 2050 (3). The prevalence of dementia is estimated to be 7% in people above the age of 65, and the rate rises exponentially with age (1, 4). Tau and β-amyloid (Aβ) deposition are considered to be the possible pathological mechanisms of dementia, AD and cognitive impairment (5, 6). However, the cause of dementia has not been expounded completely. Hypertension, diabetes, high body mass index (BMI), smoking, and air pollution are thought to be risk factors for dementia thus far (1, 7). Mild cognitive impairment (MCI) is a state between normal cognition and dementia that is considered a premonitory symptom of dementia with preserved basic cognitive abilities (8). MCI may convert into dementia when specific diagnostic criteria are used or exposed to more risk factors (8, 9). Dementia or MCI can be diagnosed mainly by history-taking, neurological examination and imaging examination (4, 10). However, there was no specific therapy for any subtype of dementia or MCI. Drugs such as cholinesterase inhibitors, N-methyl-D-aspartic acid (NMDA) receptor antagonists, selective serotonin reuptake inhibitor (SSRI) antidepressants and rehabilitation training are only used for symptomatic and supportive treatments (11–13).

With the development of modernization and industrialization, noise is ubiquitous in life. The World Health Organization (WHO) estimated that 1∼1.6 million disability-adjusted life years (DALYs) are lost each year due to noise exposure in Western European countries (14). Increasing evidence has shown that noise is associated with many diseases, such as ischemic heart disease (IHD), stroke, metabolic disorders and cognitive dysfunction (15, 16). Recent studies believe that noise may increase the risk of dementia. Animal experiments have shown that environmental noise exposure can influence cognitive performance, along with changes in Tau and β-amyloid (Aβ) at the same time (17, 18). Although it is not clear completely how noise results in dementia pathologically, some papers have reported varied relationships between them. Linares et al. (19) found that short-term exposure to noise may be associated with hospital admission for dementia. Chen et al. (20) found that roads closer to heavy traffic can increase the incidence of dementia, which may be the effect of noise and air pollution. However, Andersson et al. (21) conducted a cohort study in 2018 and reported no association with traffic noise and AD or vascular dementia (VaD). There was not enough evidence to clarify the relationship between them. Several recent reviews also suggested an ambiguous link between noise and dementia (22, 23). Studies have proposed several hypotheses about the pathogenesis mechanism of noise and dementia. Experimental evidence has shown that noise may lead to dementia by neurovascular, neuroendocrine or oxidative stress factors (7, 15, 24, 25). In addition, studies have reported an association between hearing loss and dementia but have not explained the link between hearing loss and noise (26, 27).

A systematic review conducted in 2020 summarized seven studies related to traffic noise with dementia, including Parkinson's disease, dementia, Alzheimer's disease, and hospitalizations for dementia-related illnesses (28). However, this review only focused on traffic noise and did not systematically report traffic noise and the risk of dementia. Furthermore, two reviews conducted in 2018 by the WHO focused on the effect of environmental noise on wellbeing, quality of life and cognitive impairment in multiple dimensions (14, 29). Hegewald et al. (30) carried out a meta-analysis of traffic noise and mental health (including depression, anxiety, dementia and AD) but did not perform a quantitative synthesis of dementia and AD due to different outcomes. Our research team previously summarized the experimental and epidemiological studies of the relationship between chronic noise exposure, cognitive impairment and degenerative dementia but did not conduct a systematic review (31). Some new evidence of dementia emerged after the retrieval time of the studies above so that more evidence can be included and synthesized. Thus, given that there was no quantitative meta-analysis before, in this study, we reviewed existing studies to conduct a systematic review and dose-response meta-analysis, aiming to summarize the evidence regarding the risk of dementia when exposed to chronic noise.

MATERIALS AND METHODS

We conducted this systematic review following the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines (32). This systematic review and meta-analysis was registered with PROSPERO (CRD42021249243).

Search Strategy

Two authors (LM and YZ) independently searched the PubMed, EMBASE and Cochrane Library databases to collect observational studies on chronic noise exposure and the risk of dementia. We searched all studies without language restriction from inception to 18 September 2021. We used MeSH and free text terms to identify the relevant literature. The detailed search strategy of each database is attached in the **Supplementary Material**. An additional search was performed among the references of the included studies to identify potentially eligible studies. We also manually searched the references of published reviews to collect additional relevant studies.

Inclusion and Exclusion Criteria

We used a population-intervention-comparator-outcomes-study design (PICOS) to identify the inclusion criteria. (1) Participants: People who were not diagnosed with any cognitive dysfunction (including dementia, Alzheimer's disease or MCI) without age limitations were included in the cohort study; People in case group were diagnosed with any cognitive dysfunction (including dementia, Alzheimer's disease or MCI), and the control group were not diagnosed with any cognitive dysfunction in case-control study. (2) Exposure/Control: Exposure group exposed to chronic noise, including traffic noise, community noise and occupational noise. The control group was not exposed to chronic noise. (3) Outcome: Association between chronic noise exposure and dementia. (4) Study design: Cohort studies, case-control studies or cross-sectional studies were included in this study. Because of the restrictions of research topics and ethics, no randomized controlled trials (RCTs) can be included.

The exclusion criteria for this study were as follows: (1) cohort studies included people diagnosed with any cognitive dysfunction before baseline; case-control studies included people with major psychiatric disorders or significant disease that could interfere with cognition; (2) studies that explored the association between hearing loss and dementia, cognitive dysfunction or AD but did not search for the relationship with hearing loss and noise exposure; (3) studies that could not extract data; (4) *in vitro* or animal experiments; (5) duplicate data; (6) reviews, letters, case

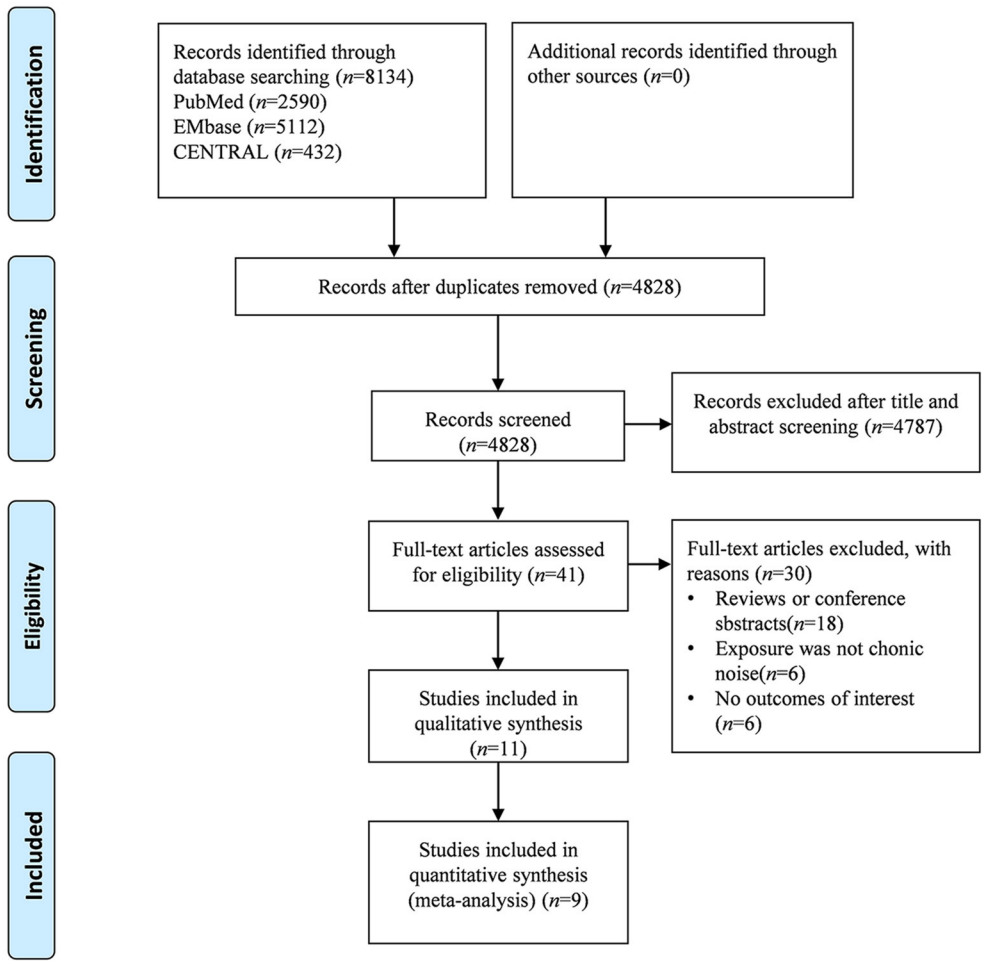

FIGURE 1 | Flowchart of literature screening.

reports, protocols, conference abstracts and any article without full text; and (7) studies were not published in English.

Data Extraction

Two authors (LM and YZ) independently selected the studies by title, abstract and full text. Disagreements were resolved by discussion or consultation with a third reviewer (LH). We extracted the data from each retained study by using a unified information extraction table. The extracted terms included first author, study country/region, publication year, study design, inclusion number, analysis number, sex, age, follow-up periods, noise (sources, level and assessment method), dementia (type and assessment method), number of patients, adjusted covariates and risk estimate indicators, including the relative risk (RR), hazard ratio (HR), odds ratio (OR) and regression coefficient (β). If the risk indicators in each study were adjusted by covariates, we extracted all the adjusted risk indicators. The main adjusted factors included (1) baseline information: age, sex, body mass index; (2) socioeconomic status: education, household income occupational environment; (3) disease and lifestyle factors: stroke, cardiovascular disease, ApoE4 level, smoking, alcohol consumption, physical activity; and (4) other exposure: NO_x, $PM_{2.5}$, PM_{10}. For the studies that had missing main data, we asked the authors for the full data by email. One author extracted the data, and another author checked the accuracy. The questions were solved through the examination of the original articles and discussion.

Risk of Bias Assessment

Two authors (LM and YZ) independently assessed the risk of bias using the Newcastle-Ottawa scale (NOS) for cohort and case-control studies and Agency for Healthcare Research and Quality (AHRQ) recommended items for cross-sectional studies. Disagreements were resolved by discussion or consultation with a third reviewer (LH). There is no recognized tool to assess the risk of bias in observational studies. The NOS is the most commonly used scale for cohort studies or case-control studies and is a semiquantitative system (33, 34). The assessment of bias of cohort studies and case-control studies included eight domains. And the details and scales of evaluation can be available in reference (30, 35). The Agency for Healthcare Research and Quality (AHRQ) recommended using 11 terms to assess the risk of bias of cross-sectional studies (https://www.ahrq.gov/). The details of the assessment method are available in reference (36).

Statistical Analysis

We used Stata 14.0 to perform the meta-analysis. Aiming to clearly research the relationship between noise exposure and dementia, we conducted a dose-response analysis between the dose of noise exposure and the risk of dementia (37–39). L_{den} is an annual long-term average A-weighted equivalent continuous noise level integrating compound metrics (40, 41). Given that the retained studies used different noise assessment methods to evaluate the exposure quantitatively, we converted the different noise metrics into the most commonly used method (L_{den}) (15, 16). The detailed conversion methods were available in the study raised by Brink et al. (40). The median dose of the category groups was considered the corresponding dose. When the boundary of the category was open, the median dose of this category was set at the lower boundary multiplied by 1.5 in the highest category, and the dose was set as zero in the lowest category (38). A one stage robust error meta-regression method (REMR) was utilized to analysis the dose-response associations. In this study, restricted cubic spline (RCS) with three knots was applied to model potentially nonlinear associations, and centering was done to eliminate the effects of inconsistency in reference group by subtracting the reference dose from the non-reference dose for each reported non-reference effect (42). R^2 was used to measure the overall fit of the regression equation, and slope equality test was applied to test for nonlinear trends (42–45). Subgroup analysis was performed by different categories of dementia. Publication bias and sensitivity analysis were not performed on analyses with <10 studies due to the low sensitivity of qualitative and quantitative tests.

RESULTS

Search Results and Study Characteristics

The study search and selection process is shown in **Figure 1**. From the three databases, we identified 4,828 publications after removing duplications. After title and abstract retrieval, 41 studies were included in the full text screening. Of the potentially eligible publications after full text screening, we excluded 18 studies because they were just reviews or systematic reviews; 6 studies because the exposure was not noise; and six studies had no outcomes of interest. Finally, 11 studies (21, 41, 46–53) were eligible for qualitative synthesis, and nine studies (21, 41, 46, 47, 49, 51–53) were eligible for quantitative data synthesis.

The baseline information of the included studies is shown in **Table 1**. The risk estimate between noise exposure and dementia outcomes from each reviewed study is attached in the **Supplementary Material**. Among the 11 eligible studies, five studies were from North America, including two studies from Canada and three from the USA. The other five studies were from Western/Northern Europe (Sweden, Germany, The UK, Danmark and Spain). Eight studies were cohort studies (21, 46, 47, 49, 50, 52, 53), two were cross-sectional studies (48, 51) and one (41) included cohort studies and case-control studies. The noise sources in the 11 studies included traffic noise and community noise. Most of the studies used L_{den}, L_{nihgt} or L_{eq} indicators to describe noise magnitude. The studies used different outcomes to describe cognitive impairment, including dementia, vascular dementia (VaD), Alzheimer's disease, mild cognitive impairment (MCI) and non-Alzheimer's dementia (NAD). In the included studies, two studies were based on the same first author and study population (53), in which the data were extracted. In addition, some studies reported multiple different indicators in one study, and we also extracted all the data severely.

Risk of Bias Assessment

With regard to different types of studies, we used different assessment methods to evaluate the risk of bias. The quality scores are presented in **Table 1**, and evaluation details are provided in the **Supplementary Material**. All the included

TABLE 1 | Characteristics of the included studies.

First author, year	Country/ Region	Study design	Total sample/ Analysis sample	Sex (male/ female)	Age (range/ mean)	Follow-up time (years)	Noise type	Noise assessment	Noise value (dB)	Types of dementia	Outcome assessment	Quality assessment (NOS/AHRQ)*
Andersson et al. (21)	Umeå, Sweden	cohort study	1721/1721	985/736	55–85/68.5 ± 9.4	7	traffic noise	Umeå Municipality Noise Survey/Leq.24 h	<55; ≥ 55	dementia	NINCDS–ADRDA[a]	7
Tyas et al. (50)	Manitoba, Canada	cohort study	1355/694	261/433	65–93/74.0 ± 5.8	5	community noise	scale	NA	AD	NINCDS–ADRDA	7
Weuve et al. (52)	Chicago, USA	cohort study	7909/5227	1986/3241	NA/73.7 ± 6.9	4.1	community noise	universal kriging model developed for the Chicago area/Leq 10:00–16:00	range 51.1–78.2	AD; MCI	NINCDS–ADRDA	8
Yuchi et al. (41)	Metro Vancouver, Canada	case–control study	13498/13498	NA	45–84/NA	5	community noise	Lden	case range 33.0–86.5; control range 4.4–92.4	AD	hospital records	7
		cohort study	633949/ 633949	300197/ 333752	45–83/NA			Lden	case range 44.5–77.2; control range 32.2–85.8	NAD	hospital records	9
Fuks et al. (49)	North Rhine–Westphalia, Germany	cohort study	834/288	0/288	67–80/74.5± 2.2	25	community noise	Lden	mean 55.9 ± 7.7	MCI	CERAD Neuropsychological Assessment Battery[b]	7
								Lnight	mean 47.2 ± 7.4	MCI	CERAD Neuropsychological Assessment Battery	
							traffic noise	ICBEN Daytime	NA	MCI	CERAD Neuropsychological Assessment Battery	
								ICBEN Night	NA	MCI	CERAD Neuropsychological Assessment Battery	
Yu et al. (53)	California Sacramento Valley, USA	cohort study	1789/1612	680/932	>60/70.2 ±6.8	6.5	traffic noise	Leq; 24-h Noise/Nighttime Noise	mean Lden=68.5 ± 8.9 Lnight=60.4 ± 8.9	MCI	3 MSE/SEVLT[c]	7
Yu et al. (53)	California Sacramento Valley, USA	cohort study	1789/1612	680/932	>60/70.2 ±6.8	6.5	traffic noise	Leq; 24-h Noise	mean Lden=68.5 ± 8.9	MCI	3 MSE/SEVLT	7
Tzivian et al. (51)	Ruhr, Germany	cross–sectional study	4157/2050	1007/1043	NA/64.1 ± 7.7	NA	traffic noise	Lden; Lnight	mean Lden=53.74 ± 9.49 Lnight=44.88 ± 9.17	MCI	International Working Group on MCI criteria	10

(Continued)

TABLE 1 | Continued

First author, year	Country/ Region	Study design	Total sample/ Analysis sample	Sex (male/ female)	Age (range/ mean)	Follow-up time (years)	Noise type	Noise assessment	Noise value (dB)	Types of dementia	Outcome assessment	Quality assessment (NOS/AHRQ)*
Carey et al. (47)	London, England	cohort study	139718/ 130978	65130/ 65848	50–79/NA	6.9	traffic noise	TRAffic Noise EXposure (TRANEX) model/ Lnight	mean 52.1 ± 4.6	AD; VaD	NA	8
Crous-Bou et al. (48)	Barcelona, Spain	cross–sectional study	228/288	136/92	NA/57.7 ±7.6	NA	traffic noise	Lden; Lnight	mean 67.7 ± 5.77	AD	MRI[d]	6
Cantuaria et al. (46)	Danmark	cohort study	1938994/ 1938994	907991/ 1031003	≥60/NA	8.5	road traffic noise	Lden max	mean 55.3	AD; VaD	clinical diagnosis	8
							railway traffic noise	Lden max	mean 44.3	AD; VaD	clinical diagnosis	
							road traffic noise	Lden min	mean 51.6	AD; VaD	clinical diagnosis	
							railway traffic noise	Lden min	mean 44.7	AD; VaD	clinical diagnosis	

*, *Newcastle–Ottawa Scale (NOS) for cohort and case–control studies and Agency for Healthcare Research and Quality (AHRQ) recommended items for cross–sectional studies.* [a], *National Institute of Neurological and Communicative Diseases and Stroke/Alzheimer's Disease and Related Disorders Association (NINCDS–ADRDA) criteria.* [b], *Consortium to Establish a Registry on Alzheimer's Disease (CERAD) Neuropsychological Assessment Battery.* [c], *Modified Mini–Mental State Examination (3 MSE) and Spanish English Verbal Learning Test (SEVLT).* [d], *Magnetic Resonance Imaging.*

cohort studies and case-control studies showed high quality scores. One cross-sectional study was rated as high quality (51), and the other was rated as moderate quality (48).

Dose-Response Meta-Analysis

A total of nine studies (21, 41, 46, 47, 49, 51–53) were included in the dose-response meta-analysis, which contained a population of 2,728,317. Heterogeneity analysis results showed that there was heterogeneity among the studies ($I^2 = 78.7\%$, $P < 0.001$). The noise range of the studies we analyzed was from 22.5 dB to 100.3 dB. From the dose-response analysis of all studies (**Figure 2A**), we found a positive linear association between the noise increment and dementia risk ($R^2 = 0.58$, slope test: F= 1.53, P= 0.223). When noise exposure increased 57 dB, the RR of dementia was 1.47 (95% CI: 1.21–1.78). Seven cohort studies including 2,078,820 participants reported chronic noise exposure and the risk of dementia (21, 46, 47, 49, 52, 53). Heterogeneity analysis results showed that there was heterogeneity among the studies ($I^2 = 50.0\%$, $P < 0.001$). When only these cohort studies were included (**Figure 2B**), we also found a non-linear association between noise increment and the risk of dementia ($R^2 = 0.58$, slope test: $F = 39.38$, $P < 0.001$). When noise exposure increased by 25 dB, the RR of dementia was 1.16 (95% CI: 1.12–1.20).

Subgroup Analysis

Five studies (21, 41, 46, 47, 52) reported noise exposure and the risk of AD ($I^2 = 78.1\%$, $P < 0.001$). The outcome showed a nonlinear positive correlation ($R^2 = 0.68$, slope test: $F = 5.11$, $P = 0.047$). When the noise increment was within 10 dB, as the noise increased, the relative risk of AD rose steeply (**Figure 3A**). In contrast, the curve rises slowly in the range of more than 10 dB. When noise exposure increased by 25 dB, the RR was 1.18 (95% CI: 1.14–1.23).

From the outcome subgroup analysis of AD and dementia, we discovered a nonlinear positive association ($R^2 = 0.67$, slope test: $F = 4.90$, $P = 0.049$). Five studies (21, 41, 46, 47, 52) were included in this subgroup ($I^2 = 75.0\%$, $P < 0.001$). The characteristics of the dose-response analysis were similar to those of the AD subgroup (**Figure 3B**). When noise exposure increased by 25 dB, the RR was 1.19 (95% CI: 1.14–1.23). Regarding the subgroup analysis of VaD and NAD, four studies (21, 41, 46, 47) were included ($I^{2=}$68.5%, P<0.001). The dose-response analysis presented a nonlinear correlation ($R^2 = 0.58$, slope test: $F = 0.21$, $P = 0.660$), with a 25 dB noise increment RR of 1.17 (95% CI: 1.06-1.30) (**Figure 3C**).

We also conducted a subgroup analysis of the cohort studies. Four cohort studies (21, 46, 47, 52) were included in the AD subgroup ($I^{2=}$67.1%, $P < 0.01$). The dose-response analysis showed a nonlinear association ($R^{2=}$ 0.74, slope test: F=19.72, P=0.004) (**Figure 3D**). When noise increased by 25 dB, the RR was 1.17 (95% CI: 1.12–1.21). In the four cohort studies (21, 46, 47, 52) of the subgroup of AD and dementia ($I^2 = 63.0\%$, $P <$ 0.01), the dose–response analysis showed a nonlinear association ($R^2 = 0.71$, slope test: $F = 13.69$, $P = 0.008$) (**Figure 3E**). When noise increased by 25 dB, the RR was 1.17 (95% CI: 1.12–1.22). Three cohort studies (21, 46, 47) were included in

the subgroup of VaD and NAD ($I^2 = 8.9\%$, $P = 0.35$). There was also a nonlinear association between noise exposure and outcomes ($R^{2=}$ 0.43, slope test: $F = 39.48$, $P = 0.008$) (**Figure 3F**). When noise exposure increased by 25 dB, the RR was 1.13 (95% CI: 0.99–1.28).

The above three subgroups of cohort studies presented similar dose response analysis characteristics. When the noise exposure increment was under 15 dB, the RR increased steeply as the noise increased. The curve began to level off in the range of more than 15 dB.

Sensitivity Analysis and Publication Bias

Considering that there were fewer than 10 eligible studies, sensitivity analysis and Begg's and Egger's tests for publication bias were not used for the meta-analysis due to the low efficiency of qualitative and quantitative tests.

DISCUSSION

From the meta-analysis of the nine observational studies, we found a significant positive association between the noise exposure increment and the risk of dementia, which was identical to previous experimental studies. For the subgroup analysis, the groups were divided based on the different subtypes of dementia to explore the source of heterogeneity. The AD group, the AD and dementia group, and the VaD and NAD group showed similar positive associations between noise exposure and the risk of outcomes. The subgroup analysis of cohort studies presented similar dose-response analysis characteristics. When the noise exposure increment was under 15 dB, the RR increased steeply as the noise increased.

In this study, we used REMR method to conduct the dose-response meta-analysis. The REMR method is a special case of the one-stage generalized least-squares trend (GLST) approach, which may have better error estimation and a better visual fit to the data than GLST model (42). The execution of the model need not extract the covariation from the data. It has been reported that the model can be an alternative method for the synthesis of dose-response data from different studies when there is lack of independence in estimates from the same study (42, 45).

From our study, we found associations between noise exposure and the risk of dementia. Few large cohort studies have focused on this field, and no high-quality evidence (meta-analysis) has been conducted in the past. To the extent of our knowledge, two systematic reviews have focused on noise exposure and cognitive impairment (14, 28). The two studies were supported by the WHO and delivered as part of the evidence that made up the Environmental Noise Guidelines for the European Region. One of them evaluated the quality of evidence on the effect of environmental noise on cognition, which was based on the population of children (14). Thirty-four eligible studies were included, and most of them were cross-sectional studies. The study found that aircraft noise was associated with poorer long-term memory, reading and oral comprehension in children, with moderate quality of evidence. The study indicated that the lack of evidence from longitudinal and intervention studies was the key limitation of the available study (14). Another study reviewed the associations of environmental noise with mental health, including cognition, dementia and other neurodegenerative outcomes (28). Similar to the study above, the evidence was strong enough to estimate the effects of environmental noise on children's reading comprehension, but the effects on dementia and neurodegenerative results still lacked evidence.

Apart from noise, other environmental factors were concerned with influencing cognitive function. The Lancet Commission published a report about dementia prevention, intervention and care in 2020, in which air pollution was considered a potential risk factor for dementia (54). A systematic review in 2019 similarly pointed toward an association between exposure to pollutants and increased risk of dementia, particularly for $PM_{2.5}$ and NO_x (55). Similar to noise, Aβ deposition is considered a potential mechanism of the effect of air pollution on dementia. Some publications have explored the association between air pollution and cognitive function (21, 41, 47, 48, 51). Although most of the studies adjusted the outcome risk factors for air pollution or conducted sensitivity analyses, the results of air pollution and noise may still have cross influence. Furthermore, Tzivian et al. (51) mentioned that the results of noise adjusted by air pollution were more robust than air pollution adjusted by noise. They also concluded that air pollution and traffic noise may synergistically influence cognitive function in adults in another analysis (56). Few studies have focused on the association effect between air pollution and noise. Whether the effect was enhanced or weakened still needs further research wshen air pollution and noise act on cognitive function at the same time.

Nevertheless, some studies have concluded that there was no significant association between noise exposure and dementia risk. Andersson et al. (21) suggested that road traffic noise would not cause dementia and that the effect may be mainly influenced by air pollution exposure. Similar results were found in another study (41). However, both studies considered the effect of air pollution as combined exposure on the outcome. Given that there were great differences in environmental factors and noise values, future studies should be conducted that consider the potential joint effects of multiple environmental exposures.

Although population-based ecological studies or observational studies are limited, some scholars still explore the mechanism of the effect of noise based on animal or cytological experiments. Noise exposure can influence brain structure directly and then damage cognitive function. A study conducted in 1998 reported that exposure to noise stress could significantly damage the cognitive function of the prefrontal cortex by conducting primate experiments (57). In addition, the experiment also suggested that the molecular mechanism of impairment may be associated with a hyperdopaminergic mechanism (57). Similarly, some experimental studies have explored the influence of noise stress and brain structures, indicating that chronic traffic noise can reduce the brain volume in the prefrontal cortex and lead to thickness in the hippocampus and amygdala area, which may activate stress reactions to increase the release of dopamine (58–61). Furthermore, studies have reported that long-term railway noise exposure may damage

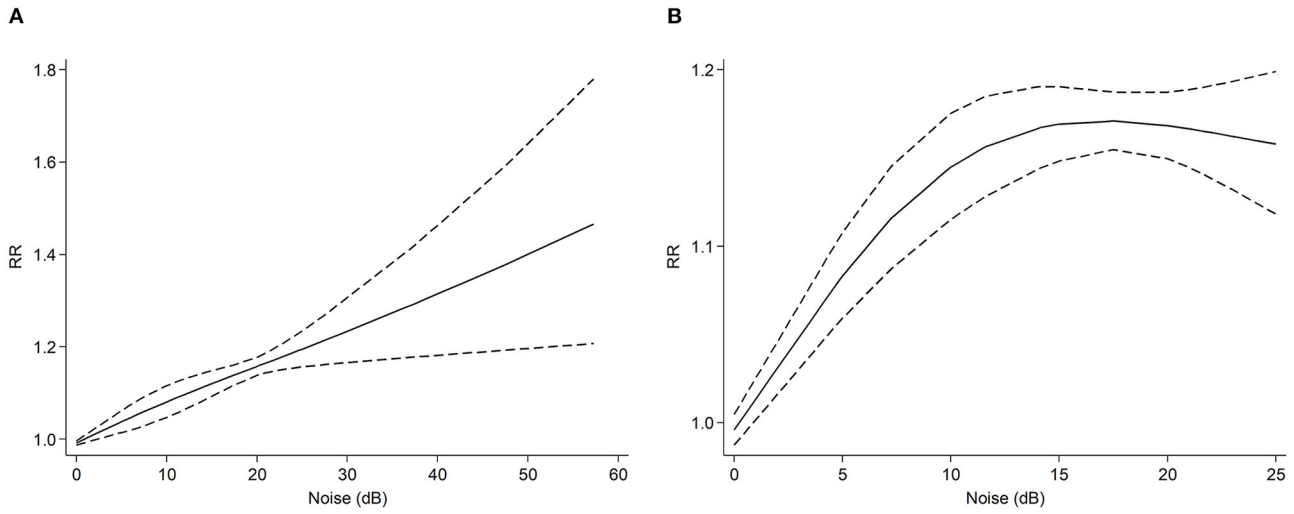

FIGURE 2 | Dose-response meta-analysis between noise exposure and risk of all kinds of dementia. The horizontal axis represents the noise increment, and the vertical axis represents the relative risk of noise increase to dementia. The two dashed lines represent the upper and lower limits of the 95% confidence interval for relative risk. **(A)** dose-response analysis of all studies. **(B)** dose-response analysis of cohort studies.

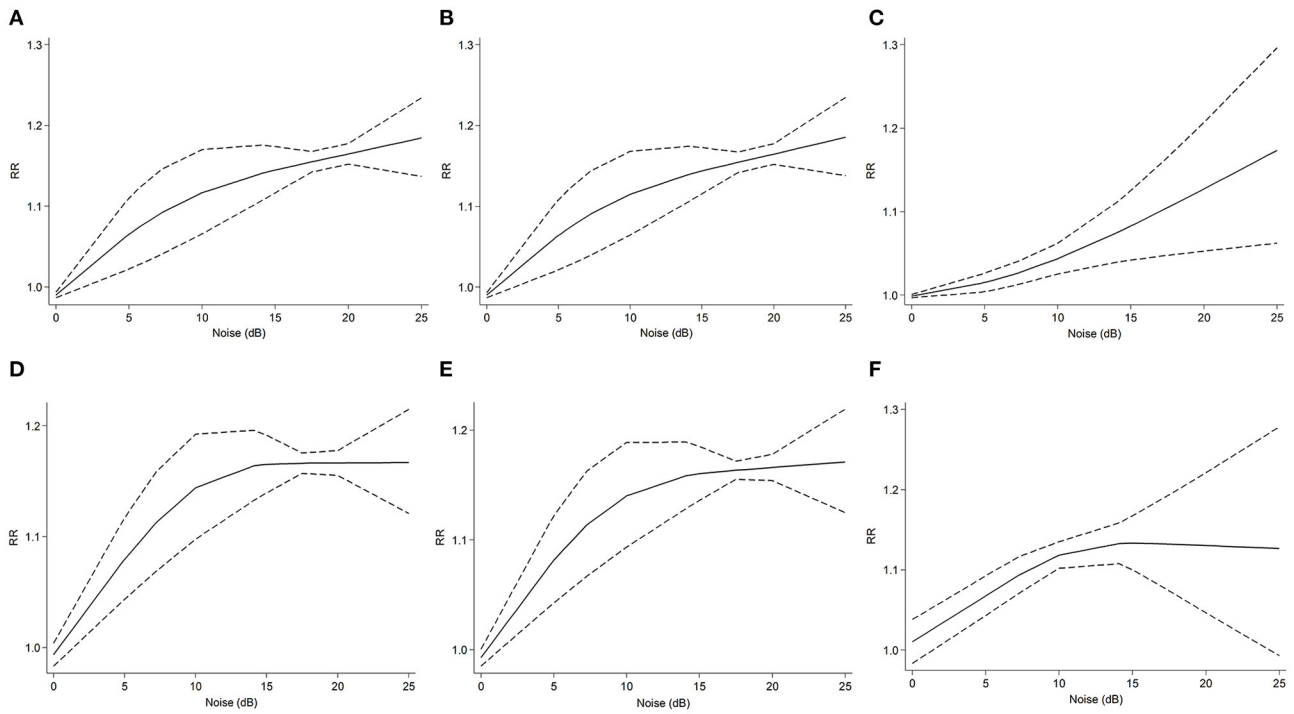

FIGURE 3 | Dose-response meta-analysis between noise exposure and risk of different types of dementia. The horizontal axis represents the noise increment, and the vertical axis represents the relative risk of noise increase to dementia. The two dashed lines represent the upper and lower limits of the 95% confidence interval for relative risk. **(A)** dose-response analysis of noise increment and AD risk in all studies. **(B)** dose-response analysis of noise increment and risk of AD and dementia in all studies. **(C)** dose-response analysis of noise increment and risk of VaD and NAD in all studies. **(D)** dose-response analysis of noise increment and AD risk in cohort studies. **(E)** dose-response analysis of noise increment and risk of AD and dementia in cohort. **(F)** dose-response analysis of noise increment and risk of VaD and NAD in cohort studies.

learning and memory function in the temporal lobe (62, 63). We noticed that several studies mentioned the dopamine release mechanism, indicating that it may be a pathway of the cause of dementia.

Moreover, noise exposure may damage cognitive function via the neuroendocrine pathway. More studies have indicated that noise can activate the hypothalamic–pituitary–adrenal axis (HPA), thus increasing the secretion of adrenocorticotropic

hormone, corticosterone and the catecholamine hormone system to affect learning and memory function (24, 64–66). Disorders of the neuroendocrine pathway can result in anxiety-like behavior, learning and memory impairment, and balance dysfunction (59). Aβ peptide deposition and tau protein hyperphosphorylation are considered possible pathological mechanisms of AD (5). Some evidence has also reported that chronic noise exposure may aggravate the deposition of Aβ and tau protein hyperphosphorylation, which might be regulated by the HPA axis (65, 67, 68). Furthermore, glucocorticoids and corticosterone regulated by the HPA axis may affect the oxidative stress response, in turn causing changes in Aβ and tau protein (59, 69).

A number of studies reported similar findings based on their experiments, which lays the foundation for population-based studies in the future. In general, based on previous studies, the mechanism of noise exposure and cognitive impairment is considered to be associated with changes in brain structure, stress reactions, the HPA axis and neuroendocrine factors. As the key link of possible pathways, the HPA axis has been related to other pathways, suggesting that there may be a close connection between several mechanisms.

This study has some limitations. First, although we searched the main common databases more comprehensively, the included studies still lacked quantity, so the number of studies we contained was limited. Second, from the baseline characteristics of the included studies, we found that the noise assessment methods, noise exposure types or outcome assessment methods were different from each other, which may increase the sources of heterogeneity. Limited by the number of included studies, we could not perform further analysis. Third, we included case-control studies and cross-sectional studies, which is not a relatively robust study design to explain the relationship. Compared with cohort studies, case-control studies and cross-sectional studies may likely have recall bias and selection bias, which influence the accuracy of the results. Also we have not investigated the potential sources of heterogeneity formally. However, this is the first quantitative analysis of the relationship between noise exposure and the risk of dementia, providing quantitative evidence for the etiology of dementia. Dose-response meta-analysis further elucidated the relationship of the increment of noise exposure and dementia, which provides accurate data for the prevention and treatment of dementia in the future.

