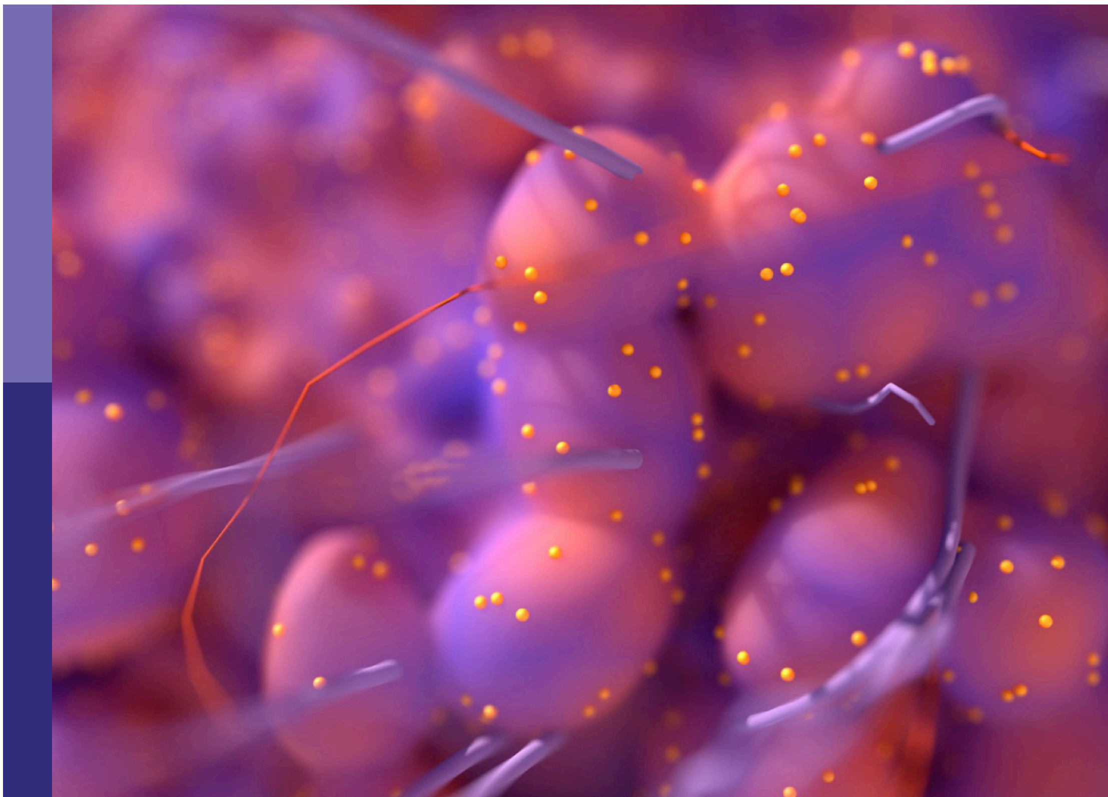
Research Topics

# Reviews in cancer imaging and image-directed interventions

**Edited by**
Omar Sultan Al-Kadi and Oliver Diaz

**Published in**
Frontiers in Oncology



ISSN 1664-8714
ISBN 978-2-8325-2061-1
DOI 10.3389/978-2-8325-2061-1

## 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

# Reviews in cancer imaging and image-directed interventions

**Topic editors**
Omar Sultan Al-Kadi — The University of Jordan, Jordan
Oliver Diaz — University of Barcelona, Spain

**Citation**
Al-Kadi, O. S., Diaz, O., eds. (2023). *Reviews in cancer imaging and image-directed interventions*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-2061-1

# Table of contents

frontiers | Frontiers in Oncology

TYPE Editorial
PUBLISHED 16 March 2023
DOI 10.3389/fonc.2023.1183302

# Editorial: Reviews in cancer imaging and image-directed interventions

Omar S. Al-Kadi[1]* and Oliver Diaz[2]*

[1]Department of Artificial Intelligence, King Abdullah II School for Information Technology, University of Jordan, Amman, Jordan, [2]Departament de Matemàtiques i Informàtica, Universitat de Barcelona, Barcelona, Spain

KEYWORDS

**cancer imaging, medical image analysis, artificial intelligence, image- guidance, cancer diagnostics, deep learning**

OPEN ACCESS

EDITED AND REVIEWED BY
Zaver Bhujwalla,
School of Medicine, Johns Hopkins University, United States

*CORRESPONDENCE
Omar S. Al-Kadi
✉ o.alkadi@ju.edu.jo
Oliver Diaz
✉ oliver.diaz@ub.edu

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 09 March 2023
ACCEPTED 13 March 2023
PUBLISHED 16 March 2023

CITATION
Al-Kadi OS and Diaz O (2023) Editorial: Reviews in cancer imaging and image-directed interventions.
*Front. Oncol.* 13:1183302.
doi: 10.3389/fonc.2023.1183302

Editorial on the Research Topic
Reviews in cancer imaging and image-directed interventions

This Research Topic is concerned with the application of medical imaging in the diagnosis and treatment of cancer. It includes the use of a range of imaging modalities, such as computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), optical coherence tomography (OCT), molecular luminescence spectroscopy, and their applications in detecting and characterizing tumors, as well as guiding interventions such as biopsies and targeted therapies. The reviews in this field aim to provide clinicians with a comprehensive understanding of the latest advances in cancer imaging and image-directed interventions, including the use of artificial intelligence (AI) and other emerging technologies. By keeping up with the latest research, clinicians can improve patient outcomes and enhance the overall quality of cancer care. The editorial article discusses 9 reviews, 7 systematic reviews, 1 mini review articles, covering organs such as liver, brain, lung, prostate, abdominal, and oral and maxillofacial regions.

Starting with the upper most part of the human body, brain cancer is a complex disease that can have a profound impact on individuals, and understanding its importance requires a comprehensive understanding of its causes, symptoms, and potential treatments. Xu et al. discusses the difficulty in diagnosing and managing glioma, and how medical imaging techniques like MRI, PET, and spectral imaging can aid physicians in treatment. The authors highlight the use of AI in medical imaging analysis, specifically in glioma diagnosis and management, such as tumor segmentation and classification, prediction of genetic markers, and treatment response and prognosis. However, the authors note that there are still issues to be solved with AI in clinical applications, such as data management, safety, and ethical and legal considerations. They suggest that interdisciplinary teamworks between clinicians and researchers are necessary to solve these issues in the future. Also, the meta-analysis in Jiang et al. evaluates the effectiveness of radiomics using non-enhanced computed tomography (NCCT) in predicting hematoma expansion in patients with spontaneous intracerebral hemorrhage. Ten articles comprising 1,525 subjects were analyzed, and the radiomics model showed an Area under the curve (AUC) of 0.80. Results revealed that the radiomics model outperformed most of the NCCT biomarkers in predicting hematoma expansion. The study suggests that the radiomics approach has the potential to predict hematoma expansion and is recommended over NCCT biomarkers. However, standardization of the radiomics pipeline is necessary for further clinical implementation.

Continuing in the upper part of the body, particularly the oral and maxillofacial region, Zhu et al. evaluated the accuracy of deep learning (DL) using the convolutional neural network VGGNet model in distinguishing benign and malignant thyroid nodules, based on ultrasound images. A total of 11 studies were included in the meta-analysis, and the overall estimates of sensitivity and specificity were 0.87 and 0.85, respectively. The results suggest that DL using the VGGNet model with ultrasound images performed good diagnostic efficacy in distinguishing benign and malignant thyroid nodules. Also, Wu et al. aimed to evaluate the diagnostic value of elastosonography for detecting salivary gland tumors and compare it to conventional ultrasound. This review analyzed 16 studies with a total of 1,105 patients and found that elastosonography had a pooled sensitivity of 0.73 and specificity of 0.64 for differentiating between benign and malignant tumors, with an AUC of 0.82. The study also found that quantitative or semi-quantitative elastosonography performed better than the qualitative one. The authors concluded that elastosonography could be considered a supplementary diagnostic technology to conventional ultrasound for detecting salivary gland tumors.

Moving down to the thoracic cavity, the review article by Van De Stadt et al. discuss the need for alternative biomarkers to predict tumor response to EGFR tyrosine kinase inhibitor therapy in NSCLC, highlighting the limitations of biopsies. PET studies using EGFR TKI-based tracers have shown promise in identifying EGFR mutational status and as a potential biomarker for tumor response. The article discusses currently investigated EGFR-directed PET biomarkers, their development process, and the advances, challenges, and opportunities for EGFR PET biomarkers to be used in routine clinical practice. Another work (Liang et al.) analyzed 19 original studies involving 2,444 patients and 3,012 subsolid pulmonary nodules (SSNs). They identified 18 clinical and CT features that correlated with SSN growth, including independent risk factors such as male sex, history of lung cancer, nodule size > 10 mm, nodule consistency, and age > 65 years. These findings can aid in establishing risk-based follow-up management strategies for SSN patients.

The abdominal cavity is the largest hollow space in the body, and detecting cancer in this area is crucial due to the vital organs located there, that are essential for proper body function. Moreover, detecting cancer in the abdominal cavity is challenging, particularly during its early stages. In Liang et al., radiomics-based MRI are evaluated for predicting microvascular invasion (MVI) in hepatocellular carcinoma (HCC) through a review of 15 studies involving 981 patients. The results show that radiomics-based MRI has high accuracy in predicting MVI in HCC, with a pooled sensitivity of 0.79, specificity of 0.81, and AUC of 0.87. Though there is heterogeneity among studies, sensitivity analysis supports the reliability of the results. The paper highlights the development of AI-based tools for liver cancer treatment. The use of interventional therapy for liver cancer is presented in Ren et al.. These tools can assist clinicians in making more precise diagnoses, treatment plans, and preventative measures for liver cancer patients, leading to more rational and personalized care. The article emphasizes that AI is bringing disruptive changes to the traditional medical model. Another systematic review article (Zhou et al.) compare two categorization systems used to diagnose hepatocellular carcinomas (HCCs) and determine their diagnostic performance. The systems compared were Contrast-enhanced ultrasound (CEUS) LI-RADS and CT/MRI LI-RADS. The study included 43 studies, and the results showed that CEUS LR-5 and CT/MRI LR-5 had similar diagnostic performance for HCCs, while CEUS LR-M had a higher proportion of HCCs and a lower proportion of non-HCC malignancies than CT/MRI LR-M. The study also found that CEUS LR-3 had a lower risk of HCCs than CT/MRI LR-3.

Moving to the undersurface of the right lobe of the liver, the work by Li et al. is interested in the application of DL in the imaging assessment of bladder cancer (BCa). DL has shown great potential in solving medical problems, particularly in the field of medical imaging. The authors provide an overview of current DL approaches used for bladder segmentation and how it helps in the diagnosis, staging, and treatment management of BCa. Also, Liu et al. discusses the application of DL in the diagnosis of gastrointestinal subepithelial lesions (SELs) using endoscopic ultrasonography (EUS). The study found that AI-assisted EUS is a promising and reliable method for distinguishing SELs, with excellent diagnostic performance, and is superior to EUS by experts. The authors recommend conducting more multicenter cohort and prospective studies to further develop AI-assisted real-time diagnostic systems and validate the superiority of AI systems. Furthermore, Liu et al. develop advanced ultrasound examination modes for diagnosing prostate cancer (PCa), including micro-Doppler, computerized-transrectal ultrasound, elastography, contrast-enhanced ultrasound, and microultrasound, collectively referred to as multiparameter ultrasound (mp-US). The combination of two or more of these modes can provide complementary information to multiparameter magnetic resonance imaging (mp-MRI) for diagnosing PCa. The authors suggest that mp-US has great potential as an imaging method for the diagnosis of PCa.

On the use of molecular luminescence for cancer imaging, the work by Chen et al. discusses the limitations of current surgical techniques for treating bone and soft tissue sarcoma and the potential for intraoperative fluorescence imaging to assist surgeons in determining tumor boundaries during surgery. The review considers the use of fluorescence imaging technology in clinical studies and assesses the potential of this technique to improve the accuracy of surgical resection. It suggests that intraoperative fluorescence imaging is a safe and straightforward technique that does not add any additional time to the surgery and has promising applications for the treatment of bone and soft tissue sarcoma. In the same realm, the work by Yang et al. shows recent advances in OCT modality, and its application in oncological diagnosis and treatment. The review highlights how OCT imaging can be used to detect and diagnose superficial and deep tumors in different types of cancers such as skin, gastrointestinal, brain, breast, bladder, and lung cancers, and how it can monitor tumor responses to treatments. Furthermore, the work in Yi et al. review molecular imaging techniques for cancer diagnosis and treatment, focusing on small-molecule inhibitors as cancer target probes. They summarize the structural designs of affinity probes based on small-molecule inhibitors and their impact on affinity and pharmacokinetics. The authors present clinical examples and provide insights for future research and clinical translations.

Others (Deng et al.) focused on the important role of NR4A1, a nuclear subfamily 4 receptor, in regulating metabolism in various cancers including melanoma, colorectal cancer, breast cancer, and hepatocellular cancer. NR4A1 has been found to mediate glycolysis, fatty acid synthesis, glutamine metabolism, and tumor immunity in cancer cells. The review suggests that regulating NR4A1 with novel ligands could be a promising approach to alter metabolism signaling pathways in cancer therapy.