CONCLUSION

In conclusion, we found a significant association between noise exposure and the risk of dementia. Based on the current evidence, exposure to noise may be a specific risk factor for dementia. Since current studies mainly focus on community noise and traffic noise, more high-quality longitudinal studies should pay attention to other noise sources, such as occupational noise exposure and risk of dementia. In addition, to better prevent dementia, rigorously designed animal experiments are also needed to explore the etiological mechanism of noise and dementia.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

LH, LM, and YZ had full access to all of the data in the study, and took responsibility for the integrity of the data and the accuracy of the data analysis and drafted the article. LH, YZ, and SZ designed the study. LH, YZ, and YL developed and tested the data collection forms. LM, YZ, SZ, and FJ acquired the data. LM, YZ, LS, and LH conducted the analysis and interpreted the data. All authors critically revised the article and read and approved the final article.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2022.832881/full#supplementary-material

REFERENCES

1. Hugo J, Ganguli M. Dementia and cognitive impairment: epidemiology, diagnosis, and treatment. *Clin Geriatr Med.* (2014) 30:421–42. doi: 10.1016/j.cger.2014.04.001
2. Garre-Olmo J. [Epidemiology of Alzheimer's disease and other dementias]. *Rev Neurol.* (2018) 66:377–86. Epidemiologia de la enfermedad de Alzheimer y otras demencias. doi: 10.33588/rn.6611.2017519
3. Wu YT, Beiser AS, Breteler MMB, Fratiglioni L, Helmer C, Hendrie HC, et al. The changing prevalence and incidence of dementia over time - current evidence. *Nat Rev Neurol.* (2017) 13:327–39. doi: 10.1038/nrneurol.2017.63
4. Gale SA, Acar D, Daffner KR. Dementia. *Am J Med.* (2018) 131:1161–9. doi: 10.1016/j.amjmed.2018.01.022
5. Ferencz B, Gerritsen L. Genetics and underlying pathology of dementia. *Neuropsychol Rev.* (2015) 25:113–24. doi: 10.1007/s11065-014-9276-3
6. Sanchez MP, Garcia-Cabrero AM, Sanchez-Elexpuru G, Burgos DF, Serratosa JM. Tau-induced pathology in epilepsy and dementia: notions from patients and animal models. *Int J Mol Sci.* (2018) 19:1092. doi: 10.3390/ijms19041092
7. Paul KC, Haan M, Mayeda ER, Ritz BR. Ambient air pollution, noise, and late-life cognitive decline and dementia risk. *Annu Rev Public Health.* (2019) 40:203–20. doi: 10.1146/annurev-publhealth-040218-044058
8. Langa KM, Levine DA. The diagnosis and management of mild cognitive impairment: a clinical review. *JAMA.* (2014) 312:2551–61. doi: 10.1001/jama.2014.13806
9. Matthews FE, Stephan BC, McKeith IG, Bond J, Brayne C, Medical Research Council Cognitive F, et al. Two-year progression from mild cognitive impairment to dementia: to what extent do different definitions agree? *J Am Geriatr Soc.* (2008) 56:1424–33. doi: 10.1111/j.1532-5415.2008.01820.x
10. Astell AJ, Bouranis N, Hoey J, Lindauer A, Mihailidis A, Nugent C, et al. Dementia Professional Interest A. Technology and dementia: the future is now. *Dement Geriatr Cogn Disord.* (2019) 47:131–9. doi: 10.1159/000497800

11. Arvanitakis Z, Shah RC, Bennett DA. Diagnosis and management of dementia: review. *JAMA*. (2019) 322:1589–99. doi: 10.1001/jama.2019.4782
12. Rubin R. Exploring the relationship between depression and dementia. *JAMA*. (2018) 320:961–2. doi: 10.1001/jama.2018.11154
13. Kalin NH. New findings related to cognition, intellectual disability, dementia, and autism. *Am J Psychiatry*. (2020) 177:473–5. doi: 10.1176/appi.ajp.2020.20040405
14. Clark C, Paunovic K, WHO. environmental noise guidelines for the european region: a systematic review on environmental noise and cognition. *Int J Environ Res Public Health*. (2018) 15:285. doi: 10.3390/ijerph15020285
15. Kempen EV, Casas M, Pershagen G, Foraster M, WHO. environmental noise guidelines for the European region: a systematic review on environmental noise and cardiovascular and metabolic effects: a summary. *Int J Environ Res Public Health*. (2018) 15:379. doi: 10.3390/ijerph15020379
16. Munzel T, Schmidt FP, Steven S, Herzog J, Daiber A, Sorensen M. Environmental Noise and the Cardiovascular System. *J Am Coll Cardiol*. (2018) 71:688–97. doi: 10.1016/j.jacc.2017.12.015
17. Cheng L, Wang SH, Chen QC, Liao XM. Moderate noise induced cognition impairment of mice and its underlying mechanisms. *Physiol Behav*. (2011) 104:981–8. doi: 10.1016/j.physbeh.2011.06.018
18. Cui B, Zhu L, She X, Wu M, Ma Q, Wang T, et al. Chronic noise exposure causes persistence of tau hyperphosphorylation and formation of NFT tau in the rat hippocampus and prefrontal cortex. *Exp Neurol*. (2012) 238:122–9. doi: 10.1016/j.expneurol.2012.08.028
19. Linares C, Culqui D, Carmona R, Ortiz C, Diaz J. Short-term association between environmental factors and hospital admissions due to dementia in Madrid. *Environ Res*. (2017) 152:214–20. doi: 10.1016/j.envres.2016.10.020
20. Chen H, Kwong JC, Copes R, Tu K, Villeneuve PJ, van Donkelaar A, et al. Living near major roads and the incidence of dementia, Parkinson's disease, and multiple sclerosis: a population-based cohort study. *Lancet*. (2017) 389:718–26. doi: 10.1016/S0140-6736(16)32399-6
21. Andersson J, Oudin A, Sundstrom A, Forsberg B, Adolfsson R, Nordin M. Road traffic noise, air pollution, and risk of dementia - results from the Betula project. *Environ Res*. (2018) 166:334–9. doi: 10.1016/j.envres.2018.06.008
22. Oudin A. Short review: air pollution, noise and lack of greenness as risk factors for Alzheimer's disease- epidemiologic and experimental evidence. *Neurochem Int*. (2020) 134:104646. doi: 10.1016/j.neuint.2019.104646
23. Robbins RN, Scott T, Joska JA, Gouse H. Impact of urbanization on cognitive disorders. *Curr Opin Psychiatry*. (2019) 32:210–7. doi: 10.1097/YCO.0000000000000490
24. Jafari Z, Kolb BE, Mohajerani MH. Noise exposure accelerates the risk of cognitive impairment and Alzheimer's disease: adulthood, gestational, and prenatal mechanistic evidence from animal studies. *Neurosci Biobehav Rev*. (2020) 117:110–28. doi: 10.1016/j.neubiorev.2019.04.001
25. Rooney RF. Preventing dementia: how lifestyle in midlife affects risk. *Curr Opin Psychiatry*. (2014) 27:149–57. doi: 10.1097/YCO.0000000000000045
26. Lin FR, Metter EJ, O'Brien RJ, Resnick SM, Zonderman AB, Ferrucci L. Hearing loss and incident dementia. *Arch Neurol*. (2011) 68:214–20. doi: 10.1001/archneurol.2010.362
27. Deal JA, Betz J, Yaffe K, Harris T, Purchase-Helzner E, Satterfield S, et al. Hearing impairment and incident dementia and cognitive decline in older adults: the health ABC study. *J Gerontol A Biol Sci Med Sci*. (2017) 72:703–9. doi: 10.1093/gerona/glw069
28. Clark C, Crumpler C, Notley H. Evidence for environmental noise effects on health for the United Kingdom policy context: a systematic review of the effects of environmental noise on mental health, wellbeing, quality of life, cancer, dementia, birth, reproductive outcomes, and cognition. *Int J Env Res Pub He*. (2020) 17:393. doi: 10.3390/ijerph17020393
29. Clark C, Paunovic K. WHO Environmental Noise Guidelines for the European Region: a systematic review on environmental noise and quality of life, wellbeing and mental health. *Int J Environ Res Public Health*. (2018) 15:2400. doi: 10.3390/ijerph15112400
30. Hegewald J, Schubert M, Freiberg A, Romero Starke K, Augustin F, Riedel-Heller SG, et al. Traffic noise and mental health: a systematic review and meta-analysis. *Int J Environ Res Public Health*. (2020) 17:6175. doi: 10.3390/ijerph17176175
31. Huang L, Zhang Y, Wang Y, Lan Y. Relationship between chronic noise exposure, cognitive impairment, and degenerative dementia: update on the experimental and epidemiological evidence and prospects for further research. *J Alzheimers Dis*. (2021) 79:1409–27. doi: 10.3233/JAD-201037
32. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. (2021) 372:n71. doi: 10.1136/bmj.n71
33. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *Eur J Epidemiol*. (2010) 25:603–5. doi: 10.1007/s10654-010-9491-z
34. Lo CK, Mertz D, Loeb M. Newcastle-Ottawa Scale: comparing reviewers' to authors' assessments. *BMC Med Res Methodol*. (2014) 14:45. doi: 10.1186/1471-2288-14-45
35. Hartling L, Milne A, Hamm MP, Vandermeer B, Ansari M, Tsertsvadze A, et al. Testing the Newcastle Ottawa Scale showed low reliability between individual reviewers. *J Clin Epidemiol*. (2013) 66:982–93. doi: 10.1016/j.jclinepi.2013.03.003
36. Rostom A, Dube C, Cranney A, Saloojee N, Sy R, Garritty C, et al. Celiac disease. in: *Evidence Report/technology Assessment*. (2004). p. 1–6.
37. Cortese S, Adamo N, Del Giovane C, Mohr-Jensen C, Hayes AJ, Carucci S, et al. Comparative efficacy and tolerability of medications for attention-deficit hyperactivity disorder in children, adolescents, and adults: a systematic review and network meta-analysis. *Lancet Psychiatry*. (2018) 5:727–38. doi: 10.1016/S2215-03661830269-4
38. Hou Y, Song A, Jin Y, Xia Q, Song G, Xing X, et al. A dose-response meta-analysis between serum concentration of 25-hydroxy vitamin D and risk of type 1 diabetes mellitus. *Eur J Clin Nutr*. (2021) 75:1010–23. doi: 10.1038/s41430-020-00813-1
39. Sadeghian M, Asadi M, Rahmani S, Akhavan Zanjani M, Sadeghi O, Hosseini SA, et al. Circulating vitamin D and the risk of gestational diabetes: a systematic review and dose-response meta-analysis. *Endocrine*. (2020) 70:36–47. doi: 10.1007/s12020-020-02360-y
40. Brink M, Schaffer B, Pieren R, Wunderli JM. Conversion between noise exposure indicators Leq24h, LDay, LEvening, LNight, Ldn and Lden: principles and practical guidance. *Int J Hyg Environ Health*. (2018) 221:54–63. doi: 10.1016/j.ijheh.2017.10.003
41. Yuchi W, Sbihi H, Davies H, Tamburic L, Brauer M. Road proximity, air pollution, noise, green space and neurologic disease incidence: a population-based cohort study. *Environ Health*. (2020) 19:8. doi: 10.1186/s12940-020-0565-4
42. Xu C, Doi SAR. The robust error meta-regression method for dose-response meta-analysis. *Int J Evid Based Healthc*. (2018) 16:138–44. doi: 10.1097/XEB.0000000000000132
43. Alexander DD, Bylsma LC, Vargas AJ, Cohen SS, Doucette A, Mohamed M, et al. Dairy consumption and CVD: a systematic review and meta-analysis. *Br J Nutr*. (2016) 115:737–50. doi: 10.1017/S0007114515005000
44. Orsini N, Li R, Wolk A, Khudyakov P, Spiegelman D. Meta-analysis for linear and nonlinear dose-response relations: examples, an evaluation of approximations, and software. *Am J Epidemiol*. (2012) 175:66–73. doi: 10.1093/aje/kwr265
45. Zhang C, Jia P, Yu L, Xu C. Introduction to methodology of dose-response meta-analysis for binary outcome: With application on software. *J Evid Based Med*. (2018) 11:125–9. doi: 10.1111/jebm.12267
46. Cantuaria ML, Waldorff FB, Wermuth L, Pedersen ER, Poulsen AH, Thacher JD, et al. Residential exposure to transportation noise in Denmark and incidence of dementia: national cohort study. *BMJ*. (2021) 374:n1954. doi: 10.1136/bmj.n1954
47. Carey IM, Anderson HR, Atkinson RW, Beevers SD, Cook DG, Strachan DP, et al. Are noise and air pollution related to the incidence of dementia? *A cohort study in London, England BMJ Open*. (2018) 8:e022404. doi: 10.1136/bmjopen-2018-022404
48. Crous-Bou M, Gascon M, Gispert JD, Cirach M, Sanchez-Benavides G, Falcon C, et al. Impact of urban environmental exposures on cognitive performance and brain structure of healthy individuals at risk for Alzheimer's dementia. *Environ Int*. (2020) 138:105546. doi: 10.1016/j.envint.2020.105546
49. Fuks KB, Wigmann C, Altug H, Schikowski T. Road traffic noise at the residence, annoyance, and cognitive function in elderly women. Int J Environ Res Public Health. 2019 16. doi: 10.3390/ijerph16101790

50. Tyas SL, Manfreda J, Strain LA, Montgomery PR. Risk factors for Alzheimer's disease: a population-based, longitudinal study in Manitoba, Canada. *Int J Epidemiol.* (2001) 30:590–7. doi: 10.1093/ije/30.3.590
51. Tzivian L, Dlugaj M, Winkler A, Hennig F, Fuks K, Sugiri D, et al. Long-term air pollution and traffic noise exposures and cognitive function: a cross-sectional analysis of the Heinz Nixdorf Recall study. *J Toxicol Environ Health A.* (2016) 79:1057–69. doi: 10.1080/15287394.2016.1219570
52. Weuve J, D'Souza J, Beck T, Evans DA, Kaufman JD, Rajan KB, et al. Long-term community noise exposure in relation to dementia, cognition, and cognitive decline in older adults. *Alzheimers Dement.* (2021) 17:525–33. doi: 10.1002/alz.12191
53. Yu Y, Haan M, Paul KC, Mayeda ER, Jerrett M, Wu J, et al. Metabolic dysfunction modifies the influence of traffic-related air pollution and noise exposure on late-life dementia and cognitive impairment: a cohort study of older Mexican-Americans. *Environ Epidemiol.* (2020) 4:e122. doi: 10.1097/EDE.0000000000001249
54. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet.* (2020) 396:413–46. doi: 10.1016/S0140-6736(20)30367-6
55. Peters R, Ee N, Peters J, Booth A, Mudway I, Anstey KJ. Air pollution and dementia: a systematic review. *J Alzheimers Dis.* (2019) 70:S145–63. doi: 10.3233/JAD-180631
56. Tzivian L, Jokisch M, Winkler A, Weimar C, Hennig F, Sugiri D, et al. Associations of long-term exposure to air pollution and road traffic noise with cognitive function-An analysis of effect measure modification. *Environ Int.* (2017) 103:30–8. doi: 10.1016/j.envint.2017.03.018
57. Arnsten AF, Goldman-Rakic PS. Noise stress impairs prefrontal cortical cognitive function in monkeys: evidence for a hyperdopaminergic mechanism. *Arch Gen Psychiatry.* (1998) 55:362–8. doi: 10.1001/archpsyc.55.4.362
58. Czeh B, Muller-Keuker JI, Rygula R, Abumaria N, Hiemke C, Domenici E, et al. Chronic social stress inhibits cell proliferation in the adult medial prefrontal cortex: hemispheric asymmetry and reversal by fluoxetine treatment. *Neuropsychopharmacology.* (2007) 32:1490–503. doi: 10.1038/sj.npp.1301275
59. Jafari Z, Kolb BE, Mohajerani MH. Chronic traffic noise stress accelerates brain impairment and cognitive decline in mice. *Exp Neurol.* (2018) 308:1–12. doi: 10.1016/j.expneurol.2018.06.011
60. Arnsten AF. Stress signalling pathways that impair prefrontal cortex structure and function. *Nat Rev Neurosci.* (2009) 10:410–22. doi: 10.1038/nrn2648
61. Guadagno A, Wong TP, Walker CD. Morphological and functional changes in the preweaning basolateral amygdala induced by early chronic stress associate with anxiety and fear behavior in adult male, but not female rats. *Prog Neuropsychopharmacol Biol Psychiatry.* (2018) 81:25–37. doi: 10.1016/j.pnpbp.2017.09.025
62. Di G, Liu G, Xu Y, Kim H. Effects of combined traffic noise on the synaptic ultrastructure and expressions of p-CaMKII and NMDAR1 in the hippocampus of young SD rats. *Environ Sci Pollut Res Int.* (2019) 26:22030–9. doi: 10.1007/s11356-019-05457-8
63. Di G, Zheng Y. Effects of high-speed railway noise on the synaptic ultrastructure and phosphorylated-CaMKII expression in the central nervous system of SD rats. *Environ Toxicol Pharmacol.* (2013) 35:93–9. doi: 10.1016/j.etap.2012.11.012
64. Raymond C, Marin MF, Majeur D, Lupien S. Early child adversity and psychopathology in adulthood: HPA axis and cognitive dysregulations as potential mechanisms. *Prog Neuropsychopharmacol Biol Psychiatry.* (2018) 85:152–60. doi: 10.1016/j.pnpbp.2017.07.015
65. Gai Z, Su D, Wang Y, Li W, Cui B, Li K, et al. Effects of chronic noise on the corticotropin-releasing factor system in the rat hippocampus: relevance to Alzheimer's disease-like tau hyperphosphorylation. *Environ Health Prev Med.* (2017) 22:79. doi: 10.1186/s12199-017-0686-8
66. Ravindran R, Rathinasamy SD, Samson J, Senthilvelan M. Noise-stress-induced brain neurotransmitter changes and the effect of Ocimum sanctum (Linn) treatment in albino rats. *J Pharmacol Sci.* (2005) 98:354–60. doi: 10.1254/jphs.FP0050127
67. Cui B, Li K, Gai Z, She X, Zhang N, Xu C, et al. Chronic noise exposure acts cumulatively to exacerbate Alzheimer's disease-like amyloid-beta pathology and neuroinflammation in the rat hippocampus. *Sci Rep.* (2015) 5:12943. doi: 10.1038/srep12943
68. Jafari Z, Okuma M, Karem H, Mehla J, Kolb BE, Mohajerani MH. Prenatal noise stress aggravates cognitive decline and the onset and progression of beta amyloid pathology in a mouse model of Alzheimer's disease. *Neurobiol Aging.* (2019) 77:66–86. doi: 10.1016/j.neurobiolaging.2019.01.019
69. Jafari Z, Mehla J, Kolb BE, Mohajerani MH. Prenatal noise stress impairs HPA axis and cognitive performance in mice. *Sci Rep.* (2017) 7:10560. doi: 10.1038/s41598-017-09799-6

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Meng, Zhang, Zhang, Jiang, Sha, Lan and Huang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Medicine

ORIGINAL RESEARCH
published: 27 June 2022
doi: 10.3389/fmed.2022.830580

The ABC-Stroke Score Refines Stroke Risk Stratification in Patients With Atrial Fibrillation at the Emergency Department

Jan Niederdöckl[1], Julia Oppenauer[1], Sebastian Schnaubelt[1], Filippo Cacioppo[1], Nina Buchtele[2,3], Alexandra-Maria Warenits[1], Roberta Laggner[4], Nikola Schütz[1], Magdalena S. Bögl[1], Gerhard Ruzicka[1], Sophie Gupta[1], Martin Lutnik[1], Safoura Sheikh Rezaei[2], Michael Wolzt[2], Harald Herkner[1], Hans Domanovits[1], Anton N. Laggner[1], Michael Schwameis[1] and Ziad Hijazi[5]*

[1] Department of Emergency Medicine, Medical University of Vienna, Vienna, Austria, [2] Department of Clinical Pharmacology, Medical University of Vienna, Vienna, Austria, [3] Department of Medicine I, Medical University of Vienna, Vienna, Austria, [4] Department of Orthopedics and Trauma-Surgery, Medical University of Vienna, Vienna, Austria, [5] Department of Medical Sciences, Cardiology, Uppsala Clinical Research Center, Uppsala University, Uppsala, Sweden

OPEN ACCESS

Edited by:
Kevin Lu,
University of South Carolina,
United States

Reviewed by:
Raphael Van Tulder,
Karl Landsteiner University of Health Sciences, Austria
Masaraf Hussain,
North Eastern Indira Gandhi Regional Institute of Health and Medical Sciences, India

***Correspondence:**
Michael Schwameis
michael.schwameis@meduniwien.ac.at

Specialty section:
This article was submitted to Family Medicine and Primary Care, a section of the journal Frontiers in Medicine

Received: 07 December 2021
Accepted: 01 June 2022
Published: 27 June 2022

Citation:
Niederdöckl J, Oppenauer J, Schnaubelt S, Cacioppo F, Buchtele N, Warenits A-M, Laggner R, Schütz N, Bögl MS, Ruzicka G, Gupta S, Lutnik M, Sheikh Rezaei S, Wolzt M, Herkner H, Domanovits H, Laggner AN, Schwameis M and Hijazi Z (2022) The ABC-Stroke Score Refines Stroke Risk Stratification in Patients With Atrial Fibrillation at the Emergency Department. Front. Med. 9:830580. doi: 10.3389/fmed.2022.830580

Aims: To evaluate the performance of the ABC (Age, Biomarkers, Clinical history) and CHA_2DS_2-VASc stroke scores under real-world conditions in an emergency setting.

Methods and Results: The performance of the biomarker-based ABC-stroke score and the clinical variable-based CHA_2DS_2-VASc score for stroke risk assessment were prospectively evaluated in a consecutive series of 2,108 patients with acute symptomatic atrial fibrillation at a tertiary care emergency department. Performance was assessed according to methods for the development and validation of clinical prediction models by Steyerberg et al. and the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis. During a cumulative observation period of 3,686 person-years, the stroke incidence rate was 1.66 per 100 person-years. Overall, the ABC-stroke and CHA_2DS_2-VASc scores revealed respective c-indices of 0.64 and 0.55 for stroke prediction. Risk-class hazard ratios comparing moderate to low and high to low were 3.51 and 2.56 for the ABC-stroke score and 1.10 and 1.62 for the CHA_2DS_2-VASc score. The ABC-stroke score also provided improved risk stratification in patients with moderate stroke risk according to the CHA_2DS_2-VASc score, who lack clear recommendations regarding anticoagulation therapy (HR: 4.35, $P = 0.001$). Decision curve analysis indicated a superior net clinical benefit of using the ABC-stroke score.

Conclusion: In a large, real-world cohort of patients with acute atrial fibrillation in the emergency department, the ABC-stroke score was superior to the guideline-recommended CHA_2DS_2-VASc score at predicting stroke risk and refined risk stratification of patients labeled moderate risk by the CHA_2DS_2-VASc score, potentially easing treatment decision-making.

Keywords: symptomatic atrial fibrillation, stroke, prediction score, performance evaluation, validation, biomarkers

INTRODUCTION

The more than 13 million strokes that occur worldwide each year drive constant efforts to optimize patient management (1, 2). Up to one-third of incident ischemic strokes are cardiothromboembolic in source and associated with atrial fibrillation (AF) (2, 3). Adequate long-term management of AF needs to be initiated early, preferably during the patient's initial presentation to the emergency department (2, 4).

Risk stratification schemes for AF patients have been continuously improving over recent years (5–7). Routine clinical risk assessment is widely based on the CHA_2DS_2-VASc score. While the CHA_2DS_2-VASc score uses clinical variables, the more-recent ABC (Age, Biomarkers, Clinical history)-stroke score incorporates age and clinical history of previous stroke/transient ischemic attack with two cardiac biomarkers: high sensitivity cardiac troponin T (hs-cTnT) and the pro-B-type natriuretic peptide N-terminal fragment (NT-proBNP) (8–11).

Risk prediction based on the ABC-stroke score appears superior to that obtained with the CHA_2DS_2-VASc score (6, 12). However, clinical presentation and medical history of patients with AF can vary greatly, which strongly impacts relevant biomarkers. Therefore, a detailed characterization of score performance at the subgroup level is essential. In the emergency setting, no optimal score for stroke prediction in patients with AF has yet been defined (13, 14). The current study evaluated the performance of the biomarker-based ABC-stroke score and the clinical variable-based CHA_2DS_2-VASc score, which is recommended by current guidelines, in patients with acute symptomatic AF in the emergency department.

METHODS

Ethical Approval

This study, which was conducted in accordance with ICH-GCP guidelines and the Declaration of Helsinki, was approved by the Ethics Committee of the Medical University of Vienna (1568/2014); Patients gave written consent before participation. The study was registered at clinicaltrials.gov (NCT03272620).

Emergency Department Cohort

This study included a consecutive series of patients with acute, symptomatic, hemodynamically stable, first-detected, or recurrent nonpermanent AF who presented to the emergency department of the Medical University of Vienna, a tertiary care academic institution, from 2011 to 2018. Only patients with acute symptoms strongly suggestive of tachyarrhythmia, such as palpitations, whose 12-lead ECG confirmed the presence of AF were eligible. Data were prospectively collected within the framework of a local, standardized AF registry that was previously described (13). Patients with permanent AF (n = 139) and cases of hemodynamic instability in which AF occurred as an epiphenomenon of the underlying disease (n = 87) were not included in the registry. Prospective enrolment into the registry required that the patient's AF be confirmed by a 12-channel ECG and that they provided informed consent. Demographic data, comorbidities, medication intake, previous electrical cardioversion attempts, CHA_2DS_2-VASc score, blood gas analysis, laboratory results, vital signs (heart rate, blood pressure, and oxygen saturation), symptoms and time of onset, and previous treatments were recorded. Treatment strategies for acute control of rate and rhythm were based on the European Society of Cardiology guidelines and considered the patient's left ventricular function, hemodynamics, medication intake (such as anticoagulants), and comorbidities (2, 4).

Laboratory Values and Measurements

Serum NT-proBNP and hs-cTnT assays were processed on a Cobas E602 Module Console (ECLIA, Roche Diagnostics GmbH, Mannheim, Germany) with coefficients of variation and reference ranges of 5.7% and 0–14 ng/L for hs-cTnT and 3.7% and 0–125 pg/mL for NT-proBNP. The limit of blank was 3 ng/L for hs-cTnT. The limits of detection were 5 ng/L for hs-cTnT and 5 pg/mL for NT-proBNP (according to CLSI EP17-A guidelines) (13).

Cohorts Used to Derive the ABC-Stroke and CHA_2DS_2-VASc Scores

A comprehensive report on the development of the ABC-stroke score based on the ARISTOTLE trial has recently been published (8). Briefly, 18,201 patients with AF at risk of stroke were randomized to receive oral anticoagulation with either warfarin or apixaban. The 14,701 patients with available biomarkers for ABC-stroke score derivation were followed for a median time of 1.9 years. The development, validation, and performance of the CHA_2DS_2-VASc score has been described in detail previously (5).

Outcome

The outcome was defined following the International Classification of Diseases (ICD), 10th Revision, German Modification (2020) and assessed in a total of 2,108 patients that were not censured for treatment or medical history. The outcome was ischemic stroke (I63.-) and included patients with haemorrhagic transformation (I.69.3) but not primary intracerebral hemorrhage (I61.-). In addition, trained study fellows used a structured telephone interview with the patient or, in cases of death or communication difficulties, the next of kin to complete a standardized follow-up questionnaire developed at our institution to systematically assess the outcome. The observation period ended on either the last date of follow up or the date of death.

Statistical Methods

We present categorized data as absolute counts and relative frequency, and continuous data as median and 25–75% interquartile range (IQR). The 1-year risk of stroke was calculated for each patient using both the ABC-stroke score and the CHA_2DS_2-VASc score, irrespective of stroke history or whether or not the patient was on anticoagulation therapy. The discriminatory capabilities of the scores were evaluated using Harrel's c-indices (15), Kaplan-Meier curves, and hazard ratios. These were applied to predefined risk classes with 0–1%, 1–2%, or >2% risk of stroke within 1 year. The predictive capabilities of the ABC-stroke score and CHA_2DS_2-VASc risk

TABLE 1 | Demographics and baseline characteristics of study patients by CHA_2DS_2-VASc risk classes.

	Total		Treatment recommendation based on CHA_2DS_2-VASc					
			No		To be considered non-sex CHA_2DS_2-VASc		Yes	
			<1		1		>1	
	n = 2,108	Available* (*n*)	*n* = 219	Available* (*n*)	*n* = 437	Available* (*n*)	*n* = 1,452	Available* (*n*)
General characteristics								
Age, years (IQR)	68 (59–76)		51 (37–58)		62 (53–68)		73 (66–80)	
Female sex, *n* (%)	911 (43)		87 (40)		163 (37)		661 (46)	
Comorbidities								
Heart failure, *n* (%)	369 (17.5)		4 (1.8)		59 (13.5)		306 (21.1)	
Hypertension, *n* (%)	1,236 (58.6)		0 (0.0)		200 (45.8)		1,036 (71.3)	
Diabetes mellitus, *n* (%)	344 (16.3)		0 (0.0)		12 (2.7)		332 (22.9)	
Prior stroke, *n* (%)	123 (5.8)		0 (0.0)		0 (0.0)		123 (8.5)	
Coronary artery disease, *n* (%)	383 (18.2)		0 (0.0)		26 (5.9)		357 (24.6)	
Prior myocardial infarction, *n* (%)	195 (9.3)		0 (0.0)		7 (1.6)		188 (12.9)	
Peripheral artery disease, *n* (%)	93 (4.4)		0 (0.0)		3 (0.7)		90 (6.2)	
AF history								
First AF episode, *n* (%)	29 (1.4)		1 (0.0)		8 (0.4)		20 (0.9)	
Heart rate, bpm (IQR)	131 (110–150)	1,097	135 (108–159)	148	134 (119–151)	208	129 (108–147)	714
Duration of AF symptoms, h (IQR)	6 (2–20)	846	4 (1–10)	121	6 (2–18)	190	7 (3–24)	535
Laboratory								
NT-proBNP, pg/ml (IQR)	636 (150–2,153)	2,108	244 (64–782)	219	465 (133–1,588)	437	932 (192–2,433)	1,452
hs-Troponin T, ng/l (IQR)	12 (5–23)	2,108	7 (3–13)	219	8 (4–16)	437	13 (6–28)	1,452
Stroke risk scores								
CHA_2DS_2-VASc (IQR)	3 (1–4)	1,984	0 (0–0)	219	1 (1–2)	437	3 (3–4)	1,328
ABC-stroke risk, 1-year (IQR)	0.9 (0.5–1.5)	2,018	0.5 (0.3–0.8)	219	0.7 (0.5–1.2)	437	1.1 (0.6–1.7)	1,452
Outcome								
Stroke, *n* (%)	61 (3.0)		1 (1.4)		14 (4.7)		46 (3.2)	
Time to event, months (IQR)	12.1 (3.5–32.2)		22.1 (1.0–33.8)		14.2 (3.9–24.4)		12.4 (2.8–32.2)	

*AF, atrial fibrillation; hs-Troponin T, high-sensitivity Troponin T; NT-proBNP, N-terminal pro-B-type natriuretic peptide. *The number of patients for whom the variable was available.*

TABLE 2 | Stroke incidence rates and hazard ratios for the ABC and CHA_2DS_2-VASc stroke risk classes.

Incidence rates and hazard ratios	*N*	Events	Incidence rate°	Hazard ratio	*p*
ABC-stroke score					
Low risk (<1%)	1,135	16	0.79 (0.49–1.30)	1.00	Ref
Medium risk (1–2%)	731	37	2.87 (2.08–3.96)	3.51 (1.95–6.31)	<0.001
High risk (>2%)	242	8	2.12 (1.06–4.23)	2.56 (1.10–5.98)	<0.030
CHA_2DS_2VASc					
Low risk (<1%*)	504	10	1.21 (0.70–2.24)	1.00	Ref
Medium Risk (1–2%**)	409	10	1.29 (0.70–2.40)	1.10 (0.46–2.64)	<0.832
High risk (>2%***)	1,195	41	1.97 (1.45–2.68)	1.62 (0.81–3.24)	<0.170

CHA_2DS_2-VASc ≤ 1.
**CHA_2DS_2-VASc = 2.*
***CHA_2DS_2-VASc > 2 according to Oldgren et al. (6).*
°*Per 100 person years.*

score were then compared (5, 6, 16). C-indices were compared utilizing 1,000-fold bootstraps. We performed sensitivity analyses taking into account anticoagulation status, history of ischemic stroke, the presence of acute heart failure and acute coronary syndrome. The net clinical benefits of the ABC-stroke score and the CHA_2DS_2-VASc score were determined with a decision curve analysis (17). In this context, net clinical benefit was defined as the relationship between the benefit of treating those who need treatment and the harm of treating those who do not need treatment. Decision curve analysis allows the evaluation of net

TABLE 3 | C-indices for the ABC-stroke and CHA_2DS_2-VASc scores.

C-indices	*N*	Events	Harrell's C
ABC-stroke score			
All patients with acute AF	2,108	61	0.64 (0.57–0.70)
Consider anticoagulation*	437	14	0.66 (0.51–0.80)
CHA_2DS_2VASc			
All patients with acute AF	2,108	61	0.55 (0.49–0.60)

C-indices with 95% confidence intervals in parentheses for the ABC-stroke and CHA_2DS_2-VASc scores in the total cohort (n = 2,108) and in patients classified as "consider anticoagulation" according to ESC guidelines (n = 437) (2, 4).
**Non-sex CHA_2DS_2-VASc = 1.*

clinical benefit of a prognostic tool over a range of threshold probabilities of having a positive outcome. Net clinical benefit is calculated as $\frac{true\ positives}{n} - \frac{false\ positives}{n} * (\frac{pt}{1-pt})$, where n is the total number of patients, and pt is the threshold probability of having a positive outcome. True and false positives are calculated using pt as the cut-off point for determining a positive or negative result. By calculating and plotting the net benefit for all reasonable thresholds, the decision curve is then determined, allowing comparison of several predictive models. The model with the highest net benefit for a given threshold is assumed to be the preferred one. This is shown in a coordinate system with possible thresholds for stroke risk on the x-axis and the net benefit per patient on the y-axis. In addition, the predictive capabilities of the ABC-stroke score, discriminatory power, and net clinical benefit, were assessed across CHA_2DS_2-VASc risk classes. The analyses were conducted in accordance with the recommendations on the derivation and validation of prediction models proposed by Steyerberg et al. (1). Reporting followed the Transparent Reporting of Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement (18). Missing data were included as separate categories for each variable as appropriate. Stata/BE 17.0 for Mac (StataCorp, College Station, TX 77845, USA) was used for data analysis. Generally, a two-sided P-value < 0.05 was considered to be statistically significant.

RESULTS

Demographics and Patient Characteristics

A consecutive series of 2,108 patients with acute symptomatic AF were included in the study. Patient characteristics are shown in **Table 1**. The median age was 68 years (IQR 59–76), 911 patients (43%) were female. The median time from onset of arrhythmia-related symptoms to presentation was 6 h (IQR 2–20). Of the 2,108 included patients, 733 patients (34.7%) were on anticoagulation therapy at presentation, and 1,344 (63.8%) were on anticoagulants at the time of discharge or transfer from the emergency department (**Supplementary Table 1**). Within this cohort, 123 patients (5.8%) had a history of ischemic stroke.

Outcome

Ischemic stroke occurred in 61 patients during the median follow-up period of 23 months (IQR: 12–39). Incidence rates of strokes per 100 person-years according to ABC-stroke score and CHA_2DS_2-VASc score risk classes are presented in **Table 2** and corresponded to an overall incidence rate of 1.66 per 100 person-years.

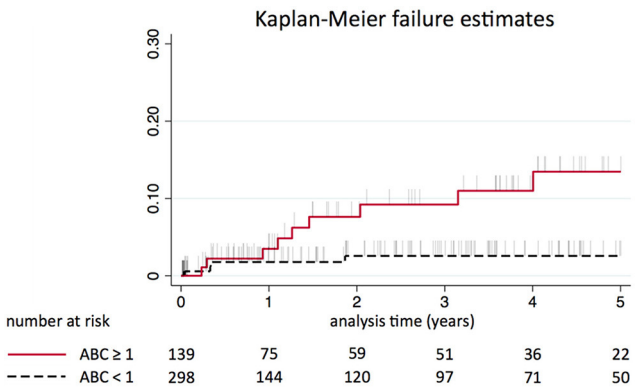

FIGURE 1 | Kaplan-Meier estimated event rate for patients labeled as moderate risk by the CHA_2DS_2-VASc score stratified by ABC-stroke risk classes (≥1 vs. <1% risk of stroke per 1 year).