Finally, the use of machine learning in cancer diagnostics, specifically focusing on the benefits of semi-supervised learning (SSL) compared to supervised learning (SL) is presented in Eckardt et al. SSL can use unlabeled samples in addition to labeled data for information abstraction, which allows for more efficient use of available data in cancer diagnostics. The article provides an overview of SSL functionalities and assumptions, and surveys key studies in image-based and non-image-based applications of SSL in cancer care, including histopathology, radiology, radiotherapy, and genomics. The authors highlight recent models and potential pitfalls in SSL study design, and suggest future directions for SSL in oncology.

We hope that this Research Topic will serve as a valuable resource for individuals interested in the important field of Cancer Imaging and Image-directed Interventions. It aims to present the most recent experimental methods used to explore fundamental concepts in Cancer Imaging and Image-directed Interventions, as well as to showcase the latest breakthroughs in the field. Furthermore, this topic underlines important areas for future research while also highlighting new clinical and therapeutic opportunities.

## Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

## 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

frontiers
in Oncology

SYSTEMATIC REVIEW
published: 29 March 2022
doi: 10.3389/fonc.2022.873913

# Risk Stratification and Distribution of Hepatocellular Carcinomas in CEUS and CT/MRI LI-RADS: A Meta-Analysis

*Yan Zhou[1,2†], Zhengyi Qin[2†], Jianmin Ding[2], Lin Zhao[2], Ying Chen[2], Fengmei Wang[1,3*] and Xiang Jing[2*]*

*[1] School of Medicine, Nankai University, Tianjin, China, [2] Department of Ultrasound, Tianjin Institute of Hepatobiliary Disease, Tianjin Key Laboratory of Extracorporeal Life Support for Critical Diseases, Artificial Cell Engineering Technology Research Center, Tianjin Third Central Hospital, Tianjin, China, [3] Department of Gastroenterology and Hepatology, Tianjin Third Central Hospital, Tianjin, China*

**Background:** CEUS LI-RADS and CT/MRI LI-RADS have been used in clinical practice for several years. However, there is a lack of evidence-based study to compare the proportion of hepatocellular carcinomas (HCCs) in each category and the distribution of HCCs of these two categorization systems.

**Purpose:** The purpose of this study was to compare the proportion of HCCs between corresponding CEUS LI-RADS and CT/MRI LI-RADS categories and the distribution of HCCs and non-HCC malignancies in each category.

**Methods:** We searched PubMed, Embase, and Cochrane Central databases from January 2014 to December 2021. The proportion of HCCs and non-HCC malignancies and the corresponding sensitivity, specificity, accuracy, diagnostic odds ratio (DOR), and area under the curve (AUC) of the LR-5 and LR-M categories were determined using a random-effect model.

**Results:** A total of 43 studies were included. The proportion of HCCs in CEUS LR-5 was 96%, and that in CECT/MRI LR-5 was 95% ($p > 0.05$). The proportion of non-HCC malignancy in CEUS LR-M was lower than that of CT/MRI LR-M (35% vs. 58%, $p = 0.01$). The sensitivity, specificity, and accuracy of CEUS LR-5 for HCCs were 73%, 92%, and 78%, respectively, and of CT/MRI LR-5 for HCCs, 69%, 92%, and 76%, respectively.

**Conclusion:** With the upshift of the LI-RADS category, the proportion of HCCs increased. CEUS LR-3 has a lower risk of HCCs than CT/MRI LR-3. CEUS LR-5 and CT/MRI LR-5 have a similar diagnostic performance for HCCs. CEUS LR-M has a higher proportion of HCCs and a lower proportion of non-HCC malignancies compared with CT/MRI LR-M.

**Keywords: contrast-enhanced ultrasound, contrast-enhanced magnetic resonance imaging, Contrast-enhanced computed tomography, hepatocellular carcinoma, Liver Imaging Reporting and Data System**

---

**Abbreviations:** HCC, hepatocellular carcinoma; LI-RADS, Liver Imaging Reporting and Data System; CEUS, contrast-enhanced ultrasonography; MRI, magnetic resonance imaging; CT, computed tomography; ACR, American College of Radiology; PPV, positive predictive value; CCRS, composite clinical reference standard.

OPEN ACCESS

***Edited by:***
*Chen Liu,*
*Army Medical University, China*

***Reviewed by:***
*Roberto Cannella,*
*University of Palermo, Italy*
*Wenwu Ling,*
*Sichuan University, China*

****Correspondence:***
*Xiang Jing*
*dr.jingxiang@aliyun.com*
*Fengmei Wang*
*wangfengmeitj@126.com*

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

***Specialty section:***
*This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology*

***Received:*** *11 February 2022*
***Accepted:*** *28 February 2022*
***Published:*** *29 March 2022*

***Citation:***
*Zhou Y, Qin Z, Ding J, Zhao L, Chen Y, Wang F and Jing X (2022) Risk Stratification and Distribution of Hepatocellular Carcinomas in CEUS and CT/MRI LI-RADS: A Meta-Analysis. Front. Oncol. 12:873913. doi: 10.3389/fonc.2022.873913*

## HIGHLIGHTS

- CEUS LR-3 has a lower proportion of HCCs than CT/MRI LR-M, while CEUS LR-M has a higher proportion of HCCs.
- Most of HCCs are in CEUS LR-5, LR-M, and LR-4, while most of HCCs are in CT/MRI LR-5 and LR-4.
- CEUS LR-M has a lower proportion of non-HCC malignancies but a higher proportion of HCCs compared with CT/MRI LR-M.

## INTRODUCTION

Contrast-enhanced computed tomography (CT), contrast-enhanced magnetic resonance imaging (MRI), and contrast-enhanced ultrasound (CEUS) were recommended by international guidelines to diagnose hepatocellular carcinomas (HCCs) (1, 2). To standardize the terminology, techniques, interpretation, reporting, and data collection of liver imaging, the American College of Radiology (ACR) released CT/MRI and the CEUS Liver Imaging Reporting and Data System (LI-RADS) (3, 4).

The contents of LI-RADS include the application of LI-RADS, techniques in different contrast-enhanced examinations, categorization, and management of lesions. According to the lesion size, major features, and ancillary features, lesions can be classified into different categories, including LR-1 to LR-5, LR-M, LR-TIV, and LR-NC. The clinical management for lesions of these categories was suggested by ACR LI-RADS based on the empirical risk of HCCs or malignancies. For example, lesions categorized into LR-5 can go through the management of HCCs without biopsy.

CT/MRI and CEUS LI-RADS were two independent systems with the same strata of categorization. Among LR-1 to LR-5 and LR-M, lesions in CEUS LI-RADS and CT/MRI LI-RADS have similar suggested management except for LR-3. This is because the positive predictive value (PPV) of HCCs is higher in CEUS LR-3 than that in CT/MRI LR-3 (3–6). Thus, the multidisciplinary discussion (MDD) was suggested additionally in CEUS LR-3. Otherwise, CT/MRI and CEUS LI-RADS are hypothesized to correspond to the same risk of HCCs in other categories. Up to now, there is a lack of evidence-based study to compare the proportion of HCCs in other categories and the distribution of HCCs of these two categorization systems (7–9). Moreover, whether the categories corresponding to similar suggestions of managements in CT/MRI LI-RADS and CEUS LI-RADS have a similar risk of HCCs is still unclear, which implies whether the suggested management is appropriate for LI-RADS categories also remains to be verified. In this meta-analysis, therefore, we aim to explore the risk and distribution of HCCs and non-HCC malignancies in each category of CEUS and CT/MRI LI-RADS and to explore the diagnostic performance of HCCs by LR-5 and of non-HCC malignancies by LR-M.

## METHOD

This meta-analysis was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) Statement (10). This study was registered at the Prospero International Prospective Register of Systematic Reviews (CRD42020175800).

### Literature Search Strategy

We searched corresponding studies from January 2014 to December 2021 in the PubMed, Embase, and Cochrane Central databases. The details of the strategy of searching are provided in **Supplementary Table 1**. Only English articles were included in this study. Case reports, reviews, letters, comments, and erratum were excluded.

### Inclusion and Exclusion Criteria

We included studies that met the following criteria: (1) patients with high risk for HCC; (2) the observations undergoing contrast-enhanced CT/MRI examination categorized according to CT/MRI LI-RADS V2014, V2017, or V2018, or the observations undergoing CEUS classified according to CEUS LI-RADS V2016 or V2017; (3) the contrast agent for CEUS being SonoVue; and (4) pathology or composite clinical reference standard (CCRS, multiple imaging or imaging follow-up) used as the reference standard. The exclusion criteria were as follows: (1) studies applied to patients without high risk for HCCs, (2) studies including duplicated data, (3) studies only including HCCs or HCCs and non-HCC malignancies, and (4) studies without sufficient data for inclusion in the pooled analysis.

### Study Selection

After excluding duplicates, two researchers independently reviewed the titles and abstracts of the articles. The full texts of the relevant articles were read to determine their inclusion. In the case of multiple studies from a center, we selected the most recent and complete one.

### Data Extraction

The following data were extracted from the included studies: (1) the characteristics of the study, including the first author, year of publication, nationality of patients, time of patient recruitment, and design (prospective or retrospective); (2) the characteristics of patients, including the number of patients, ages, and sexes; and (3) the tests to be evaluated, reference criteria, and results. The number of observations, HCCs, non-HCC malignancies, and benign lesions in each LI-RADS category was extracted from each study. If more than one data set was available in a study (e.g., different data from more than one viewer), the average data were adopted. Data extraction was conducted independently by the aforementioned two researchers, and no discrepancy was found during the process.

### Quality Assessment

The Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool (11) was used to evaluate the research

Quality. QUADAS-2 includes four aspects: patient selection, index test, reference standard, flow, and timing. For each aspect, the risk of bias was classified as high, low, or unknown. Two researchers independently assessed the risk of bias for each study, and any discrepancy was resolved by discussion with the third researcher. The results of the risk of bias assessments are shown in **Supplementary Figure 1**.

## Statistical Analysis

Random-effect models were used to evaluate the proportion of HCCs and non-HCC malignancies in each LI-RADS category, and the sensitivity, specificity, accuracy, and diagnostic odds ratio (DOR) of the LR-5 and LR-M categories, and to generate forest plots and 95% confidence intervals (95% CIs). The Q test and $I^2$ statistic were used to analyze the heterogeneity of the study, and $I^2$ >50% was considered to indicate heterogeneity (12). The variance of the logit-transformed percentage method was used to compare the differences in the proportions of HCCs and non-HCC malignancies in each category, and in the sensitivity, specificity, accuracy, and DOR of LR-5/M. The publication bias of the proportion of HCCs and non-HCC malignancies in each category was not evaluated according to the guidance of diagnostic test accuracy of systematic reviews (10). All statistical analyses were performed by the R language (v3.6.3, R Foundation for Statistical Computing, Vienna, Austria).

# RESULTS

A total of 786 studies were initially identified. 84 studies were then reviewed, and 59 studies were considered suitable for inclusion in this meta-analysis. After further excluding studies with insufficient data in the analysis, 43 studies were finally included (**Figure 1**) (13–55). Detailed information of the included and excluded studies is shown in **Supplementary Tables 2, 3**.

There were 15 studies on CEUS LI-RADS involving 6,573 patients with 7,234 lesions, including 5,387 HCCs, 624 non-HCC malignancies, and 1,223 benign lesions. There were 30 studies on CT/MRI LI-RADS involving 5,274 patients with 6,522 lesions, including 4,554 HCCs, 481 non-HCC malignancies, and 1,487 benign lesions.

## Proportions and Distribution of HCCs in Each CEUS and CT/MRI LI-RADS Category

A total of 5,387 HCCs in CEUS and 4,554 HCCs in CT/MRI can be used for the calculation of proportions and distribution of HCCs in each CEUS and CT/MRI LI-RADS category. There is no HCC in the CEUS and CT/MRI LR-1. The proportion of HCC gradually increases with the upshift of the category of both CEUS and CT/MRI LI-RADS. The proportions of HCCs of LR-2, 3, 4, and 5 were 1%, 21%, 75%, and 96% for CEUS LI-RADS and 4%, 35%, 75%, and 95% for CT/MRI LI-RADS, respectively. The proportion of HCCs in CEUS LR-3 is lower than that of CT/MRI LR-3 (21% vs. 35%, $p = 0.02$). The proportion of HCCs in CEUS LR-M is 56% (95% CI: 44%–69%), which is higher than that in CT/MRI LR-M, namely, 33% (95% CI: 22%–45%) ($p = 0.01$). The proportion of HCCs in each category is shown in **Table 1** and **Figure 2**. The forest plots of percentages of HCCs in CEUS and CT/MRI LR-5 are provided **Supplementary Figure 2**.

In CEUS LI-RADS, most of the HCCs are in LR-5 (68.5%). Most of the rest HCCs are in LR-M (15.8%) and LR-4 (10.8%). In CT/MRI LI-RADS, most of the HCCs are also in LR-5 (66.7%). Most of the remaining HCCs are in LR-4 (20.4%) but not LR-M (3.7%). There are more HCCs classified into CT/MRI LR-2, 3, and 4, compared with CEUS LR-2, 3, and 4, while there are more HCCs classified into CEUS LR-M than CT/MRI LR-M.