Performance Evaluation

The hazard ratios for the ABC-stroke score and CHA_2DS_2-VASc score are presented in **Table 2**. Cumulative event rates and c-indices for both scores are shown in **Table 3**. Overall, the ABC-stroke score had a c-index for stroke prediction of 0.64, and the CHA_2DS_2-VASc score had a c-index of 0.55. In patients with moderate stroke risk according to the CHA_2DS_2-VASc score, the ABC-stroke score provided improved risk stratification (hazard ratio: 4.35, $P = 0.001$). Results from sensitivity analyses in important clinical subgroups are available in **Supplementary Table 2**. The ABC-stroke score performed consistently in important subgroups, such as the anticoagulation status at discharge, the presence of acute heart failure, and in patients without acute coronary syndromes.

Decision Curve Analysis

The decision curve analyses showed a net benefit for using either scoring system. The benefit of using the ABC-stroke score was superior to that of the CHA_2DS_2-VASc score over the 1–5% decision threshold range and was most prominent in patients

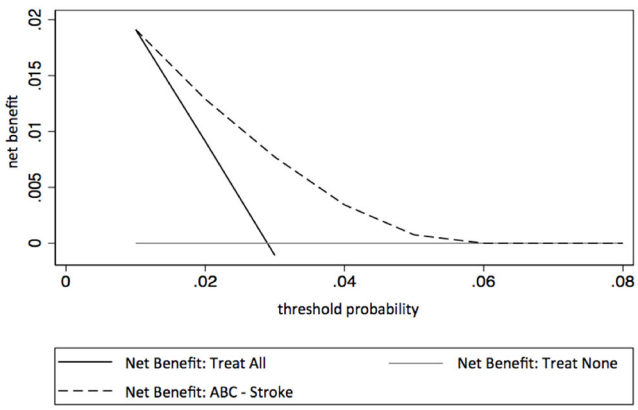

FIGURE 2 | Decision curve analysis for the ABC-stroke score. Net benefit (y-axis) reflects whether basing ischemic stroke risk prediction on the ABC-stroke score provides greater benefit than harm. The unit of net benefit is true positives (ischemic strokes) per patients. A net benefit of 0.01 means that using the ABC-stroke score increases the number of correctly predicted ischemic strokes by 1 out of 100 target patients, without changing the number of false-positive stroke predictions. Threshold probability (x-axis) refers to the cut-offs of predicted ischemic stroke risk used to decide treatment (19, 20).

labeled moderate risk by the CHA_2DS_2-VASc score, for whom anticoagulation may be considered (**Figures 1**, **2**).

The Kaplan-Meier estimated event rate stratified by CHA_2DS_2-VASc risk classes are shown in the **Supplementary Figure 1**.

DISCUSSION

We evaluated the performance of the biomarker-based ABC-stroke score compared to the widely used clinical variables-based CHA_2DS_2-VASc score in a large series of patients admitted to the Medical University of Vienna Emergency Department with acute symptomatic AF. The ABC-stroke score predicted stroke better than the CHA_2DS_2-VASc score.

Because of their overall high stroke rates, it is important that patients with AF receive adequate anticoagulation treatment whenever indicated (2, 4, 21, 22). Thus, published guidelines recommend the use of prediction tools to guide appropriate treatment (2). While the clinical CHA_2DS_2-VASc score is recommended for daily practice, biomarker-driven scores are gaining in importance (2, 6, 8–10).

The ABC-stroke score has been validated, and its superiority over other prediction models has been shown in several patient cohorts (6, 8, 12, 23, 24). Its discriminative ability is based on four variables: age, the history of previous stroke or transient ischemic attack, and two cardiac biomarkers (hs-cTnT and NT-proBNP) (8). While the ubiquitous availability of the two cardiac biomarkers represents an advantage, their sensitivity to various modifying factors should be kept in mind (6, 13, 14). Both biomarkers included in the ABC-stroke score may be affected by myocardial oxygen demand, which is altered by heart rate and tachyarrhythmia severity (11, 14, 25). Since the ABC-stroke score was developed in clinically stable patients with AF treated with oral anticoagulation and enrolled in a randomized controlled trial, its usefulness in an unselected "real-world" cohort of patients presenting with acute AF to an emergency department has not yet been determined.

To our knowledge, the current study is the first to show that the ABC-stroke score performs well in a relatively large and diverse emergency cohort. This present study cohort had a wide spectrum of comorbidities and included patients with significantly altered vital parameters, such as ongoing tachyarrhythmia with median heart rates of 128 bpm at baseline (IQR: 110–131) (13, 26). Despite such conditions potentially influencing the cardiac biomarkers in the ABC-stroke score, it outperformed the guideline-recommended CHA_2DS_2-VASc score that is based solely on categorical variables (6). The wide inclusion criteria in the study also permitted clinically important subanalyses. The ABC-stroke score performed consistently in important subgroups, irrespective of anticoagulation status and heart failure. However, the ABC-stroke score performed less well in patients with concomitant acute coronary syndrome. However, this is expected because such a condition substantially alters cardiac troponin levels. Therefore, acute coronary syndrome should be ruled out before stroke risk stratification. The ABC-stroke score yielded a higher stroke risk in moderate-risk patients than high-risk patients. This was most likely due to a lower proportion of patients in the moderate-risk group taking oral anticoagulants (27). Nevertheless, decision curve analysis indicated superior treatment decision-making with the ABC-stroke score compared to the CHA_2DS_2-VASc score. Additionally, the ABC-stroke score improved risk prediction within the subgroup of patients classified as moderate risk by the CHA_2DS_2-VASc score. The ABC-stroke score also more-accurately identified patients at a low risk of stroke. Given the higher precision of the ABC-stroke score, its use in this important acute cardiac setting may be improve treatment decisions (2, 28).

The dynamic nature of the cardiac biomarkers included in the ABC-stroke score may also be advantageous because changes in these biomarkers may reflect differences in individual stroke risk that may be modified by various factors (29, 30). The latter could be of particular importance since continuous re-evaluation of both stroke and bleeding risk are key for assessing the risks and benefits of anticoagulant therapy (2). However, it should be noted that the benefit-risk ratio of oral anticoagulation in patients labeled moderate risk by the CHA_2DS_2-VASc has not been well studied. In some individuals, the anticoagulation risks may outweigh the benefits. Therefore, in acute AF, the decision for long-term anticoagulation must carefully and simultaneously balance individual stroke and bleeding risks. Dynamic bleeding risk assessment could include biomarker-based tools, such tools as the ABC-bleeding score (31).

The stroke incidence rate was 1.66 per 100 person-years in this study. This was lower than expected, especially compared to elderly patients with AF. It is therefore possible that the improved stroke prediction provided by the ABC-stroke score could have an even greater impact in elderly populations, who have a substantially higher risk of adverse events. Despite this

lower incidence, our study clearly demonstrated that the ABC-stroke score successfully predicts stroke risk in an emergency department, particularly in cases where the CHA_2DS_2-VASc score provides weak anticoagulation guidance. The ABC-stroke score may therefore refine risk stratification. Another issue encountered in routine clinical care is the turnaround time for biomarker tests. However, these biomarkers are usually routinely available in emergency departments with turnaround times of 1–1.5 h and around-the-clock availability. The online availability of the ABC-stroke score nomogram and its calculator also facilitates its clinical application (https://www.ucr.uu.se/en/services/abc-risk-calculators) (6, 8).

Strengths and Limitations

Clear strengths of the present ABC-stroke score validation are its real-world emergency cohort of patients with AF, the long combined observation period of 3,686 years, several clinically relevant subgroups, and sensitivity analyses that included the ABC-stroke score recalibration based on a recent study of patients with AF not treated with oral anticoagulation therapy (27). In addition, this evaluation utilized the principles and methods for validating and reporting clinical prediction models described by Steyerberg et al. and the TRIPOD consensus statement (17, 18). However, a significant limitation of this study was the lack of information regarding the cause of death in deceased patients.

Finally, it should be noted that the study does not allow conclusions to be drawn regarding patients with classic valvular AF. These patients are only very sparsely represented in our AF registry and therefore could not be investigated separately in the present study.

CONCLUSION

In a large series of patients with acute AF treated in the emergency department, the ABC-stroke score provided superior stroke prediction over the CHA_2DS_2-VASc score. It also improved the risk stratification of patients labeled moderate risk by the CHA_2DS_2-VASc score, thereby easing anticoagulation treatment decision-making.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of the Medical University of Vienna (1568/2014). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JN, MS, and ZH developed the study idea and wrote the final version of the manuscript in critical review. JO, SSc, FC, NB, A-MW, RL, NS, MB, GR, SG, ML, and SSR carried out the main part of the study and helped in writing the manuscript as well as in creating the figures, tables, and graphical and video abstract. HH carried out the statistical tests and analyses and critically revised the final manuscript. MW, HD, and AL participated in the project as senior consultants and critically revised the final manuscript. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2022.830580/full#supplementary-material

REFERENCES

1. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. (2010) 21:128–38. doi: 10.1097/EDE.0b013e3181c30fb2
2. GBD 2015. Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet*. (2016) 388:1659–724. doi: 10.1016/S0140-6736(16)31679-8
3. Hindricks G, Potpara T, Dagres N, Arbelo E, Bax JJ, Blomström-Lundqvist C, et al. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association of Cardio-Thoracic Surgery (EACTS). *Eur Heart J*. (2021) 42:373–498. doi: 10.1093/eurheartj/ehaa612
4. Yiin GSC, Li L, Bejot Y, Rothwell PM. Time trends in atrial fibrillation-associated stroke and premorbid anticoagulation. *Stroke*. (2018) 50:STROKEAHA118022249. doi: 10.1161/STROKEAHA.118.022249
5. Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *Eur Heart J*. (2016) 37:2893–962. doi: 10.5603/KP.2016.0172
6. Lip GYH, Frison L, Halperin JL, Lane DA. Identifying patients at high risk for stroke despite anticoagulation: a comparison of contemporary stroke risk stratification schemes in an anticoagulated atrial fibrillation cohort. *Stroke*. (2010) 41:2731–8. doi: 10.1161/STROKEAHA.110.590257
7. Oldgren J, Hijazi Z, Lindbäck J, Alexander JH, Connolly SJ, Eikelboom JW, et al. Performance and validation of a novel biomarker-based stroke risk score for atrial fibrillation. *Circulation*. (2016) 134:1697–707. doi: 10.1161/CIRCULATIONAHA.116.022802
8. The SPAF III. Writing Committee for the Stroke Prevention in Atrial Fibrillation Investigators. Patients With Nonvalvular Atrial Fibrillation at Low Risk of Stroke During Treatment With Aspirin: Stroke Prevention in Atrial Fibrillation III Study. *JAMA*. (1998) 279:1273–7. doi: 10.1001/jama.279.16.1273
9. Hijazi Z, Lindbäck J, Alexander JH, Hanna M, Held C, Hylek EM, et al. The ABC (age, biomarkers, clinical history) stroke risk score: a biomarker-based risk score for predicting stroke in atrial fibrillation. *Eur Heart J*. (2016) 37:1582–90. doi: 10.1093/eurheartj/ehw054
10. Hijazi Z, Wallentin L, Siegbahn A, Andersson U, Alexander JH, Atar D, et al. High-sensitivity troponin T and risk stratification in patients with atrial

fibrillation during treatment with apixaban or warfarin. *J Am Coll Cardiol.* (2014) 63:52–61. doi: 10.1016/j.jacc.2013.07.093
11. Hijazi Z, Wallentin L, Siegbahn A, Andersson U, Christersson C, Ezekowitz J, et al. N-terminal pro-B-type natriuretic peptide for risk assessment in patients with atrial fibrillation: insights from the ARISTOTLE Trial (Apixaban for the Prevention of Stroke in Subjects With Atrial Fibrillation). *J Am Coll Cardiol.* (2013) 61:2274–84. doi: 10.1016/j.jacc.2012.11.082
12. Hijazi Z, Oldgren J, Andersson U, Connolly SJ, Ezekowitz MD, Hohnloser SH, et al. Cardiac biomarkers are associated with an increased risk of stroke and death in patients with atrial fibrillation: a randomized evaluation of long-term anticoagulation therapy (RE-LY) substudy. *Circulation.* (2012) 125:1605–16. doi: 10.1161/CIRCULATIONAHA.111.038729
13. Berg DD, Ruff CT, Jarolim P, Giugliano RP, Nordio F, Lanz HJ, et al. Performance of the ABC scores for assessing the risk of stroke or systemic embolism and bleeding in patients with atrial fibrillation in ENGAGE AF-TIMI 48. *Circulation.* (2019) 139:760–71. doi: 10.1161/CIRCULATIONAHA.118.038312
14. Niederdöckl J, Simon A, Schnaubelt S, Schuetz N, Laggner R, Sulzgruber P, et al. Cardiac biomarkers predict mortality in emergency patients presenting with atrial fibrillation. *Heart.* (2019) 105:482–8. doi: 10.1136/heartjnl-2018-313145
15. Niederdöckl J, Simon A, Cacioppo F, Buchtele N, Merrelaar A, Schütz N, et al. Predicting spontaneous conversion to sinus rhythm in symptomatic atrial fibrillation: the ReSinus score. *Eur J Intern Med.* (2021) 83:45–53. doi: 10.1016/j.ejim.2020.07.022
16. Harrell FE. *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis.* 2nd ed. Cham Heidelberg New York: Springer (2015). p. 582 (Springer series in statistics).
17. Lip GYH, Nieuwlaat R, Pisters R, Lane DA, Crijns HJGM. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach. *Chest.* (2010) 137:263–72. doi: 10.1378/chest.09-1584
18. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *J Clin Epidemiol.* (2015) 68:112–21. doi: 10.1016/j.jclinepi.2014.11.010
19. Vickers AJ, Cronin AM, Elkin EB, Gonen M. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. *BMC Med Inform Decis Mak.* (2008) 8:53. doi: 10.1186/1472-6947-8-53
20. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making.* (2006) 26:565–74. doi: 10.1177/0272989X06295361
21. Johnson CO, Nguyen M, Roth GA, Nichols E, Alam T, Abate D, et al. Global, regional, and national burden of stroke, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol.* (2019) 18:439–58.
22. Pisters R, Lane DA, Marin F, Camm AJ, Lip GYH. Stroke and thromboembolism in atrial fibrillation. *Circ J.* (2012) 76:2289–304. doi: 10.1253/circj.CJ-12-1036
23. Kearon C. In AF, ABC scores predicted stroke or major bleeding better than CHA2DS2-VASc and HAS-BLED scores, respectively. *Ann Intern Med.* (2019) 170:JC71. doi: 10.7326/ACPJ201906180-071
24. Rivera-Caravaca JM, Roldán V, Esteve-Pastor MA, Valdés M, Vicente V, Lip GYH, et al. Long-term stroke risk prediction in patients with atrial fibrillation: comparison of the ABC-stroke and CHA2DS2-VASc scores. *J Am Heart Assoc.* (2017) 6:490. doi: 10.1161/JAHA.117.006490
25. Avezum A, Lopes RD, Schulte PJ, Lanas F, Gersh BJ, Hanna M, et al. Apixaban in comparison with warfarin in patients with atrial fibrillation and valvular heart disease: findings from the apixaban for reduction in stroke and other thromboembolic events in atrial fibrillation (ARISTOTLE) trial. *Circulation.* (2015) 132:624–32. doi: 10.1161/CIRCULATIONAHA.114.014807
26. Andrade J, Khairy P, Dobrev D, Nattel S. The clinical profile and pathophysiology of atrial fibrillation: relationships among clinical features, epidemiology, and mechanisms. *Circ Res.* (2014) 114:1453–68. doi: 10.1161/CIRCRESAHA.114.303211
27. Benz AP, Hijazi Z, Lindbäck J, Connolly SJ, Eikelboom JW, Oldgren J, et al. Biomarker-based risk prediction with the ABC-AF scores in patients with atrial fibrillation not receiving oral anticoagulation. *Circulation.* (2021) 143:1863–73. doi: 10.1161/CIRCULATIONAHA.120.053100
28. Shin SY, Han S, Kim J, Im SI, Shim J, Ahn J, et al. Identification of markers associated with development of stroke in "clinically low-risk" atrial fibrillation patients. *JAHA.* (2019) 8:012697. doi: 10.1161/JAHA.119.012697
29. Hijazi Z, Oldgren J, Andersson U, Connolly SJ, Ezekowitz MD, Hohnloser SH, et al. Importance of persistent elevation of cardiac biomarkers in atrial fibrillation: a RE-LY substudy. *Heart.* (2014) 100:1193–200. doi: 10.1136/heartjnl-2013-304872
30. Eckman MH, Singer DE, Rosand J, Greenberg SM. Moving the tipping point: the decision to anticoagulate patients with atrial fibrillation. *Circ Cardiovasc Qual Outcomes.* (2011) 4:14–21. doi: 10.1161/CIRCOUTCOMES.110.958108
31. Hijazi Z, Oldgren J, Lindbäck J, Alexander JH, Connolly SJ, Eikelboom JW, et al. The novel biomarker-based ABC (age, biomarkers, clinical history)-bleeding risk score for patients with atrial fibrillation: a derivation and validation study. *Lancet.* (2016) 387:2302–11. doi: 10.1016/S0140-6736(16)00741-8

Conflict of Interest: ZH reports personal fees from Boehringer Ingelheim, Bristol-Myers Squibb, Pfizer and Roche Diagnostics for lectures, personal fees from Boehringer Ingelheim, Bristol-Myers Squibb, Pfizer, and Roche Diagnostics for consulting, and grants from the Swedish Society for Medical Research (S17-0133) and the Swedish Heart-Lung Foundation (20170718), outside the submitted work.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Niederdöckl, Oppenauer, Schnaubelt, Cacioppo, Buchtele, Warenits, Laggner, Schütz, Bögl, Ruzicka, Gupta, Lutnik, Sheikh Rezaei, Wolzt, Herkner, Domanovits, Laggner, Schwameis and Hijazi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Medicine

TYPE Original Research
PUBLISHED 24 August 2022
DOI 10.3389/fmed.2022.834497

Check for updates

OPEN ACCESS

EDITED BY
Kevin Lu,
University of South Carolina,
United States

REVIEWED BY
Yu Ching Lin,
National Cheng Kung University,
Taiwan
Wen Chung Tsai,
Taoyuan Chang Gung Memorial
Hospital, Taiwan
Robet L. Drury,
ReThink Health, United States

*CORRESPONDENCE
Szu-Fu Chen
szufuchen@yahoo.com.tw
Fu-Cheng Wang
fcw@ntu.edu.tw

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Family Medicine and Primary Care, a section of the journal Frontiers in Medicine

RECEIVED 13 December 2021
ACCEPTED 28 July 2022
PUBLISHED 24 August 2022

CITATION
Liu Y-T, Lin A-C, Chen S-F, Shih C-J, Kuo T-Y, Wang F-C, Lee P-H and Lee AP (2022) Superior gait performance and balance ability in Latin dancers.
Front. Med. 9:834497.
doi: 10.3389/fmed.2022.834497

COPYRIGHT
© 2022 Liu, Lin, Chen, Shih, Kuo, Wang, Lee and Lee. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Superior gait performance and balance ability in Latin dancers

Yen-Ting Liu[1], Ang-Chieh Lin[1], Szu-Fu Chen[1,2*†], Chih-Jen Shih[3], Tien-Yun Kuo[3], Fu-Cheng Wang[3*†], Pei-Hsin Lee[4] and Adeline Peiling Lee[5]

[1]Department of Physical Medicine and Rehabilitation, Cheng Hsin General Hospital, Taipei, Taiwan, [2]Department of Physiology and Biophysics, National Defense Medical Center, Taipei, Taiwan, [3]Department of Mechanical Engineering, National Taiwan University, Taipei, Taiwan, [4]Power and Health Physical Medicine and Rehabilitation Center, Taipei, Taiwan, [5]Taipei Love Dance Company, Taipei, Taiwan

Background: Latin dance consists of various fast and stability-challenging movements that require constant body adjustments to maintain proper posture and balance. Although human gaits are assumed to be symmetrical, several factors can contribute to asymmetrical behavior of the lower extremities in healthy adults. These include lower limb dominance, ground reaction forces, lower limb muscle power, foot placement angle, and range of joint motion. Gait impairment can lead to a high risk of falling, diminished mobility, and even cognition impairment. We hypothesized that Latin dancers might have a more symmetric gait pattern and better balance ability than healthy non-dancer controls.

Methods: We investigated the impact of Latin dance training on gait behaviors and body balance. We recruited twenty Latin dancers and 22 normal healthy subjects to conduct walking experiments and one-leg stance tests, and we measured their kinematic data by inertial measurement units. We then defined four performance indexes to assess gait performance and body stability to quantify the potential advantages of dance training.

Results: We found that the two gait asymmetric indexes during the walking test and the two performance indexes during the one-leg stance tests were better in Latin dancers compared with the healthy control group. The results confirmed the superiority of Latin dancers over the healthy control group in gait symmetry and balance stability. Our results suggest that Latin dancing training could effectively strengthen lower limb muscles and core muscle groups, thereby improving coordination and enhancing gait performance and balance.

Conclusion: Latin dance training can benefit gait performance and body balance. Further studies are needed to investigate the effect of Latin dance training on gait and balance outcomes in healthy subjects and patients with gait disorders.

KEYWORDS

gait, performance, symmetry, balance, stability, Latin dance, IMU, muscle

Introduction

The maintenance of stable movements involves a complex interaction of the nervous, musculoskeletal, and cardiorespiratory systems (1). For this reason, gait can be a sensitive indicator of overall health and functional status, even in healthy adults without underlying pathological conditions (2). For example, walking speed has even been recommended as the "6th vital sign" (3). Gait impairment can lead to a high risk of falling (4), diminished mobility (5), and even cognition impairment. More than 30% of people over 65 years of age fall at least once a year because of loss of balance and muscle weakness (6). Thus, gait impairment may greatly affect the quality of life, functional independence, and health status (7). The observation of gait performance may therefore provide information on many factors, such as musculoskeletal condition, neurological status, dynamic balance, and postural control ability (8).

The mode and efficiency of a person's gait depend on core and lower limb muscle strength, coordination, and free joint mobility (9, 10). Accumulating evidence suggests that exercise interventions, such as muscle strength, endurance, and balance training, can improve gait performance in healthy adults, especially in elderly subjects (11). Exercise can also help to reduce the rate of falls by approximately 25% and the number of older adults who experience falls by 15% (6). Dance is an easily accessible exercise that combines strength and coordination practice and provides high levels of enjoyment, thereby increasing exercise program adherence (12). Dance-based exercises can improve balance capability and gait performance in both young and old subjects (13, 14).

Latin dance is a type of ballroom dancing enjoyed by a part of the general population for its physical and psychological benefits. Latin dance consists of a complex dance sequence involving three-dimensional motions. Technical requirements include proper body posture, alignment, and rhythm which require an effort to maintain static and dynamic balance. The dance movement involves multiple segments of the body, with continuous alternations between one-leg and double-leg stance phases, such as hopping, sideways steps, or crossing one foot behind the other, thus requiring strong postural control. As a result, dancers become skilled at weight shifts that constantly challenge their postural control systems to maintain balance. This complex balance training enhances muscular strength, endurance, and coordination. It may also be a beneficial activity for gait stability since Latin dancing movements involve a complex motion of core muscles and lower limb muscles. The limited literature that has examined the effects of Latin dance indicates that Latin dance practice can improve mobility, strength, balance, and gait stability in older adults (15–17). However, information regarding the actual quantitative gait performance and balance ability of Latin dancers is lacking.

Human gait is a symmetrical and rhythmic periodic motion. A gait cycle typically consists of about 60% stance phase and 40% swing phase, and both legs are usually symmetrical in healthy persons. Therefore, gait performance can be assessed by the symmetry of leg movements. Many methods are available for obtaining gait information. For instance, Knutsson and Richards (18) applied electromyogram signals to identify muscle activation patterns. Optical motion capture systems, such as VZ4000 (19) and VICON (20), used high-speed cameras to record the positions of the markers attached to subjects. For example, Wang et al. (21) applied VZ4000 to detect stroke patients' motions and automatically repeat the therapists' interventive rehabilitation. Much research applied force plates to obtain gait information. For instance, Wong et al. (22) implemented load sensors to analyze foot contact patterns and evaluate walking abilities. Daliri (23) and El Maachi et al. (24) used the ground reaction forces underneath both feet to detect subjects with Parkinson's disease (PD). However, the optical systems and the force plates can be costly, and their measurements are restricted to particular locations. In addition, the optical systems need to process markers' position data into corresponding velocity and acceleration data. Therefore, Bilro et al. (25) used wearable optical sensors to monitor their subjects' gaits. Díaz et al. (26) discussed the use of wearable sensor technologies in gait analysis and motion research. In the present paper, we applied inertial measurement units (IMUs) to measure kinematic data and to evaluate the influences of Latin dance training on gait behaviors and balance performance.

We recruited Latin dancers and age-matched healthy subjects as a non-dancer control group to conduct experiments. During the walking experiments, we attached two IMUs to these subjects' limbs to obtain their gait information. We also conducted the one-leg stance tests and applied an IMU to the subjects' waists to measure their body movements. The recorded IMU data were then used to evaluate the subjects' gait performance and balance ability. The results showed that Latin dancers tended to have better gait performance and body stability compared to the normal control group.

Materials and methods

Experimental design and data collection

This study was approved by the Institutional Review Board of Cheng Hsin General Hospital, Taiwan [(876)110-22]. All gait behaviors and balance performance data were collected from participants who provided written informed consent.

We recruited 20 Latin dancers from a Latin dance studio and 22 age-matched healthy subjects as a non-dancer control group to conduct our experiments. Inclusion criteria of the dancers were the following: (a) overall good health; (b) between the ages of 20 and 60 years old; (c) practicing 3–6 days per week and at least 5 years of experience in Latin dance; (d) no

10.3389/fmed.2022.834497

limb or leg discrepancy; (e) no history of surgery on the lower limbs or spine; and (f) no history of musculoskeletal injury over the lower back in the past 6 months. The non-dancer control subjects were 22 age- and gender-matched healthy participants, who had no experience with dance training and were free from injury (no musculoskeletal injuries or neurological conditions). The non-dancer control subjects participated in regular exercise at a frequency of 2 or 3 times a week. Our subjects' data are illustrated in Table 1. These subjects took walking and stability tests, and we measured their kinematic data to compare the gait performance and body balance between Latin dancers and non-dancer control subjects. For the walking experiments, we

TABLE 1 Basic data of the subjects.

Subject	Age	Gender	Height (cm)	Weight (kg)	BMI	Dancing experience (year)
D1	21	F	161	47	18.1	11
D2	35	F	166	51	18.5	10
D3	36	M	172	68	23	25
D4	29	F	161	51	19.7	6
D5	33	M	176	75	24.2	9
D6	28	F	161	50	19.3	10
D7	38	M	173	69	23.1	18
D8	25	F	162	53	20.2	10
D9	29	M	177	70	22.3	13
D10	27	M	173	70	23.4	14
D11	38	M	196	68	17.7	18
D12	30	M	176	71	22.9	19
D13	34	F	162	50	19.1	12
D14	26	M	175	60	19.6	10
D15	31	M	176	78	25.2	10
D16	35	F	158	48	19.2	9
D17	32	M	178	74	23.4	6
D18	42	F	162	51	19.4	21
D19	26	M	168	77	27.3	15
D20	27	F	156	49	20.1	14
						Exercise experience
N1	35	F	182	95	28.7	Basketball, weight training
N2	33	F	163	75	28.2	Aerobics
N3	39	M	167	68	24.4	Walking
N4	40	F	151	58	25.4	Jogging, weight training
N5	25	F	154	48	20.2	Walking
N6	23	M	165	60	22.0	Football, weight training
N7	44	F	162	58	22.1	Pilates
N8	25	M	177	95	30.3	Jogging, yoga
N9	28	M	172	78	26.4	Walking
N10	29	M	171	67	22.9	Walking
N11	29	F	160	53	20.7	Walking, weight training
N12	31	M	174	75	24.8	Basketball, weight training
N13	28	M	170	58	20.0	Basketball, jogging
N14	34	F	156	70	28.8	Waking, mountain climbing
N15	28	F	168	46	16.3	Walking
N16	38	M	186	115	33.2	Walking
N17	38	M	164	75	27.9	Walking
N18	40	M	173	67	22.4	Basketball, mountain biking
N19	38	M	175	64	20.9	Jogging, walking
N20	33	F	160	64	25.0	Jogging, walking
N21	36	F	150	60	26.6	Jogging, walking
N22	38	F	161	59	22.7	Jogging, walking

attached two IMUs to these subjects' limbs to obtain their gait information. We also conducted the one-leg stance tests and applied an IMU to the subjects' waists to measure their body movements. The recorded IMU data were then used to evaluate the subjects' gait performance and balance ability.

Inertial measurement unit system

We applied the OPAL IMU system (27) with wearable IMUs to measure the kinematic data of these subjects during the experiments. The specifications of the OPAL IMU system are illustrated in Table 2. Two IMUs were attached to the subjects' lower limbs and one IMU was attached to the subjects' waist, as shown in Figure 1, to record the subjects' kinematic data with a sampling rate of 128 Hz.

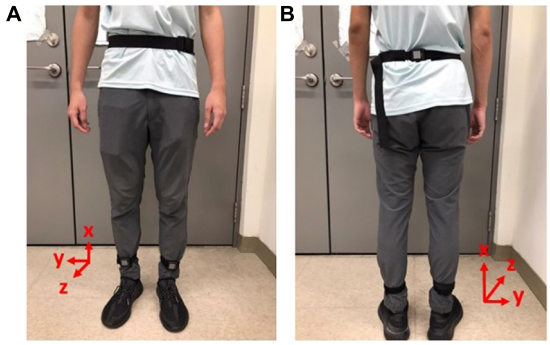

FIGURE 1
Gait measurements by IMUs. Two IMUs were attached to the subjects' lower limbs and one IMU was attached to the subjects' waist to record the subjects' kinematic data with a sampling rate of 128 Hz. **(A)** Front view. **(B)** Rear view.

Evaluation of gait performance

All subjects were required to walk in a straight line at their most comfortable pace. We applied two IMUs to measure the subjects' angular velocities of the shanks on the sagittal plane (28) (i.e., the y-axis in Figure 1) to investigate the gait performance. For example, the angular velocities ω_y of D3 and N9 on the sagittal plane are shown in Figure 2, where the important gait events are marked. All subjects' angular velocities are shown in Appendix 1.

We can further divide the gait responses into individual gait cycles, because human gaits are regular and periodic. A complete gait cycle is composed of a stance phase and a swing phase, containing three important gait events in one gait cycle: the mid-swing (MS), the heel strike (HS), and the toe-off (TO), as shown in Figure 3 (29). The MS occurs when the angular velocity of the shank on the sagittal plane reaches the maximum in the gait cycle; the HS event occurs when the heel touches the ground, where the angular velocity has the first negative trough after the mid-swing; and the TO occurs when the toes leave the ground (29). The stance phase is defined as the time interval from the HS to the TO and accounts for about 60% of the gait cycle. The swing phase is defined as the time interval from the TO to the next HS and takes about 40% of the gait cycle. Because gaits are periodic, a single gait cycle can be measured from any gait event to the next same event on the same foot. In this paper, we applied the IMU to measure gait information and used the angular velocity of the shanks on the sagittal plane to recognize these gait events (28, 30, 31). Because subjects might have varying walking speeds, we marked the HS events on the gait responses and divided the measured IMU data into individual gait cycles. All subjects' gait cycles are illustrated in Appendix 2.

TABLE 2 Specifications of the OPAL IMU system (27).

Dimensions	**43.7 × 39.7 × 13.7 mm**
Weight	< 25 g (with battery)
Resolutions	17.5 bits
Sampling rates	20–128 Hz
Transmission range	30 m line of sight
Ranges of the accelerometer	± 200 g
Ranges of the gyroscope	± 2,000 deg/s
Ranges of the magnetometer	± 8 Gauss

Differentiating the gait performance between dancers and the healthy controls from their gait cycles is difficult because of their similar appearance. Therefore, we proposed two indexes to quantify the gait performance of these subjects and to discuss the benefits of Latin dance training in improving gait performance. Healthy persons tend to have regular and periodic gaits, so the swing times of their two legs should be similar. Suppose the time interval of one complete gait cycle (from HS to HS) is represented as T_{gait} and the swing time (from TO to HS) of that gait cycle is T_{SW}. The swing phase is then defined as the proportion of the swing time in the complete gait cycle, as in the following:

$$P_{SW} = \frac{T_{SW}}{T_{gait}}.$$

Therefore, we can define the asymmetry of the swing phases as follows: (29, 32)

$$Asym_{SP} = \frac{\left|(P_{SW})_R - (P_{SW})_L\right|}{\min\left\{(P_{SW})_R\ ,\ (P_{SW})_L\right\}} \times 100\%,$$

where $(P_{SW})_R$ and $(P_{SW})_L$ represent the swing phase on the right foot and the left foot, respectively. Gait symmetry has been assumed in healthy persons, and the swing time on both sides takes about 40% of the complete gaits. Therefore, the

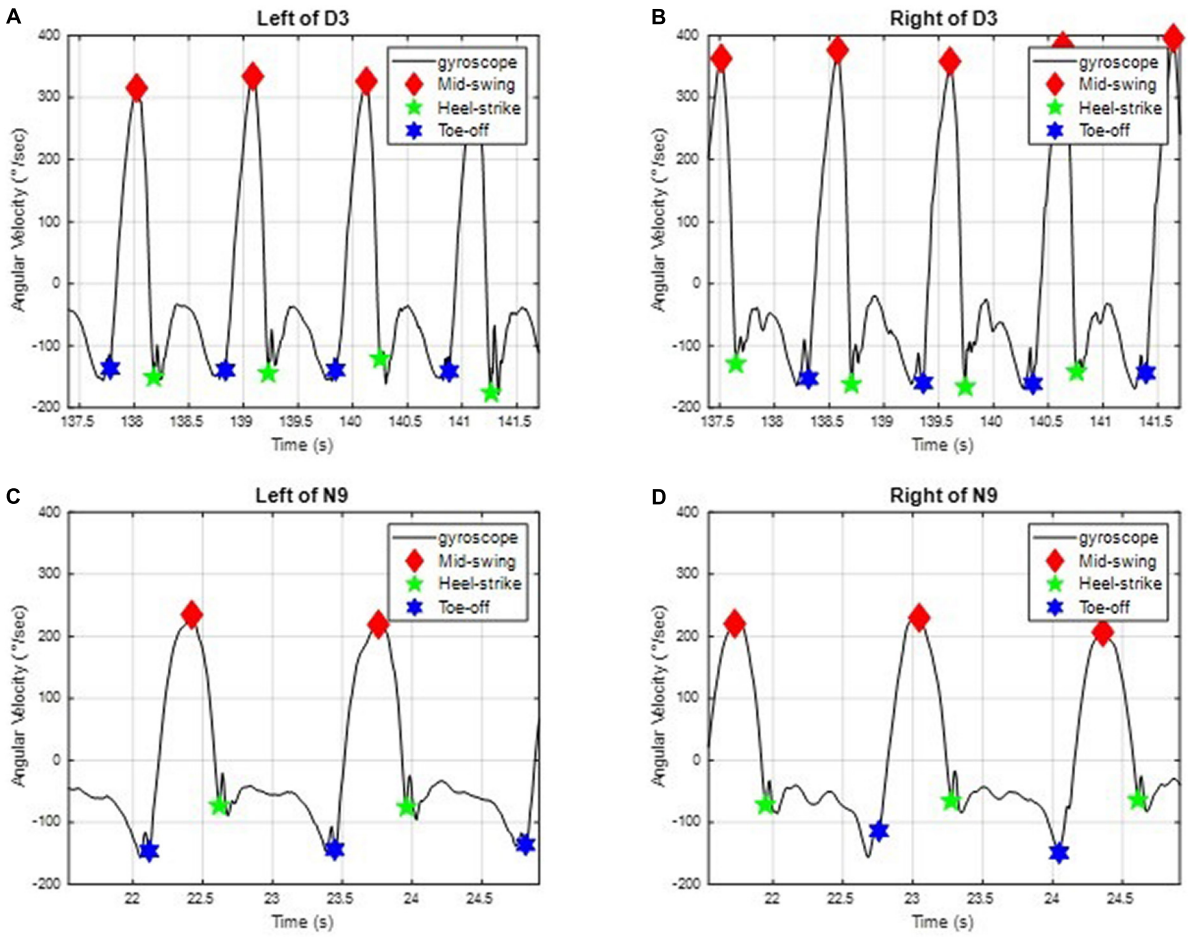

FIGURE 2
The angular velocities of D3 and N9 of the shank on the sagittal plane. The important gait events, such as the mid-swing, the heel strike, and the toe-off, are marked. **(A)** On the left leg of D3. **(B)** On the right leg of D3. **(C)** On the left leg of N9. **(D)** On the right leg of N9.

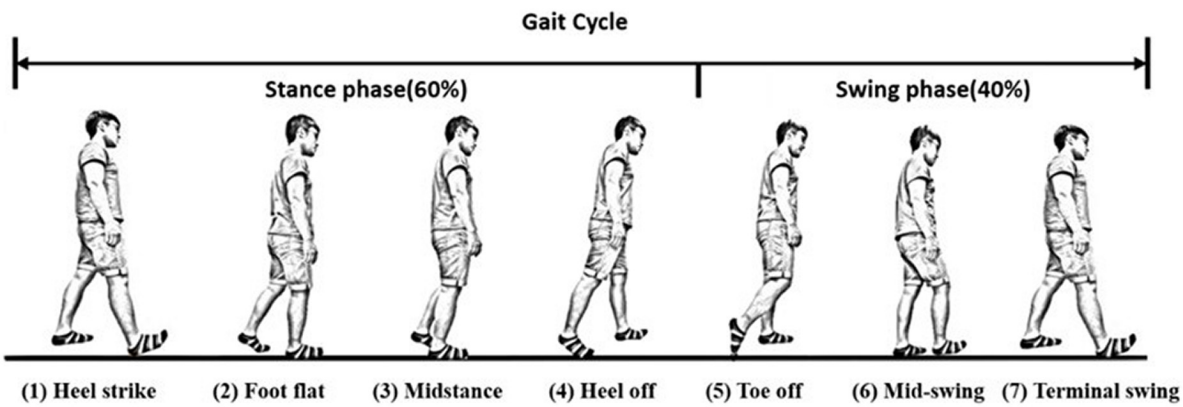

FIGURE 3
A complete gait cycle. A complete gait cycle is composed of a stance phase and a swing phase, containing three important gait events in one gait cycle: the mid-swing (MS), the heel strike (HS), and the toe-off (TO). The stance phase is defined as the time interval from the HS to the TO and accounts for about 60% of the gait cycle. The swing phase takes about 40% of the gait cycle and is defined as the time interval from the TO to the next HS.

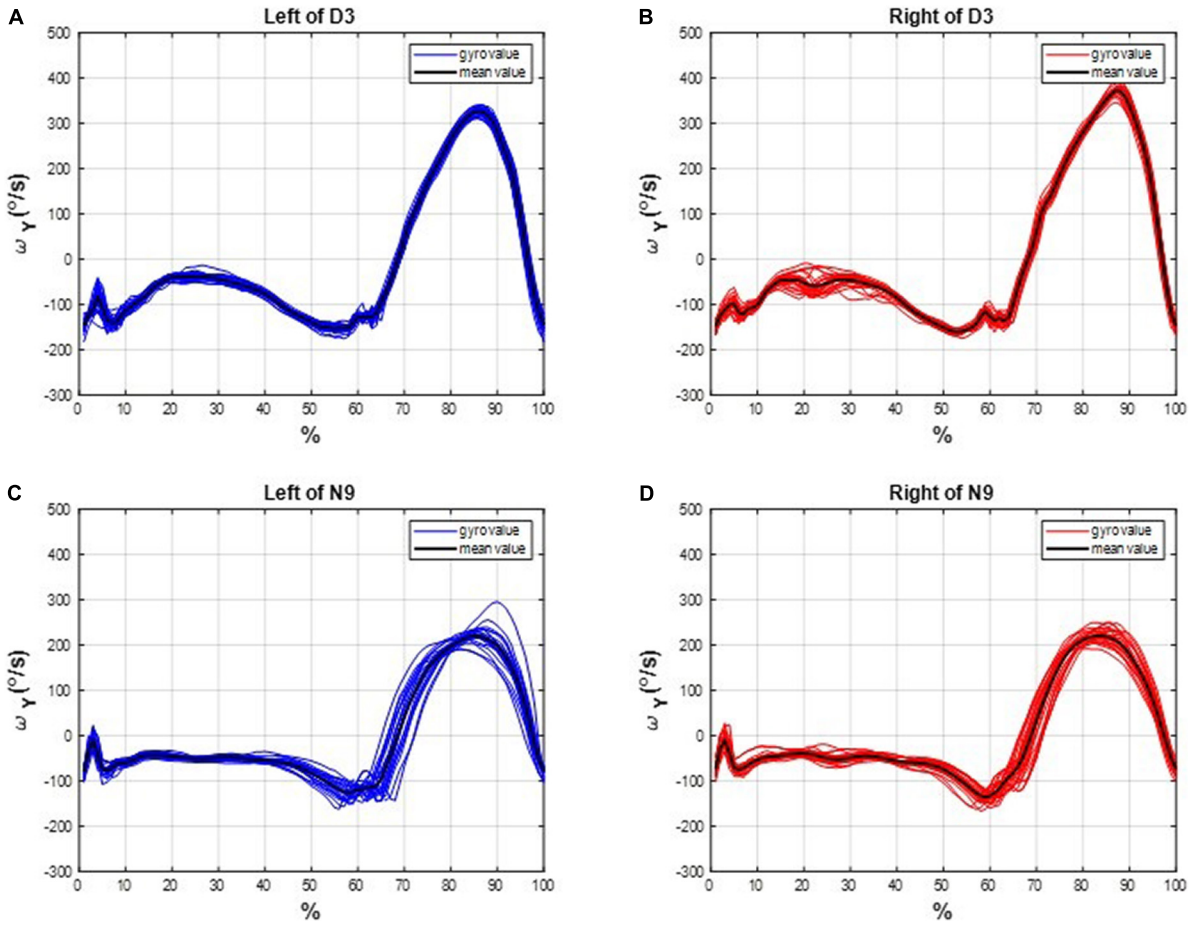

FIGURE 4
The gait cycles of D3 and N9. The dancer tended to have more uniform gait cycles than the non-dancer control, which indicates a superiority of dancing training on gait performance. **(A)** On the left leg of D3. **(B)** On the right leg of D3. **(C)** On the left leg of N9. **(D)** On the right leg of N9.

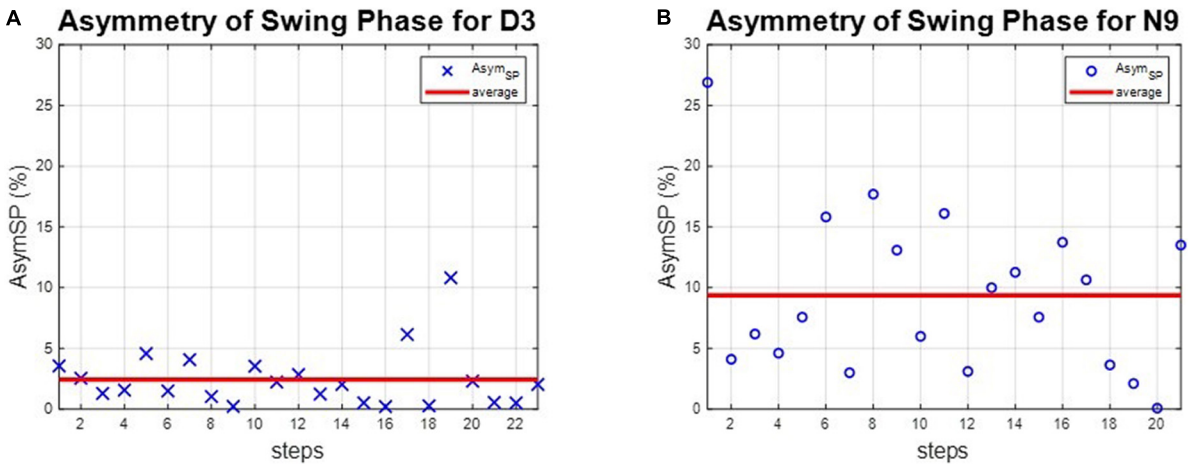

FIGURE 5
The asymmetry of the swing phases of D3 and N9. The dancer tended to have smaller $Asym_{SP}$ than the non-dancer control, because of dance training. **(A)** The dancer D3. **(B)** The non-dancer control N9.

ideal $Asym_{SP}$ should be zero. However, gait asymmetry has been reported in healthy adults. For instance, asymmetrical behavior of the lower extremities was observed in spatio-temporal and kinematic parameters, such as lower limb dominance, ground reaction forces (33), electromyographic activities (34), lower limb muscle power (35), stride length (36), foot placement angle (37), and range of joint motion (38) in people without impairments. Clinically, gait asymmetry may be associated with a number of negative consequences, such as inefficient walking, impaired balance control, and increased risk of musculoskeletal injury to the dominant lower limb (39, 40). Apart from pathological conditions, such as stroke, amputation, or cerebral palsy, gait asymmetry is often observed in older adults and is a risk factor for falls (41). Hence, we can use $Asym_{SP}$ to evaluate the gait performance of our dancers and non-dancer subjects. All subjects' asymmetries of swing phases $Asym_{SP}$ are illustrated in **Appendix 3**.

We can also define the ratio of swing time, denoted as R_{SW}, as follows (29, 42):

$$R_{SW} = \frac{\max\left\{(T_{SW})_R\ ,\ (T_{SW})_L\right\}}{\min\left\{(T_{SW})_R\ ,\ (T_{SW})_L\right\}} \times 100\%,$$

where $(T_{SW})_R$ and $(T_{SW})_L$ represent the swing time on the right side and the left side, respectively. Gait in healthy adults has been assumed to be symmetric, so the swing time on both legs should be similar. Thus, the ideal value of R_{SW} should be 100% (43). Nevertheless, previous work has reported differences in gait parameters between the limbs (33–38). Therefore, we applied R_{SW} to evaluate the gait performance of these subjects. All subjects' ratios of swing time are illustrated in **Appendix 4**.

Because the ideal value of $Asym_{SP}$ is zero and the ideal value R_{SW} is 100%, we were able to further quantify the gait performance by the following indexes:

$$J_1 = \left(\frac{1}{N}\sum_{i=1}^{N}\left|Asym_{SP}(i)\right|^2\right)^{1/2},$$

$$J_2 = \left(\frac{1}{N}\sum_{i=1}^{N}\left|R_{SW}(i) - 100\right|^2\right)^{1/2},$$

where N is the number of gait cycles, and i represents the i-th gait cycle.

Evaluation of balance ability

To investigate the balance ability, all subjects were required to conduct the following one-leg stance tests:

(a) standing on the dominant foot with eyes open;
(b) standing on the dominant foot with eyes closed;
(c) standing on the non-dominant foot with eyes open;
(d) standing on the non-dominant foot with eyes closed.

The subjects were asked to stand on one foot for about 50 s at each action if possible, but they were free to stand on both feet if body balance could not be maintained. They could take a rest of about 20 s between each action.

We implemented one IMU on the subjects' waist to record their angular velocities and linear accelerations during the actions. When we stand on one foot, our bodies usually need to adjust position to maintain balance, like an inverted pendulum.

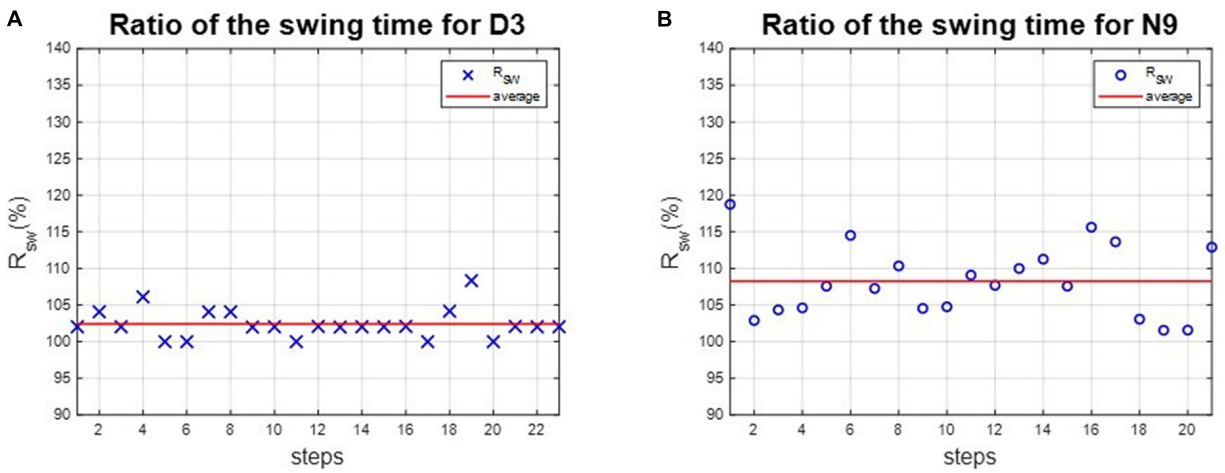

FIGURE 6
The ratio of swing time of D3 and N9. The dancer tended to have better R_{SW} than the non-dance control, because of dance training. **(A)** The dancer D3. **(B)** The non-dancer control N9.

Therefore, we were able to estimate the balance ability of these subjects from their kinematic data.

We defined the following two indexes to quantify the balance ability of the subjects:

$$J_3 = \frac{1}{N}\sum_{i=1}^{N}|\omega(i)| = \frac{1}{N}\sum_{i=1}^{N}\left(\omega_x^2(i) + \omega_y^2(i) + \omega_z^2(i)\right)^{1/2},$$

$$J_4 = \frac{1}{N}\sum_{i=1}^{N}|a(i)| = \frac{1}{N}\sum_{i=1}^{N}\left(a_x^2(i) + a_y^2(i) + a_z^2(i)\right)^{1/2},$$

where N is the total number of samples, $\omega_x(i)$, $\omega_y(i)$, and $\omega_z(i)$are the three-axial angular velocities, and $a_x(i)$, $a_y(i)$, and $a_z(i)$ are the three-axial linear accelerations measured from the IMU on the waist at the i-th sample. That is, $\omega(i)$ and a(i) represent the absolute angular velocity and acceleration of the waist at sample i. Because subjects might have different balance durations, we defined the average absolute angular velocities and the average linear accelerations to estimate their balance ability. The subject's balance ability is considered better if the average angular velocity and acceleration are smaller (i.e., the subject did not need much effort to maintain balance).

Results

Evaluation of gait performance

We marked the HS events on the subjects' gait responses and divided the measured IMU data into individual gait cycles. For example, the gait cycles of D3 and N9 are shown in Figure 4, where the dancer D3 tended to have more uniform gait cycles than the non-dancer control N9.

The asymmetries of the swing phases $Asym_{SP}$ and the ratio of swing time R_{SW} were applied to evaluate the gait performance of dancers and healthy controls. For instance, the asymmetries of the swing phases of D3 and N9 are shown in Figure 5, while the ratio of swing time of D3 and N9 are shown in Figure 6. From Figures 5, 6, the dancer D3 tended to have better $Asym_{SP}$ and R_{SW} than the non-dancer control N9. All subjects' asymmetries of swing phases $Asym_{SP}$ are illustrated in Appendix 3, while all subjects' ratios of swing time are illustrated in Appendix 4.

We calculated each subject's average asymmetry of the swing phase $Asym_{SP}$ and average ratio of swing time R_{SW} in each gait cycle, as illustrated in Table 3 and Figure 7.

First, the average $Asym_{SP}$ of the dancers ranged between 2.19% (D9) and 6.79% (D8), while the average $Asym_{SP}$ of the non-dancer controls ranged between 2.45% (N20) and 14.53 (N12). Second, the average R_{SW} of dancers ranged between 102.29% (D16) and 106.74% (D8), while the average R_{SW} of healthy controls ranged between 102.35% (N20) and 113.52% (N12). The dancers' average $Asym_{SP}$ was closer to zero compared to the non-dancer control subjects, and the dancers' average R_{SW} was closer to 100% compared to the non-dancer control subjects.

Table 3 also illustrates the two indexes, J_1 and J_2, to quantify the gait symmetry of the subjects. The results showed that, first,

TABLE 3 Average $Asym_{SP}$ data and average R_{SW} for the test subjects.

Subject	Average $Asym_{SP}$	Average R_{SW}	J_1	J_2
D1	5.05	105.39	6.11	6.62
D2	6.15	106.69	7.42	8.11
D3	2.43	102.41	3.37	3.12
D4	6.24	105.46	7.32	6.98
D5	2.73	103.00	3.53	3.65
D6	3.12	104.02	4.40	4.84
D7	3.20	103.70	4.01	4.26
D8	6.79	106.74	7.89	7.82
D9	2.19	102.31	2.99	3.00
D10	4.33	104.36	5.31	5.04
D11	4.85	104.81	5.27	5.81
D12	5.46	106.19	6.63	7.89
D13	4.11	105.45	5.83	7.26
D14	4.09	105.31	6.61	7.56
D15	3.29	103.24	4.00	3.87
D16	2.46	102.29	3.22	2.76
D17	5.43	104.32	6.22	5.25
D18	3.44	103.68	4.52	4.93
D19	3.71	103.40	5.02	4.88
D20	2.92	103.78	4.10	4.96
N1	8.46	110.33	10.65	12.19
N2	9.57	110.08	12.47	13.80
N3	6.15	106.30	7.96	7.72
N4	8.62	109.29	10.11	11.54
N5	11.14	111.81	13.07	13.50
N6	12.50	110.26	13.97	11.78
N7	3.44	102.80	4.63	3.14
N8	8.97	109.81	12.42	13.26
N9	9.37	108.27	11.32	9.53
N10	6.21	106.16	6.45	6.51
N11	12.01	110.54	13.31	12.68
N12	14.53	113.52	16.65	15.63
N13	5.78	105.08	7.16	6.52
N14	7.44	107.68	9.96	9.61
N15	6.11	108.11	8.11	10.21
N16	10.28	109.39	10.49	10.23
N17	4.82	104.61	7.91	8.28
N18	3.35	104.45	4.43	6.94
N19	4.57	104.64	6.68	6.60
N20	2.45	102.35	3.27	3.24
N21	8.20	107.70	9.07	9.25
N22	5.64	106.39	8.49	8.69

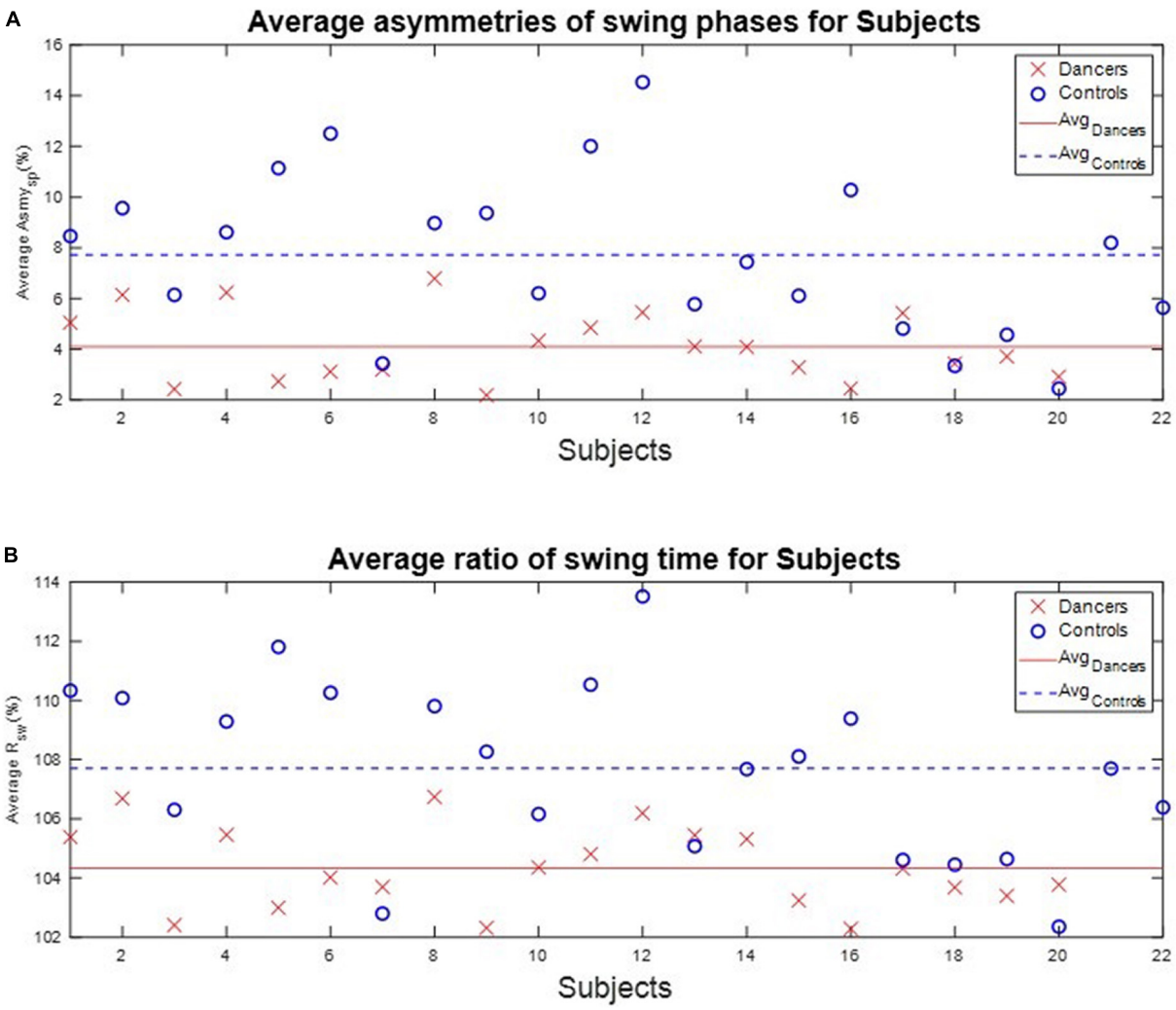

FIGURE 7
The average $Asym_{SP}$ and R_{SW} of subjects. The average $Asym_{SP}$ of the dancers ranged between 2.19% (D9) and 6.79% (D8), while the average $Asym_{SP}$ of the non-dancer controls ranged between 2.45% (N20) and 14.53 (N12). The average R_{SW} of dancers ranged between 102.29% (D16) and 106.74% (D8), while the average R_{SW} of healthy controls ranged between 102.35% (N20) and 113.52% (N12). The dancers tended to have better gait symmetry than the non-dancer controls. **(A)** The average asymmetry of the swing phases $Asym_{SP}$ of each subject. **(B)** The average ratio of swing time R_{SW} of each subject.

the average J_1 is 5.19%, with a standard deviation of 1.51%, for the dancers. By contrast, for the non-dancer control group, the average J_1 is 9.48% with a standard deviation of 3.39%. Second, the average J_2 is 5.43%, with a standard deviation of 1.74%, for the dancers, while it is 9.58%, with a standard deviation of 3.32%, for the non-dancer control group. That is, the dancers tended to have better gait symmetry than the control subjects.

Evaluation of balance ability

We implemented one IMU on the subjects' waist to record their angular velocities and linear accelerations to estimate the balance ability of these subjects from their kinematic data during the one-leg stance tests. For example, the three-axial angular velocities and three-axial linear accelerations of D3 and N9, when standing on the dominant foot with eyes open (action (a)), are shown in Figure 8. All subjects' IMU data during the one-leg stance tests are illustrated in Appendix 5. The dancers tended to make smaller movements than the control group to maintain body balance. The subject's balance ability is considered better if the angular velocity and acceleration are smaller (i.e., the subject did not need much effort to maintain balance). For example, Figure 9 shows the absolute angular velocity $\omega(t)$ and the absolute linear acceleration a(t) of D3 and N9 when standing on the dominant foot with eyes open [action (a)]. Apparently, the dancer D3 made much smaller body movements than the non-dancer control N9 to maintain stability.

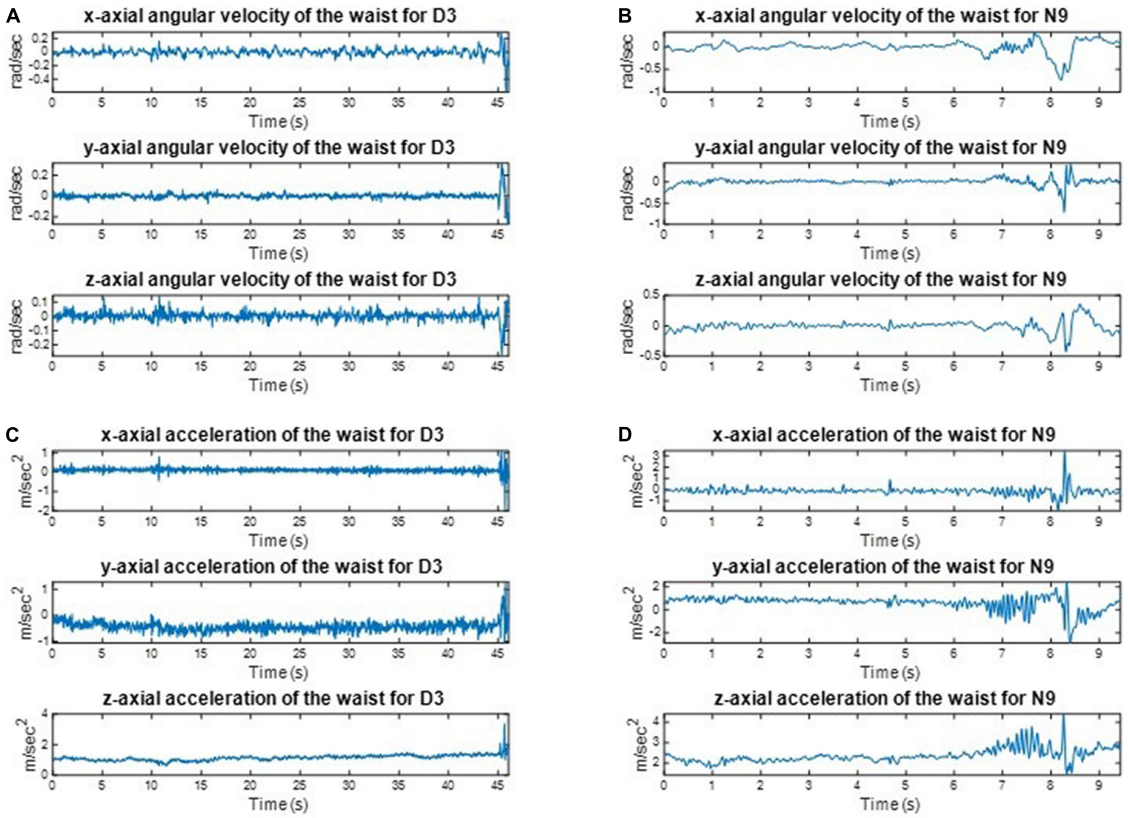

FIGURE 8
The recorded kinematic data of D3 and N9 when standing on the dominant foot with eyes open. The dancers tended to make smaller movements than the control group to maintain body balance. **(A)** The angular velocities of D3. **(B)** The angular velocities of N9. **(C)** The linear accelerations of D3. **(D)** The linear accelerations of N9.

All subjects' angular velocity and acceleration of the waist during the one-leg stance tests are illustrated in **Appendix 6**. The subjects' average J_3 and J_4 during the one-leg stance tests are shown in **Figures 10**, **11**, respectively, which show that the average J_3 and J_4 of the dancers were much lower than those of the normal controls. The statistical data of **Figures 10**, **11** are shown in **Table 4**. The dancers' angular velocity index J_3 were about 80.68, 58.52, 60.66, and 77.34%, respectively, of those of the healthy controls during the four actions. The dancers' linear acceleration index J_4 were about 79.39, 65.20, 80.18, and 77.61%, respectively, of the healthy controls during the four actions. That is, the dancer group needed less body adjustment to maintain postural stability.

Discussion

Evaluation of gait performance

In this paper, we assess gait performance by the symmetry of gaits. Gait symmetry has been assumed in healthy persons, and the swing time on both sides takes about 40% of the complete gaits. Therefore, the ideal asymmetry of the swing phase $Asym_{SP}$ should be zero and the ideal ratio of swing time R_{SW} should be 100%. However, as we described above, gait asymmetry has been reported in healthy adults (33–38). Clinically, gait asymmetry may be associated with a few negative consequences, such as inefficient walking, impaired balance control, and increased risk of musculoskeletal injury to the dominant lower limb (39, 40). Gait asymmetry is often observed in older adults and is a risk factor for falls (41). Hence, $Asym_{SP}$ and R_{SW} can be applied to evaluate the gait performance of the dancers and non-dancer subjects. The average $Asym_{SP}$ was 4.10% for dancers, but 7.71% for healthy controls, while the average R_{SW} was 104.33% for dancers, but 107.71% for healthy controls. Our results showed that the asymmetries of the swing phases of dancers are smaller than those of non-dancer controls. Latin dance consists of a complex dance sequence involving motion of core and hip muscles. During dancing, core muscles, including the transverse abdominis, multifidus, rectus abdominis, and erector spinae, work in coordination to maintain proper body posture and alignment. The unique hip motion, characterized by a rhythmic rotation of the hips around the spine, is essential for Latin

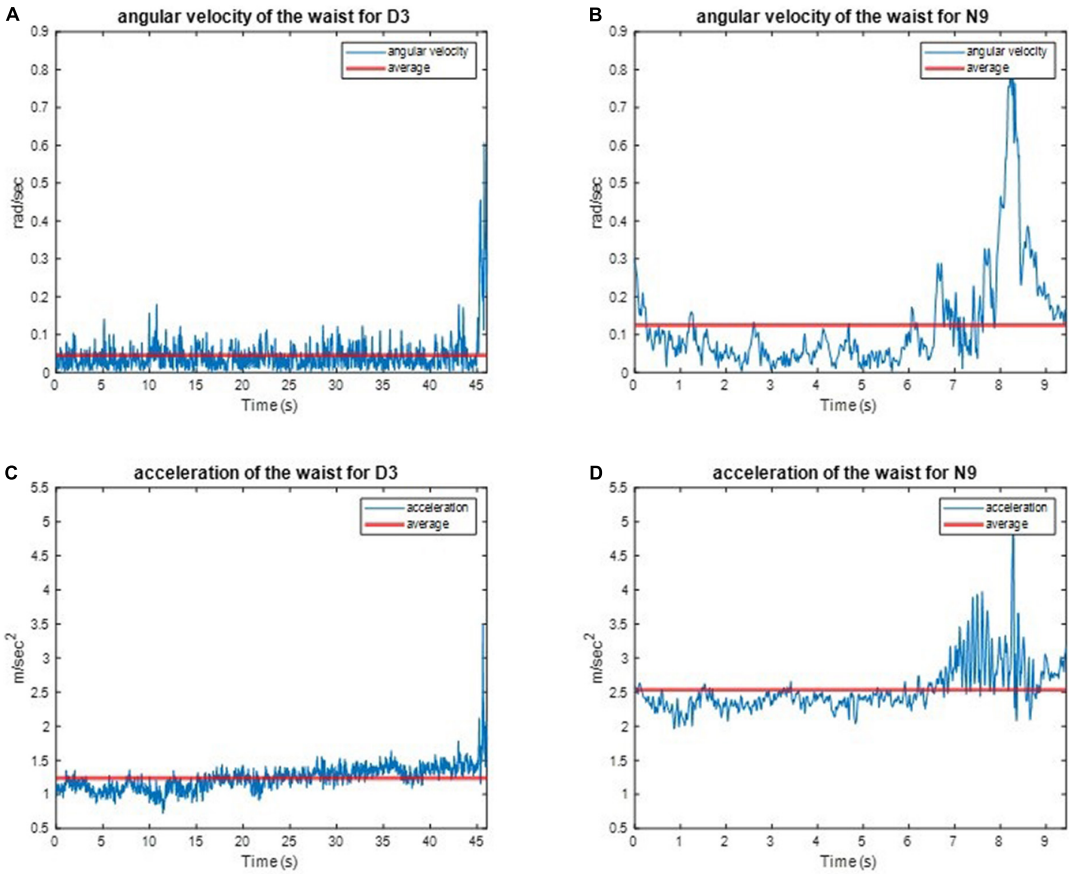

FIGURE 9
The absolute angular velocity and acceleration of D3 and N9 when standing on the dominant foot with eye open. The dancer made much smaller body movements than the non-dancer control maintain stability. **(A)** The absolute angular velocity of D3. **(B)** The absolute angular velocity of N9. **(C)** The absolute linear acceleration of D3. **(D)** The absolute linear acceleration of N9.

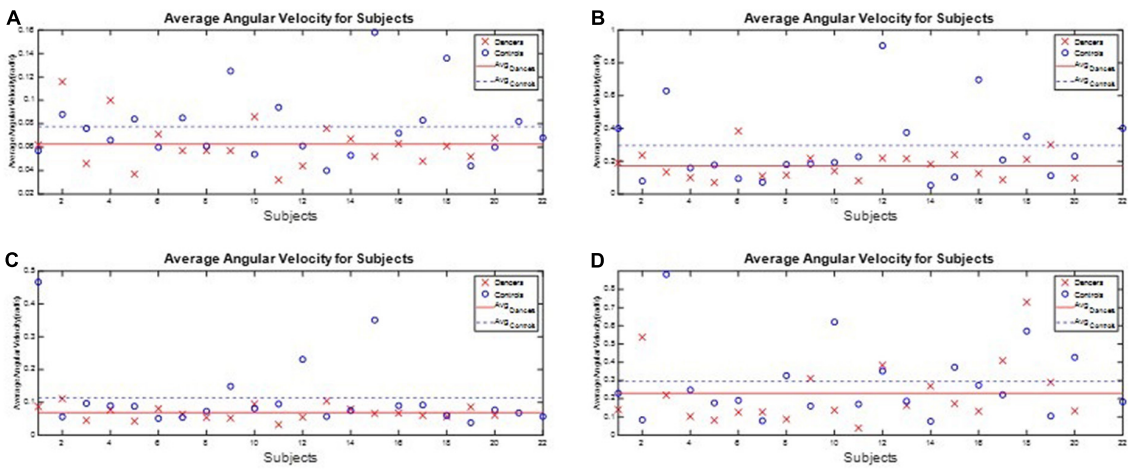

FIGURE 10
The performance index J_3 during the one-leg stance tests. The dancers' angular velocity index J_3 were about 80.68, 58.52, 60.66, and 77.34%, respectively, of those of the healthy controls during the four actions. **(A)** Standing on the dominant foot with eyes open. **(B)** Standing on the dominant foot with eyes closed. **(C)** Standing on the non-dominant foot with eyes open. **(D)** Standing on the non-dominant foot with eyes closed.