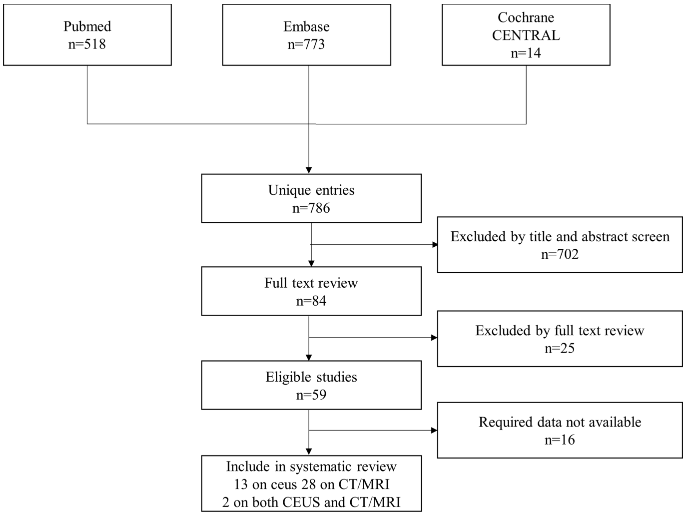


**FIGURE 1** | Flow diagram of the selection of studies.

**TABLE 1** | Proportions of HCCs in each CEUS and CT/MRI LI-RADS category.

| | CEUS | | | CT/MRI | | | |
|---|---|---|---|---|---|---|---|
| | **HCC (95% CI)** | **Observations** | **$I^2$, %** | **HCC (95% CI)** | **Observations** | **$I^2$, %** | **p** |
| LR-2 | 1 (0–6) | 134 | 0 | 4 (1–9) | 297 | 57 | 0.33 |
| LR-3 | 21 (13–31) | 670 | 78 | 35 (29–43) | 835 | 73 | 0.02 |
| LR-4 | 75 (61–85) | 735 | 88 | 75 (65–82) | 1299 | 89 | 0.99 |
| LR-5 | 96 (94–98) | 3858 | 89 | 95 (93–97) | 3205 | 83 | 0.46 |
| LR-M | 56 (44–69) | 1361 | 93 | 33 (22–45) | 490 | 80 | 0.01 |
| LR-TIV | 97 (77–100) | 100 | 0 | 72 (58–83) | 103 | 13 | 0.03 |

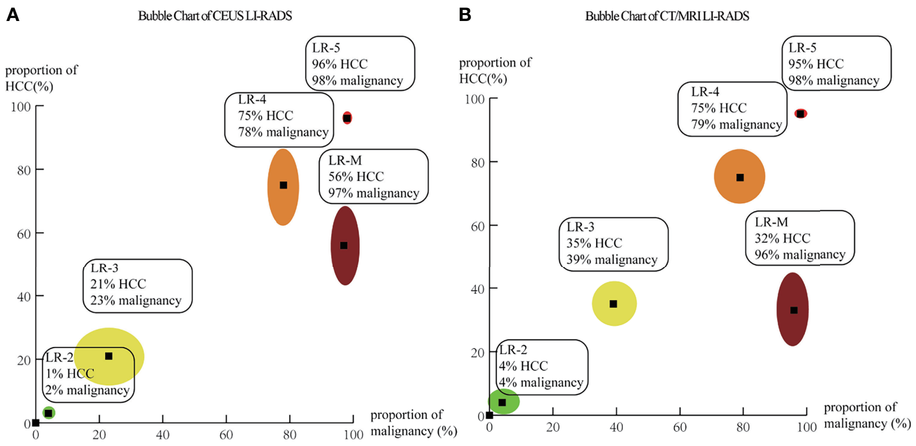


**FIGURE 2** | Bubble chart based on pooled percentage of HCCs and non-HCC malignancies for each category of CEUS **(A)** and CT/MRI **(B)** LI-RADS. The points at the centers of bubbles correspond to the pooled percentages of HCCs and non-HCC malignancies. The outer bubble margins correspond to 95% CIs for percentages of HCCs (y-axis) and non-HCC malignancies (x-axis).

The distributions of HCCs among different categories are shown in **Figure 3** and **Supplemental Table 4**.

## Proportions and Distribution of Non-HCC Malignancies in Each CEUS and CT/MRI LI-RADS Category

A total of 624 non-HCC malignancies in CEUS and 481 non-HCC malignancies in CT/MRI can be used for the calculation of proportions and distributions of non-HCC malignancies in each CEUS and CT/MRI LI-RADS category. There is no non-HCC malignancy in the CEUS and CT/MRI LR-1. The proportions of non-HCC malignancies in CEUS and CT/MRI LR-2 to LR-5 range from 1% to 5%. The proportion of non-HCC malignancies in CEUS LR-M is 35%, significantly lower than that of CT/MRI LR-M (58%, $p$ = 0.01). The proportions of non-HCC malignancies among different categories are shown in **Table 2**. The forest plots of percentages of non-HCC malignancies in CEUS and CT/MRI are depicted in **Supplementary Figure 3**.

In CEUS LI-RADS, most of the non-HCC malignancies are in LR-M (78.7%), most of the remaining non-HCC malignancies are in LR-5 (15.1%). In CT/MRI LR-RADS, most of the non-HCC malignancies are in LR-M (61.7%), and most of the remaining non-HCC malignancies are in LR-5 (18.3%) and LR-4 (9.9%). The distributions of non-HCC malignancies among different categories are shown in **Figure 4** and **Supplementary Table 5**.

## Meta-Regression for the Proportion of HCCs in Each LI-RADS Category

We analyzed the population (Asian or non-Asian), the reference standard (pathological or CCRS), and the version of LI-RADS used in these studies. The meta-regression results show that the proportion of HCCs in LR-3 for Asians is lower than that of LR-3 for non-Asians (14.3% vs. 32.3%, $p$ = 0.02). The proportion of HCCs in CEUS LR-M for Asian is higher than that of LR-M for non-Asian (67.3% vs. 35.7%, $p$ < 0.01).

The meta-regression results for CT/MRI LI-RADS show that the proportions of HCCs in CT/MRI LR-4 and CT/MRI LR-5 using pathology as the reference standard are lower than those using CCRS as the reference standard (for LR-4: 63.4% vs. 81.2%, $p$ = 0.03; for LR-5: 92.3% vs. 97%, $p$ = 0.01). The proportions of HCCs in CT/MRI LR-M using LI-RADS V2017 are lower than that using V2018 (6% vs. 42.2%, $p$ = 0.04).

## Diagnostic Performance of LR-5 for HCCs

The pooled sensitivity, specificity, and accuracy of CEUS LR-5 for HCC are comparable to those of CT/MRI LR-5 (**Table 3**). The DOR and the area under the summary receiver operating

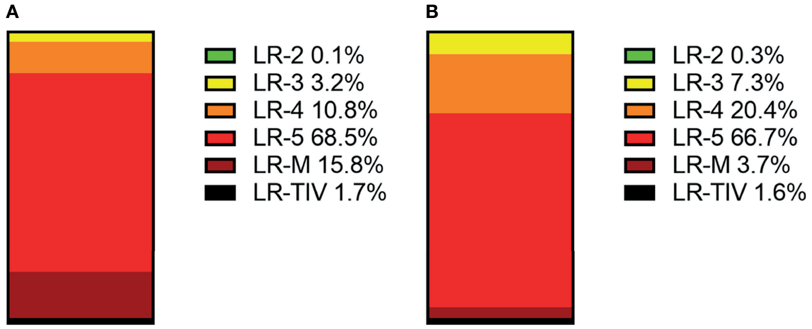


**FIGURE 3** | Distributions of HCCs in each CEUS category **(A)** and CT/MRI category **(B)**.

**TABLE 2** | Proportions of non-HCC malignancies in each CEUS and CT/MRI LI-RADS category.

| | CEUS | | | CT/MRI | | | |
|---|---|---|---|---|---|---|---|
| | **Malignancy (95% CI)** | **Observations** | **$I^2$, %** | **Malignancy (95% CI)** | **Observations** | **$I^2$, %** | **p** |
| LR–2 | 4 (1-11) | 126 | 0 | 5 (3–9) | 294 | 0 | 0.73 |
| LR–3 | 5 (2–11) | 644 | 52 | 4 (3–6) | 766 | 0 | 0.70 |
| LR–4 | 1 (0–6) | 629 | 80 | 3 (1–4) | 1266 | 58 | 0.40 |
| LR–5 | 2 (1–3) | 3594 | 82 | 2 (1–3) | 2993 | 68 | 0.65 |
| LR–M | 35 (26–45) | 1323 | 90 | 58 (43–72) | 441 | 85 | 0.01 |
| LR–TIV | 3 (0-23) | 100 | 62 | 22 (11–40) | 102 | 43 | 0.08 |

characteristic (SROC) curve for CEUS LR-5 are 28.0 and 0.74, and for CT/MRI LR-5, 23.9 and 0.75, as depicted in **Figure 5**.

## Diagnostic Performance of LR-M for Non-HCC Malignancies

The pooled sensitivity for non-HCC malignancies in CEUS LR-M (83%) is higher than that of CT/MRI LR-M (65%), while the pooled specificity for non-HCC malignancies in CT/MRI LR-M (92%) is similar to that of CEUS LR-M (92%) (**Table 4**). The DOR and the area under the SROC curve for CEUS LR-M are 36.5 and 0.87, for CT/MRI LR-M 46.6 and 0.73, respectively, as depicted in **Figure 6**.

## Meta-Regression for the Diagnostic Performance

The meta-regression for CT/MRI LR-5 shows that the specificity of the studies using pathology as the reference standard is lower than that using CCRS (85.4% vs. 95.0%, $p < 0.01$).

CEUS LR-M V2017 has a higher sensitivity (85.1% vs. 78.0%, $p < 0.01$), lower specificity (85.8% vs. 96.0%, $p < 0.01$), and lower accuracy (84.9% vs. 95.0%, $p < 0.01$) for the diagnosis of non-HCC malignancies compared with that of CEUS LR-M V2016. The sensitivity, specificity, and accuracy of CEUS LR-M for the diagnosis of non-HCC malignancies in the studies using pathology as the reference standard are lower compared with those using CCRS as the reference (sensitivity: 60.0% vs. 69.0%, $p = 0.01$; specificity: 96.0% vs. 97.2%, $p = 0.02$; accuracy: 90.1% vs. 94.63%, $p < 0.01$). Meta-regression for specificity shows that CT/MRI LR-M V2017 has a higher specificity than CT/MRI LR-M V2018 (99.5% vs. 93.4%, $p < 0.01$) for the diagnosis of non-HCC malignancies. CT/MRI LR-M V 2017 has a higher accuracy than CT/MRI LR-M V2014 (96.9% vs. 93.2%, $p = 0.02$) and V2018 (96.9% vs. 90.1%, $p < 0.01$).

# DISCUSSION

To our best knowledge, this work represents the first systematic review of the comparison of the percentages and distributions of HCCs and non-HCC malignancies between the CEUS and CT/MRI LI-RADS. The upshift of LI-RADS categories from LR-1 to LR-5 mirrors monotonically greater proportions of HCCs. We found that the proportion of HCCs in CEUS LR-3 is lower than that of CT/MRI LR-3. However, the proportions of HCCs in CEUS LR-M are higher than those of CT/MRI LI-RADS, while the percentage of non-HCC malignancies in CEUS LR-M is lower than that of CT/MRI LR-M. Furthermore, CEUS LR-M has a higher sensitivity in the diagnosis of non-HCC malignancies than CT/MRI LR-M.

ACR LI-RADS aims at stratifying the risk of HCCs and recommending the clinical management of each category (3, 4). The risk of HCCs in each category is the basis of clinical management. Evidence-based studies and feedback from clinical practice can help advise on the classification and management of lesions. ACR published two LI-RADS systems, namely, CEUS LI-RADS and CT/MRI LI-RADS. Although CEUS LI-RADS and CT/MRI LI-RADS have the same categories, the two LI-RADS systems have differences among the criteria and managements of classifications (3, 4). Understanding whether corresponding

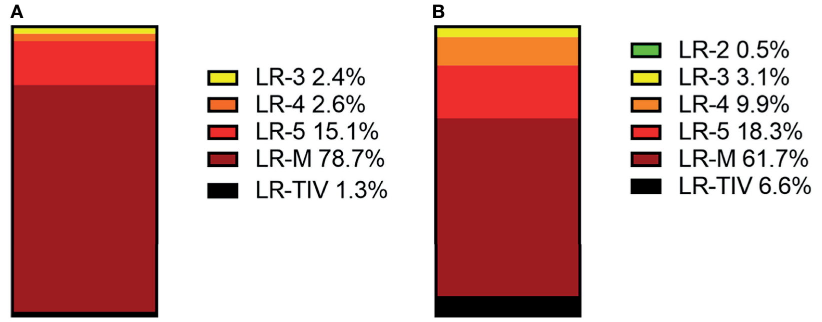


**FIGURE 4** | Distributions of non-HCC malignancies in each CEUS category **(A)** and CT/MRI category **(B)**.

**TABLE 3** | Diagnostic performance of CEUS and CT/MRI LR-5 for HCCs.

| | | CEUS | $I^2$,% | CT/MRI | $I^2$,% | p |
|---|---|---|---|---|---|---|
| LR-5 | Sensitivity(95% CI) | 73% (67–78) | 87 | 69% (64–74) | 92 | 0.32 |
| | Specificity(95% CI) | 92% (86–95) | 75 | 92% (88–94) | 86 | 0.96 |
| | Accuracy(95% CI) | 78% (71–84) | 90 | 76% (72–79) | 93 | 0.54 |
| | DOR (95% CI) | 28.0 (14.2–55.3) | 79 | 23.9 (15.8–36.3) | 87.3 | 0.70 |
| | AUC | 0.74 | | 0.75 | | |

categories of the two LI-RADS systems lead to distinct risk stratification of HCCs and whether the recommended management is appropriate for each category is of paramount importance. However, there is no evidence-based systematic review to address the issues mentioned above.