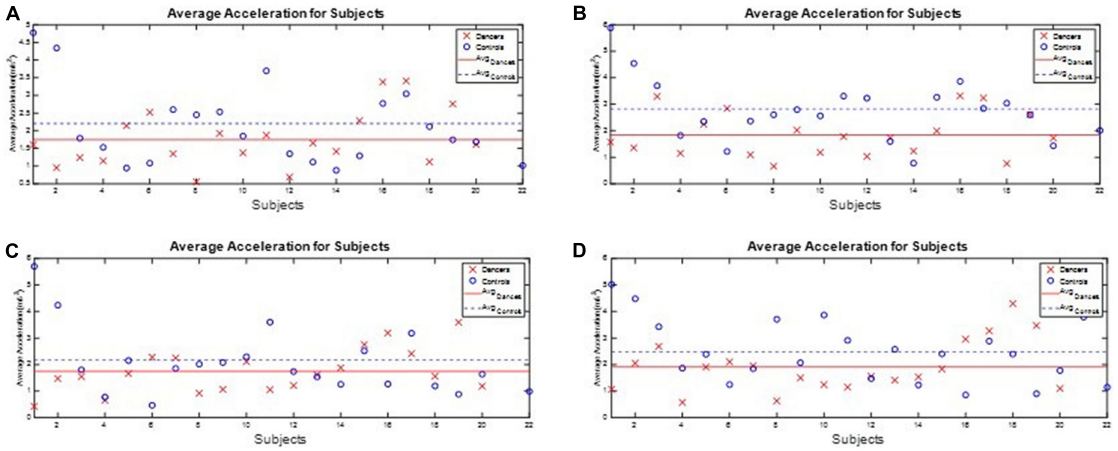

FIGURE 11
The performance index J_4 during the one-leg stance tests. The dancers' linear acceleration index J_4 were about 79.39, 65.20, 80.18, and 77.61%, respectively, of the healthy controls during the four actions. **(A)** Standing on the dominant foot with eyes open. **(B)** Standing on the dominant foot with eyes closed. **(C)** Standing on the non-dominant foot with eyes open. **(D)** Standing on the non-dominant foot with eyes closed.

dance, and the core and hip muscles work sequentially and symmetrically. In addition, as Latin dancing is predominantly based on walking actions, repeatedly performing dance steps works the major thigh muscles involved in walking, such as the quadriceps and hamstring muscles. Our data imply that the Latin dancing might improve gait symmetry through the enhancement of dynamic postural control and the strengthening of the core and lower limb muscles. Another possible reason for the Latin dancers' gait superiorities is the body awareness required in Latin dancing. Previous work has demonstrated that ballroom dancing can provide benefits to body perception (44). Latin dancing is a complex sensorimotor activity that integrates many skills, such as rhythm, synchronization movements with music, balance, coordination, and spatial sense. Dancers need to perceive their anatomical body space and the planes of movements, and then interact precisely *via* integration of tactile, proprioceptive, kinesthetic, and environmental information. Thus, Latin dancers may direct body awareness toward greater positional body symmetry during walking.

Among the non-dancer controls, N7, N18, and N20 had better gait symmetry compared with the others, possibly because of their regular exercise programs. For example, subject N7 has been practicing Pilates regularly for more than 10 years. Pilates is a combination of stretch and strengthening exercises performed in a neutral spine position (45). Through stretching and isometric, eccentric, and concentric muscle strengthening, Pilates improves balance skills, proprioception, coordination, and postural control. Therefore, Pilates exercises can also improve various motor abilities and are effective in enhancing postural stability, dynamic balance, flexibility, muscle strength, and walking capabilities (46, 47). Subject N18 is a mountain biker who has been riding mountain bikes regularly for more than 10 years. Mountain biking is a dynamic activity that requires participants to constantly change their body position to adapt to the bumpy road (48). Both the abdominal and core muscles play an important role in this form of exercise to maintain good control of the bike, so that mountain bikers have better balance and coordination skills compared with the general population. Subject N20 conducts regular walking and jogging, which are similar to the activities of the other control participants. The regimen of jogging and walking, genetic factors, or other environmental factors, such as nutrition, may possibly play a role in maintaining gait symmetry.

Evaluation of balance ability

The one-leg stance is a frequently used clinical tool for assessment of balance in persons with various balance disorders (49). The subject's balance ability is considered better if the angular velocity and acceleration are smaller (i.e., the subject did not need much effort to maintain balance). We found that the dancers' linear acceleration index J_4 were about 79.39, 65.20, 80.18, and 77.61%, respectively, of the healthy controls during the four actions. That is, the dancer group needed less body adjustment to maintain balance. Our results showed that Latin dancers had smaller trunk movements while conducting the one-leg stance test, indicating that Latin dancers have better postural steadiness compared with the non-dancer controls. As mentioned above, Latin dancers tend to have stronger core muscle strength, which plays a crucial role in maintaining both posture and trunk stability. In contrast to gait asymmetry, the one-leg stance is an easy clinical balance test that can assess a person's postural steadiness in a relatively static condition. Our finding of improved leg stance in Latin dancers is important clinically, since poor one-leg balance is a significant predictor of

TABLE 4 The performance indexes J_3 and J_4 of the subjects during the one-leg stance tests.

	J_3				J_4			
	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(d)
Dancers								
D1	0.062	0.190	0.087	0.140	1.603	1.569	0.422	1.074
D2	0.116	0.237	0.111	0.538	0.954	1.354	1.472	2.052
D3	0.046	0.134	0.045	0.220	1.240	3.292	1.546	2.685
D4	0.100	0.100	0.077	0.103	1.146	1.146	0.648	0.573
D5	0.037	0.070	0.043	0.082	2.145	2.244	1.660	1.904
D6	0.071	0.384	0.081	0.126	2.522	2.846	2.277	2.100
D7	0.057	0.110	0.064	0.127	1.350	1.093	2.247	1.949
D8	0.057	0.115	0.054	0.087	0.554	0.663	0.916	0.625
D9	0.057	0.218	0.052	0.311	1.924	2.023	1.063	1.499
D10	0.086	0.141	0.095	0.137	1.373	1.178	2.109	1.237
D11	0.032	0.081	0.032	0.039	1.874	1.781	1.049	1.153
D12	0.044	0.219	0.055	0.384	0.690	1.031	1.205	1.554
D13	0.076	0.216	0.104	0.163	1.650	1.732	1.614	1.409
D14	0.067	0.182	0.080	0.270	1.416	1.234	1.880	1.531
D15	0.052	0.240	0.066	0.173	2.283	1.993	2.752	1.833
D16	0.063	0.125	0.067	0.131	3.382	3.315	3.192	2.959
D17	0.048	0.087	0.060	0.409	3.409	3.239	2.412	3.270
D18	0.061	0.211	0.057	0.730	1.121	0.762	1.560	4.301
D19	0.052	0.301	0.086	0.290	2.758	2.609	3.596	3.473
D20	0.068	0.098	0.061	0.133	1.611	1.723	1.183	1.097
Average	0.063	0.173	0.069	0.230	1.750	1.841	1.740	1.913
Healthy controls								
N1	0.057	0.400	0.467	0.230	4.773	5.881	5.704	5.023
N2	0.088	0.079	0.056	0.084	4.341	4.543	4.239	4.485
N3	0.076	0.629	0.097	0.882	1.787	3.699	1.800	3.426
N4	0.066	0.159	0.090	0.248	1.529	1.820	0.774	1.864
N5	0.084	0.177	0.088	0.177	0.945	2.348	2.151	2.389
N6	0.060	0.094	0.051	0.191	1.081	1.217	0.462	1.242
N7	0.085	0.072	0.054	0.079	2.598	2.364	1.852	1.847
N8	0.061	0.180	0.073	0.327	2.451	2.604	2.013	3.708
N9	0.125	0.183	0.149	0.160	2.534	2.794	2.076	2.069
N10	0.054	0.193	0.081	0.621	1.849	2.563	2.294	3.869
N11	0.094	0.227	0.095	0.170	3.695	3.306	3.601	2.913
N12	0.061	0.905	0.231	0.352	1.348	3.231	1.736	1.473
N13	0.040	0.375	0.057	0.187	1.115	1.595	1.537	2.576
N14	0.053	0.053	0.075	0.076	0.881	0.775	1.253	1.229
N15	0.158	0.103	0.351	0.373	1.291	3.264	2.516	2.402
N16	0.072	0.697	0.090	0.274	2.773	3.865	1.266	0.850
N17	0.083	0.208	0.092	0.222	3.046	2.840	3.181	2.885
N18	0.136	0.352	0.060	0.571	2.119	3.039	1.191	2.396
N19	0.044	0.112	0.038	0.105	1.745	2.606	0.879	0.897
N20	0.060	0.230	0.077	0.427	1.687	1.433	1.637	1.775
N21	0.082	0.673	0.068	0.594	3.898	4.331	4.600	3.793
N22	0.068	0.401	0.057	0.183	1.015	2.009	0.986	1.145
Average	0.078	0.296	0.114	0.297	2.204	2.824	2.170	2.466

injurious falls, especially in the elderly population (50). Vision also plays a crucial role in processing and integrating other sensory information involving postural control and balance (51). We avoided the impact of visual feedback by conducting the one-leg stance test with eyes both closed and open. The dancers had significantly better balance with one-leg standing with eyes open, as a less challenging level, but they also had better balance with eyes closed, as a more challenging level, through improved proprioception. That is, the Latin dancers could still maintain better postural stability without visual dependence. Our findings suggest that Latin dancing is a promising exercise training for fall prevention.

However, this study has some limitations. One was that only a small number of Latin dancers and non-dancer healthy participants were recruited, so the data may not be generalizable to Latin dancers at large. Another was that our study design was a cross-sectional observational study, so further interventional research is necessary to clarify the efficacy of Latin dance training. A further limitation is that although we recruited non-dancer controls who were physically active and exercised regularly, the physical activity was reported by questionnaires and the exercise types differed significantly among participants. Further studies should focus on participants who engage in the same sports to provide a better comparison with Latin dancers. In the future, we plan to collect more kinematic data and to consider other performance indexes to investigate the potential advantages of Latin dance training in improving postural stability, balance, and gait performance in patients with neurological or orthopedic disorders.

Conclusion

This paper investigated the potential benefits of Latin dance training on gait behaviors and balance ability. Several factors, such as lower limb dominance, ground reaction forces, lower limb muscle power, foot placement angle, and range of joint motion, can contribute to asymmetrical behavior of the lower extremities in people without impairments. Because Latin dancing movements involve a complex motion of core muscles and lower limb muscles, this type of exercise may be a beneficial activity for improved gait stability. In this paper, we recruited twenty dancers and twenty-two age-matched normal healthy subjects to conduct walking experiments and one-leg stance tests. Their gait responses and body movements were recorded by IMUs and then used to evaluate gait performance and body stability. We defined four performance indexes to quantify gait performance and balance ability. The results showed that Latin dancers had better gait symmetry and body balance than the healthy control group, suggesting that gait performance and stability might be improved by Latin dance training. The possible contributing factors are enhancing dynamic postural control, strengthening the core and lower limb muscles, and better body awareness. Further studies are needed to investigate the effect of Latin dance training on gait and balance outcomes in healthy subjects and patients with gait disorders. This paper applied the commercial OPAL IMU system because of its high sampling rates and resolutions. In the future, the smartphone

IMU system might be compatible as technologies advance and can be used for performance evaluation in clinical practice.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

Ethics statement

The studies involving human participants were reviewed and approved by his study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of Cheng Hsin General Hospital, Taiwan [no. (876)110-22; date of approval: 7 Oct, 2021]. Informed consent was obtained from all subjects involved in this study. The patients/participants provided their written informed consent to participate in this study.

Author contributions

Y-TL, S-FC, and F-CW contributed to conception and design of the study. Y-TL, A-CL, C-JS, T-YK, P-HL, and AL contributed to experiments and data collection. C-JS and T-YK contributed to software development. C-JS, T-YK, and F-CW conducted data analyses. Y-TL, S-FC, C-JS, and F-CW wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This research was financially supported in part by the Ministry of Science and Technology of Taiwan (Grands MOST 107-2634-F-002-018, MOST 108-2634-F-002- 016-, MOST 109-2634-F-002- 027-, MOST 110-2634-F-002 -042-), the National Taiwan University, Center for Artificial Intelligence and Advanced Robotics. This research was also financially supported in part by the Cheng Hsin General Hospital (CHNDMC-110-8).

Acknowledgments

We appreciate the supports from the Ministry of Science and Technology of Taiwan, and the Center for Artificial Intelligence and Advanced Robotics in National Taiwan University. We are grateful to the staff of Cheng Hsin General Hospital and Taipei Love Dance Company for their assistance in helping the experiments.

Conflict of interest

AL was employed by the Taipei Love Dance Company.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Walter P, Katzenschlager R. Gait disorders in adults and the elderly. *Wiener klin Wochenschrift.* (2016) 2016:81–95. doi: 10.1007/s00508-016-1096-4

2. Guralnik JM, Ferrucci L, Simonsick EM, Salive ME, Wallace RB. Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. *N Engl J Med.* (1995) 332:556–62. doi: 10.1056/NEJM199503023320902

3. Fritz S, Lusardi M. White paper: "walking speed: the sixth vital sign". *J Geriatr Phys Ther.* (2009) 32:46–9.

4. Verghese J, Holtzer R, Lipton RB, Wang C. Quantitative gait markers and incident fall risk in older adults. *J Gerontol A Biol Sci Med Sci.* (2009) 64:896–901. doi: 10.1093/gerona/glp033

5. Montero-Odasso M, Schapira M, Varela C, Pitteri C, Soriano ER, Kaplan R, et al. Gait velocity in senior people. An easy test for detecting mobility impairment in community elderly. *J Nutr Health Aging.* (2004) 8:340–3.

6. Sherrington C, Fairhall NJ, Wallbank GK, Tiedemann A, Michaleff ZA, Howard K, et al. Exercise for preventing falls in older people living in the community. *Cochr Datab Syst Rev.* (2019) 1:CD012424. doi: 10.1002/14651858.CD012424.pub2

7. Mahlknecht P, Kiechl S, Bloem BR, Willeit J, Scherfler C, Gasperi A, et al. Prevalence and burden of gait disorders in elderly men and women aged 60-97 years: a population-based study. *PLoS One.* (2013) 8:e69627. doi: 10.1371/journal.pone.0069627

8. Paulson S, Gray M. Parameters of gait among community-dwelling older adults. *J Geriatr Phys Ther.* (2015) 38:28–32. doi: 10.1519/JPT.0000000000000018

9. Choi M, Ahn N, Park J, Kim K. 12-Week exercise training of knee joint and squat movement improves gait ability in older women. *Int J Environ Res Public Health.* (2021) 18:18041515. doi: 10.3390/ijerph18041515

10. Lee J, Jeon J, Lee D, Hong J, Yu J, Kim J. Effect of trunk stabilization exercise on abdominal muscle thickness, balance and gait abilities of patients with hemiplegic stroke: a randomized controlled trial. *NeuroRehabilitation.* (2020) 47:435–42. doi: 10.3233/NRE-203133

11. Hortobagyi T, Lesinski M, Gabler M, VanSwearingen JM, Malatesta D, Granacher U. Effects of three types of exercise interventions on healthy old adults' gait speed: a systematic review and meta-analysis. *Sports Med.* (2015) 45:1627–43. doi: 10.1007/s40279-015-0371-2

12. Hwang PW, Braun KL. The effectiveness of dance interventions to improve older adults' health: a systematic literature review. *Altern Ther Health Med.* (2015) 21:64–70.

13. Federici A, Bellagamba S, Rocchi MB. Does dance-based training improve balance in adult and young old subjects? A pilot randomized controlled trial. *Aging Clin Exp Res.* (2005) 17:385–9. doi: 10.1007/BF03324627

14. Joung HJ, Lee Y. Effect of creative dance on fitness, functional balance, and mobility control in the elderly. *Gerontology.* (2019) 65:537–46. doi: 10.1159/000499402

15. Abreu M, Hartley G. The effects of Salsa dance on balance, gait, and fall risk in a sedentary patient with Alzheimer's dementia, multiple comorbidities, and recurrent falls. *J Geriatr Phys Ther.* (2013) 36:100–8. doi: 10.1519/JPT.0b013e318267aa54

16. Aguinaga S, Marquez DX. Impact of latin dance on physical activity, cardiorespiratory fitness, and sedentary behavior among latinos attending an adult day center. *J Aging Health.* (2019) 31:397–414. doi: 10.1177/0898264317733206

17. Sofianidis G, Dimitriou AM, Hatzitaki V. A comparative study of the effects of pilates and latin dance on static and dynamic balance in older adults. *J Aging Phys Act.* (2017) 25:412–9. doi: 10.1123/japa.2016-0164

18. Knutsson E, Richards C. Different types of disturbed motor control in gait of hemiparetic patients. *Brain.* (1979) 102:405–30. doi: 10.1093/brain/102.2.405

19. Phoenix Technologies Inc. *VZ4000 Tracker*. (2021). Available online at: http://www.ptiphoenix.com/?prod-trackers-post=vz4000 (accessed December 9,2021)

20. VICON. *VICON Products*. (2021). Available online at: https://www.vicon.com/ (accessed December 9,2021)

21. Wang F-C, Lin Y-Y, Li Y-C, Chen P-Y, Yu C-H. Development of an automated assistive trainer inspired by neuro-developmental treatment. *Sens Mater.* (2020) 32:3019–37. doi: 10.18494/SAM.2020.2708

22. Wong AM, Pei Y-C, Hong W-H, Chung C-Y, Lau Y-C, Chen CP. Foot contact pattern analysis in hemiplegic stroke patients: an implication for neurologic status determination. *Arch Physical Med Rehabil.* (2004) 85:1625–30. doi: 10.1016/j.apmr.2003.11.039

23. Daliri MR. Chi-square distance kernel of the gaits for the diagnosis of Parkinson's disease. *Biomed Signal Process Control.* (2013) 8:66–70. doi: 10.1016/j.bspc.2012.04.007

24. El Maachi I, Bilodeau G-A, Bouachir W. Deep 1D-Convnet for accurate Parkinson disease detection and severity prediction from gait. *Expert Syst Applicat.* (2020) 143:113075. doi: 10.1016/j.eswa.2019.113075

25. Bilro L, Pinto JL, Oliveira J, Nogueira R. Gait monitoring with a wearable plastic optical sensor. *SENSORS, 2008 IEEE*. Piscataway, NJ: IEEE (2008). doi: 10.1109/ICSENS.2008.4716559

26. Díaz S, Stephenson JB, Labrador MA. Use of wearable sensor technology in gait, balance, and range of motion analysis. *Appl Sci.* (2020) 10:234. doi: 10.3390/app10010234

27. APDM. *Opal User Guide*. (2021). Available online at: https://www.apdm.com/ (accessed December 9,2021)

28. Aminian K, Najafi B, Büla C, Leyvraz P-F, Robert P. Spatio-temporal parameters of gait measured by an ambulatory system using miniature gyroscopes. *J Biomech.* (2002) 35:689–99. doi: 10.1016/s0021-9290(02)00008-8

29. Wang F-C, Li Y-C, Wu K-L, Chen P-Y, Fu L-C. Online gait detection with an automatic mobile trainer inspired by neuro-developmental treatment. *Sensors.* (2020) 20:3389. doi: 10.3390/s20123389

30. Yun X, Calusdian J, Bachmann ER, McGhee RB. Estimation of human foot motion during normal walking using inertial and magnetic sensor measurements. *IEEE Transact Instrument Measur.* (2012) 61:2059–72.

31. Rebula JR, Ojeda LV, Adamczyk PG, Kuo AD. Measurement of foot placement and its variability with inertial sensors. *Gait Posture.* (2013) 38:974–80.

32. Epomedicine. *Physical Examination: Gait*. (2021). Available online at: http://epomedicine.com/clinical-medicine/physical-examination-gait/ (accessed December 9,2021)

33. Herzog W, Nigg BM, Read LJ, Olsson E. Asymmetries in ground reaction force patterns in normal human gait. *Med Sci Sports Exerc.* (1989) 21:110–4. doi: 10.1249/00005768-198902000-00020

34. Arsenault AB, Winter DA, Marteniuk RG. Bilateralism of EMG profiles in human locomotion. *Am J Phys Med.* (1986) 65:1–16.

35. Damholt V, Termansen N. Asymmetry of plantar flexion strength in the foot. *Acta Orthopaed Scand.* (1978) 49:215–9. doi: 10.3109/17453677809005754

36. Chodera J. Analysis of gait from footprints. *Physiotherapy.* (1974) 60:179–81.

37. Sadeghi H, Allard P, Prince F, Labelle H. Symmetry and limb dominance in able-bodied gait: a review. *Gait Posture.* (2000) 12:34–45. doi: 10.1016/s0966-6362(00)00070-9

38. Stefanyshyn DJ, Engsberg JR. Right to left differences in the ankle joint complex range of motion. *J Biomech.* (1994) 27:816–816.

39. Lewek MD, Osborn AJ, Wutzke CJ. The influence of mechanically and physiologically imposed stiff-knee gait patterns on the energy cost of walking. *Arch Phys Med Rehabil.* (2012) 93:123–8. doi: 10.1016/j.apmr.2011.08.019

40. Wong DW, Lam WK, Lee WC. Gait asymmetry and variability in older adults during long-distance walking: implications for gait instability. *Clin Biomech.* (2020) 72:37–43. doi: 10.1016/j.clinbiomech.2019.11.023

41. Bautmans I, Jansen B, Van Keymolen B, Mets T. Reliability and clinical correlates of 3D-accelerometry based gait analysis outcomes according to age and fall-risk. *Gait Posture.* (2011) 33:366–72. doi: 10.1016/j.gaitpost.2010.12.003

42. Patterson KK, Parafianowicz I, Danells CJ, Closson V, Verrier MC, Staines WR, et al. Gait asymmetry in community-ambulating stroke survivors. *Arch Phys Med Rehabil.* (2008) 89:304–10. doi: 10.1016/j.apmr.2007.08.142

43. Chen G, Patten C, Kothari DH, Zajac FE. Gait differences between individuals with post-stroke hemiparesis and non-disabled controls at matched speeds. *Gait Posture.* (2005) 22:51–6. doi: 10.1016/j.gaitpost.2004.06.009

44. Fonseca CC, Thurm BE, Vecchi RL, Gama EF. Ballroom dance and body size perception. *Percept Mot Skills.* (2014) 119:495–503. doi: 10.2466/25.PMS.119c26z1

45. Byrnes K, Wu PJ, Whillier S. Is Pilates an effective rehabilitation tool? A systematic review. *J Bodyw Mov Ther.* (2018) 22:192–202. doi: 10.1016/j.jbmt.2017.04.008

46. Irez GB, Ozdemir RA, Evin R, Irez SG, Korkusuz F. Integrating Pilates exercise into an exercise program for 65+ year-old women to reduce falls. *J Sports Sci Med.* (2011) 10:105.

47. Pata RW, Lord K, Lamb J. The effect of Pilates based exercise on mobility, postural stability, and balance in order to decrease fall risk in older adults. *J Bodywork Mov Therap.* (2014) 18:361–7. doi: 10.1016/j.jbmt.2013.11.002

48. Roberts L, Jones G, Brooks R. Why do you ride?: a characterization of mountain bikers, their engagement methods, and perceived links to mental health and well-being. *Front Psychol.* (2018) 9:1642. doi: 10.3389/fpsyg.2018.01642

49. Vellas BJ, Rubenstein LZ, Ousset PJ, Faisant C, Kostek V, Nourhashemi F, et al. One-leg standing balance and functional status in a population of 512 community-living elderly persons. *Aging.* (1997) 9:95–8. doi: 10.1007/BF03340133

50. Vellas BJ, Wayne SJ, Romero L, Baumgartner RN, Rubenstein LZ, Garry PJ. One-leg balance is an important predictor of injurious falls in older persons. *J Am Geriatr Soc.* (1997) 45:735–8. doi: 10.1111/j.1532-5415.1997.tb01479.x

51. Schmid M, Nardone A, De Nunzio AM, Schmid M, Schieppati M. Equilibrium during static and dynamic tasks in blind subjects: no evidence of cross-modal plasticity. *Brain.* (2007) 130:2097–107. doi: 10.1093/brain/awm157

APPENDIX

Appendix 1

The angular velocities ω_y of all subjects are available at http://140.112.14.7/∼sic/PaperMaterial/Dance_Gait_Responses.zip.

Appendix 2

The gait cycles of all subjects are available at http://140.112.14.7/∼sic/PaperMaterial/Dance_IMU_Data.zip.

Appendix 3

The asymmetry of swing phases ($Asym_{SP}$) of all subjects are available at: http://140.112.14.7/∼*sic*/PaperMaterial/Dance_Asym_Data.zip.

Appendix 4

The ratio of swing time (R_{SW}) of all subjects are available at: http://140.112.14.7/∼*sic*/PaperMaterial/Dance_Ratio_Data.zip.

Appendix 5

The IMU data on waist of all subjects are available at: http://140.112.14.7/∼*sic*/PaperMaterial/Dance_Gait_Responses.zip.

Appendix 6

The angular velocity and acceleration of the waist for all subjects are available at: http://140.112.14.7/∼*sic*/PaperMaterial/Subjects' angular velocity and acceleration of the waist.zip.

frontiers | Frontiers in Pharmacology

TYPE Systematic Review
PUBLISHED 20 October 2022
DOI 10.3389/fphar.2022.835991

Check for updates

OPEN ACCESS

EDITED BY
Kevin Lu,
University of South Carolina,
United States

REVIEWED BY
Laurent Misery,
Université de Bretagne Occidentale,
France
Joao Massud,
Independent Researcher, São Paulo,
Brazil

*CORRESPONDENCE
Rensong Yue,
yuerensong163@hotmail.com

SPECIALTY SECTION
This article was submitted to Drugs Outcomes Research and Policies, a section of the journal Frontiers in Pharmacology

RECEIVED 15 December 2021
ACCEPTED 27 September 2022
PUBLISHED 20 October 2022

CITATION
Xu C, Yue R, Lv X, Wang S and Du M (2022), Efficacy and safety of pharmacological interventions for pruritus in primary biliary cholangitis: A systematic review and meta-analysis. *Front. Pharmacol.* 13:835991. doi: 10.3389/fphar.2022.835991

COPYRIGHT
© 2022 Xu, Yue, Lv, Wang and Du. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Efficacy and safety of pharmacological interventions for pruritus in primary biliary cholangitis: A systematic review and meta-analysis

Chenyi Xu[1], Rensong Yue[1]*, Xuelian Lv[2], Shengnan Wang[1] and Mengmeng Du[3]

[1]Hospital of Chengdu University of Traditional Chinese Medicine, Chengdu, China, [2]Xinjin Hospital of Traditional Chinese Medicine, Chengdu, China, [3]Qing Dao NO.6 People's Hospital, Qing Dao, China

Background and objective: Pruritus is a common complication in patients with primary biliary cholangitis (PBC). The pathogenesis is not clear, and also the precise therapeutic measures remain alluring. In order to systematically evaluate the efficacy and safety of drug interventions in the treatment of pruritus associated with PBC, this systemic review and meta-analysis was conducted.

Methods: The randomized controlled trials (RCTs) on drug interventions in the treatment of pruritus associated with primary cholangitis were searched in the electronic databases of PubMed, EMBASE, Cochrane Library, Web of Science, and ClinicalTrials.gov. Two researchers independently screened the literature, extracted and integrated the data, and assessed the bias risk of the selected literature, according to the *Cochrane handbook*. Finally, the STATA 15. 0 software was used for the meta-analysis.

Results: A total of 23 RCTs involving 2,194 patients were studied, that included 12 pharmacological interventions. In terms of itching relief, compared with placebo, UDCA, methotrexate and GSK2330672 had a definite effect in improving pruritus (pruritus remission rate before and after treatment, $p < 0.05$). In terms of serum indexes, compared with placebo group, UDCA, OCA, rifampicin, cyclosporine, NGM282, seladelpar and colchicine may improve blood alkaline phosphatase (ALP) ($p < 0.05$), but only rifampicin showed low heterogeneity. UDCA, bezafibrate, OCA, rifampicin, NGM282 and others may improve blood γ-glutamyl transpeptidase (γ-GGT) ($p < 0.05$), but due to the high heterogeneity and the limitation of research samples, a clear conclusion cannot be drawn. In terms of adverse events, except high (>15 mg/kg/day) and low doses (<13 mg/kg/day) of UDCA increased the incidence of adverse events, there were no risk of increasing the incidence of adverse events compared with placebo ($p > 0.05$), and a moderate dose of UDCA (13–15 mg/kg/day) and malotilate (1,500 mg/day) may also help in reducing the incidence of adverse events ($p < 0.05$).

Conclusion: UDCA, methotrexate and GSK2330672 may relieve itching in patients with PBC, but there is a lack of robust evidence to support their effect on ALP or γ-GGT. Due to the heterogeneity in the published studies, based on the present review, we cannot explicitly recommend any specific drug for the treatment of PBC-related pruritus.

Systematic Review Registration: link-https://osf.io/2g8ya, identifier 10.17605/OSF.IO/2G8YA

KEYWORDS

primary biliary cholangitis, pruritus, pharmacological interventions, alkaline phosphatase, γ-glutamyl transpeptidase

Introduction

Primary biliary cholangitis (PBC), formerly known as primary biliary cirrhosis, is an autoimmune liver disease, which is predominantly seen in women. The interaction of specific anti-mitochondrial antibodies with specific autoantigens, accompanied by pathophysiological processes such as bile duct injury, cholestasis, liver fibrosis and even liver cirrhosis, are the main features of the pathological progress of PBC (Lee et al., 2019; Gulamhusein and Hirschfield, 2020). At present, the incidence of PBC is geographically varied, and the number is on the rise too. According to a survey, there are 118.75 cases of prostate cancer per million people in the Asia-Pacific region (Zeng et al., 2019), 218.1 cases per million people in North America, 145.9 cases per million people in Europe, and 189.0 cases per million in Victoria, Australia (J et al., 2020). It was reported to be 346.0 cases per million in Sweden (HU et al., 2019), 149.0 cases per million in Slovakia (Drazilova et al., 2020) and, 279.0 cases per million in Italy (Rajaobelina et al., 2019). Environmental and genetic factors play an important role in the occurrence of the disease (Carey et al., 2015). The levels of serum alkaline phosphatase (ALP) and γ-glutamyl transpeptidase (γ-GGT) are generally increased in patients with PBC.

Fatigue and pruritus are the most common symptoms in patients with PBC. Scratching, sleep deprivation, depression and, even suicidal thoughts caused by itching affect 20%–70% of patients (Beuers et al., 2014; Carey et al., 2015), seriously affecting their quality of life. The accumulation and deposition of bile acid and bile salt, the regulation of lysophosphatidic acid, the abnormality of endogenous opioid receptors and the effects of serotonin and substance P are considered to be the main pathophysiological mechanisms of cholangitis pruritus (Bray et al., 2018). Even so, the pathogenesis of cholestatic pruritus has not been well described because of its diversity and complexity (Tajiri and Shimizu, 2017; SP et al., 2019).

A cross-sectional study of 2194 PBC pruritus patients in the United Kingdom shows that there is a lack of adequate understanding, management and guidelines for the disease, and there is insufficient evidence on the recommended treatment (Hegade et al., 2019). There exist few evidence-based guidelines for treating PBC pruritus, but the limitation is that there are specific clinical conditions where these guidelines cannot be applied. Ursodeoxycholic acid (UDCA) and OCA are drugs approved by the U.S. Food and Drug Administration (FDA) for the treatment of PBC, which can improve the liver biochemical indexes, prolong survival time and delay the development of esophageal varices. However, UDCA cannot completely cure this disease, and its effectiveness in the treatment of cholangitis pruritus is still controversial (JS et al., 2012). Interestingly, another potential anti-PBC drug, OCA, was reported to increase the risk of itching (Trauner et al., 2019). The search for an effective treatment for PBC-related pruritus interventions has never stopped. Newer pharmacological interventions have been reported such as cholestyramine (TB et al., 1961), rifampicin (CN and SG, 1988), sertraline (MJ et al., 2007), ondansetron (JW et al., 2005), maralixibat (Mayo et al., 2019), ileal apical sodium bile acid transporter (ASBT) inhibitors such as GSK2330672 (Hegade et al., 2017). In order to determine the efficacy and safety of the available drugs for the treatment of PBC-associated pruritus, we conducted a systematic review and meta-analysis of pharmacological interventions in PBC-related pruritus.

Materials and methods

We followed a predetermined protocol and the principles of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) for this systemic review and meta-analysis (DG et al., 2009).

Search strategy

To systematically evaluate the efficacy of interventions, we searched PubMed, EMBASE, Cochrane Library, Web of Science, and ClinicalTrials.gov from inception up to June 2021. The search strategy was implemented by an experienced medical librarian.

Strategies were selected using a combination of medical subject headings (MeSH) and text words, and search terms included "primary biliary cholangitis (cirrhosis)," "pruritus" or "pruritis" or "itching," and "randomized controlled trial." The language was limited to English, and the publication status was not restricted.

Selection criteria

The inclusion criteria were as follows

(1) The study was a randomized controlled trial (RCT).
(2) The diagnosis of PBC was based on at least two of the following: the presence of anti-mitochondrial antibody (AMA), cholestasis with an elevation of ALP activity, histopathologic evidence of nonsuppurative cholangitis and destruction of small or medium-sized bile ducts (KD et al., 2019), and patients of primary biliary cholangitis with persistent pruritus (course of disease≥3 months).
(3) All pharmacological interventions related to the treatment of pruritus in PBC.
(4) Outcomes related to the efficacy and safety of pruritus in PBC.

Exclusion criteria

(1) The interventions are not suitable (more than three interventions or the study interventions beyond our study).
(2) PBC associated with other underlying diseases.
(3) Unavailable data (there is no data we need in the study).
(4) Absence of a clear basic information about the study subjects.
(5) Duplicate publications.
(6) Studies with ambiguous diagnostic criteria.

Study selection and data extraction

Two reviewers (Chenyi Xu and Xuelian Lv) independently extracted basic information about the articles (article title, first author, year of publication, sample size, country or region), trial design (participants, interventions, time span of the trial, follow-up time), clinical efficacy and adverse events.

Clinical efficacy measures included pruritus scores; 0–10 numerical rating scale (NRS), quality of life scale for primary biliary cirrhosis (cholangitis) (PBC-40), and 5-D itch scale (or the pruritus relief rate before and after treatment). Secondary outcomes were laboratory parameters such as the changes in serum alkaline phosphatase (ALP) and gamma-glutamyl transpeptidase (γ–GGT). We considered both iatrogenic and non-iatrogenic adverse events, and carried out a quantitative analysis of these events.

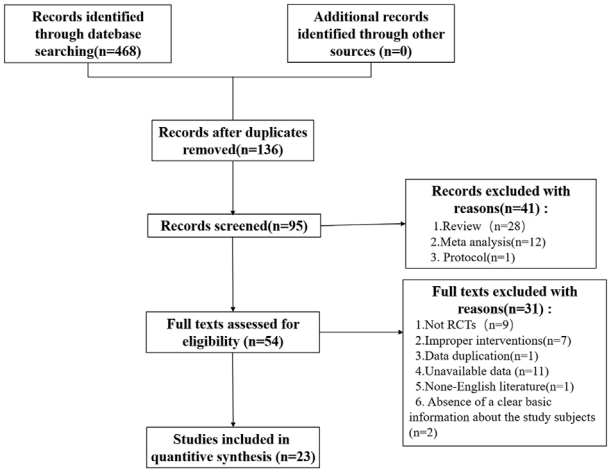

FIGURE 1
Flow chart of trail selection.

Risk of bias assessment

Two investigators evaluated each of the included RCTs and recorded the following six items, as per the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins and Green, 2008): the methods of blinding, the generation of information, data, distribution of randomized control sequences, selective reports and other possible problems. The risk of biases was marked as high, uncertain, or low.

Statistical analysis

STATA15.0 (STATA statistical software: Release 15.0 College Station, TX: Stata Corp LP) was used to analyze the data. Successive mean differences in pruritus scores were reported as standardized mean differences (SMD), and binary variables used risk ratios (RR) or odds ratio (OR), providing a 95% confidence interval (CI) for each effect. Heterogeneity was evaluated by I^2, $p < 0.05$ was considered to be statistically significant. The studies with high heterogeneity ($p \leq 0.10$ and $I^2 \geq 50\%$) were analyzed by random effect model, the studies with low heterogeneity ($p > 0.10$ and $I^2 < 50\%$) were analyzed by fixed effect model. The stability of the results was evaluated by subgroup analysis and sensitivity analysis. If enough studies were included in meta-analysis ($n \geq 10$), funnel chart analysis was used for evaluate publication bias.

TABLE 1 Basic characteristics of 23 included studies.