In this systematic review, we found that there was no statistical significance in the proportions of HCCs of the corresponding CEUS and CT/MRI LI-RADS categories, except CEUS LR-3 and LR-M and the CT/MRI counterparts. On the one hand, there is no HCC or non-HCC malignancy in the CEUS and CT/MRI LR-1, which is consistent with the definition of LR-1, i.e., definite benign. On the other hand, there are 96% HCCs in CEUS LR-5 and 95% HCCs in CT/MRI LR-5, which is also consistent with the definition of LR-5, definite HCCs. Thus, the lesions in CEUS LR-5 or CT/MRI LR-5 can go through clinical management of HCC without biopsy or MDD, as suggested by ACR. The management of CEUS LR-3, however, was different from that of CT/MRI LR-3. The suggested management for CT/MRI LR-3 is alternative or repeating diagnostic imaging in 3–6 months. By comparison, the suggested management for CEUS LR-3 is alternative or repeating diagnostic imaging in ≤6 months, with consideration for MDD. The recommended management for CEUS LR-3 is based on retrospective studies, which demonstrate that the percentage of CEUS LR-3 is 60%, higher than that of CT/MRI LR-3 (56–58). In the present study, the pooled proportion of HCCs in CEUS LR-3 is lower than that of CT/MRI LR-3 (21% vs. 35%, $p = 0.02$), which implies that there is still space for future improvement of the suggested management for CEUS and CT/MRI LR-3.

One of the main goals of LI-RADS LR-M is to avoid misdiagnosis of hepatocellular carcinoma for cholangiocarcinoma. In this study, we found that the percentage of HCCs in CEUS LR-M is higher than that of CT/MRI LR-M (56% vs. 33%). This difference may be induced by the differences in the diagnostic criteria of LR-M between CEUS and CT/MRI LI-RADS. Lesions with rim APHE or early washout or marked washout are classified into CEUS LR-M. Part of HCCs, especially the moderately and poorly differentiated HCCs, can present imaging features of LR-M (59). Compared with the suggested management of LR-5, MDD, alternative or repeating imaging, biopsy, or treatment is needed additionally for LR-M. Currently, lesions in CEUS LR-M have the same recommended management as those in CT/MRI LR-M. Thus, part of HCCs in CEUS LR-M, which can go through treatment, still needs an additional examination or MDD in practice. Despite the higher proportion of HCCs and the lower proportion of non-HCC malignancies in CEUS LR-M, the PPV and specificity of HCCs in CEUS LR-5 and CT/MRI LR-5 were comparable, which means that CEUS LR-M can avoid misdiagnosis of HCCs for cholangiocarcinoma. Still, in order to reduce the proportions of HCCs with additional examination or MDD, a previous study aimed at withdrawing HCCs in CEUS LR-M to LR-5 without decreasing the positive predictive value and specificity of HCCs in CEUS LR-5 (17).

LI-RADS LR-5 is used as the diagnostic criteria for HCCs, and LR-M is used as the diagnostic criteria for non-HCC malignancies in some studies (59, 60). The results of our systematic review show that CEUS LR-5 and CT/MRI LR-5 have comparable diagnostic performance for HCCs, namely, similar sensitivity, specificity, and accuracy. Our results are consistent with the result of previous studies (61). For non-HCC malignancies, however, CEUS LR-M has a different risk and sensitivity compared with CT/MRI LR-M. At the beginning

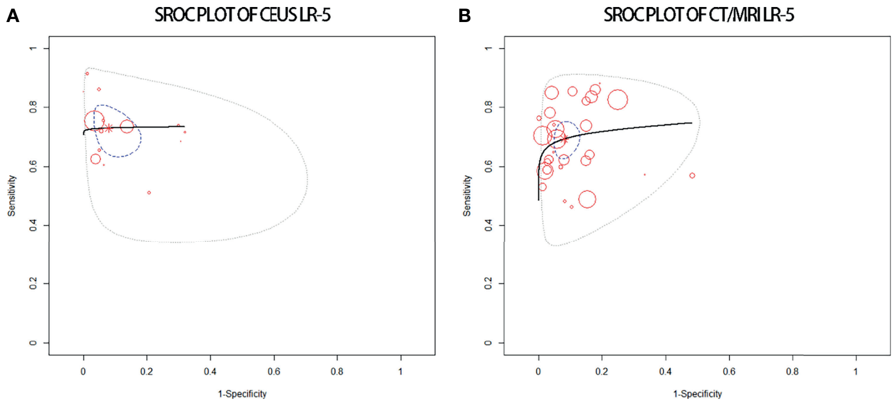


**FIGURE 5** | sROC plots of CEUS LR-5 **(A)** and of CT/MRI LR-5 **(B)** for the diagnosis of HCCs.

**TABLE 4** | Diagnostic performance of CEUS and CT/MRI LR-M for non-HCC malignancies.

| | | CEUS | $I^2$,% | CT/MRI | $I^2$,% | p |
|---|---|---|---|---|---|---|
| LR-M | Sensitivity (95% CI) | 83% (73–90) | 53 | 65% (56–73) | 78 | 0.01 |
| | Specificity (95% CI) | 92% (86–95) | 75 | 92% (88–94) | 86 | 0.96 |
| | Accuracy (95% CI) | 78% (70–84) | 90 | 76% (72–79) | 93 | 0.54 |
| | DOR (95% CI) | 36.5 (16.6–80.0) | 96 | 46.6 (24.9–88.2) | 86 | 0.64 |
| | AUC | 0.87 | | 0.73 | | |

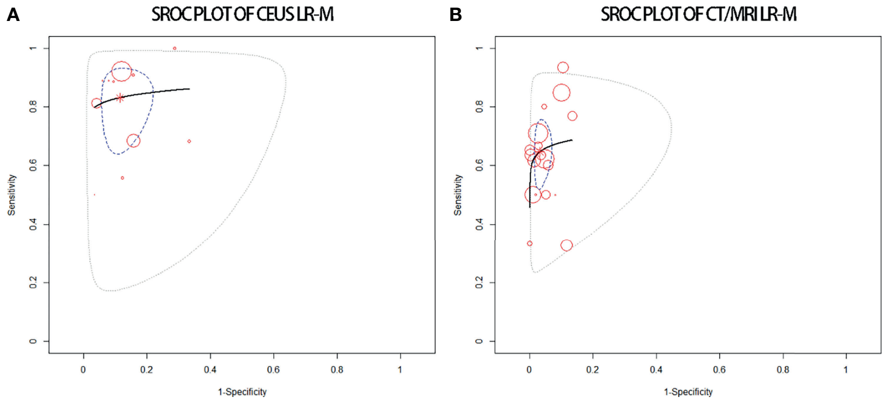


**FIGURE 6** | sROC plots of CEUS LR-M **(A)** and of CT/MRI LR-M **(B)** for the diagnosis of non-HCC malignancies.

of the application of CEUS LI-RADS, some studies focused on the PPV of LR-M in the diagnosis of non-HCC malignancies and found that CEUS LR-M has lower PPV than CT/MRI LR-M. They concluded that CT/MRI LR-M has higher differential diagnostic performance for non-HCC malignancies compared with CEUS LR-M (20, 53). Hu et al. (62) compared the diagnostic performance of non-HCC malignancies between CEUS LR-M and CT/MRI LR-M and demonstrated that the two LI-RADS systems had similar performance and sensitivity. However, a meta-analysis from the same group found that CEUS LR-M has a high sensitivity (84%) and specificity (90%) for non-HCC malignancies, while the CT/MRI counterpart has a moderate sensitivity (63%) and high specificity (95%) (61). In this study, CT/MRI LR-M has a higher percentage of non-HCC malignancies compared with CEUS LR-M, in agreement with previous studies (9, 63). CEUS LR-M, however, has higher sensitivity of non-HCC malignancies compared with CT/MRI LR-M. Thus, we conclude that both of the two LI-RADS systems have their advantages for the differential diagnosis of HCCs and non-HCC malignancies. Further studies are needed to explore the diagnostic performance for non-HCC malignancies.

This study has several limitations. First, we aimed to compare the risk of HCCs for CEUS LI-RADS classifications and the CT/MRI counterparts. However, few paired studies are available for this review. Second, the heterogeneity of the distribution and diagnostic performance of HCCs cannot be well explained by the meta-regression analysis. Last, the effects of tumor size on the classification by LI-RADS were not explored.

In conclusion, the proportions of HCCs increase with the upshift of LI-RADS categories from LR-1 to LR-5. CEUS LR-3 has a lower proportion of HCCs than CT/MRI LR-3, while CEUS LR-M has a higher proportion of HCCs. CEUS LR-M has a lower proportion of non-HCC malignancies than CT/MRI LR-M. CEUS LR-5 and CT/MRI LR-5 show comparable diagnostic performances of HCC, while CEUS LR-M has a higher sensitivity of non-HCC malignancies compared with CT/MRI LR-M.

## 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

YZ and ZQ designed the study and wrote the manuscript. JD, YC, and LZ collected the data. XJ and FW supervised the findings of this study. All authors contributed to the article and approved the submitted version.

## FUNDING

The present work was supported by the Tianjin Health and Health Committee (Nos. MS20017, KJ20170, ZD20014, NQ20033) and founded by the Tianjin Key Medical Discipline (Specialty) Construction Project.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.873913/full#supplementary-material

## REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660
2. Marrero JA, Kulik LM, Sirlin CB, Zhu AX, Finn RS, Abecassis MM, et al. Diagnosis, Staging, and Management of Hepatocellular Carcinoma: 2018 Practice Guidance by the American Association for the Study of Liver Diseases. *Hepatology* (2018) 68(2):723–50. doi: 10.1002/hep.29913
3. Liver Reporting & Data System (LI-RADS). Available at: https://www.acr.org/Clinical-Resources/Reportingand-Data-Systems/LI-RADS (Accessed April 24, 2020).
4. CEUS LI-RADS® v2017. Available at: https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/LI-RADS/CEUS-LI-RADS-v2017 (Accessed April 24, 2020).
5. Wilson SR, Lyshchik A, Piscaglia F, Cosgrove D, Jang HJ, Sirlin C, et al. CEUS LI-RADS: Algorithm, Implementation, and Key Differences From CT/MRI. *Abdom Radiol* (2018) 43(1):127–42. doi: 10.1007/s00261-017-1250-0
6. Schellhaas B, Hammon M, Strobel D, Pfeifer L, Kielisch C, Goertz RS, et al. Interobserver and Intermodality Agreement of Standardized Algorithms for non-Invasive Diagnosis of Hepatocellular Carcinoma in High-Risk Patients: CEUS-LI-RADS Versus MRI-LI-RADS. *Eur Radiol* (2018) 28(10):4254–64. doi: 10.1007/s00330-018-5379-1
7. Lee S, Kim SS, Roh YH, Choi JY, Park MS, Kim MJ. Diagnostic Performance of CT/MRI Liver Imaging Reporting and Data System V2017 for Hepatocellular Carcinoma: A Systematic Review and Meta-Analysis. *Liver Int* (2020) 40(6):1488–97. doi: 10.1111/liv.14424
8. Liang Y, Xu F, Guo Y, Lai L, Jiang X, Wei X, et al. Diagnostic Performance of LI-RADS for MRI and CT Detection of HCC: A Systematic Review and Diagnostic Meta-Analysis. *Eur J Radiol* (2021) 134:109404. doi: 10.1016/j.ejrad.2020.109404
9. Li J, Yang L, Ma L, Lu Q, Luo Y. Diagnostic Accuracy of Contrast-Enhanced Ultrasound Liver Imaging Reporting and Data System (CEUS LI-RADS) for Differentiating Between Hepatocellular Carcinoma and Other Hepatic Malignancies in High-Risk Patients: A Meta-Analysis. *Ultraschall der Med* (2021) 42(02):187–93. doi: 10.1055/a-1309-1568
10. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PMthe PRISMA-DTA Group, et al. Preferred Reporting Items for a Systematic-Review and Meta Analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. *J Am Med Assoc* (2018) 319:388–96. doi: 10.1001/jama.2017.19163
11. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. *Ann Intern Med* (2011) 155:529–36. doi: 10.7326/0003-4819-155-8-201110180-00009
12. Ryan RCochrane Consumers and Communication Group. *Heterogeneity and Subgroup Analyses in Cochrane Consumers and Communication Group Reviews: Planning the Analysis at Protocol Stage*. Available at: http://cccrg.cochrane.org (Accessed June 15, 2018).
13. Schellhaas B, Görtz RS, Pfeifer L, Kielisch C, Neurath MF, Strobel D. Diagnostic Accuracy of Contrast-Enhanced Ultrasound for the Differential Diagnosis of Hepatocellular Carcinoma: ESCULAP Versus CEUS-LI-RADS. *Eur J Gastroen Hepate* (2017) 29(9):1036–44. doi: 10.1097/MEG.0000000000000916
14. Ghiuchici AM, Dănilă M, Popescu A, Sirli R, Moga T, Topan M, et al. Contrast-Enhanced Ultrasound Algorithm (ACR CEUS LI-RADSv 2017)- A Valuable Tool for the Noninvasive Diagnosis of Hepatocellular Carcinoma in Patients With Chronic Liver Disease. *[J]. Med Ultrason* (2021) 23(4):383–9. doi: 10.11152/mu-2887
15. Terzi E, Iavarone M, Pompili M, Veronese L, Cabibbo G, Fraquelli M, et al. Contrast Ultrasound LI-RADS LR-5 Identifies Hepatocellular Carcinoma in Cirrhosis in a Multicenter Restropective Study of 1,006 Nodules. *J Hepatol* (2018) 68(3):485–92. doi: 10.1016/j.jhep.2017.11.007
16. Tan Z, Teoh WC, Wong KM, Wansaicheong GK, Sandrasegaran K. Analysis of Comparative Performance of CEUS and CECT/MR LI-RADS Classification: Can CEUS Dichotomize LI-RADS Indeterminate Lesions on CT or MRI? *Clin Imaging* (2020) 62:63–8. doi: 10.1016/j.clinimag.2020.02.002
17. Zheng W, Li Q, Zou XB, Wang JW, Han F, Li F, et al. Evaluation of Contrast-Enhanced US LI-RADS Version 2017: Application on 2020 Liver Nodules in Patients With Hepatitis B Infection. *Radiology* (2020) 294(2):299–307. doi: 10.1148/radiol.2019190878
18. Wang JY, Feng SY, Xu JW, Li J, Chu L, Cui XW, et al. Usefulness of the Contrast-Enhanced Ultrasound Liver Imaging Reporting and Data System in Diagnosing Focal Liver Lesions by Inexperienced Radiologists. *J Ultrasound Med* (2020) 39(8):1537–46. doi: 10.1002/jum.15242
19. Li J, Ling W, Chen S, Ma L, Yang L, Lu Q, et al. The Interreader Agreement and Validation of Contrast-Enhanced Ultrasound Liver Imaging Reporting

and Data System. *Eur J Radiol* (2019) 120:108685. doi: 10.1016/j.ejrad.2019.108685

20. Ding J, Long L, Zhang X, Chen C, Zhou H, Zhou Y, et al. Contrast-Enhanced Ultrasound LI-RADS 2017: Comparison With CT/MRI LI-RADS. *Eur Radiol* (2021) 31(2):847–54. doi: 10.1007/s00330-020-07159-z
21. Abd Alkhalik Basha M, Abd El Aziz El Sammak D, El Sammak AA. Diagnostic Efficacy of the Liver Imaging-Reporting and Data System (LI-RADS) With CT Imaging in Categorising Small Nodules (10–20 Mm) Detected in the Cirrhotic Liver at Screening Ultrasound. *Clin Radiol* (2017) 72(10):901.e1–e11. doi: 10.1016/j.crad.2017.05.019
22. Cha DI, Jang KM, Kim SH, Kang TW, Song KD. Liver Imaging Reporting and Data System on CT and Gadoxetic Acid-Enhanced MRI With Diffusion-Weighted Imaging. *Eur Radiol* (2017) 27(10):4394–405. doi: 10.1016/j.crad.2017.05.019
23. Choi SH, Byun JH, Kim SY, Lee SJ, Won HJ, Shin YM, et al. Liver Imaging Reporting and Data System V2014 With Gadoxetate Disodium-Enhanced Magnetic Resonance Imaging: Validation of LI-RADS Category 4 and 5 Criteria. *Invest Radiol* (2016) 51(8):483–90. doi: 10.1097/RLI.0000000000000258
24. Fraum TJ, Tsai R, Rohe E, Ludwig DR, Salter A, Nalbantoglu I, et al. Differentiation of Hepatocellular Carcinoma From Other Hepatic Malignancies in Patients at Risk: Diagnostic Performance of the Liver Imaging Reporting and Data System Version 2014. *Radiology* (2018) 286 (1):158–72. doi: 10.1148/radiol.2017170114
25. Kim BR, Lee JM, Lee DH, Yoon JH, Hur BY, Suh KS, et al. Diagnostic Performance of Gadoxetic Acid-Enhanced Liver MR Imaging Versus Multidetector CT in the Detection of Dysplastic Nodules and Early Hepatocellular Carcinoma. *Radiology* (2017) 285(1):134–46. doi: 10.1148/radiol.2017162080
26. Kim YY, An C, Kim S, Kim MJ. Diagnostic Accuracy of Prospective Application of the Liver Imaging Reporting and Data System (LI-RADS) in Gadoxetate-Enhanced MRI. *Eur Radiol* (2018) 28(5):2038–46. doi: 10.1007/s00330-017-5188-y
27. Liu W, Qin J, Guo R, Xie S, Jiang H, Wang X, et al. Accuracy of the Diagnostic Evaluation of Hepatocellular Carcinoma With LI-RADS. *Acta Radiol* (2018) 59(2):140–6. doi: 10.1177/0284185117716700
28. An C, Lee CH, Byun JH, Lee MH, Jeong WK, Choi SH, et al. Intraindividual Comparison Between Gadoxetate-Enhanced Magnetic Resonance Imaging and Dynamic Computed Tomography for Characterizing Focal Hepatic Lesions: A Multicenter, Multireader Study. *Korean J Radiol* (2019) 20 (12):1616–26. doi: 10.3348/kjr.2019.0363
29. Qi Z, Li H. *The Utility of LI-RADS V2014 to Characterize the Nodules Detected During Hepatocarcinogenesis in HBV-Related Liver Cirrhosis: A Comparative Study of MR Imaging and Pathology* (2016). Radiological Society of North America: Scientific Assembly and Annual Meeting. Available at: http://archive.rsna.org/2016/16019077.html (Accessed June 15, 2018).
30. Joo I, Lee JM, Lee DH, Jeon JH, Han JK. Retrospective Validation of a New Diagnostic Criterion for Hepatocellular Carcinoma on Gadoxetic Acid-Enhanced MRI: Can Hypointensity on the Hepatobiliary Phase be Used as an Alternative to Washout With the Aid of Ancillary Features? *Eur Radiol* (2019) 29(4):1724–32. doi: 10.1007/s00330-018-5727-1
31. Rosiak G, Podgorska J, Rosiak E, Cieszanowski A. Comparison of LI-RADS V.2017 and ESGAR Guidelines Imaging Criteria in HCC Diagnosis Using MRI With Hepatobiliary Contrast Agents. *BioMed Res Int* (2018) 2018:7465126. doi: 10.1155/2018/7465126
32. Renzulli M, Biselli M, Brocchi S, Granito A, Vasuri F, Tovoli F, et al. New Hallmark of Hepatocellular Carcinoma, Early Hepatocellular Carcinoma and High-Grade Dysplastic Nodules on Gd-EOB-DTPA MRI in Patients With Cirrhosis: A New Diagnostic Algorithm. *Gut* (2018) 67(9):1674–82. doi: 10.1136/gutjnl-2017-315384
33. Min JH, Kim JM, Kim YK, Kang TW, Lee SJ, Choi GS, et al. Prospective Intraindividual Comparison of Magnetic Resonance Imaging With Gadoxetic Acid and Extracellular Contrast for Diagnosis of Hepatocellular Carcinomas Using the Liver Imaging Reporting and Data System. *Hepatology* (2018) 68 (6):2254–66. doi: 10.1002/hep.30122
34. Alhasan A, Cerny M, Olivié D, Billiard JS, Bergeron C, Brown K, et al. LI-RADS for CT Diagnosis of Hepatocellular Carcinoma: Performance of Major and Ancillary Features. *Abdom Radiol* (2019) 44(2):517–28. doi: 10.1007/s00261-018-1762-2
35. Zhang T, Huang ZX, Wei Y, Jiang HY, Chen J, Liu XJ, et al. Hepatocellular Carcinoma: Can LI-RADS V2017 With Gadoxetic-Acid Enhancement Magnetic Resonance and Diffusion-Weighted Imaging Improve Diagnostic Accuracy? *World J Gastroenterol* (2019) 25(5):622–31. doi: 10.3748/wjg.v25.i5.622
36. Basha MAA, Refaat R, Mohammad FF, Khamis MEM, El-Maghraby AM, El Sammak AA, et al. The Utility of Diffusion-Weighted Imaging in Improving the Sensitivity of LI-RADS Classification of Small Hepatic Observations Suspected of Malignancy. *Abdom Radiol* (2019) 44(5):1773–84. doi: 10.1007/s00261-018-01887-z
37. Ren AH, Zhao PF, Yang DW, Du JB, Wang ZC, Yang ZH. Diagnostic Performance of MR for Hepatocellular Carcinoma Based on LI-RADS V2018, Compared With V2017. *J Magn Reson Imaging* (2019) 50(3):746–55. doi: 10.1002/jmri.26640
38. Lee S, Kim MJ, Kim SS, Shin H, Kim DY, Choi JY, et al. Retrospective Comparison of EASL 2018 and LI-RADS 2018 for the Noninvasive Diagnosis of Hepatocellular Carcinoma Using Magnetic Resonance Imaging. *Hepatol Int* (2020) 14(1):70–9. doi: 10.1007/s12072-019-10002-3
39. Chernyak V, Flusberg M, Berman J, Fruitman KC, Kobi M, Fowler KJ, et al. Liver Imaging Reporting and Data System Version 2018: Impact on Categorization and Hepatocellular Carcinoma Staging. *Liver Transplant* (2019) 25(10):1488–502. doi: 10.1002/lt.25614
40. De Gaetano AM, Catalano M, Pompili M, Marini MG, Rodríguez Carnero P, Gullì C, et al. Critical Analysis of Major and Ancillary Features of LI-RADS V2018 in the Differentiation of Small (≤ 2 Cm) Hepatocellular Carcinoma From Dysplastic Nodules With Gadobenate Dimeglumine-Enhanced Magnetic Resonance Imaging. *Eur Rev Med Pharmacol Sci* (2019) 23 (18):7786–801. doi: 10.26355/eurrev_201909_18988
41. Jiang H, Liu X, Chen J, Wei Y, Lee JM, Cao L, et al. Man or Machine? Prospective Comparison of the Version 2018 EASL, LI-RADS Criteria and a Radiomics Model to Diagnose Hepatocellular Carcinoma. *Cancer Imaging* (2019) 19(1):84. doi: 10.1186/s40644-019-0266-9
42. Lee SM, Lee JM, Ahn SJ, Kang HJ, Yang HK, Yoon JH. LI-RADS Version 2017 Versus Version 2018: Diagnosis of Hepatocellular Carcinoma on Gadoxetate Disodium-Enhanced MRI. *Radiology* (2019) 292(3):655–63. doi: 10.1148/radiol.2019182867
43. Barabino M, Gurgitano M, Fochesato C, Angileri SA, Franceschelli G, Santambrogio R, et al. LI-RADS to Categorize Liver Nodules in Patients at Risk of HCC: Tool or a Gadget in Daily Practice? *Radiol Med* (2020) 126:5–13. doi: 10.1007/s11547-020-01225-8
44. Cannella R, Vernuccio F, Sagreiya H, Choudhury KR, Iranpour N, Marin D, et al. Liver Imaging Reporting and Data System (LI-RADS) V2018: Diagnostic Value of Ancillary Features Favoring Malignancy in Hypervascular Observations ≥ 10 Mm at Intermediate (LR-3) and High Probability (LR-4) for Hepatocellular Carcinoma. *Eur Radiol* (2020) 30(7):3770–81. doi: 10.1007/s00330-020-06698-9
45. Kierans AS, Song C, Gavlin A, Roudenko A, Lu L, Askin G, et al. Diagnostic Performance of LI-RADS Version 2018, LI-RADS Version 2017, and OPTN Criteria for Hepatocellular Carcinoma. *Am J Roentgenol* (2020) 215(5):1085–92. doi: 10.2214/AJR.20.22772
46. Ko A, Park HJ, Lee ES, Park SB, Kim YK, Choi SY, et al. Comparison of the Diagnostic Performance of the 2017 and 2018 Versions of LI-RADS for Hepatocellular Carcinoma on Gadoxetic Acid Enhanced MRI. *Clin Radiol* (2020) 75(4):319.e1–9. doi: 10.1016/j.crad.2019.11.004
47. Makoyeva A, Kim TK, Jang HJ, Medellin A, Wilson SR. Use of CEUS LI-RADS for the Accurate Diagnosis of Nodules in Patients at Risk for Hepatocellular Carcinoma: A Validation Study. *Radiol Imaging Cancer* (2020) 2(2):e190014. doi: 10.1148/rycan.2020190014
48. Ciocalteu A, Iordache S, Cazacu SM, Urhut CM, Sandulescu SM, Ciurea AM, et al. Role of Contrast-Enhanced Ultrasonography in Hepatocellular Carcinoma by Using LI-RADS and Ancillary Features: A Single Tertiary Centre Experience. *Diagn (Basel)* (2021) 11(12):2232. doi: 10.3390/diagnostics11122232
49. Xian MF, Huang Y, Xie WX, Pan KM, Zeng D, Huang H, et al. LR-M Observations on Contrast-Enhanced Ultrasound: Detection of Hepatocellular Carcinoma Using Additional Features in Comparison With Current LI-RADS Criteria [Published Online Ahead of Print, 2021 Dec 15]. *AJR Am J Roentgenol* (2021). doi: 10.2214/AJR.21.26837
50. Ding J, Qin Z, Zhou Y, Zhou HY, Zhang Q, Wang YD, et al. Impact of Revision of the LR-M Criteria on the Diagnostic Performance of Contrast-

Enhanced Ultrasound LI-RADS. *Ultrasound Med Biol* (2021) 47(12):3403–10. doi: 10.1016/j.ultrasmedbio.2021.08.007