Study and year	Design	Population	Sample size	Intervention	Outcomes	Follow-up	Concomitant treatment
UDCA							
Hiroshi OKA 1990	RCT	PBC. Mean age 59 years, 89% female	22 UDCA 23 Placebo	UDCA 8–12 mg/kg/day	(1) (2) (3) (4)	2 years	None detailed
Poupon RE 1991	RCT	PBC. Mean age 56 years, 92% female	73 UDCA 73 Placebo	UDCA 13–15 mg/kg/day	(1) (2) (3) (4)	2 years	None detailed
Matti 1995	RCT	PBC. Mean age 54.5 years, 92% female	30 UDCA 31 Placebo	UDCA 12–15 mg/kg/day	(1) (2) (3) (4)	2 years	None detailed
Albert Parés 2000	RCT	PBC. Mean age 54.1 years, 93% female	99 UDCA 93 Placebo	UDCA 14–16 mg/kg/day	(1) (2) (4)	2 years	Cholestyramine at least 2 h after the intake of UDCA or Placebo
E.Jenny heathcote 1994	RCT	PBC.Mean age 56.4 years, 93% female	111 UDCA 111 Placebo	UDCA 14 mg/kg/day	(1) (3) (4)	2 years	Cholestyramine, in the morning or at least 4 h before the trial capsules
P.M.Battezzati1993	RCT	PBC.Mean age 54.5 years, 89% female	44 UDCA 44 Placebo	UDCA 8.7 mg/kg/day	(1) (3) (4)	1 year	Cholestyramine take at least 4 h before or after the study drug
Poupon RE 1990	RCT	PBC.Mean age 56.5 years, 91% female	70 UDCA 68 Placebo	UDCA 13–15 mg/kg/day	(1) (2) (3) (4)	2 years	None detailed
K. D. Lindor1994	RCT	PBC.Mean age 53 years, 89% female	89 UDCA 91Placebo	UDCA 13–15 mg/kg/day	(1) (3) (4)	2 years	Cholestyramine were asked to take drug 2 h after their study drug
Study and year	Design	Population	Sample size	Intervention	Outcomes	Follow-up	Concomitant treatment
Rifampicin							
Laura Bachs 1989	RCT	PBC. Mean age 49.7 years, 100% female	21 Rifampicin 18 Phenobarbitone	Rifampicin 10 mg/kg/day, Phenobarbitone 3 mg/kg	(1) (2) (3) (4)	14 days	None detailed
Ana Podesta 1991	RCT	PBC. Mean age 43 years, 93% female	14 Rifampicin 14 Placebo	Rifampicin 300 mg twice a day	(1) (3) (4)	3 months	None detailed
GSK2330672							
Vinod S Hegade 2017	RCT	PBC. Mean age 52.9 years, 90% female	22 Placebo run-in 21 GSK2330672 21 Placebo	GSK2330672 45 mg twice per day on days 1–3, and 90 mg twice daily on days4–14	(1) (2) (3) (4)	28 days	19 patients taking UDCA (14 mg/kg/day) during study period
Maralixibat							
M. J. Mayo 2019	RCT	PBC. Mean age 53.4 years, 91% female	21 Maralixibat 10 mg 21 Maralixibat 20 mg 24 Placebo	Maralixibat 10 mg,10 mg/day Maralixibat 20 mg, 20 mg/day	(1) (2) (3) (4)	13 weeks	None detailed
Study and year	Design	Population	Sample size	Intervention	Outcomes	Follow-up	Concomitant treatment
Obeticholic Acid							
F. Nevens 2016	RCT	PBC. Mean age 56 years, 91% female	70 OCA5-10 mg 73 OCA 10 mg 73 Placebo	OCA 5–10 mg/day or 10 mg/day	(1) (3)	1 year	standard-of-care UDCA (13–15 mg/kg/day) or not
Kris V. Kowdley 2018	RCT	PBC. Mean age 54 years, 85% female	20 OCA 10 mg 16 OCA 50 mg 23 Placebo	OCA10 mg/day or 50 mg/day	(1) (3) (4)	3 months	None detailed

(Continued on following page)

TABLE 1 (*Continued*) Basic characteristics of 23 included studies.

Study and year	Design	Population	Sample size	Intervention	Outcomes	Follow-up	Concomitant treatment
G. M Hirschfield 2014	RCT	PBC. Mean age 55.1 years, 95% female	38 OCA 10 mg 48 OCA 25 mg 41 OCA 50 mg 38 Placebo	OCA10 mg/day or 25 mg/day or 50 mg/day	(2) (3) (4)	3 months	UDCA (15.6–16.3 mg/kg/day)
Bezafibrate							
C. Corpechot 2018	RCT	PBC.Mean age 53 years, 95% female	50 Bezafibrate 50 Placebo	Bezafibrate 400 mg/day	(1) (2) (3) (4)	2 years	UDCA (13–15 mg/kg/day)
Tatsuo Kanda 2003	RCT	PBC.Mean age 56 years, 86% female	11 UDCA with Bezafibrate 11 UDCA	Bezafibrate 400 mg/day	(1) (2) (3) (4)	0.5 years	UDCA (600 mg/day)
Study and year	Design	Population	Sample size	Intervention	Outcomes	Follow-up	Concomitant treatment
NGM282							
M. J. mayo 2018	RCT	PBC. Mean age 56.3 years, 91% female	14 NGM282 0.3 mg 16 NGM282 3 mg 15 Placebo	NGM282 0.3mg, 0.3 mg/day NGM282 3mg, 3 mg/day	(1) (2) (3) (4)	28 days	None detailed
Seladelpar (MBX-8025)							
David Jones 2017	RCT	PBC. Mean age 56 years, 95% female	13 Seladelpar50 mg 13Seladelpar200 mg 12 Placebo	seladelpar50 mg,50 mg/day Seladelpar200 mg,200 mg/day	(1) (2) (3) (4)	12 weeks	Continue UDCA at the same dose
Malotilate							
A European Multi--centre Study Group 1993	RCT	PBC. Mean age 54.4 years, 96% female	52 Malotilate 49 Placebo	Malotilate:500 mg, three times a day	(1) (2) (3) (4)	1.5 years	None detailed
Cyclosporine							
R.H.Wiesner 1990	RCT	PBC. Mean age 46.8 years, 97% female	19 Cyclosporine 10 Placebo	cyclosporine 4 mg/kg/day	(1) (2) (3) (4)	2 years	None detailed
Study and year	Design	Population	Sample size	Intervention	Outcomes	Follow-up	Concomitant treatment
Colchicine							
P. L. Almasio 2000	RCT	PBC. Mean age 54.4 years, 90% female	46 Colchicine plus UDCA 44 UDCA	Colchicine, 1 mg/daily UDCA (5.4–11.6 mg/kg/day)	(1) (2) (3) (4)	3 years	Cholestyramine, no more than 8 g/day
Methotrexate							
M.T. Hendrickse1999	RCT	PBC. Mean age 57 years, 92% female	30 MTX 30 Placebo	MTX, 7.5 mg/wk (on Friday, Saturday, and Sunday of each week, 2.5 mg/day)	(1) (2) (3) (4)	6 years	None detailed

Abbreviations: PBC, Primary biliary cholangitis; MTX, Methotrexate; OCA, Obeticholic Acid; UDCA, Ursodeoxycholic acid; (1) ALP, Alkaline phosphatase; (2) γ-GGT, Gamma-glutamyltranspeptidase; (3)Number with pruritus, grade of pruritus; (4)Adverse event rate(%).

Results

Study selection

A total of 468 records were retrieved from the electronic database (PubMed $n = 90$, EMBASE $n = 80$, Cochrane Library $n = 151$, Web of Science $n = 106$, ClinicalTrials.gov $n = 41$), and 332 records were excluded after duplicates removed. After browsing the titles and abstracts, 41 articles were excluded based on article type (review $n = 28$, meta-analysis $n = 12$, protocol $n = 1$). After that, 54 articles were assessed as full text and 31 were excluded for the following reasons: not RCTs ($n = 9$), improper intervention ($n = 7$), data duplication ($n = 1$), unavailable data ($n = 11$), non-English literature ($n = 1$), absence of a clear basic information about the study subjects ($n = 2$). Finally, 23 studies were obtained. The details are shown in the flowchart (Figure 1). Of the 23 studies included, eight were RCTs studies of UDCA (Poupon et al., 1990; Oka et al., 1991; Poupon et al., 1991; Battezzati et al., 1993; Lindor et al., 1994; Heathcote et al., 1995; Vuoristo et al., 1995; Parés et al., 2000).

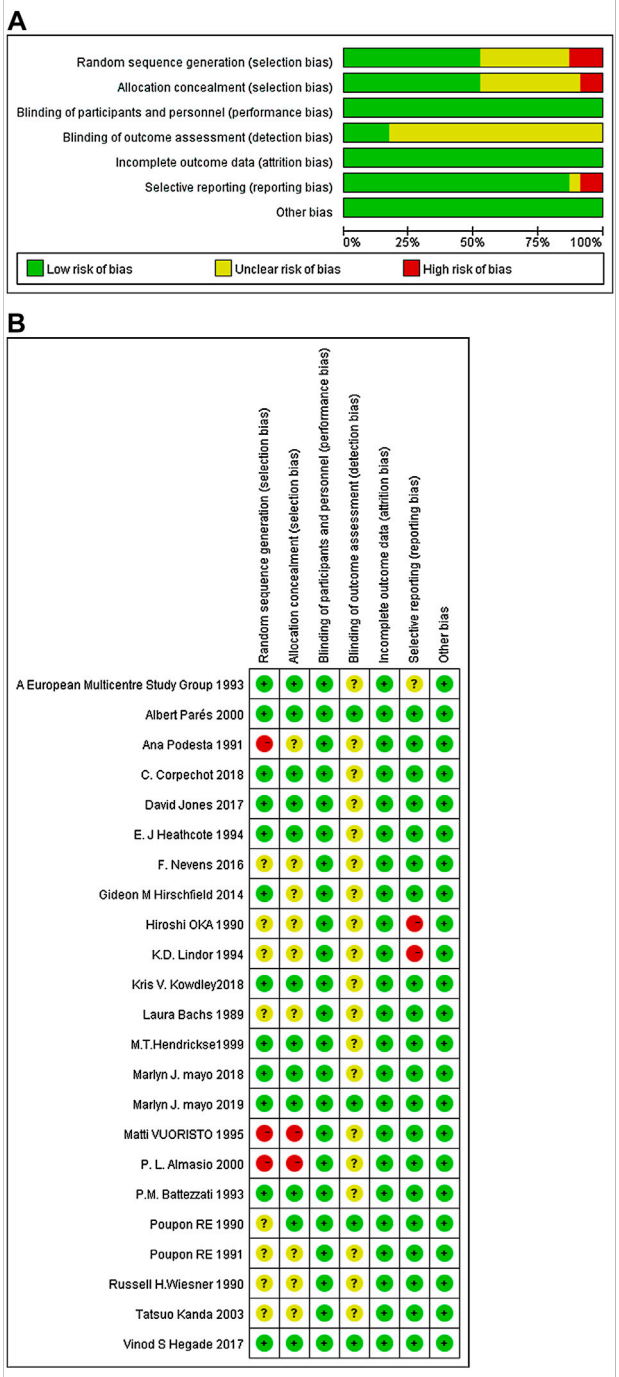

FIGURE 2
The quality assessment of included trials based on the Cochrane risk assessment tool. **(A)** Each risk of bias item presented as percentages across all included studies; **(B)** Each risk of bias item for each included study.

The remaining 15 RCTs contained three studies of obeticholic acid in PBC (Hirschfield et al., 2015; Nevens et al., 2016; Kowdley et al., 2018), two studies of bezafibrate in PBC (Kanda et al., 2003; Corpechot et al., 2018), and two studies of rifampicin (Bachs et al., 1989; Podesta et al., 1991), and one RCT each evaluating the ileal bile acid transporter inhibitor GSK2330672 (Hegade et al., 2017), FGF19 analog NGM282 (Mayo et al., 2018), selective PPAR-δ agonist seladelpar (MBX-8025) (Jones et al., 2017), cyclosporine (Wiesner et al., 1990), colchicine (Almasio et al., 2000), methotrexate (Hendrickse et al., 1999), maralixibat (Mayo et al., 2019) and malotilate (Listed, 1993).

The 23 RCTs included 2,194 patients, 92% of the study population was women, and the average age was 54.9 years. The studies included in our systematic review are shown in Table 1.

Study quality

Of all the 23, 11 studies mentioned the randomization techniques (including random sampling with computer (Battezzati et al., 1993; Hendrickse et al., 1999; Hirschfield et al., 2015; Hegade et al., 2017; Jones et al., 2017; Corpechot et al., 2018; Kowdley et al., 2018; Mayo et al., 2019), or interactive voice/web response system (Mayo et al., 2018), or random numbers (Listed, 1993; Heathcote et al., 1995)), while two studies were considered "high risk" because the random methods were inappropriate (Podesta et al., 1991; Vuoristo et al., 1995). Twelve studies followed allocation concealment (including third-party random allocation (Poupon et al., 1990; Battezzati et al., 1993; Lindor et al., 1994; Jones et al., 2017; Corpechot et al., 2018; Mayo et al., 2018; Mayo et al., 2019), sequence number (Listed, 1993; Heathcote et al., 1995; Hendrickse et al., 1999; Hegade et al., 2017), or envelope (Parés et al., 2000)). Of all the included studies, two followed single-blind (Oka et al., 1991; Heathcote et al., 1995), and the rest followed double-blind designs. The data evaluator was blinded in five studies. All the 23 studies reported complete data and there was no risk of other biases. The Cochrane deviation risk assessment tool was used for risk assessment, and the results are shown in Figures 2A,B.

Pruritus (relief rate)

Most trials reported only the number of patients with pruritus pre-treatment and post-treatment, so we considered the relief rate as the primary outcome. Among the eight studies (Poupon et al., 1990; Oka et al., 1991; Poupon et al., 1991; Battezzati et al., 1993; Lindor et al., 1994; Heathcote et al., 1995; Vuoristo et al., 1995; Parés et al., 2000), of UDCA seven documented itching. The results showed that UDCA combined with or without cholestyramine, compared to placebo (RR = 1.85, 95%CI (1.40, 2.45), $p < 0.001$, $I^2 = 0.0\%$), had a significant difference in relieving pruritus (Figure 3A). Among these, two studies recorded the itching score in detail (Battezzati et al., 1993; Parés et al., 2000). Compared with placebo, UDCA combined

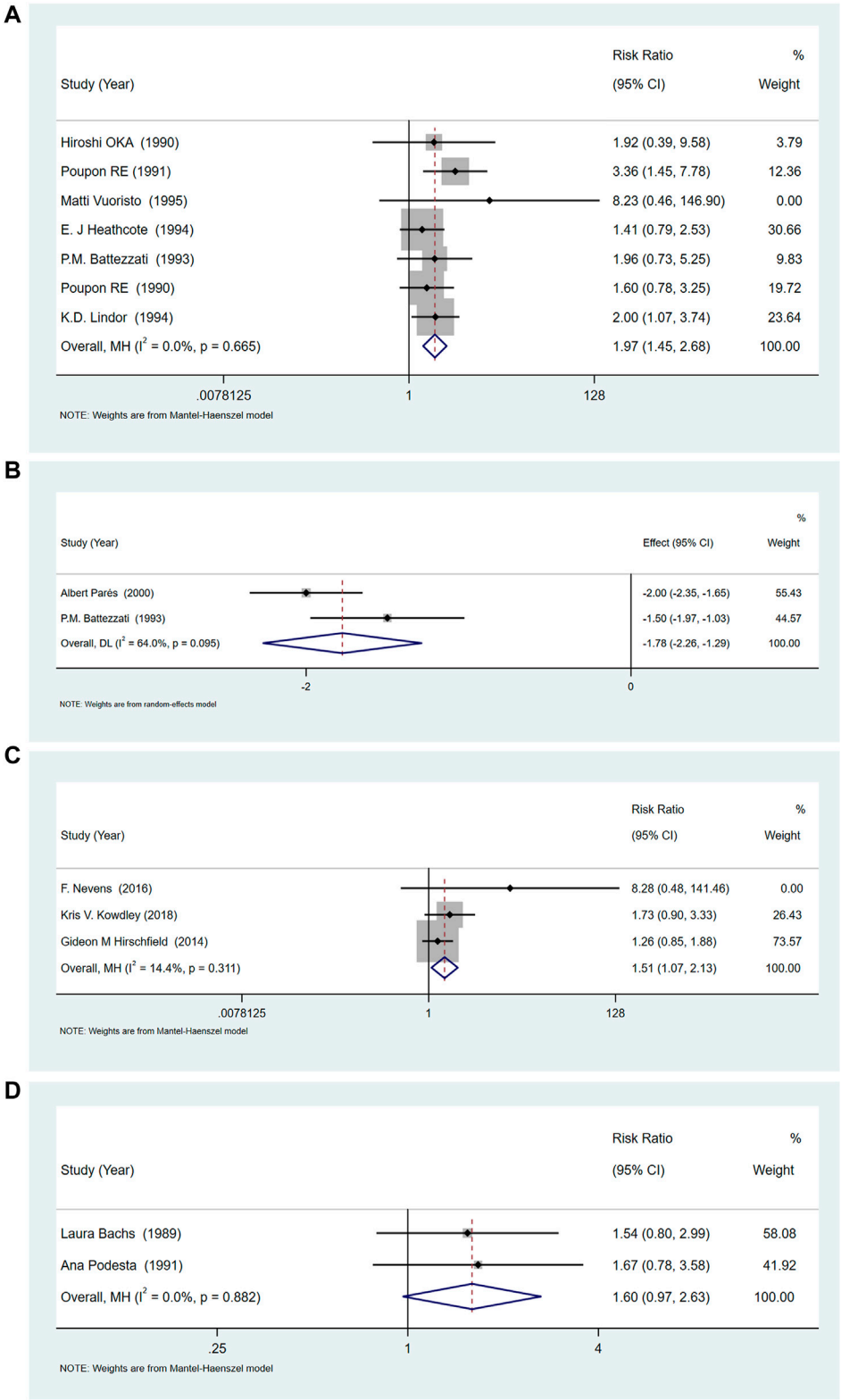

FIGURE 3
(A) The effect of UDCA on pruritus relief compared with placebo. **(B)** Two studies of UDCA in pruritus scores compared with placebo. **(C)** OCA increases the risk of pruritus compared with placebo. **(D)** The effect of Rifampicin on the rate of pruritus relief compared with placebo.

with cholestyramine significantly reduced the itching score (SMD = −1.78, 95% CI (−2.26, −1.29), $p < 0.001$, $I^2 = 64\%$), but had a high heterogeneity ($p = 0.095$, $I^2 = 64.0\%$) (Figure 3B).

Three studies (Hirschfield et al., 2015; Nevens et al., 2016; Kowdley et al., 2018) compared the relief of pruritus after OCA treatment, and the heterogeneity among studies was low ($p = 0.311$, $I^2 = 14.4\%$). Meta-analysis showed that OCA increased the incidence of pruritus, compared with placebo (RR = 1.511, 95% CI (1.07, 2.12), $p = 0.018$) (Figure 3C).

Two studies (Bachs et al., 1989; Podesta et al., 1991) compared the relief of pruritus after rifampicin treatment, and the heterogeneity between the studies was low ($p = 0.882$, $I^2 = 0.0\%$). The results of meta-analysis showed that rifampicin had no significant effect on itching compared with placebo or control drugs [RR = 1.595, 95%CI (0.97, 2.63), $p = 0.067$] (Figure 3D).

One study (Wiesner et al., 1990) reported the incidence of pruritus in patients with PBC, before and after treatment, with cyclosporine and placebo. The results showed that cyclosporine did not significantly reduce the incidence of pruritus compared with placebo [RR = 0.726, 95%CI (0.51, 1.03), $P = 0.072$]

One study (Mayo et al., 2018) reported a comparison between NGM282 and placebo. The results suggested that NGM282 had no significant effect on 5-D itch score [SMD = −0.429, 95%CI (−1.06, 0.20), $p = 0.179$] and VAS score [SMD = 0.335, 95%CI (−0.29, 0.96), $p = 0.293$] in patients with PBC.

One study (Listed, 1993) reported a comparison between malotilate and placebo, which suggested that malotilate did not significantly reduce the incidence of pruritus [RR = 1.083, 95%CI (0.43, 2.73), $p = 0.865$].

One study (Jones et al., 2017) reported the comparison of seladelpar and placebo after treatment. The results suggested that seladelpar did not significantly reduce pruritus in patients with PBC. The 5-D itch score [SMD = 0.000, 95%CI (−0.79, 0.79), $p = 1.000$] and VAS score [SMD = 0.138, 95%CI(−0.65, 0.92), $p = 0.731$] showed no significant alteration.

One study (Mayo et al., 2019) reported a comparison between maralixibat and placebo. The results suggested that maralixibat did not significantly reduce the pruritus 5-D itch score in patients with PBC [SMD = −0.09, 95%CI (−0.59, 0.41), $p = 0.725$].

A study (Hendrickse et al., 1999) on methotrexate showed that, compared with placebo, it significantly reduced pruritus scores [SMD = 1.000, 95%CI (−1.54, −0.46), $p = 0.000$].

In a comparative study (Vuoristo et al., 1995) between colchicine plus UDCA and UDCA, it was found that colchicine had no significant effect on reducing the incidence of pruritus [RR = 0.964, 95%CI (0.42, 2.24), $p = 0.931$].

Another study (Hegade et al., 2017) on GSK2330672 used different scoring systems to evaluate the change in pruritus score before and after treatment. The percentage changes from baseline itch scores were -57% [95% CI (−73, −42), $p < 0.0001$] in NRS, −31% (−42~-20, $p < 0.0001$) in PBC-40 itching, and −35% (−45~−25, $p < 0.0001$) in 5-D itch score.

One study (Parés et al., 2000) of UDCA could not be analyzed because of the lack of data. Indicators changes before and after treatment for all included studies are listed in Table 2.

Alkaline phosphatase

Compared with placebo, UDCA, OCA and rifampicin could reduce serum ALP levels, UDCA [SMD = -2.91, 95%CI (−4.37, −1.44), $p = 0.000$], OCA [SMD = −5.56, 95%CI (−9.82, −1.30), $p = 0.011$], rifampicin [SMD = −0.53, 95%CI (−1.02, −0.04), $p = 0.033$]. However, the change brought about by bezafibrate [SMD = −4.98, 95%CI (−12.09, 2.16), $p = 0.172$] was not statistically significant. But the ALP results showed a high degree of heterogeneity in UDCA ($p = 0.000$, $I^2 = 98.7\%$), OCA ($p < 0.01$, $I^2 = 97.5\%$), and bezafibrate ($p < 0.01$, $I^2 = 98.8\%$) (Figures 4A–C). However, the heterogeneity of the results of rifampicin was low ($p = 0.704$, $I^2 = 0.0\%$) (Figure 4D). We conducted sensitivity analysis (Supplementary Material S1) and subgroup analysis (according to UDCA dose, low: <13 *mg/kg/d*, medium: 13–15 *mg/kg/d*, high:>15 *mg/kg/d*) (Supplementary Material S2), study area (Asia, Europe, America) (Supplementary Material S3), and if cholestyramine was used as a combination (Supplementary Material S4).

A study (Wiesner et al., 1990) comparing cyclosporine with placebo found that, cyclosporine was superior to placebo in reducing ALP [SMD = −5.36, 95%CI (−6.98, −3.74), $p = 0.000$].

One study (Mayo et al., 2018) reported a comparison between NGM282 and placebo. The results suggested that NGM282 could significantly reduce the level of ALP in patients with PBC [SMD = −1.205, 95%CI (−1.98, −0.44), $p = 0.002$].

One study (Listed, 1993) reported a comparison between malotilate and placebo. The results showed no significant difference in the reduction of ALP between the two [SMD = −0.236, 95%CI (−0.63, 0.16), $p = 0.238$].

One study (Jones et al., 2017) reported a comparison between seladelpar and placebo, which suggested that seladelpar significantly reduced ALP [SMD = −2.224, 95% CI(−3.24, −1.21], $p = 0.000$).

One study (Mayo et al., 2019) on comparison between maralixibat and placebo concluded that maralixibat did not significantly reduce serum ALP [SMD = −0.183, 95%CI (−0.77, 0.40), $p = 0.540$].

A study (Vuoristo et al., 1995) on colchicine plus UDCA against UDCA alone showed that, compared with UDCA, colchicine plus UDCA significantly reduced ALP [SMD = −0.183, 95%CI (−0.77, 0.40), $p = 0.540$]. The other studies on methotrexate (Listed, 1993), colchicine (Almasio et al.,

TABLE 2 Summary of results for studies were not included in the meta-analysis.

Study ID	Intervention	Change in pruritus (event/noevent)	Change in ALP(U/L) (MD ± SD)	Change in γ-GGT (U/L) (MD ± SD)	Chang in adverse events (event/noevent)
F. Nevens 2016	OCA 10 mg	6/50	−130 ± 15	Not reported	Not reported
	Placebo	−19/28	−14 ± 15		
K.V. Kowdley 2018	OCA 10 mg	Not reported	−159 ± 67	Not reported	33/3
	Placebo		1 ± 17		21/2
Hirschfield 2014	OCA 10 mg	Not reported	Not reported	−92 ± 31	34/4
	Placebo			−1 ± 31	32/6
C. Corpechot 2018	Bezafibrate	Not reported	−60 ± 4	−38 ± 9	43/7
	Placebo		0 ± 9	7 ± 16	45/5
Tatsuo Kanda 2003	Bezafibrate + UDCA	1/6	−310 ± 234	−31 ± 20	1/10
	UDCA	0/5	−5 ± 219	−1 ± 14	0/11
Poupon RE 1990	UDCA	18/20	−417 ± 403	−413 ± 394	Not reported
	Placebo	10/20	20 ± 328	−32 ± 512	
Albert Parés 2000	UDCA	Not reported	−466 ± 56	−256 ± 33	9/90
	Placebo		−97 ± 64	−17 ± 30	6/87
P.M.Battezzati1993	UDCA	11/29	Not reported	Not reported	4/40
	Placebo	5/34			1/43
E.J Heathcote 1994	UDCA	24/63	−42 ± 99	Not reported	12/99
	Placebo	16/63	3 ± 98		19/92
K. D. Lindor1994	UDCA	27/21	−711 ± 811	Not reported	21/68
	Placebo	12/39	−300 ± 633		43/48
Laura Bachs 1989	Rifampicin	19/2	−316 ± 647	−99 ± 187	1/20
	Phenobarbitone	8/10	125 ± 795	173 ± 294	3/15
Study ID	Intervention	Change in pruritus (event/noevent)	Change in ALP (U/L) (MD ± SD)	Change in γ-GGT (U/L) (MD ± SD)	Chang in Adverse events (event/noevent)
Ana Podesta 1991	Rifampicin	14/0	−383 ± 683	Not reported	0/14
	Placebo	6/8	−92 ± 697		0/14
M. J. mayo 2018	NGM282 3 mg	5-D itch:-2.1 ± 4.7, VAS:-6.2 ± 15.6	−66 ± 57.7	−50.8 ± 71.9	22/8
	Placebo	5-D itch:-0.2 ± 3.8, VAS:-12.6 ± 24.9	3.3 ± 57.3	−5.6 ± 47.3	12/3
R. H.Wiesner 1990	Cyclosporine	16/1	−143 ± 623	Not reported	15/4
	Placebo	3/6	273 ± 345		5/5
A European Multi--centre study group	Malotilate	8/16	−61 ± 111	Not reported	13/39
	Placebo	6/20	−30 ± 150		2/47
David Jones 2017	Seladelpar200 mg (continue UDCA)	4/1	−156 ± 100	Not reported	12/12
	Placebo (continue UDCA)	3/1	4.7 ± 3.5		6/7

(Continued on following page)

10.3389/fphar.2022.835991

TABLE 2 (*Continued*) Summary of results for studies were not included in the meta-analysis.

Study ID	Intervention	Change in pruritus (event/noevent)	Change in ALP(U/L) (MD ± SD)	Change in γ-GGT (U/L) (MD ± SD)	Chang in adverse events (event/noevent)
M.J. mayo 2019		5-D itch score			
	Maralixibat 10 mg	−6.9 ± 6.8	−7.4 ± 80.4	28.6 ± 249.8	15/6
	Placebo	−6.3 ± 6.4	7.3 ± 80.0	45 ± 231	11/13
M.T.Hendrickse 1999		Pruritus scores	Not reported	Not reported	
	Methotrexate	−0.1 ± 0.2			21/9
	Placebo	0.1 ± 0.2			19/11
P. L. Almasio 2000	Colchicine + UDCA	9/16	Not reported	Not reported	6/40
	UDCA	9/18			11/33
V.S Hegade 2017		5-D itch scale	Not reported	Not reported	
	GSK2330672	6.5 ± 22.0			17/4
	Placebo	—			17/4

Abbreviations: MD, Mean Deviation; SD, Standard Deviation; VAS, Visual Analogue Score; Ursodeoxycholic acid; ALP, Alkaline phosphatase; γ-GGT, Gamma-glutamyltranspeptidase.

2000)and GSK2330672 (Hegade et al., 2017) did not report any change of ALP level after drug treatment.

Gamma-glutamyl transpeptidase

Blood γ-GGT levels were reported in 5 RCTs of UDCA (Poupon et al., 1990; Oka et al., 1991; Poupon et al., 1991; Vuoristo et al., 1995; Parés et al., 2000)and two of bezafibrate (Kanda et al., 2003; Corpechot et al., 2018). The results showed that UDCA and bezafibrate could reduce serum γ-GGT levels (UDCA [SMD = −2.18, 95%CI (−2.43, −1.93), p = 0.000], bezafibrate [SMD = −2.65, 95%CI (−4.34, 0.96), p = 0.002]) (Figures 5A,B). However, there was obvious heterogeneity for UDCA (p < 0.01, I^2 = 98.9%) and bezafibrate (p < 0.01, I^2 = 88.8%). Therefore, we conducted a subgroup analysis based on the dose of UDCA (Supplementary Material S5), study area (Supplementary Material S6) and whether cholestyramine was combined with UDCA (Supplementary Material S7), to evaluate its effect on serum γ-GGT. Sensitivity analysis showed that the results were consistent (Supplementary Material S8).

A study (Hirschfield et al., 2015) on the effect of OCA in serum γ-GGT showed that OCA reduced serum γ-GGT than placebo [SMD = −2.935, 95%CI (−3.59, −2.28), p = 0.000].

One study (Bachs et al., 1989) reported a comparison between rifampicin and placebo. The results suggested that the former could significantly reduce the level of serum γ-GGT [SMD = −1.123, 95%CI (−1.80, −0.44), p = 0.001].

One study (Mayo et al., 2018) reported a comparison between NGM282 and placebo. The results suggested that NGM282 can reduce the level of γ-GGT [SMD = −0.738, 95% CI (−1.47, −0.01), p = 0.047] in patients with PBC.

A study (Mayo et al., 2019) on the effect of maralixibat on serum γ-GGT, compared with placebo, showed no significant difference between the two drugs [SMD = −0.068, 95% CI (−0.65, 0.52), p = 0.819].

The other two studies of OCA (Nevens et al., 2016; Kowdley et al., 2018), three of UDCA (Battezzati et al., 1993; Lindor et al., 1994; Heathcote et al., 1995), one of rifampicin (Podesta et al., 1991), one of cyclosporine (Wiesner et al., 1990), one of malotilate (Listed, 1993), one of seladelpar (Jones et al., 2017), one of methotrexate (Hendrickse et al., 1999), one of colchicine (Almasio et al., 2000) and one of GSK2330672 (Hegade et al., 2017) did not report any change of serum γ-GGT level after treatment.

Adverse events

Compared with placebo, the incidences of adverse events with UDCA were lower [OR = 0.61, 95%CI (0.42, 0.89), p = 0.011], and there was no significant difference in OCA (OR = 1.03, 95%CI (0.61, 1.75), p = 0.901) and bezafibrate (OR = 0.99,

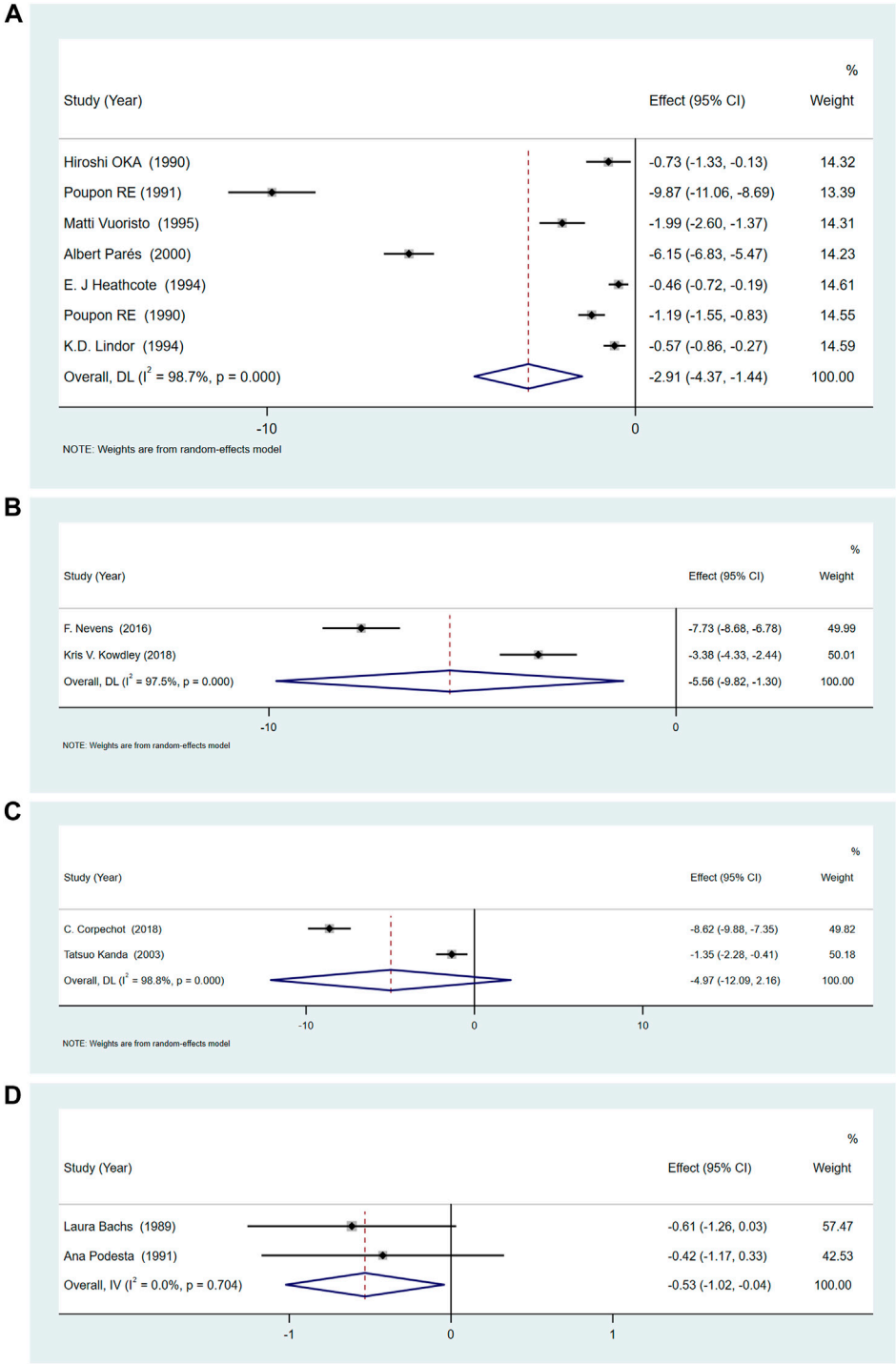

FIGURE 4
(A) The effect of UDCA in serum ALP. **(B)** The effect of OCA on serum ALP. **(C)** The effect of Bezafibrate on serum ALP. **(D)** The effect of Rifampicin on serum ALP.

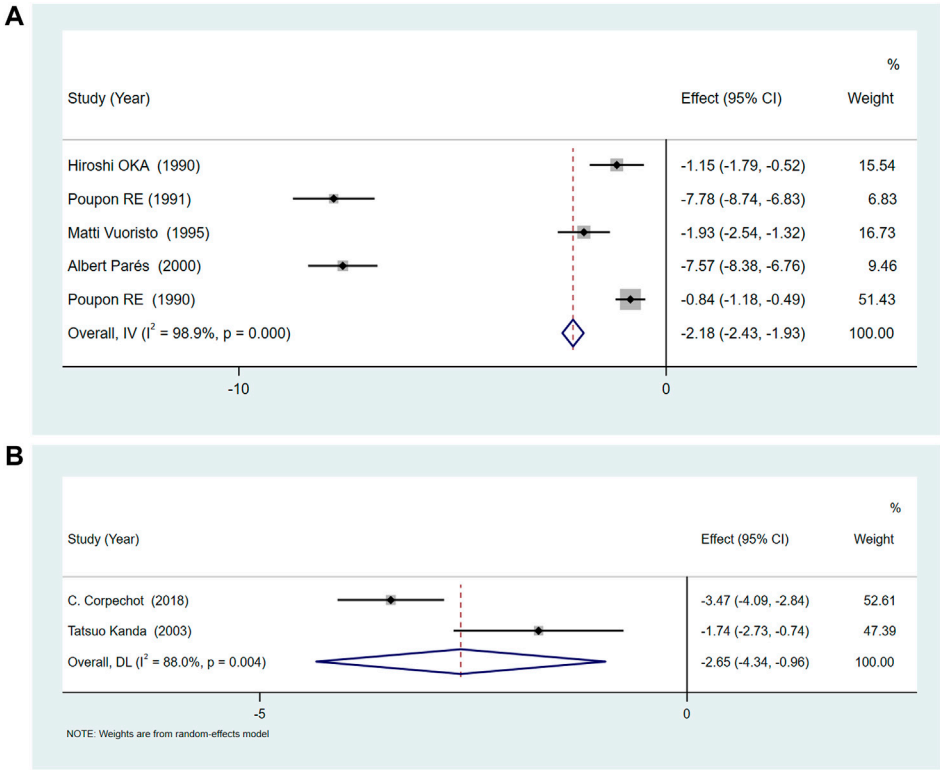

FIGURE 5
(A) The effect of UDCA on serum γ-GGT. **(B)** The effect of Bezafibrate on serum γ-GGT.

95%CI (0.56, 1.74), p = 0.967). The results showed that the heterogeneity was low, (for UDCA: p = 0.195 and I^2 = 32.0%, for OCA: p = 0.892 and I^2 = 0.0%, and for bezafibrate: p = 0.504, I^2 = 0.0%) (Figures 6A–C).

Sensitivity analysis of UDCA indicated that the results were consistent (Supplementary Material S9). Subgroup analysis based on UDCA dose (Supplementary Material S10), study area (Supplementary Material S11), year of publication (Supplementary Material S12) and whether UDCA was combined with cholestyramine (Supplementary Material S13), showed that the occurrence of adverse events was dose-dependent. Both high (>15 *mg/kg/day*) and low doses (<13 *mg/kg/day*) of UDCA increased the incidence of adverse events, while the middle dose (13–15 *mg/kg/day*) of UDCA did not increase the incidence of adverse events.

A study (Mayo et al., 2018) on the comparison of adverse reactions between NGM282 and placebo showed no significant difference [OR = 0.917, 95%CI (0.36, 2.34), p = 0.856].

One study (Wiesner et al., 1990) reported no significant difference in the incidence of adverse events when cyclosporine and placebo were compared [OR = 1.579, 95%CI (0.44, 5.62), p = 0.481].

A study (Listed, 1993) compared malotilate with placebo, showed that malotilate was superior than placebo in reducing adverse events [OR = 6.125, 95%CI (1.31, 28.52), p = 0.021].

Two separate studies (Jones et al., 2017; Mayo et al., 2019) reported no significant difference in the reduction of adverse events between seladelpar (MBX-8025) and placebo groups [OR = 1.820, 95%CI (0.59, 5.62), p = 0.298], and between Maralixibat and placebo [OR = 1.558, 95%CI (0.59, 4.13), p = 0.372].

Similarly, studies (Hendrickse et al., 1999; Almasio et al., 2000; Hegade et al., 2017) reported no significant differences for the adverse events between methotrexate and placebo [OR = 1.105, 95%CI (0.50, 2.46), p = 0.806], colchicine and placebo [OR = 0.522, 95%CI (0.18, 1.53), p = 0.236], and GSK2330672 and placebo [OR = 1.000, 95%CI (0.41, 2.47), p = 1.000].

Two studies (Bachs et al., 1989; Podesta et al., 1991) reported no significant difference in the reduction of adverse events between rifampicin and placebo [OR = 0.286, 95%CI (0.03, 2.99), p = 0.296].

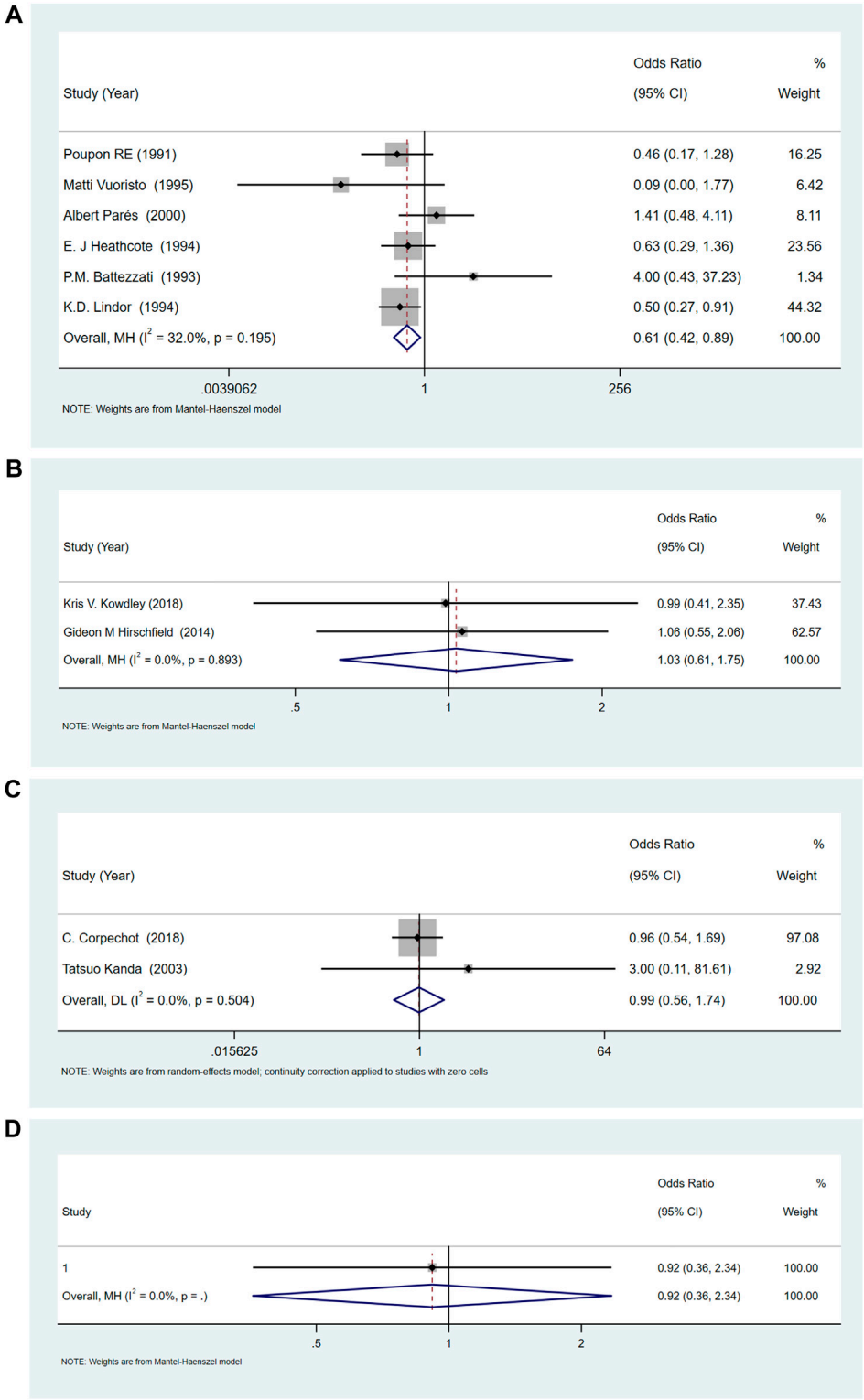

FIGURE 6
(A) The adverse events for UDCA. **(B)** The adverse events for OCA. **(C)** The adverse events for Bezafibrate. **(D)** The adverse events for NGM282.

Discussion

The rising incidence of pruritus in PBC and the lack of effective treatment methods were serious concerns. One study found that the incidence of pruritus can growing from 19% at the beginning to 80% in 10 years later (Prince et al., 2002). Pruritus in PBC was belong to cholestatic pruritus. Unlike histamine-related pruritus, cholestatic pruritus had no evidence-based guidelines because the underlying pathogenesis was unclear. Some studies had shown that itching symptoms often occured in female patients with PBC, accompanied by increasing levels of ALP and γ-GGT, but the mechanism behind it was still unknown. Existing studies had found that PBC-related pruritus was mainly related to bile acids, lysophosphatidic acid A (LPA), G-protein coupled bile acid receptor1 (GPBAR1), endogenous opioids, 5-hydroxytryptamine (5-HT), nitric oxide and substance P. But since the circulation level of these substances was not well correlated with the severity of pruritus, there may be complex interactions among multiple pruritus substances (Quarneti et al., 2015). Firstly, for the treatment of PBC-related pruritus, we need to rule out renal failure, psoriasis, idiopathic dermatitis and other diseases that can cause pruritus. Secondly, drugs treatment and non-drugs treatment measures should be adopted according to the severity of the disease. The first-line treatment drugs were mainly bile acid-binding resin (Gideon et al., 2017), such as colesevelam hydrochloride, which has better therapeutic effect than cholestyramine. Second-line therapeutic drugs such as rifampicin, opioid antagonists (naltrexone) and modulators of 5-HT receptor pathway (sertraline) can be used as supplements to first-line drugs. In addition, most of the non-drug treatments belong to invasive treatment strategies, and liver transplantation was the last choice for patients with intractable pruritus when all treatment strategies were ineffective. Unfortunately, due to lack of evidence-based evidence, the role of many drug interventions in the treatment of cholestatic pruritus, including PBC was uncertain. This result has also been confirmed in a recent related study (Dervout et al., 2022).

Moreover, drugs for PBC were indispensable in the treatment of pruritus. In recent years, newer target drugs were introduced, which mainly focus on reducing cholestasis and reduced bile acid toxicity, and are immunomodulatory in action and antifibrotic in nature (Shah and Kowdley, 2020). However, the efficacy of these drugs remains to be further evaluated. For example, OCA was a semisynthetic chenodeoxycholic acid analogue, which can inhibit bile acid synthesis and stimulate bile secretion to protect hepatocytes by activating farnesol X receptor. Amazingly, it was found that pruritus increased dose-dependently with OCA treatment (Gong et al., 2008; Trauner et al., 2019), and so it using in patients with PBC associated pruritus was limited. Systematic reviews and meta-analyses on UDCA showed that it improved the levels of serum ALP and γ-GGT, but there were no significant effects on pruritus, fatigue, or reduction in adverse events in PBC (Gong and Gluud, 2005). Similar results were obtained for methotrexate (Guo et al., 2015), fenofibrate (Zhang et al., 2015), and bezafibrate (Gong et al., 2007). Cyclosporin A may significantly improving pruritus, but it had significant side effects compared with placebo (Gerussi et al., 2021). However, we found that these studies did not consider pruritus as a primary outcome measure, and we need an effective and safe pharmacological intervention for pruritus associated with PBC was yet unmet.

Therefore, we conduct this systematic review and meta-analysis to find the effective and safety pharmacological intervention for managing pruritus in PBC.

Pruritus (relief rate)

This meta-analysis of 23 RCTs found that UDCA, methotrexate and GSK2330672 improved pruritus (comparing the pruritus relief rate before and after treatment). OCA may increase the risk of pruritus, hence it was not recommended for patients suffering from PBC pruritus. However, due to the limited number of studies, it was suggested that more RCTs be conducted to understand their role in improving the symptom of itching (Specific data are summarized in Table 2).

Serum alkaline phosphatase and γ-glutamyl transpeptidase

Serum ALP and γ-GGT were two important indicators for the diagnosis and prognosis of PBC[57], while these also serve as important markers to diagnose the existence of cholestasis. We analyzed the effects of drug intervention on ALP and γ-GGT. All the included studies reported that UDCA, OCA, rifampicin, cyclosporine, NGM282, seladelpar and colchicine may improve blood ALP. Further, UDCA, bezafibrate, OCA, rifampicin and NGM282 may improve blood γ-GGT. We found that rifampicin can significantly reduce the blood ALP level with low heterogeneity, while UDCA and OCA have high heterogeneity in reducing the level of ALP. Although UDCA and bezafibrate reduce the level of blood γ-GGT, they have high heterogeneity. Rifampicin and NGM282 were equally efficacious in reducing blood ALP and γ- GGT, which was an area worthy of further study.

Adverse events

The existing evidence showed that a medium dose of UDCA (13–15 *mg/kg/day*) and malotilate may have the benefit of reducing the incidence of adverse events ($p < 0.05$), and these studies showed low heterogeneity. It was also found that the other drugs that inculded did not significantly raising the incidence of adverse events compared with placebo ($p > 0.05$).

Conclusion

In this study, we attempted to evaluate the efficacy and safety of drug interventions in the treatment of PBC-associated pruritus. It was found that UDCA, methotrexate, and GSK2330672 may improve pruritus, but due to the existence of literature quality and heterogeneity of the included drugs, we cannot recommend some therapeutic drugs in line with clinical practice. Notably, with the continuous accumulation of high-quality clinical trial evidence of some emerging drugs such as bezafibrate and ileal apical bile acid transporter inhibitors (GSK2330672), perhaps in the near future, PBC-related pruritus will get more high-quality evidence and standardized treatment, and the rate of utilization of liver transplantation will become lower.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

Conceptualization: CX and RY. Date extracted: CX and XL. Assessed evidence: XL and CX. Date analyzed: CX and XL. Writing-review and editing: CX and MD. Interpreted the results: SW.

Funding

This project is funded by the National Natural Science Foundation of China (No. 81774279). The sponsors are not involved in the design, execution, or writing the study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2022.835991/full#supplementary-material

References

Almasio, P., Floreani, A., Chiaramonte, M., Provenzano, G., Battezzati, P., CrosignAni, A., et al. (2000). Multicentre randomized placebo-controlled trial of ursodeoxycholic acid with or without colchicine in symptomatic primary biliary cirrhosis. *Aliment. Pharmacol. Ther.* 14 (12), 1645–1652. doi:10.1046/j.1365-2036.2000.00869.x

Bachs, L., Parés, A., Elena, M., Pares, A., and Rodes, J. (1989). Comparison of rifampicin with phenobarbitone for treatment of pruritus in biliary cirrhosis. *Lancet* 1 (8638), 574–576. doi:10.1016/s0140-6736(89)91608-5

Battezzati, P. M., Podda, M., Bianchi, F. B., NaccaRato, R., OrlandiF.Surrenti, C., et al. (1993). Ursodeoxycholic acid for symptomatic primary biliary cirrhosis. Preliminary analysis of a double-blind multicenter trial. Italian Multicenter Group for the Study of UDCA in PBC. *J. Hepatol.* 17 (3), 332–338. doi:10.1016/S0168-8278(05)80214-4

Beuers, U., Kremer, A., Bolier, R., and Elferink, R. P. J. O. (2014). Pruritus in cholestasis: Facts and fiction. *Hepatology* 60 (1), 399–407. doi:10.1002/hep.26909

Bray, G., Heisel, W., Afshin, A., Jensen, M. D., Dietz, W. H., Long, M., et al. (2018). The science of obesity management: An endocrine society scientific statement. *Endocr. Rev.* 39 (2), 79–132. doi:10.1210/er.2017-00253

Carey, E. J., Ali, A. H., and Lindor, K. D. (2015). Primary biliary cirrhosis. *Lancet* 386 (10003), 1565–1575. doi:10.1016/s0140-6736(15)00154-3

Cn, G., and Sg, C. (1988). Treatment of pruritus in primary biliary cirrhosis with rifampin. Results of a double-blind, crossover, randomized trial. *Gastroenterology* 94 (2), 488–493. doi:10.1016/0016-5085(88)90442-8

Corpechot, C., Chazouillères, O., Rousseau, A., Le Gruyer, A., Habersetzer, F., Mathurin, P., et al. (2018). A placebo-controlled trial of bezafibrate in primary biliary cholangitis. *N. Engl. J. Med.* 378 (23), 2171–2181. doi:10.1056/NEJMoa1714519

Dervout, C., Boulais, N., Barnetche, T., Nousbaum, J. B., Brenaut, E., and Misery, L. (2022). Efficacy of treatments for cholestatic pruritus: A systemic review and meta-analysis. *Acta Derm. Venereol.* 102, adv00653. doi:10.2340/actadv.v102.310

Dg, A., J, T., Tetzlaff, J., Mulrow, C., Gotzsche, P. C., Ioannidis, J. P. A., et al. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration. *BMJ Clin. Res. ed.)* 339, b2700. doi:10.1136/bmj.b2700

Drazilova, S., Babinska, I., Gazda, J., Halanova, M., Janicko, M., Kucinsky, B., et al. (2020). Epidemiology and clinical course of primary biliary cholangitis in Eastern Slovakia. *Int. J. Public Health* 65 (5), 683–691. doi:10.1007/s00038-020-01391-6

Gerussi, A., Bernasconi, D., O'Donnell, S., Lammers, W. J., Van Buuren, H., Hirschfield, G., et al. (2021). Measurement of gamma glutamyl transferase to determine risk of liver transplantation or death in patients with primary biliary cholangitis. *Clin. Gastroenterol. Hepatol.* 19 (8), 1688–1697. e1614. doi:10.1016/j.cgh.2020.08.006

Gideon, H., Ulrich, B., Christophe, C., Pietro, I., David, J., Marco, M., et al. (2017). EASL Clinical Practice Guidelines: The diagnosis and management of patients with primary biliary cholangitis. *J. Hepatol.* 67 (1), 145–172. doi:10.1016/j.jhep.2017.03.022

Gong, Y., Christensen, E., and Gluud, C. (2007). Cyclosporin A for primary biliary cirrhosis. *Cochrane Database Syst. Rev.* 3, CD005526. doi:10.1002/14651858.CD005526.pub2

Gong, Y., and Gluud, C. (2005). Methotrexate for primary biliary cirrhosis. *Cochrane Database Syst. Rev.* 3 (3), CD004385. doi:10.1002/14651858.CD004385.pub2

Gong, Y., Zhi, B. H., Christensen, E., and Gluud, C. (2008). Ursodeoxycholic acid for primary biliary cirrhosis. *Cochrane Database Syst. Rev.* 3 (3), CD000551. doi:10.1002/14651858.CD000551.pub2

Gulamhusein, A. F., and Hirschfield, G. M. (2020). Primary biliary cholangitis: Pathogenesis and therapeutic opportunities. *Nat. Rev. Gastroenterol. Hepatol.* 17 (2), 93–110. doi:10.1038/s41575-019-0226-7

Guo, C., Zhang, Y., Lei, H., Wang, F., Chen, K., Li, J., et al. (2015). Combination therapy of fenofibrate and ursodeoxycholic acid in patients with primary biliary cirrhosis who respond incompletely to UDCA monotherapy: A meta-analysis. *Drug Des. devel. Ther.* 9, 2757–2766. doi:10.2147/DDDT.S79837

Heathcote, E. J., Cauch-Duder, K., Walker, V., Bailey, R. J., Blendis, L. M., Ghent, C. M., et al. (1995). The Canadian multi-center double-blind randomized controlled trial of ursodeoxycholic acid in primary biliary cirrhosis. *Hepatology* 19, 1149–1156. doi:10.1016/0270-9139(95)90074-8

Hegade, V. S., Kendrick, S., Dobbins, R. L., Miller, S. R., Thompson, D., Richards, D., et al. (2017). Effect of ileal bile acid transporter inhibitor GSK2330672 on pruritus in primary biliary cholangitis: A double-blind, randomised, placebo-controlled, crossover, phase 2a study. *Lancet* 389 (10074), 1114–1123. doi:10.1016/S0140-6736(17)30319-7

Hegade, V. S., Mells, G. F., Fisher, H., Kendrick, S., DiBello, J., Gilchrist, K., et al. (2019). Pruritus is common and undertreated in patients with primary biliary cholangitis in the United Kingdom. *Clin. Gastroenterol. Hepatol.* 17 (7), 1379–1387. e1373. doi:10.1016/j.cgh.2018.12.007

Hendrickse, M. T., Rigney, E., Giaffer, M. H., Soomro, I., Triger, D. R., Underwood, J. C., et al. (1999). Low-dose methotrexate is ineffective in primary biliary cirrhosis: Long-term results of a placebo-controlled trial. *Gastroenterology* 117 (2), 400–407. doi:10.1053/gast.1999.0029900400

Higgins, J. P., and Green, S. (2008). *Cochrane handbook for systematic reviews of interventions*. Wiley-Blackwell. doi:10.1002/9780470712184.ch1

Hirschfield, G. M., Mason, A., Luketic, V., Lindor, K., Gordon, S. C., Mayo, M., et al. (2015). Efficacy of obeticholic acid in patients with primary biliary cirrhosis and inadequate response to ursodeoxycholic acid. *Gastroenterology* 148 (4), 751–761. doi:10.1053/j.gastro.2014.12.005

Hu, M., I, H., S, L., Soderdahl, F., Thuresson, M., Wahlin, S., et al. (2019). Incidence, prevalence, and outcome of primary biliary cholangitis in a nationwide Swedish population-based cohort. *Sci. Rep.* 9 (1), 11525. doi:10.1038/s41598-019-47890-2

J, F., I, v. d. M., S, S., Ng, J., Angus, P., Lubel, J., et al. (2020). Increasing prevalence of primary biliary cholangitis in Victoria, Australia. *J. Gastroenterol. Hepatol.* 35 (4), 673–679. doi:10.1111/jgh.14924

Jones, D., Boudes, P. F., Swain, M. G., Bowlus, C. L., Galambos, M. R., Bacon, B. R., et al. (2017). Seladelpar (MBX-8025), a selective PPAR-δ agonist, in patients with primary biliary cholangitis with an inadequate response to ursodeoxycholic acid: A double-blind, randomised, placebo-controlled, phase 2, proof-of-concept study. *Lancet. Gastroenterol. Hepatol.* 2 (10), 716–726. doi:10.1016/S2468-1253(17)30246-7

Js, R., Mn, K., Bjelakovic, G., and Gluud, C. (2012). Ursodeoxycholic acid for primary biliary cirrhosis. *Cochrane Database Syst. Rev.* 12, CD000551. doi:10.1002/14651858.CD000551.pub3

Jw, O. D., Sp, P., Ac, A., Haigh, C. G., Wilkinson, J. R., and Williams, R. (2005). A controlled trial of ondansetron in the pruritus of cholestasis. *Aliment. Pharmacol. Ther.* 21 (8), 1041–1045. doi:10.1111/j.1365-2036.2005.02430.x

Kanda, T., Yokosuka, O., Imazeki, F., and Saisho, H. (2003). Bezafibrate treatment: A new medical approach for PBC patients? *J. Gastroenterol.* 38 (6), 573–578. doi:10.1007/s00535-002-1102-7

Kd, L., Cl, B., J, B., Levy, C., and Mayo, M. (2019). Primary biliary cholangitis: 2018 practice guidance from the American association for the study of liver diseases. *Hepatol. Baltim. Md.)* 69 (1), 394–419. doi:10.1002/hep.30145

Kowdley, K., Luketic, V., Chapman, R., Hirschfield, G. M., Poupon, R., Schramm, C., et al. (2018). A randomized trial of obeticholic acid monotherapy in patients with primary biliary cholangitis. *Hepatology* 67 (5), 1890–1902. doi:10.1002/hep.29569

Lee, J. Y., Danford, C. J., Trivedi, H. D., Tapper, E. B., Patwardhan, V. R., and Bonder, A. (2019). Treatment of fatigue in primary biliary cholangitis: A systematic review and meta-analysis. *Dig. Dis. Sci.* 64 (8), 2338–2350. doi:10.1007/s10620-019-5457-5

Lindor, K. D., Dickson, E. R., Baldus, W. P., Jorgensen, R. A., Ludwig, J., Murtaugh, P. A., et al. (1994). Ursodeoxycholic acid in the treatment of primary biliary cirrhosis. *Gastroenterology* 106 (5), 1284–1290. doi:10.1016/0016-5085(94)90021-3

Listed, N. A. (1993). The results of a randomized double blind controlled trial evaluating malotilate in primary biliary cirrhosis. A European multicentre study group. *J. Hepatol.* 17 (2), 227–235. doi:10.1016/s0168-8278(05)80043-1

Mayo, M. J., Pockros, P. J., Jones, D., Bowlus, C. L., Levy, C., Patanwala, I., et al. (2019). A randomized, controlled, phase 2 study of maralixibat in the treatment of itching associated with primary biliary cholangitis. *Hepatol. Commun.* 3 (3), 365–381. doi:10.1002/hep4.1305

Mayo, M. J., Wigg, A. J., Leggett, B. A., Arnold, H., Thompson, A. J., Weltman, M., et al. (2018). NGM282 for treatment of patients with primary biliary cholangitis: A multicenter, randomized, double-blind, placebo-controlled trial. *Hepatol. Commun.* 2 (9), 1037–1050. doi:10.1002/hep4.1209

Mj, M., I, H., S, S., Jacobe, H., Getachew, Y., and Rush, A. J. (2007). Sertraline as a first-line treatment for cholestatic pruritus. *Hepatology* 45 (3), 666–674. doi:10.1002/hep.21553

Nevens, F., Andreone, P., Mazzella, G., Strasser, S. I., Bowlus, C., Invernizzi, P., et al. (2016). A placebo-controlled trial of obeticholic acid in primary biliary cholangitis. *N. Engl. J. Med.* 375 (7), 631–643. doi:10.1056/NEJMoa1509840

Oka, H., Toda, G., Ikeda, Y., HashimotoN.Hasumura, Y., Kamimura, T., et al. (1991). A multi-center double-blind controlled trial of ursodeoxycholic acid for primary biliary cirrhosis. *Gastroenterol. Jpn.* 25 (6), 774–780. doi:10.1007/BF02779195

Parés, A., Caballería, L., Rodés, J., Bruguera, M., Rodrigo, L., García-Plaza, A., et al. (2000). Long-term effects of ursodeoxycholic acid in primary biliary cirrhosis: Results of a double-blind controlled multicentric trial. UDCA-cooperative group from the Spanish association for the study of the liver. *J. Hepatol.* 32 (4), 561–566. doi:10.1016/s0168-8278(00)80216-0

Podesta, A., Lopez, P., Terg, R., Villamil, F., Flores, D., Mastai, R., et al. (1991). Treatment of pruritus of primary biliary cirrhosis with rifampin. *Dig. Dis. Sci.* 36 (2), 216–220. doi:10.1007/BF01300759

Poupon, R. E., Balkau, B., Eschwège, E., and Poupon, R. (1991). A multicenter, controlled trial of ursodiol for the treatment of primary biliary cirrhosis. UDCA-PBC Study Group. *N. Engl. J. Med.* 324 (22), 1548–1554. doi:10.1056/NEJM199105303242204

Poupon, R. E., Eschwège, E., and Poupon, R. (1990). Ursodeoxycholic acid for the treatment of primary biliary cirrhosis: Interim analysis of a double-blind multicentre randomized trial. *J. Hepatol.* 11 (1), 16–21. doi:10.1016/0168-8278(90)90265-s

Prince, M., Chetwynd, A., Newman, W., Metcalf, J. V., and James, O. F. W. (2002). Survival and symptom progression in a geographically based cohort of patients with primary biliary cirrhosis: Follow-up for up to 28 years. *Gastroenterology* 123 (4), 1044–1051. doi:10.1053/gast.2002.36027

Quarneti, C., Muratori, P., Lalanne, C., Fabbri, A., Menichella, R., Granito, A., et al. (2015). Fatigue and pruritus at onset identify a more aggressive subset of primary biliary cirrhosis. *Liver Int.* 35 (2), 636–641. doi:10.1111/liv.12560

Rajaobelina, K., Dow, C., Romana Mancini, F., Dartois, L., Boutron-Ruault, M. C., Balkau, B., et al. (2019). Population attributable fractions of the main type 2 diabetes mellitus risk factors in women: Findings from the French E3N cohort. *J. Diabetes* 11 (3), 242–253. doi:10.1111/1753-0407.12839

Shah, R. A., and Kowdley, K. V. (2020). Current and potential treatments for primary biliary cholangitis. *Lancet. Gastroenterol. Hepatol.* 5 (3), 306–315. doi:10.1016/S2468-1253(19)30343-7

Sp, P., C, V., B, H., Meixiong, J., Dong, X., and Kwatra, S. G. (2019). Cholestatic pruritus: Emerging mechanisms and therapeutics. *J. Am. Acad. Dermatol.* 81 (6), 1371–1378. doi:10.1016/j.jaad.2019.04.035

Tajiri, K., and Shimizu, Y. (2017). Recent advances in the management of pruritus in chronic liver diseases. *World J. Gastroenterol.* 23 (19), 3418–3426. doi:10.3748/wjg.v23.i19.3418

Tb, V. I., Sa, H., Rs, C., and Tennent, D. M. (1961). The treatment of pruritus and hypercholesteremia of primary biliary cirrhosis with cholestyramine. *N. Engl. J. Med.* 265, 469–474. doi:10.1056/nejm196109072651004

Trauner, M., Nevens, F., Shiffman, M., Drenth, J. P. H., Bowlus, C. L., Vargas, V., et al. (2019). Long-term efficacy and safety of obeticholic acid for patients with primary biliary cholangitis: 3-year results of an international open-label extension study. *Lancet. Gastroenterol. Hepatol.* 4 (6), 445–453. doi:10.1016/s2468-1253(19)30094-9

Vuoristo, M., Farkkil, M., Karvonen, A. L., Leino, R., Lehtola, J., Makinen, J., et al. (1995). A placebo-controlled trial of primary biliary cirrhosis treatment with colchicine and ursodeoxycholic acid. *Gastroenterology* 108 (5), 1470–1478. doi:10.1016/0016-5085(95)90696-7

Wiesner, R. H., Ludwig, J., Lindor, K. D., Jorgensen, R. A., Baldus, W. P., Homburger, H. A., et al. (1990). A controlled trial of cyclosporine in the treatment of primary biliary cirrhosis. *N. Engl. J. Med.* 322 (20), 1419–1424. doi:10.1056/NEJM199005173222003

Zeng, N., Duan, W., Chen, S., Wu, S., Ma, H., Ou, X., et al. (2019). Epidemiology and clinical course of primary biliary cholangitis in the asia–pacific region: A systematic review and meta-analysis. *Hepatol. Int.* 13 (1333–1341), 788–799. doi:10.1007/s12072-019-09984-x

Zhang, Y., Chen, K., Dai, W., Xia, Y., Wang, F., Shen, M., et al. (2015). Combination therapy of bezafibrate and ursodeoxycholic acid for primary biliary cirrhosis: A meta-analysis. *Hepatol. Res.* 45 (1), 48–58. doi:10.1111/hepr.12373

frontiers | Frontiers in Medicine

TYPE Original Research
PUBLISHED 10 November 2022
DOI 10.3389/fmed.2022.827821

Check for updates

OPEN ACCESS

EDITED BY
Kevin Lu,
University of South Carolina,
United States

REVIEWED BY
Martin Dawes,
University of British Columbia, Canada
Therese Zink,
Brown University, United States

*CORRESPONDENCE
Luis González-de Paz
gonzalezdepaz@hotmail.com

SPECIALTY SECTION
This article was submitted to
Family Medicine and Primary Care,
a section of the journal
Frontiers in Medicine

RECEIVED 02 December 2021
ACCEPTED 25 October 2022
PUBLISHED 10 November 2022

CITATION
González-de Paz L, Kostov B, Freixa X, Herranz C, Lagarda E, Ortega M, Pérez E, Porcar S, Sánchez E, Serrato M, Vidiella I and Sisó-Almirall A (2022) Cost-accuracy and patient experience assessment of blood pressure monitoring methods to diagnose hypertension: A comparative effectiveness study. *Front. Med.* 9:827821.
doi: 10.3389/fmed.2022.827821

COPYRIGHT
© 2022 González-de Paz, Kostov, Freixa, Herranz, Lagarda, Ortega, Pérez, Porcar, Sánchez, Serrato, Vidiella and Sisó-Almirall. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Cost-accuracy and patient experience assessment of blood pressure monitoring methods to diagnose hypertension: A comparative effectiveness study

Luis González-de Paz[1,2]*, Belchin Kostov[3], Xavier Freixa[1], Carmen Herranz[1,2], Elena Lagarda[1], María Ortega[1], Elisa Pérez[1], Silvia Porcar[1], Eva Sánchez[1], Montserrat Serrato[1], Ingrid Vidiella[1] and Antoni Sisó-Almirall[1,2,4]

[1]Consorci d'Atenció Primària de Salut Barcelona Esquerra (CAPSBE), Barcelona, Spain, [2]Primary Healthcare Transversal Research Group, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain, [3]Department of Statistics and Operational Research, Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, [4]School of Medicine and Health Sciences, University of Barcelona, Barcelona, Spain

Objectives: Studies of the diagnosis of hypertension have emphasized long-term cost-effectiveness analysis, but the patient experience and costs of blood pressure monitoring methods at the diagnosis stage remain unclear. We studied four diagnostic methods: a new 1 h-automated office blood pressure (BP) monitoring, office BP measurement, home BP monitoring, and awake-ambulatory BP monitoring.

Methods: We carried out a comparative effectiveness study of four methods of diagnosing hypertension in 500 participants with a clinical suspicion of hypertension from three primary healthcare (PHC) centers in Barcelona city (Spain). We evaluated the time required and the intrinsic and extrinsic costs of the four methods. The cost-accuracy ratio was calculated and differences between methods were assessed using ANOVA and Tukey's honestly significant difference (HSD) *post-hoc* test. Patient experience data were transformed using Rasch analysis and re-scaled from 0 to 10.

Results: Office BP measurement was the most expensive method (€156.82, 95% CI: 156.18–157.46) and 1 h-automated BP measurement the cheapest (€85.91, 95% CI: 85.59–86.23). 1 h-automated BP measurement had the best cost-accuracy ratio (€ 1.19) and office BP measurement the worst (€ 2.34). Home BP monitoring (8.01, 95% CI: 7.70–8.22), and 1 h-automated BP measurement (7.99, 95% CI: 7.80–8.18) had the greatest patient approval: 66.94% of participants would recommend 1 h-automated BP measurement as the first or second option.

Conclusion: The relationship between the cost-accuracy ratio and the patient experience suggests physicians could use the new 1 h-automated BP measurement as the first option and awake-ambulatory BP monitoring in complicated cases and cease diagnosing hypertension using office BP measurement.

KEYWORDS

blood pressure, hypertension diagnosis, patient experience, cost analysis, primary care, cost-effectiveness research

Introduction

The diagnosis of hypertension requires accurate blood pressure (BP) monitoring (1). Currently, international guidelines accept four methods: the reference 24-h ambulatory blood pressure monitoring (24 h-ABPM), home blood pressure monitoring (HBPM), office blood pressure measurement (OBPM), and automated office blood pressure measurement (AOBP) (2–4).