51. Yang D, Hu H, Li R, Tang CL, Ma KS, Guo DY. The Diagnostic Value of Contrast-Enhanced Ultrasound LI-RADS for Hepatocellular Carcinoma in Patients With Cirrhosis and Chronic Hepatitis B. *Abdom Radiol (NY)* (2022) 47(2):630–9. doi: 10.1007/s00261-021-03345-9
52. Zeng D, Xu M, Liang JY, Cheng MQ, Huang H, Pan JM, et al. Using New Criteria to Improve the Differentiation Between HCC and non-HCC Malignancies: Clinical Practice and Discussion in CEUS LI-RADS 2017. *Radiol Med* (2022) 127(1):1–10. doi: 10.1007/s11547-021-01417-w
53. Lv K, Cao X, Dong Y, Geng D, Zhang J. CT/MRI LI-RADS Version 2018 Versus CEUS LI-RADS Version 2017 in the Diagnosis of Primary Hepatic Nodules in Patients With High-Risk Hepatocellular Carcinoma. *Ann Transl Med* (2021) 9(13):1076. doi: 10.21037/atm-21-1035
54. Zhang Y, Tang W, Xie S, Chen J, Zhang L, Rong D, et al. The Role of Lesion Hypointensity on Gadobenate Dimeglumine-Enhanced Hepatobiliary Phase MRI as an Additional Major Imaging Feature for HCC Classification Using LI-RADS V2018 Criteria. *Eur Radiol* (2021) 31(10):7715–24. doi: 10.1007/s00330-021-07807-y
55. Xie S, Zhang Y, Chen J, Jiang T, Liu W, Rong D, et al. Can Modified LI-RADS Increase the Sensitivity of LI-RADS V2018 for the Diagnosis of 10-19 Mm Hepatocellular Carcinoma on Gadoxetic Acid-Enhanced MRI? *Abdom Radiol (NY)* (2022) 47(2):596–607. doi: 10.1007/s00261-021-03339-7
56. Terzi E, De Bonis L, Leoni S, Benevento F, Granito A, Tovoli F, et al. CEUS LI-RADS are Effective in Predicting the Risk Hepatocellular Carcinoma of Liver Nodules. *Dig Liv Dis* (2017) 49: Suppl 1:e22. doi: 10.1016/j.dld.2017.01.046
57. Choi JY, Cho HC, Sun M, Kim HC, Sirlin CB. Indeterminate Observations (Liver Imaging Reporting and Data System Category3) on MRI in the Cirrhotic Liver: Fate and Clinical Implications. *Am J Roentgenol* (2013) 201 (5):993–1001. doi: 10.2214/AJR.12.10007
58. Tanabe M, Kanki A, Wolfson T, Costa EA, Mamidipalli A, Ferreira MP, et al. Imaging Outcomes of Liver Imaging Reporting and Data System Version 2014 Category 2, 3, and 4 Observations Detected at CT and MR Imaging. *Radiology* (2016) 281(1):129–39. doi: 10.1148/radiol.2016152173
59. Huang JY, Li JW, Lu Q, Luo Y, Lin L, Shi YJ, et al. Diagnostic Accuracy of CEUS LI-RADS for the Characterization of Liver Nodules 20 Mm or Smaller in Patients at Risk for Hepatocellular Carcinoma. *Radiology* (2020) 294 (2):329–39. doi: 10.1148/radiol.2019191086
60. Ronot M, Fouque O, Esvan M, Lebigot J, Aubé C, Vilgrain V. Comparison of the Accuracy of AASLD and LI-RADS Criteria for the Non-Invasive Diagnosis of HCC Smaller Than 3 Cm. *J Hepatol* (2018) 68(4):715–23. doi: 10.1016/j.jhep.2017.12.014
61. Li L, Hu Y, Han J, Li Q, Peng C, Zhou J. Clinical Application of Liver Imaging Reporting and Data System for Characterizing Liver Neoplasms: A Meta-Analysis. *Diagnostics* (2021) 11(2):323. doi: 10.3390/diagnostics11020323
62. Hu YX, Shen JX, Han J, Mao SY, Mao RS, Li Q, et al. Diagnosis of non-Hepatocellular Carcinoma Malignancies in Patients With Risks for Hepatocellular Carcinoma: CEUS LI-RADS Versus CT/MRI LI-RADS. *Front Oncol* (2021) 11:641195:641195. doi: 10.3389/fonc.2021.641195
63. Van der Pol CB, Lim CS, Sirlin CB, McGrath TA, Salameh JP, Bashir MR, et al. Accuracy of the Liver Imaging Reporting and Data System in Computed Tomography and Magnetic Resonance Image Analysis of Hepatocellular Carcinoma or Overall Malignancy—A Systematic Review. *Gastroenterology* (2019) 156(4):976–86. doi: 10.1053/j.gastro.2018.11.020

**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.

frontiers | Frontiers in Oncology

REVIEW
published: 27 June 2022
doi: 10.3389/fonc.2022.905087

# Application of Multiple Ultrasonic Techniques in the Diagnosis of Prostate Cancer

*Yushan Liu*[1], *Shi Zeng*[1] *and Ran Xu*[2]*

[1] *Department of Ultrasound, The Second Xiangya Hospital of Central South University, Changsha, China,* [2] *Department of Urology, The Second Xiangya Hospital of Central South University, Changsha, China*

Methods for diagnosing prostate cancer (PCa) are developing in the direction of imaging. Advanced ultrasound examination modes include micro-Doppler, computerized-transrectal ultrasound, elastography, contrast-enhanced ultrasound and microultrasound. When two or more of these modes are used in PCa diagnosis, the combined technique is called multiparameter ultrasound (mp-US). Mp-US provides complementary information to multiparameter magnetic resonance imaging (mp-MRI) for diagnosing PCa. At present, no study has attempted to combine the characteristics of different ultrasound modes with advanced classification systems similar to the PIRADS system in mpMRI for the diagnosis of PCa. As an imaging method, mp-US has great potential in the diagnosis of PCa.

**Keywords: multiparameter ultrasound, prostate cancer, contrast - enhanced ultrasound, elastography, ultrasound molecular imaging**

OPEN ACCESS

**Edited by:**
*Xuefeng Qiu,*
*Nanjing Drum Tower Hospital, China*

**Reviewed by:**
*Pietro Pepe,*
*Cannizzaro Hospital, Italy*
*Massimo Lazzeri,*
*Humanitas Research Hospital, Italy*
*Hodo Haxhimolla,*
*University of Canberra, Australia*

***Correspondence:**
*Ran Xu*
*xuran@csu.edu.cn*

**Specialty section:**
*This article was submitted to*
*Cancer Imaging and*
*Image-directed Interventions,*
*a section of the journal*
*Frontiers in Oncology*

**Received:** *28 March 2022*
**Accepted:** *12 May 2022*
**Published:** *27 June 2022*

**Citation:**
*Liu Y, Zeng S and Xu R*
*(2022) Application of Multiple*
*Ultrasonic Techniques in the*
*Diagnosis of Prostate Cancer.*
*Front. Oncol. 12:905087.*
*doi: 10.3389/fonc.2022.905087*

## HIGHLIGHTS

1. This article is a review of the application and development of various ultrasound techniques in the diagnosis of PCa.
2. Multiparameter ultrasound is a new combined mode of several ultrasound techniques, which is similar to multiparameter magnetic resonance imaging and it can significantly improve the diagnosis rate of PCa.
3. A complete ultrasound examination scoring system will have important clinical application value in improving PCa diagnosis and follow-up.

## INTRODUCTION

Prostate cancer (PCa) is the most common genitourinary system tumor in middle-aged and elderly men, and it is common in most Northern and Western countries. With the “Westernization” of lifestyles, the rapid aging of the population and the development of metabolic syndrome, the incidence and mortality of prostate cancer in our country have gradually increased in recent years

(1, 2). The onset of PCa is insidious and lacks typical clinical manifestations. Most patients are already in the middle and late stages when they are diagnosed. Therefore, the early clinical diagnosis and treatment of PCa are of great significance in improving the survival rate of patients and their quality of life.

At present, early diagnostic tests of PCa mainly include on digital rectal examination (DRE), serum prostate specific antigen (PSA) measurement and conventional transrectal ultrasound (TRUS) (3). DRE is limited to palpation of the posterior area of the prostate, which can cause physical discomfort, rectal bleeding and even syncope. The level of PSA can indicate the risk of prostate cancer, but its sensitivity (SE) is high and specificity (SP) is low. The PSA of prostate cancer patients can even be in the normal range. Acute prostatitis and benign prostatic hypertrophy can also lead to an increase in PSA levels.Twelve-core systematic TRUS-guided biopsy for patients with serum PSA levels> 4.0 ng/mL is currently the gold standard for diagnosing PCa. Its SE is low, however, and the detection rate is only 27%-40.3% (4, 5). Additionally, the false negative rate of systemic prostate biopsy ranges from 17% to 21% (6, 7). Increasing the number of core biopsies can increase the detection rate of PCa and help better evaluate GS score (8, 9). The main disadvantage of systematic biopsy is that it is invasive, and can cause various complications such as prostatitis, hematuria, hematochezia, urinary retention and hematospermia (10). Additionally, it cannot detect small, low-risk, and clinically atypical cancers. Thus, it can lead to misdiagnosis, missed diagnosis, too many false negatives, and excessive puncture.

Therefore, an increasing number of researchers are dedicated to exploring imaging technologies with high SE, SP, and noninvasiveness. PCa imaging research focuses on two platforms: magnetic resonance imaging (MRI) and ultrasound (US). Multiparameter MRI (mp-MRI) is currently an important imaging method for PCa detection and localization and guidance of needle biopsy. The more commonly used sequences are T2-weighted imaging (T2WI), diffusion weighted imaging (DWI), dynamic contrast-enhanced MRI (DCE-MRI) and three-dimensional MR spectral imaging (11). However, MRI is not appropriate for claustrophobic patients, patients with pacemaker implantation and patients with metal pelvic implants. US is highly cost-effective and has wide applicability and strong practicability. Advanced US modalities include micro-Doppler, computerized-transrectal ultrasound, elastography, contrast-enhanced ultrasound and microultrasound. When different modes are used in combination, it is called multiparametric ultrasound (mp-US). This is a novel US examination mode similar to mpMRI, that can significantly improve the diagnosis rate of prostate cancer. This article introduces the basic principles and performance of different ultrasound-based modes and reports the clinical effects of combining them in mp-US.

## GREYSCALE TRUS

Currently, conventional TRUS is commonly used for prostate cancer detection, guided systematic biopsy, and guided radiotherapy particle placement (12). Because prostate cancer tissue and normal prostate tissue have similar backscatter signals and heterogeneity in the prostate transition zone, traditional TRUS has limitations in detecting PCa. Moreover, the higher frequency of the transrectal transducer can cause attenuation artifacts in the examination, especially when there are more calcifications in the prostate tissue. Approximately 60% of PCa lesions are hypoechoic on TRUS (13), and approximately 35–39% are isoechoic (14). Some nonmalignant diseases of the prostate, such as prostate inflammation and benign prostatic hyperplasia, can also appear hypoechoic on ultrasound images, leading to false positive test results.

The SE of TRUS in diagnosing PCa is between 8% and 88%, and the SP is between 42.5% and 99% (15–17). Taverna et al. observed that the PCa detection rate of 13-core TRUS-guided biopsy was 29% in 100 patients (18). A study by Klein et al. showed that TRUS has poor SP for early PCa, with a false negative rate for pathological results of systematic biopsy guided by TRUS of up to 30% (19). Hwang et al. noted that increasing the number of transrectal ultrasound-guided prostate punctures and the number of needles can increase the detection rate of PCa but would also increase the incidence of puncture complications (20). Therefore, targeted biopsy methods have arisen as the newest direction of research, as it can reduce the number of puncture needles and increase the detection rate of PCa. A study showed that the SP of TRUS-guided targeted biopsy in detecting PCa is better than that of mpMRI (41% vs. 96%) (21).

According to the European Urology Association (EAU) guidelines, standard grayscale TRUS remains the standard technique for biopsy guidance (22). The current US imaging system mostly uses nonlinear imaging. Its main advantage is the high contrast resolution of the tissue and low clutter in the inspection. The current trend in prostate ultrasound diagnosis is to increase the frequency of the probe and use broadband single-crystal piezoelectric elements to provide higher contrast and spatial resolution. In recent years, transrectal three-dimensional ultrasound (3D-TRUS) has been developed to provide more information for the diagnosis of PCa. Long et al. found that the accuracy and repeatability of needle biopsy guided by real-time three-dimensional ultrasound are better than those of two-dimensional ultrasound (23). Zhao et al. (24) and Guo et al. (25) showed that 3D-TRUS can help identify targeted puncture sites and increase the positive rate of PCa examination.

As a new high-resolution imaging method to guide prostate biopsy, microultrasound has received extensive clinical attention. Compared with mpMRI, microultrasound has potential advantages, such as relatively low cost and ease of operation. Several studies have reported the use of microultrasound in diagnosing PCa. In a meta-analysis of 769 patients, Zhang et al. found that microultrasound had a pooled SE, SP, diagnostic odds ratio (DOR), and area under curve (AUC) of 0.91, 0.49, 10, and 0.82, respectively (26). In 104 patients with suspected PCa, Lughezzani et al. found that the sensitivity and detection rate of microultrasound for the detection of csPCa were 94% and 54%, respectively (27). This study suggests that microUS can be used as an auxiliary diagnostic tool for MRI in diagnosis csPCa. PCa is a lower proportion of MRI-diagnosed

PI-RADS 3 lesions, but prostate biopsy is still required. In a study of 111 patients, Pier Paolo Avolio et al. found that microUS detected 100% of csPCa in patients with a PI-RADS 3 lesion at mpMRI, while reducing the detection rate of ncsPCa to 23.8% (28). Sountoulides et al. (29) found that microultrasound-guided prostate biopsy has a PCa diagnosis rate equivalent to that of mpMRI-guided prostate biopsy. Laura Wiemer et al. found that the positive predictive value of micro-ultrasound for diagnosing csPCa was significantly higher than that of mpMRI in 159 patients (30). Microultrasound can be used as an inexpensive and convenient alternative to mpMRI in diagnosing PCa. Based on the findings of other studies, microultrasound can satisfactorily diagnose clinically significant prostate cancer (csPCa). The high SE of microultrasound in the diagnosis of PCa can further improve MRI/US-targeted biopsy and avoid unnecessary system biopsy. The diagnostic value of microultrasound still needs to be more comprehensively analyzed with more clinical data.

The latest EAU guideline (2022) recommend to perform, if possible, transperineal instead transrectal approach. In 200 patients with persistently elevated PSA, Pietro Pepe et al. found that transperineal MRI/TRUS cognitive targeted biopsy had a higher detection rate of csPCa in the anterior zone of prostate compared with transrectal MRI/TRUS fusion targeted biopsy (93.3 vs. 25%) (31). The transperineal approach provides relatively easy access to the anterior region, thus it reduces the patient's risk of sepsis (32). In a study of 3000 patients with suspected PCa, Pietro Pepe et al. found that the detection rate of PCa by transperineal prostate biopsy was 38.3%, and 40.2% of the patients underwent biopsy without sepsis, only urinary tract infection and urinary retention (33). There are several studies focusing on fusion targeted biopsy and cognitive targeted biopsy, which one is more suitable for the detection of csPCa. In a study of 200 patients with persistently elevated PSA, Pietro Pepe et al. found that the sensitivity, specificity, positive predictive value, negative predictive value, and diagnostic accuracy of transperineal cognitive targeted biopsy in detection rate of csPCa were slightly higher than those of transrectal fusion targeted biopsy (97.2% vs. 66.7%, 78.2% vs. 71.8%, 59% vs. 42.1%, 97.2% vs. 87.5%, 68.9% vs. 57.5%, respectively.) (31). However, several studies have presented higher accuracy of MRI/TRUS fusion targeted biopsy compared with cognitive targeted biopsy, because the latter is operator-dependent (5, 34). In the past few years, mpMRI/TRUS fusion-targeted biopsy has improved the diagnostic accuracy of csPCa, especially in patients with repeat biopsies (35). More researches are still needed to present their respective advantages in diagnosing csPCa.

## COMPUTER-ASSISTED DIAGNOSTIC SYSTEM

In artificial neural network analysis/computerized-transrectal ultrasound (ANNA/C-TRUS), the doctor performs routine grayscale TRUS examinations on the patient before radical prostatectomy (RP). The images are sent to the ANNA/C-TRUS server through an internet platform. The C-TRUS system uses the ANNA algorithm to analyze the ultrasound images, then colors suspicious areas and returns them to the user terminal. This system is a convenient method for clinicians to performed guided, targeted tumor biopsy (36).

Among 132 patients with elevated PSA or abnormal DRE, 66 cases of cancer were found through C-TRUS targeted biopsy (37). Another study compared the tumor localization of the C-TRUS system before surgery and the pathological results after RP in 28 patients and found that the SE, SP, negative predictive value (NPV), positive predictive value (PPV) and total accuracy of the ANNA/C-TRUS system in detecting cancer were 83.1%, 63.9%, 68.4%, 80.1% and 76.2%, respectively (38). Moreover, the ANNA/C-TRUS system can better predict tumor differentiation than random systemic biopsy. In 164 preoperative patients undergoing RP, the SE of ANNA/C-TRUS in predicting the RP Gleason classification of the index lesions was 85% (39).When performing traditional TRUS, the additional use of C-TRUS can assist in the detection and biopsy of cancerous lesions. The combination of C-TRUS and MR can increase the detection rate for high-risk PCa patients (40). The use of ANNA/C-TRUS can improve the accuracy of PCa diagnosis, but a larger multicenter study is still needed to assess its clinical value.

Histoscanning (HS) is an ultrasound-based tissue characterization technology that can be used for PCa detection and localization. TRUS is used to first perform a full scan of the prostate to obtain three-dimensional grayscale data. Then the examiner uses HS software to color-code suspicious area and determine the tumor volume. This technique has shown encouraging results in the detection of csPCa.

In a study of 32 preoperative patients with RP, the SE, SP, PPV, and NPV of HS in detecting PCa were 93.5%, 79.5%, 67.35%, and 96.5%, respectively (41). HS can assist in diagnosing patients through prostate biopsy diagnosis. It has a higher detection rate for cancer lesions with a volume of ≥0.50 mL (42) and a diameter of ≥0.1cm (43). In 43 patients, the cancer detection rate of transrectal ultrasound biopsy with a standard 12-core system guided by prostate tissue scanning targeting (PHS-TT) was 46.5%, and the length of the PHS-TT cores was significantly higher than that of the systematic cores (55.4% vs. 37.5%, $p < 0.05$) (44). PHS-TT can be used as an effective tool for the clinical guidance of prostate biopsy in real time.

In a study of 14 preoperative prostate HS in patients with RP, there was a significant correlation between tumor diameter and final pathology ($r=0.95$, $p<0.001$) (45). Simmons et al. observed a good correlation between tumor volume and final pathology ($r=0.7$) in a study of 27 patients, and the SE and SP of PHS in localizing of lesions ≥ 0.2 mL within a sextant were 90% and 72%, respectively (46). However, some studies arrived at different conclusions. A study of 148 PCa patients indicated that there was no significant correlation between the tumor volume measured by PHS and obtained in the final pathology ($r = -0.0083$, $p = 0.9$) (47). Javed S et al. also showed that the tumor volume measured by PHS was not correlated with the pathological volume after RP ($r = -0.096$) (48).

HS-targeted biopsy of the prostate is gradually being applied in clinical practice, but it still cannot replace the important role of systematic biopsy in detecting PCa. Compared with those of TRUS-guided prostate biopsy and transperineal template prostate biopsy (TTB), the overall cancer detection rates of PHS-targeted biopsy and TRUS-guided systemic biopsy are 38.1% and 62.5%, respectively (48). The total cancer detection rates of PHS-targeted biopsy and standard TTB were 13.4% and 54.4%, respectively (48). The SE and SP of PHS in the posterior gland were 100% and 13%, respectively, and those in the anterior gland were 6% and 82%, respectively (48). Therefore, it is currently not recommended to use HS to reliably identify and characterize PCa. The potential of PHS in assisting in the detection of PCa is considerable, and a larger patient population is still needed to further verify its clinical value.

## COLOR DOPPLER/POWER DOPPLER

Several studies have reported the added value of Doppler technology over grayscale ultrasound (GSU) (9, 49, 50). Color Doppler ultrasound (CDU) and power Doppler ultrasound (PDU) can be used to detect invisible lesions on the GSU by revealing abnormal blood vessels in the tissue. CDU describes the speed and direction of blood flow by detecting the frequency changes when the signal is reflected by red blood cells (51). If the lesion is located in the peripheral zone of the prostate with nodular or clustered hypoechogencity, CDU manifests an intralesional vascular hyperplasia. Then, the lesion is likely to be malignant. Conventional CDU can improve the PCa detection rate (51).

PDU is another method of displaying blood flow in color, but it is more sensitive to perfusion than CDU. However, PDU cannot describe the direction of blood flow. PDU can detect low-velocity blood flow in blood vessels with an inner diameter as small as 1 mm. Okihara et al. used PDU to examine 107 men with high serum PSA levels. The results showed that the SE, SP, PPV and NPV of PDU in detecting of lesions were 98%, 78%, 59% and 99%, respectively (52). Sauvain et al. found that the SE and SP of PDU in detecting low-risk PCa in 243 patients were 45% and 74%, respectively (53). Eisenberg et al. compared GSU and PDU with 620 RP postoperative specimens and reported that adding PDU to GSU increased the SP from 47% to 74%, although the SE was reduced from 58% to 47% (15).

Both CDU and PDU can help identify vascular tissue, and the latter is more sensitive, but neither is sufficient to detect early PCa. Tumor growth and progression within the prostate are usually accompanied by angiogenesis, which may significantly increase the microvessel density (MVD) in the lesions. An increase in MVD is associated with a higher tumor grade and a worse prognosis (51). The limited resolution of Doppler ultrasound can detect blood vessels in the millimeter range, while the angiogenesis of malignant tumors can generate blood vessels as small as 10-50 microns in diameter (51). Therefore, the Doppler technique may be effective only in detecting increased blood flow in large vessels that are found in larger, advanced, high Gleason-grade tumors. Another potential disadvantage of Doppler and other blood flow-based ultrasound techniques is that the left-side lying position often used by patients may result in an asymmetrical distribution of blood flow in the prostate tissue. Harper et al. found that CDU and PDU showed a significant difference in blood flow in tissues ($P<0.002$) that are beneficial to the left side of the prostate instead of the right side (54).

## ELASTOGRAPHY

Ultrasound elastography (UE) can reveal stiff lesions that are not visible on traditional TRUS (17, 55). The main methods for the UE diagnosis of PCa include transrectal real-time tissue elastography (TRTE) and shear-wave elastography (SWE). The index for evaluation with TRTE is the ratio of the stress on the material to the structural deformation caused by the stress, and the index for evaluation with SWE is expressed as the shear wave velocity and Young's modulus.

### Transrectal Real-Time Tissue Elastography

In TRTE, the rectal probe cyclically compresses the suspicious prostate tissue and monitors the degree of elastic strain. The speckle comparison caused by each cycle of compression and decompression will generate a color-coded map, which is then overlaid on the grayscale image of the prostate. Note that the tissue deformation is homogeneous over the imaging plane, and the region of interest (ROI) should cover the entire gland and surrounding tissues to obtain a qualified and reproducible elastogram. Finally, the operator compares the tissue strain ratios of the two ROIs, with one considered "normal" and one considered "abnormal", on the elastogram. On the elastic chart of the TRTE examination, low strain is highlighted by color coding in blue, and the corresponding high strain soft tissue is coded in red. Blue hypoechoic lesions of the prostate are suspected of malignancy. Normally, the stiffness of the glands in the prostate increases with age. PCa tissue is harder than normal prostate tissue due to increased cellular density, microvascularization and stromal reaction combined with collagen deposition in the surrounding prostate parenchyma (56). Thus, the organization of PCa tumors often involves partial or no obvious compression during TRTE inspection. The detection rate of prostate anterior parenchyma is lower than that in the posterior areas, and that of the base of the prostate is also lower than that of the apical regions in TRTE examination (55, 56).

Most studies on prostate elastography have used TRTE. A meta-analysis of 6 studies by Salomon G et al. showed that for TRTE targeted biopsy for PCa detection, the SE and SP per patient were 62% and 79%, respectively, and the SE and SP per core were 34% and 93%, respectively (57). Zhang B et al. compared TRTE with histopathological results after RP in a meta-analysis of 508 patients, and the pooled SE and SP of TRTE in diagnosing PCa were 0.72 and 0.76, respectively (58). Miyanaga et al. analyzed 29 patients with PCa before RP. The results showed that the SE of TRTE, GSU, and DRE in diagnosing PCa were 93%, 59%, and 55%, respectively (59).

Aigner et al. reported that in 94 patients, the SE, SP, PPV, and NPV of TRTE targeted biopsy were 74.0%, 60.0%, 39.0%, and 93.0%, respectively. Furthermore, TRTE-targeted biopsy was better than systemic biopsy, and the detection rate of PCa was 4.7 times higher (60). A comparative study of TRTE targeted guided needle biopsy and systemic biopsy by Brock et al. showed that TRTE had a higher positive rate for prostate needle biopsy than TRUS, but TRTE targeted guided needle biopsy was still unable to replace systemic needle biopsy (55). Therefore, we believe that TRTE-guided targeted biopsy can complement traditional systematic biopsy.

A study of 33 patients showed that the PCa detection rate of TRTE is basically equivalent to that of MRI. The SE and NPV of TRTE were 84.6% and 86.7%, respectively, while those of mpMRI were 84.6% and 83.3%, respectively (61). Pelzer et al. found that the SE and SP of TRTE in diagnosing PCa in 46 patients were 44.1–58.9% and 83.0–74.8%, respectively, while those of MRI were 36.7–43.1% and 85.9–79.8%, respectively (62). TRTE has advantages in the apical and middle parts of the prostate, while MRI has advantages in the gland base and TZ. The combination of the two detection methods can increase the total PCa detection rate (62). A study involving 41 patients showed that lesions on ventral prostate sectors were easier to detect by MRI, while TRTE more easily detected lesions in dorsal and apical sectors. The combination of MRI-TRTE significantly increased the area under the mpMRI curve from 0.65 to 0.75 (63). Brock et al. found that the SE and SP of the combined MRI/TRTE in detecting PCa were 77.8% and 77.3%, respectively (64).

Among the limitations of TRTE are that it performs a semiquantitative analysis of tissue elasticity; it cannot provide uniform compression for the whole gland; it has a low detection rate for small and low-grade prostates (65); and insufficient image acquisition and low reproducibility of the operation when the probe slips off the prostate, as shown for 32% of patients (66). Real-time balloon inflatable elastography (RBIE) has been adopted by clinics as a new technology for supplementing TRTE. It uses a pistol syringe connected to the balloon on the rectal probe to apply force to the prostate through inflation and deflation instead of manual compression. RBIE can more sensitively detect tumors with higher Gleason scores and hard-to-reach tumors in the prostate area. RBIE provides stable elastic motion images and improves the ability of TRTE to detect prostate cancer (67).

## Shear–Wave Elastography

In recent years, SWE has been primarily used for the diagnosis of thyroid, breast and liver diseases. SWE evaluates the hardness of the tissue by measuring the propagation speed of a shear wave delivered to the tissue. It is a quantitative technique that standardizes the detection of prostate cancer. The SWE measurements have excellent in-observer repeatability (ICC = 0.876) (68). However, SWE is plane-dependent, and the hardness of the sagittal image of the prostate is higher than that of the axial image (69); the shear waves attenuate significantly in larger glands; and for larger prostates, it is difficult to perform SWE without prepressurization.

What distinguishes SWE from TRTE is that the former avoids putting pressure on the rectal wall. The color rendering mode of SWE is opposite that of TRTE; low strain is highlighted in red, and soft tissues are shown in blue. Red hypoechoic areas are suspicious of malignant lesions. In young men without prostate hyperplasia, the area around and in the center of the prostate is uniformly displayed in blue, and the stiffness value ranges from 15 to 25 kPa. As prostate hyperplasia develops, the central area of the prostate becomes an uneven red with stiffness values ranging from 30 to 180 kPa, while the surrounding area still maintains a more uniform blue color (70). While attempting not to compress the prostate during SWE examination, the prostate is scanned from base to apex to obtain the original elastic image containing each plane. Then, the operator calculates the elasticity measure (mean, min and max) of each ROI, as well as the ratio between the quantitative box (Q-box) placed in the suspicious prostate area and the adjacent normal surrounding area.

SWE is a commonly used ultrasound imaging method for PCa diagnosis in the clinic and shows good diagnostic value. In a prospective study of 53 patients, a Young's modulus value of 37 kPa was used as the cutoff value between benign and malignant prostate tissues. The SE, SP, PPV and NPV of SWE in detecting PCa were 96.2%, 96.2%, 69.4% and 99.6%, respectively (71). The meta-analysis results of Sang et al. showed that the pooled SE and SP of SWE in diagnosing PCa were 0.844 and 0.860, respectively (72). Boehm K et al. used 50 kPa as the Young's modulus threshold for benign and malignant prostate tissues, and the SE and SP of SWE in detecting PCa were 80.9% and 69.1%, respectively (70). At present, the results of some studies using SWE show that the critical value for distinguishing benign and malignant lesions is in the range of 35 to 43.9 kPa (71, 73, 74).

The increase in PCa tissue stiffness is related to the GS (75) and disease severity (76). The average Young's modulus of prostate cores with a Gleason score of 7 (163 ± 63 kPa) was higher than that of prostate cores with a Gleason score of 6 (95 ± 28.5 kPa; $P = 0.007$) (77). Woo et al. reported that Young's modulus was significantly correlated with the Gleason score ($r = 0.343$, $P = 0.002$) ($r = 0.898$, $P < 0.0001$) (73, 78, 79). Similarly, there is a correlation between the strain index (SD) and the Gleason score. The mean elastic strain index SD (3.26~1.77) of malignant focal lesions was found to be significantly higher than that of benign focal lesions (2.16~1.52; $P<0.008$), and the strain index was moderately linearly correlated with the Gleason score ($r=0.441$; $P=0.013$) (55). This finding may be attributed to the higher cell density and stiffness associated with higher grades of prostate cancer.

Rui et al. reported a new 11-point scoring system based on SWE and other clinical parameters (TRUS, DRE, and free PSA/total PSA ratio), and the results showed that when scoring based on SWE and clinical parameters, the AUC of the system (0.911) was higher than that of SWE alone (0.842) or of clinical parameters (0.868) alone (80). Recently, research has been conducted on the efficacy of 3D SWE in detecting prostate cancer. When the critical value of tissue elasticity of 41 kPa was combined with the PI-RADS score, the SE, SP, PPV and NPV of cancer detection were 70%, 98%, 91% and 92%, respectively (79). In the future, 3D SWE may have the potential to improve the detection of major prostate cancer.

### Acoustic Radiation Force Pulse Imaging

Acoustic radiation force pulse imaging (ARFI), another mode of UE, shows promise in the diagnosis and treatment of PCa. In ARFI, a short-term high-intensity focused ultrasound beam is transmitted to the prostate tissue to displace it. Zhai et al. successfully distinguished benign hyperplastic nodules, calcifications and cancerous lesions in the prostate using ARFI imaging (81). Wang et al. noted that a high-intensity ultrasound pulse can separate prostate cancer tissue from normal tissue, a potential noninvasive prostate cancer resection technique that and has therapeutic value (82).

## CONTRAST–ENHANCED ULTRASOUND

A large number of microvessels are generated inside PCa tumors, which provide the necessary nutrients for tumor proliferation, metastasis and invasion. The density of microvessels in a PCa tumor is significantly higher than that of normal prostate tissue. In contrast-enhanced ultrasound (CEUS), an intravenous injection of ultrasound contrast agent (UCA) with a diameter close to red blood cells is made to observe the blood perfusion of the lesion and adjacent tissues in real time. CEUS can detect blood flow signals in microvessels with a diameter of 1–10 mm (83). The main component of the UCA is microbubbles (MBs), the incidence of allergic reactions is much lower than that of iodine contrast agents (84), and there is no nephrotoxicity. After intravenous injection of the UCA, one ROI is delineated in the suspicious area, and another is drawn in the enhanced normal parenchyma as a reference. The signal intensity change of the contrast agent in the prostate ROI area is plotted over time, which is called the time intensity curve (TIC). PCa tissue shows higher peak enhancement, and a shorter rise time and peak time than normal parenchyma (85).

In a prospective study of 65 patients with elevated PSA, Zhao et al. found that the SE and SP of CEUS in diagnosing PCa were 79.3% and 86.1%, respectively (86). In a meta-analysis of 16 studies with a total of 2624 patients, Li et al. found that the SE, SP, and DOR of CEUS imaging in detecting prostate cancer were 0.70, 0.74 and 9.09, respectively (87). Sedelaar et al. performed three-dimensional contrast-enhanced Doppler ultrasound (3D CE-PDU) on 7 patients with PCa confirmed by biopsy and found that the MVD on the "enhanced" side was 1.93 times that on the "unenhanced" side (88). Using 3D CE-PDU, 86% of cancer foci were found in 70 patients with PCa who planned to undergo RP (89). 3D CE-PDU has the ability to visualize lesions with high MVD.

CEUS-guided prostate targeted biopsy is widely used in clinical PCa detection. In a study of 1,776 men, Mitterberger et al. found that the PCa detection rate of CEUS–targeted biopsy was significantly higher than that of systematic biopsy (10.8% vs. 5.1%) (90). In a study involving 690 patients, Strazdina et al. found that CEUS–guided targeted biopsy had good SE in the detection of PCa with high Gleason scores (6 or higher) (91). Some studies have shown that targeted needle biopsy guided by CE-TRUS can not only improve the diagnostic SE of PCa but also increase the positive rate of needle biopsy (92–94). However, several studies have instead shown that there is no significant difference in the detection rate of PCa between the CEUS guided targeted puncture method and the systematic puncture method (95). CEUS is a promising tool for detecting PCa, but it still cannot completely replace systematic biopsy under existing circumstances.

Compared with other ultrasound modes, CEUS shows good diagnostic value in the diagnosis of PCa. Among 115 men with a serum PSA level greater than 4.0 ng/ml, a study showed that the SE, SP and accuracy of CEUS in diagnosing PCa were 65%, 83% and 73%, respectively, which were higher than those of TRUS and PDU (96). However, Taverna et al. reported that CEUS did not significantly increase the detection rate of PCa over PDU or GSU (18). Some recent studies used CEUS in combination with other ultrasound modes to detect PCa. Halpern et al. and Matsumoto et al. found that the total SE of CEUS and GSU in the diagnosis of PCa in 12 and 50 prostate cancer patients was 42% and 40%, respectively (97, 98). The combination of multiple ultrasound modes can significantly improve the ability to detect PCa clinically.

Contrast-enhanced ultrasound diffusion imaging (CUDI) is a very promising new technique for prostate cancer imaging developed in recent years. It analyzes the time evolution of the UCA concentration in the neovasculature of cancer foci to generate quantitative maps of perfusion parameters to better characterize microvascular structure. Jung et al. measured ultrasound contrast perfusion quantitative parameters in 20 PCa patients, including the early irrigation rate (EIR), mean transit time (MTT) and rise time (RT). The results showed that the SE, SP, NPV and PPV of PCa were 88%, 100%, 60% and 90%, respectively (99). This preliminary study shows that the quantitative analysis of CEUS perfusion parameters can help visualize the microvascular blood circulation and preoperative location of prostate cancer. In a study of 82 patients, Francesco M. Drudi et al. found that the sensitivity of mpMRI and quantitative analysis of contrast-enhanced ultrasound (CEUS) for detecting PCa were 91.3% and 40%, respectively, and the specificity were 66.7% and 97.2%, respectively (100). CUDI has also been studied in three dimensions. In a study using 3D CUDI to detect the PCa tumors in 43 patients, perfusion parameters were significantly different between benign and malignant tissues, including correlation (r) and wash-in time (WIT). The SE and SP of r in detecting PCa were 94% and 50%, and those of WIT were 53% and 81% (101).

Ultrasound molecular imaging is a new direction in the field of the early diagnosis of tumors. Due to the size limitation of MBs, CEUS is limited to the vasculature where MBs accumulate in the tumor. Only particles with a diameter of less than 700 nm can penetrate the tumor blood vessel wall and enter the tumor interstitium (102). Prostate-specific membrane antigen (PSMA) is a type II glycoprotein that is mainly distributed in prostate epithelial cells. It is highly expressed in prostate intraepithelial neoplasia, hormone-dependent or hormone-independent prostate cancer, and metastatic cancer (103) but expressed at low levels in normal prostate epithelial cells. This feature makes it one of the most important biomarkers in the diagnosis and treatment of PCa. Therefore, some studies focused on the construction of targeted nanobubbles (NBs) with a diameter of less than 700 nm to achieve specific ultrasound-enhanced imaging of prostate cancer cells (102–104). At present, PSMA-targeting, indocyanine green (ICG)-loaded nanobubbles (NBs) (102) and PSMA single-chain variable fragment (scFv)–loaded NBs have been reported (104). These new targeted NBs have been proven to be excellent US contrast agents that extend the signal enhancement time and have stronger penetrating ability and higher specificity (105, 106). If the NBs are loaded with drugs, targeted therapy of PCa can also be achieved.

## MRI/US FUSION IMAGING

When a lesion is detected on MRI, MRI/US fusion can be helpful (**Figure 1**). A number of studies have demonstrated that MRI/US fusion technology-guided biopsy improves the detection rate of PCa. Brock et al. found that using MRI/TRUS fusion targeted biopsy in 121 men, the SE and SP in the detection of PCa were 77.8% and 77.3%, respectively, and the detection rate per core for combined targeted biopsy (14.7%) was higher than the detection rate per core of system biopsy (6.5%, p <0.001) (107). In a retrospective study of 135 patients, MRI combined with 3D TRUS targeted needle biopsy was performed before RP, and the SE of the detection of prostate index tumors was 95% (108). Siddiqui MM et al. compared 1003 patients with MRI/US combined with prostate targeted biopsy and standard biopsy. The results showed that the accuracy of targeted biopsy was 30% higher than that of standard biopsy in diagnosing high-risk cancers (109). Tewes S et al. reported the SE, SP, and NPV of MRI/TRUS combined-guided targeted biopsy in detecting prostate lesions with PI-RADS scores ≥ 4 were 85%, 82% and 92%, respectively (110). US and MRI have advantages in the diagnosis of PCa, and when combined, the detection of PCa is obviously improved.

## MPUS

Transrectal multimodal ultrasound refers to a combination of GSU, CDU, PDU, TRTE, and CEUS. GSU shows the anatomical location of the prostate lesion (**Figure 2A**). Doppler ultrasound shows the blood flow in the larger hyperplastic vessels in the lesion (**Figure 2B**). Elastography shows the hardness of the lesion tissue to infer properties about its nature (**Figure 2C**). CEUS shows new microvessels in the lesion (**Figure 2D**). Clinically, the combination of different ultrasound modes can improve the detection rate of PCa. At present, there are few studies on the combination of ultrasound modes.

Nelson et al. used GSU, PDU, and TRTE for the targeted biopsy of prostate lesions in 137 patients. The results showed that in 106 positive sextant sites, the positive rates of GSU, CDU, TRTE, and combined ultrasound modes were 16%, 29%, 25%, and 46%, respectively, indicating that combined ultrasound methods with different modes can improve the detection rate of PCa (17). Xie et al. conducted transrectal GSU, PDU and their combination with a third mode (DCE-US) to detect PCa in 150 patients. The results showed that the sensitivities of the combined mode (GSU+PDU), GSU, PDU and DCE-US, were 73%, 51%, 48%, and 63%, respectively (111). In a retrospective study involving 133 men with elevated serum levels of PSA (≥1. 25 ng/mL), the PCa detection rate of CEUS and TRTE combined targeted biopsy was