In daily clinical practice, 24 h-ABPM is not feasible in all cases due to the cost, the limited availability of devices, and possible patient distress. Therefore, HBPM or OBPM is often used to diagnose hypertension (5, 6). OBPM results in a considerable proportion of false positive and false negative diagnoses in patients with white coat syndrome or masked hypertension. HBPM is a good alternative, although most patients do not have a device at home before the diagnosis and carry out BP readings in unsuitable conditions (7). Automated BP devices, such as AOBP, are considered unattended office BP monitoring (8). AOBP reduces the white coat effect and measurements are close to mean daytime pressures obtained with ABPM or HBPM. Physicians accept AOBP because it offers measurements in a relatively controlled environment, and the results are obtained rapidly (9). At present, the Canadian hypertension guidelines and the American Heart Association accept AOBP as a method of diagnosing hypertension (10, 11). However, AOBP is not uniform (12): the first described AOBP method recommended three to four BP readings over 5 min, while others suggest more (13).

A recent paper reported on 1-h Automated Office Blood Pressure (1 h-AOBP) measurement consisting of BP measurement for 1 h every 5 min using the 24-h ABPM device. 1 h-AOBP has a sensitivity of 76.6% (95% CI: 71.1–81.5), a specificity of 64.8% (95% CI: 57–72.1), with a diagnostic accuracy greater than OBP and HBPM (14). 1 h-AOBP showed similar results to HBPM and a reasonable variability compared with awake-ABPM. ABPM is the gold standard; however, it requires a device for the patient for 24 h, and the high incidence of hypertension does not allow all patients to undergo screening with this method (15). HBPM at home or the pharmacy is a good alternative; however, clinicians often suspect patients' technique is incorrect (2). OBPM has been reported to result in overdiagnosis in 15–30% of cases and a higher risk of white coat hypertension (5). Therefore, 1 h-AOBP, which uses the same devices as ABPM for 1 h, allows multiple daily screenings and avoids common disadvantages.

Physician decision-making when choosing a hypertension diagnosis method is likely to vary during hypertension follow-up (16). In primary healthcare (PHC) clinical practice or outpatient clinics, where patients are highly independent and the hypertension diagnosis is opportunistic when exploring risk factors, physician choices are mostly discretional. Family physicians and PHC nurses might prefer BP measurement methods according to individual patient choices when they are unable or unwilling to carry out a specific BP measurement method (17). However, there is a lack of evidence on patient preferences at the time of the diagnosis of hypertension. Physician decision-making in the choice of hypertension methods is also hampered by the lack of evidence on the direct costs because studies have focused on long-term cost-effectiveness (18, 19). Therefore, the objective of this study was to evaluate the factors that most influence physicians and patients when choosing a BP measurement method: the accuracy of the method, the costs, and the patient experience. We conducted a comparative effectiveness study to evaluate four hypertension diagnostic methods, a 1-h automated office blood pressure (1 h-AOBP), awake-ABPM, HBPM, and OBPM.

Materials and methods

Study context

We designed a comparative effectiveness study using the view of the National Academies of Science to examine the cost-accuracy and patient experience outcomes of the hypertension diagnosis (20). This comparative effectiveness study was part of a clinical trial carried out in three PHC centers in Barcelona (Spain). The results of the clinical trial are published (14), and

details of the study, the patient sample, and the characteristics of the protocol are described elsewhere (21).

The four BP methods are described elsewhere (22), and the BP methods studied were carried out according to clinical guideline recommendations (12). Participants underwent 24 h-ABPM, 1 h-AOBP, HBPM, and OBPM. 1 h-AOBP used a 24-h ABPM device set up to measure BP every 5 min for 1 h. During the hour, the patient remained in the waiting room or in a quiet consulting room in the PHC center, without walking actively, eating, or smoking. After 1 h, the device and arm cuff were removed. HBPM was carried out according to daily clinical practice: Participants were required to measure three BP readings over 3 days, but if they did not have a valid BP device at home, they were asked to record the BP in the pharmacy following the same requirements as at home. A summary of the characteristics of BP methods and requirements are in **Supplementary Table 1** and all BP diagnostic methods are described.

A total of 500 patients referred by family physicians for a hypertension diagnosis underwent the four BP methods. Exclusion criteria were severe physical or cognitive limitations, episodes of any arrhythmia, any disease-causing permanent tremor, arm circumference > 42 cm, arterial-venous fistula in the arm, mental disorders or intolerance to the BP measurement method, inability to attend the study at the PHC or programmed hospitalization during the study. Data collection commenced in June 2017 and the completion date was December 2019.

Data collection and outcomes

The time horizon of the study was determined as the time when the family physician or nurse required a BP test for hypertension to the obtention of the results. After each BP method, participants completed a questionnaire on their experience. This included comfort, time the test required, and the degree of recommendation. Five-point visual scales were used to assess comfort and time with higher points indicating better patient experience. The degree of recommendation was an item to rank the order of the four BP methods if patients had to recommend them to other patients. The patient experience data collection sheet is reported elsewhere, and further details of the data collection are published and public available (21). Accuracy figures were used as reported, using Awake-ABPM as the reference method, with a diagnostic accuracy of 100% (14). Accuracy reflected the probability that an individual would be correctly classified by the method.

The cost of the visit was calculated according to Catalan public healthcare system costs (23). Expenses reported by participants (if any) and the cost of travel to the PHC were calculated. In all cases, we considered a round trip for each visit. Public transport (metro and bus) costs were estimated (24). Taxi costs were calculated according to the price per minute, established by the Barcelona metropolitan area as € 0.56 per minute (25), and the cost of traveling in private cars was estimated per minute according to the scheme of allocation for travel expense of the Spanish Tax Agency (26). The cost-accuracy ratio estimated how much it cost to gain a unit of diagnostic accuracy. Additionally, the time was estimated in minutes as the time used to reach the PHC reported by participants added to an estimated time of 10 min per consultation. In the case of 1 h-AOBP we added 60 min because the test takes place during this time in the PHC.

Statistical analysis

We carried out a cost-accuracy study with a similar perspective to a cost-effectiveness analysis (27). Categorical variables were presented as absolute frequencies and percentages and continuous variables as mean and standard deviation (SD). The time necessary was calculated by adding the transport times and consultation times per patient; costs were calculated by adding the consultation rates, transportation costs, and patient expenses. The patient experience results from the questionnaire were transformed into interval scaling using a Rasch analysis (28), which first required verification of the goodness-of-fit of the data to the model. Patient experience logit units from the Rasch analysis were re-scaled to 0–10, where 10 meant the

TABLE 1 Participant characteristics.

Characteristics	*N* = 500
Sex, female	244 (48.8)
Age (years)	59.7 ± 14.1
Ethnicity	
Caucasian	434 (86.8)
Latin	53 (10.6)
Other	13 (2.6)
Educational level	
Without studies	15 (3.0)
Primary school	56 (11.2)
High-school (up to 16 years)	71 (14.2)
High school or vocational studies (up to 18 years)	129 (25.8)
University studies or higher	229 (45.8)
Time used to reach the primary health care facility (minutes)	**12.3 ± 8.9**
Transport used	
Walking	430 (86.0)
Private car	30 (6.0)
Public transport (metro, bus)	36 (7.2)
Taxi	4 (0.8)
Had a sphygmomanometer at home	**269 (53.8)**
Went to a pharmacy to check blood pressure (HBPM)	**242 (48.4)**
Mean price paid at pharmacy	**€0.6 ± 1.3**

Data are presented as *n* (%) or mean ± standard deviation.

best comfort and time experience and 0 the worst experience. Patient experiences of BP methods were studied by examining the overall mean experience and patient characteristics using ANOVA. We used sociodemographic variables to examine between-group differences: age, sex, educational level, and ethnicity. We analyzed the proportion of the rank order of the four BP methods with charts where the participant had to recommend each BP method to other patients.

The cost study was carried out using ANOVA tests with the mean costs and time of each BP method. In all analyses, the results were expressed as mean and 95% confidence intervals (95% CI). Differences between BP methods were examined using Tukey's honestly significant difference (HSD) post hoc test. The analysis was made using Winsteps Rasch software and R version 3.6.0 for Windows (29, 30).

Ethics

The study was approved by the Ethics Committee of the Hospital Clinic of Barcelona (ref. number HCB/2014/0615) and registered at ClinicalTrials.gov (NCT03147573). All participants were informed of the study aims and provided written informed consent.

Results

Of the 500 patients included, 244 (48.8%) were female, with a mean age of 59.7 years (SD 14.1 years). Most participants had a tertiary education and reached the PHC center by walking < 13 min; they were not diagnosed with hypertension but 53.8% of participants reported having a sphygmomanometer at home. Participant characteristics are shown in **Table 1**.

The costs, diagnostic accuracy, and time study in **Table 2** shows that 1 h-AOBP was the cheapest (€85.91, 95% CI: €85.59–€86.23), while OBPM was the most expensive with a mean cost of €156.82.91 (95% CI: €156.18–€157.46). On average, HBPM required a shorter time to diagnosis than the other methods. The cost-accuracy ratio showed that HBPM, 1 h-AOBP and awake-ABPM were very similar, while the OBPM had an almost twofold higher cost-accuracy ratio.

According to patient experience, the highest rated BP methods were HBPM and 1 h-AOBP with 8 out of 10, and the least valued was awake-ABPM with a mean score of 5.19 (95% CI: 4.98–5.40). Pairwise comparison analysis showed greater differences (> 2 points) in comparisons of methods with awake-ABPM. The results are shown in **Table 3**. ANOVA tests of the mean experience by BP method and group characteristics showed significant but not relevant differences (<1 point) by age group in OBP, HBPM, and 1 h-AOBP. The results of this analysis are shown in **Supplementary Table 2**.

Figure 1 shows the differences in the distribution of the patient experience by BP method. The HBPM and 1 h-AOBP produced almost the same experience, while the worst rated was awake-ABPM. The stacked bar chart in the **Figure 2** shows that 66.94% of participants would recommend 1 h-AOBP to other patients as the first or second option. **Figure 3** shows the relationship between the cost-accuracy ratio, and the patient experience. BP methods with better cost-accuracy figures (close to 1 in the x axis) were 1 h-AOBP and HBPM, with a similar ratio to the gold-standard of awake-ABPM. However, 1 h-AOBPM and HBPM showed the best relationship according to the cost-accuracy ratio and high patient experience (y axis).

Discussion

We evaluated four methods of diagnosing hypertension according to cost-accuracy and patient experience. The results showed 1 h-AOBP had the best cost-accuracy, and the patient experience was similar to HBPM. 1 h-AOBP and HBPM showed the best outcome indicators compared with OBPM and awake-ABPM.

TABLE 2 Cost and time study of diagnosing hypertension.

Method	Total transport costs [95% CI]	Time necessary in minutes (consultations and transport) [95% CI]	Total cost [95% CI]	Cost-accuracy ratio [95% CI]
Awake-ABPM	€ 1.37 [0.89–1.85]	103.33 [98.60–108.06]	€ 121.38 [120.89–121.85]	1.21 [1.24–1.25]
OBPM.	€ 1.82 [1.18–2.46]	137.78 [131.46–144.09]	€ 156.82 [156.18–157.46]	2.34 [2.33–2.34]
HBPM	€ 0.91 [0.59–1.23]	68.88 [65.73–72.04]	€ 86.23 [85.90–86.56]	1.25 [1.24–1.25]
1 h-AOBP.	€ 0.91 [0.59–1.23]	118.88 [115.73–122.04]	€ 85.91 [85.59–86.23]	1.19 [1.24–1.25]

Data are presented as mean [95% confidence intervals]. Awake-ABPM, Awake ambulatory blood pressure monitoring; OBPM, Office blood pressure measurement; HBPM, Home blood pressure monitoring; 1 h-AOBP, 1 h- automated office blood pressure measurement.

TABLE 3 ANOVA test and pairwise comparisons (Tukey HSD tests) to examine the patient experience of each BT method.

BP test	Patient experience (handling and time 0–10 scale) mean and [95% CI]	*F*	*P*-value
Awake-ABPM	5.19 [4.98–5.40]	151.8	<0.001
OBPM.	7.32 [7.55–7.09]		
HBPM.	8.01 [7.70–8.22]		
1 h-AOBP.	7.99 [7.80–8.18]		
Pairwise comparisons	**Mean difference and 95% CI**		***P*-value**
OBPM vs. awake-ABPM	2.14 [1.74–2.53]		<0.001
1 h-AOBP vs. awake-ABPM	2.80 [2.41–3.19]		<0.001
1 h-AOBP vs. OBPM	0.67 [0.27–1.06]		<0.001
HBPM vs. awake-ABPM	2.82 [2.43–3.22]		<0.001
HBPM vs. OBPM	0.68 [0.29–1.08]		<0.001
HBPM vs. 1 h-AOBP	0.02 [−0.38 to 0.41]		0.999

Patient experience

Patient preferences for handling and the time required was clearly favorable to 1 h-AOBP and HBPM. The lower patient endorsement of OBPM might be influenced by the number of consultations and the time required (30). In contrast, awake-ABPM, although it only requires two consultations, has the drawback of wearing the BP measurement cuff and device for 1 day, meaning patients cannot carry out their daily activities (e.g., ride a motorcycle or any activities involving physical effort) (31). Patients clearly endorsed 1 h-AOBP and HBPM over other BP methods, and this was not associated with patient characteristics. Therefore, 1 h-AOBP and HBPM may be the best methods for physicians based on the patient experience. Studies of patient experiences in the diagnosis of hypertension show that HBPM is better accepted than OBPM, mainly because of patient confidence in the accuracy of BP measurement (32). However, a nationwide physician survey in Canada—where AOBP is most established, since the Canadian Hypertension guidelines included it as a valid method—showed that the preferred method of hypertension screening remains OBPM carried out with mercury devices (54.2%) compared with AOBP (38.8%) (33). A similar trend was observed in a survey carried out in a large sample of physicians in Hong Kong with a lower proportion of AOBP use (13.7%) and no use of HBPM (34). In Europe, a recent BP monitoring and measurement method guideline incorporated AOBP, although with relatively little information on it (35). While OBPM has the worst indicators and HBPM at the diagnostic stage might cause suspicion in clinicians, we suggest that our results might persuade healthcare

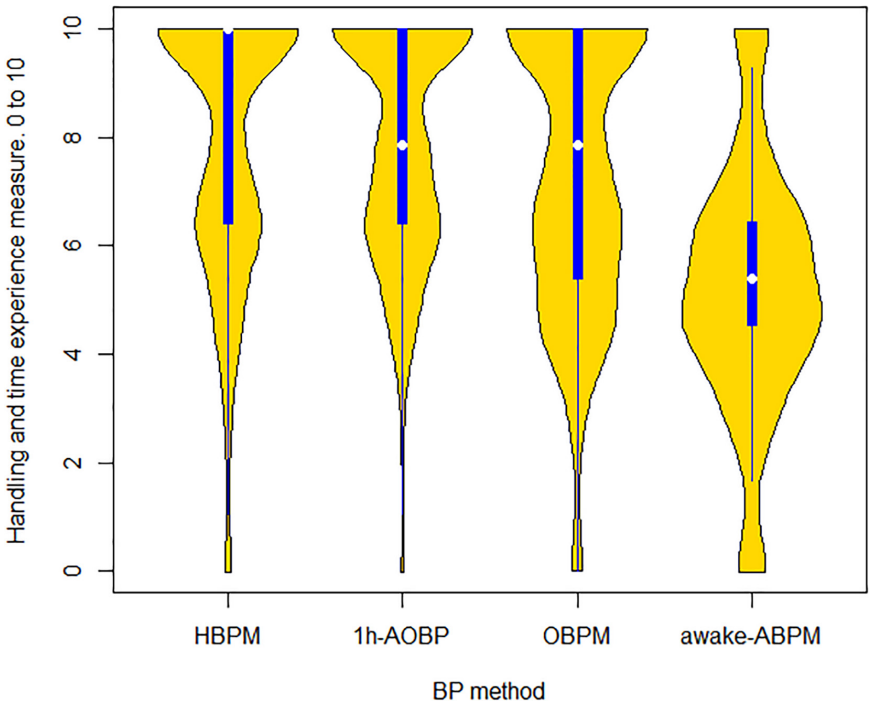

FIGURE 1
Distribution plot with the experience (0–10 scale) by BP method. The yellow shape shows density of the data at different values of patient experience, and the blue line and white dot are the inter-quartile rank and median, respectively.

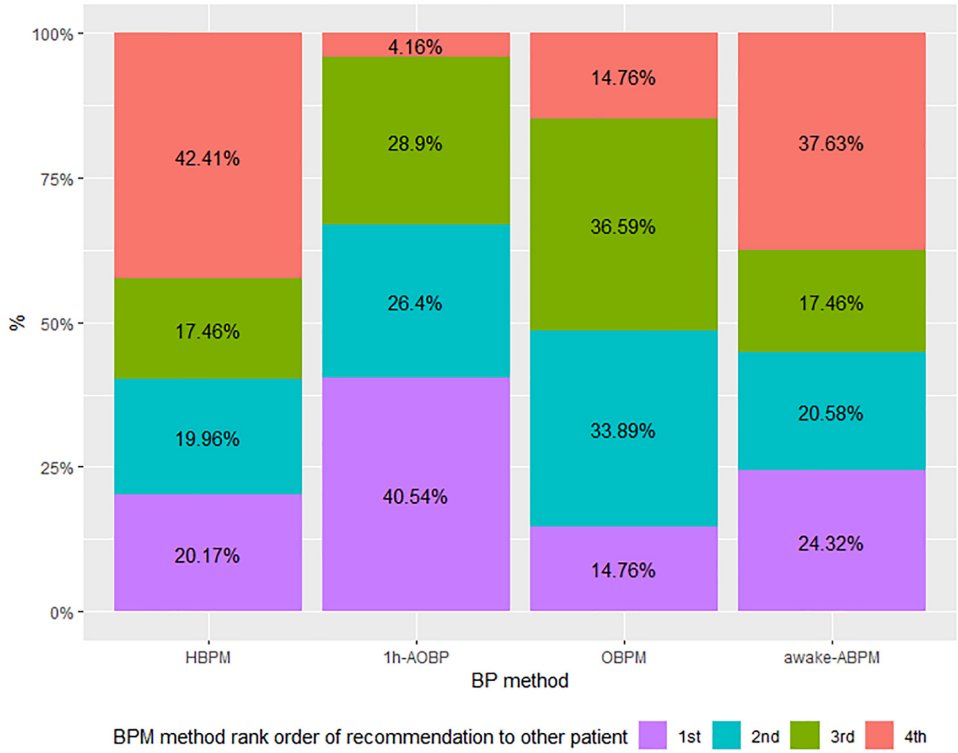

FIGURE 2
Lower part: staked bar chart with participants rank order of recommendation of each BP method to other patients. HBPM, Home blood pressure monitoring. 1 h-AOBP, 1 h- automated office blood pressure measurement; OBPM, Office Blood Pressure measurement; Awake-ABPM, Awake ambulatory blood pressure monitoring.

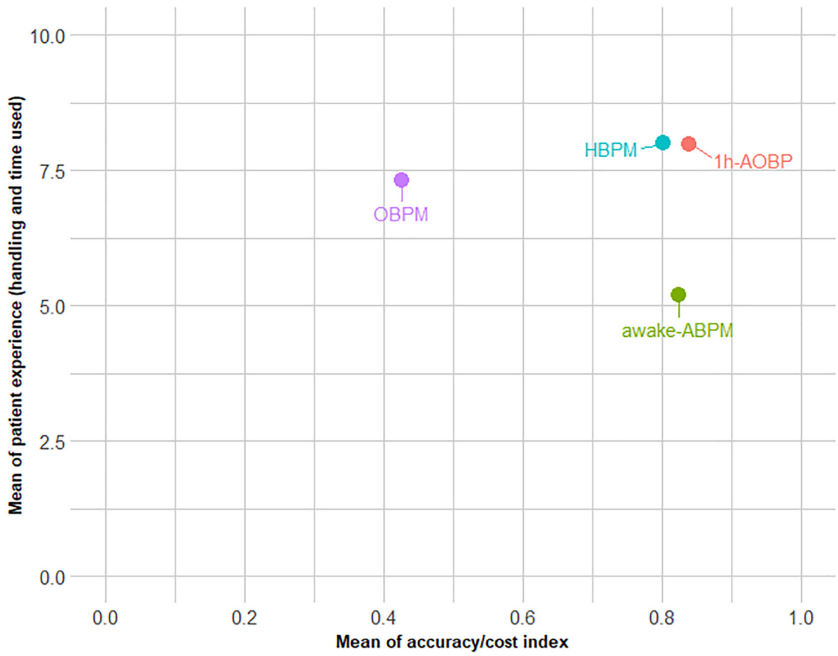

FIGURE 3
Relationship between the cost-accuracy ratio and the patient experience. HBPM, Home blood pressure monitoring. 1 h-AOBP, 1 h- automated office blood pressure measurement; OBPM, Office Blood Pressure measurement; Awake-ABPM, Awake ambulatory blood pressure monitoring.

providers and physicians to opt for new AOBP methods when selecting the best monitoring to diagnose hypertension.

Cost of diagnosing hypertension

The cost study showed 1 h-AOBP had results similar to HBPM, and the most expensive method was OBPM. Likewise, the cost-accuracy analysis showed OBPM was almost twice as good as the other methods, while 1 h-AOBP and awake-ABPM had the best cost-accuracy values. Some reports advise AOBP in all patients due to time or cost factors (31). Our results show that AOBP is cheaper than other diagnostic methods. To our knowledge, this approach to the effectiveness of BP measurement methods has not been studied, as researchers have prioritized cost-effectiveness (18). The evaluation of diagnostic methods often stops after quantifying the accuracy of the method compared with the gold-standard test (36). However, in the diagnosis of hypertension, BP measurement is linked to the patient experience and direct and indirect patient costs (37), and, therefore, incorporating the patient experience within the cost-accuracy framework is relevant in the evaluation of replacing OBPM by 1 h-AOBP in daily clinical practice: 1 h-AOBP favors the three domains (cost, patient outcomes and diagnostic accuracy).

The difference between the most expensive and the cheapest BP method (OBPM vs. 1 h-AOBP) was €-70.91. These results reflect real costs and have not been inflated but emerged from the reality of the well-established Spanish PHC system. In Spain, PHC is a core element of the health system, and is provided by specialist family physicians and nurses (38). PHC is managed, with few exemptions, by public trusts funded from the public budget according to block grants, including acute and chronic care, health promotion and prevention activities (39). While the benefits are not subject to patient cost-sharing, our results show that patients share indirect costs (e.g., number of visits). The results on effectiveness, in terms of real costs, of a PHC system according to the case study of diagnosing hypertension with different BP methods meant we adapted the number of visits to daily clinical practice differences in reimbursements, which might be higher but might not alter the ranking of costs of BP methods. We propose a feasible strategy to use 1 h-AOBP according to the results and the sensibility and specificity of 1 h-AOBP (12): (1) If hypertension is suspected—not restricted to opportunistic screening in the office—the physician should program 1 h-AOBP, asking the patient to bring the HBPM results. (2) Once the physician has the 1 h-AOBP and HBPM results they will decide whether hypertension can be diagnosed or, (3) in case of doubt, program 24 h-ABPM as a confirmatory method. We suggest this strategy could deal with some of the problems derived from false negatives and false positives and reduce the difference in cost-utility with awake or 24 h-ABPM (15).

Limitations and strengths of the study

The study has some limitations. We carried out a cost-accuracy and patient experience study to diagnose hypertension, but not the follow up, and only studied direct costs. A health technology assessment modeling study might answer the question of the indirect costs, but this was not the aim of our study. We aimed to motivate implementation of the most cost-accurate BP method to diagnose hypertension and suggest this is the first step to plan a cost-effectiveness study of 1 h-AOBPM. However, there is abundant literature on the follow up of hypertension, and the diagnostic stage is critical for subsequent monitoring. The reference test to diagnose hypertension is the 24 h-ABPM has limitations with the availability of the devices. However, this study examined the cost-accuracy and patient experience indicators of the four most important BP methods to diagnose hypertension. The results provide evidence of all advantages and drawbacks to help clinicians make better-informed decisions. The associated costs correspond to PHC and may seem low. However, we suggest the relative differences between BP methods can be adapted to other healthcare provision, including hospitals. Although the diagnosis of hypertension is highly dependent on the physician, our results come from real-life conditions and therefore physicians and healthcare managers could use the results to establish standardized procedures.

Conclusion

We found 1 h-AOBP had the best cost-accuracy ratio, with patient time and comfort indicators similar to HBPM. The 1 h-AOBP indicators of cost-accuracy and patient experience might impact the speed of the diagnosis. Our findings are important with respect to clinical strategies, as OBPM as a diagnostic strategy for hypertension in PHC could cease, as it has the highest costs, the lowest accuracy and is not well accepted by patients. 1 h-AOBP might reduce healthcare diagnostic costs and improve patients' healthcare experience.

Data availability statement

The raw data supporting the conclusions of this article will available by the authors upon reasonable request.

Ethics statement

The studies involving human participants were reviewed and approved by the Hospital Clinic of Barcelona (Spain). The patients/participants provided their written informed consent to participate in this study.

Author contributions

LG-dP, BK, and AS-A: conceptualization, methodology, and writing—review and editing. BK: formal analysis. XF, CH, EL, MO, EP, SP, ES, MS, IV, LG-dP, and BK: resources and data curation. LG-dP: writing—original draft preparation. LG-dP and AS-A: funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding

This project received research grants from: The Carlos III Institute of Health, Ministry of Economy and Competitiveness (Spain), awarded on the 2016 call under the Health Strategy Action 2013–2016, within the National Research Program oriented to Societal Challenges, within the Technical, the Scientific and Innovation Research National Plan 2013–2016, with reference PI16/00660, co-funded with European Union ERDF funds (European Regional Development Fund). The Department of Health of the Generalitat de Catalunya, in the call corresponding to the year 2019 of the Strategic Plan of Research and Innovation in Health (PERIS) 2016–2020, with file code SLT008/18/00013.

Acknowledgments

We thank all patients participating in the study and David Buss for technical assistance.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Author disclaimer

The views expressed in this article are those of the authors and do not necessarily reflect the views of the funder agency.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2022.827821/full#supplementary-material

References

1. Stergiou GS, Palatini P, Asmar R, Bilo G, de La Sierra A, Head G, et al. Blood pressure monitoring: theory and practice. European society of hypertension working group on blood pressure monitoring and cardiovascular variability teaching course proceedings. *Blood Press Monit.* (2018) 23:1–8. doi: 10.1097/MBP.0000000000000301

2. Unger T, Borghi C, Charchar F, Khan NA, Poulter NR, Prabhakaran D, et al. 2020 international society of hypertension global hypertension practice guidelines. *Hypertension* (2020) 75:1334–57. doi: 10.1161/HYPERTENSIONAHA.120.15026

3. Whelton PK, Carey RM, Aronow WS, Casey DE, Collins KJ, Himmelfarb CD, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: executive summary: a report of the American college of cardiology/American heart association task. *Hypertension.* (2018) 71:1269–324. doi: 10.1161/HYP.0000000000000066

4. Williams B, Mancia G, Spiering W, Rosei EA, Azizi M, Burnier M, et al. 2018 ESC/ESH Guidelines for the management of arterial hypertension. *Eur Heart J.* (2018) 39:3021–104. doi: 10.1093/eurheartj/ehy339

5. Kronish IM, Kent S, Moise N, Shimbo D, Safford MM, Kynerd RE, et al. Barriers to conducting ambulatory and home blood pressure monitoring during hypertension screening in the United States. *J Am Soc Hypertens.* (2017) 11:573–80. doi: 10.1016/j.jash.2017.06.012

6. Carter EJ, Moise N, Alcántara C, Sullivan AM, Kronish IM. Patient barriers and facilitators to ambulatory and home blood pressure monitoring: a qualitative study. *Am J Hypertens.* (2018) 31:919–27. doi: 10.1093/ajh/hpy062

7. Stergiou GS, Kario K, Kollias A, McManus RJ, Ohkubo T, Parati G, et al. Home blood pressure monitoring in the 21st century. *J Clin Hypertens.* (2018) 20:1116–21. doi: 10.1111/jch.13284

8. O'Brien E, Dolan E, Stergiou GS. Achieving reliable blood pressure measurements in clinical practice: it's time to meet the challenge. *J Clin Hypertension.* (2018) 20:1084–8. doi: 10.1111/jch.13323

9. Myers MG, Asmar R, Staessen JA. Office blood pressure measurement in the 21st century. *J Clin Hypertens.* (2018) 20:1104–7. doi: 10.1111/jch.13276

10. Nerenberg KA, Zarnke KB, Leung AA, Dasgupta K, Butalia S, McBrien K, et al. Hypertension Canada's 2018 guidelines for diagnosis, risk assessment, prevention, and treatment of hypertension in adults and children. *Can J Cardiol.* (2018) 34:506–25. doi: 10.1016/j.cjca.2018.02.022

11. Muntner P, Shimbo D, Carey RM, Charleston JB, Gaillard T, Misra S, et al. Measurement of blood pressure in humans: a scientific statement from the American heart association. *Hypertension.* (2019) 73:E35–66. doi: 10.1161/HYP.0000000000000087

12. Stergiou GS, Palatini P, Parati G, O'Brien E, Januszewicz A, Lurbe E, et al. 2021 European society of hypertension practice guidelines for office and out-of-office blood pressure measurement. *J Hypertens.* (2021) 39:1293–302. doi: 10.1097/HJH.0000000000002843

13. Vischer AS, Hug R, Socrates T, Meienberg A, Mayr M, Burkard T. Accuracy of abbreviated protocols for unattended automated office blood pressure measurements, a retrospective study. *PLoS One.* (2021) 16:e0248586. doi: 10.1371/journal.pone.0248586

14. Sisó-Almirall A, Kostov B, Blat E, García N, de Andres B, Roura S, et al. Validity and reliability of 1-h automated office blood pressure measurement for the diagnosis of hypertension. *J Hypertens.* (2022) 40:453–61. doi: 10.1097/HJH.0000000000003029

15. Stergiou GS, Parati G, Vlachopoulos C, Achimastos A, Andreadis E, Asmar R, et al. Methodology and technology for peripheral and central blood pressure and blood pressure variability measurement. *J Hypertens.* (2016) 34:1665–77. doi: 10.1097/HJH.0000000000000969

16. Andrews JA, Weiner K, Will CM, Henwood F, Dickson JM. Healthcare practitioner views and experiences of patients self-monitoring blood pressure: a vignette study. *BJGP Open.* (2020) 4:1–9. doi: 10.3399/bjgpopen20X101101

17. Tompson AC, Schwartz CL, Fleming S, Ward AM, Hobbs FDR, Heneghan CJ, et al. Patient experience of home and waiting room blood pressure measurement: a qualitative study of patients with recently diagnosed hypertension. *Br J Gen Pract.* (2018) 68:E835–43. doi: 10.3399/bjgp18X699761

18. Beyhaghi H, Viera AJ. Comparative cost-effectiveness of clinic, home, or ambulatory blood pressure measurement for hypertension diagnosis in US adults: a modeling study. *Hypertension.* (2019) 73:121–31. doi: 10.1161/HYPERTENSIONAHA.118.11715

19. Costa D, Peixoto Lima R. Cost-effectiveness of ambulatory blood pressure monitoring in the management of hypertension. *Rev Port Cardiol.* (2017) 36:129–39. doi: 10.1016/j.repc.2016.09.007

20. Institute of Medicine. *Initial National Priorities for Comparative Effectiveness Research*. Washington, DC: National Academies Press (2009). 227 p. doi: 10.17226/12648

21. González-De Paz L, Kostov B, Alvira-Balada MDC, Colungo C, García N, Roura S, et al. Effectiveness of a new one-hour blood pressure monitoring method to diagnose hypertension: a diagnostic accuracy clinical trial protocol. *BMJ Open.* (2019) 9:e029268. doi: 10.1136/bmjopen-2019-029268

22. Xxxx. *Description of Blood Pressure Monitoring Methods*. (2022). Available online at: https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fcdn-links.lww.com%2Fpermalink%2Fhjh%2Fb%2Fhjh_2021_10_04_almirall_jh-d-21-00741_sdc1.docx&wdOrigin=BROWSELINK (accessed June 12, 2022).

23. Generalitat de Catalunya. Departament de Salut,. Ordre SLT/71/2020, de 2 de Juny, per la Qual es Regulen els Supòsits i conceptes Facturables i S'aproven els Preus Públics Corresponents als Serveis que Presta l'Institut Català de la Salut. (2020). Catalonia: Diari Oficial de la Generalitat de Catalunya

24. Transports Metropolitans de Barcelona. *T-casual Barcelona Metro Bus Tickets | Transports Metropolitans de Barcelona*. (2022). Available online at: https://www.tmb.cat/en/barcelona-fares-metro-bus/single-and-integrated/t-casual (accessed June 6, 2022)

25. Àrea Metropolitana de Barcelona [AMB]. *Fares – Taxi*. (2022). Available online at: https://taxi.amb.cat/en/usuari/tarifes (accessed June 6, 2022).

26. Spanish Tax Agency. *Allocations for Travel Expenses – Tax Agency*. (2022). Available online at: https://sede.agenciatributaria.gob.es/Sede/en_gb/manuales/ejercicio-2018/modelo-100/7-cumplimentacion-irpf/7_2-rendimientos-trabajo-personal/7_2_1-rendimientos-integros/7_2_1_2-dietas-gastos-viaje/asignaciones-gastos-locomocion.html (accessed June 6, 2022).

27. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *BMJ* (2013) 346:f1049. doi: 10.1136/bmj.f1049

28. Smith EV, Smith RM. *Introduction to Rasch Measurement: Theory, Models and Applications*. Maple Grove, MN: JAM Press (2004). 689 p.

29. Linacre, JM. *Winsteps® Rasch measurement computer program, V.5.2.3*. (2022). Portland, OR: Winsteps.com.

30. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing (2021).

31. Boonyasai RT, McCannon EL, Landavaso JE. Automated office-based blood pressure measurement: an overview and guidance for implementation in primary care. *Curr Hypertens Rep.* (2019) 21:29. doi: 10.1007/s11906-019-0936-9

32. Khaddage SJ, Patterson JA, Sargent LJ, Price ET, Dixon DL. Sex and age differences in ambulatory blood pressure monitoring tolerability. *Am J Hypertens.* (2021) 34:335–8. doi: 10.1093/ajh/hpaa182

33. Doane J, Buu J, Jason Penrod M, Bischoff M, Conroy MB, Stults B. Measuring and managing blood pressure in a primary care setting: a pragmatic implementation study. *J Am Board Fam Med.* (2018) 31:375–88. doi: 10.3122/jabfm.2018.03.170450

34. Kaczorowski J, Myers MG, Gelfer M, Dawes M, Mang EJ, Berg A, et al. How do family physicians measure blood pressure in routine clinical practice? National survey of Canadian family physicians. *Can Fam Phys.* (2017) 63:e193–9.

35. Lee EKP, Choi RCM, Liu L, Gao T, Yip BHK, Wong SYS. Preference of blood pressure measurement methods by primary care doctors in Hong Kong: a cross-sectional survey. *BMC Fam Pract.* (2020) 21:95. doi: 10.1186/s12875-020-01153-6

36. Barrett BJ, Fardy JM. *"Evaluation of Diagnostic Tests.," Methods in Molecular Biology*. Totowa, NJ: Humana Press Inc (2021). p. 319–33. doi: 10.1007/978-1-0716-1138-8_18

37. van den Bruel A, Cleemput I, Aertgeerts B, Ramaekers D, Buntinx F. The evaluation of diagnostic tests: evidence on technical and diagnostic accuracy, impact on patient outcome and cost-effectiveness is needed. *J Clin Epidemiol.* (2007) 60:1116–22. doi: 10.1016/j.jclinepi.2007.03.015

38. Bernal-Delgado E, Garcia-Armesto S, Oliva J, Sanchez Martinez FI, Repullo JR, Pena-Longobardo LM, et al. Spain: health system review. *Health Syst Transit.* (2018) 20:1–179.

39. OECD. *Expenditure on primary health care*. Paris: OECD (2020). doi: 10.1787/c050dd46-en

Frontiers in Pharmacology

Explores the interactions between chemicals and living beings

The most cited journal in its field, which advances access to pharmacological discoveries to prevent and treat human disease.

Discover the latest Research Topics

See more →

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact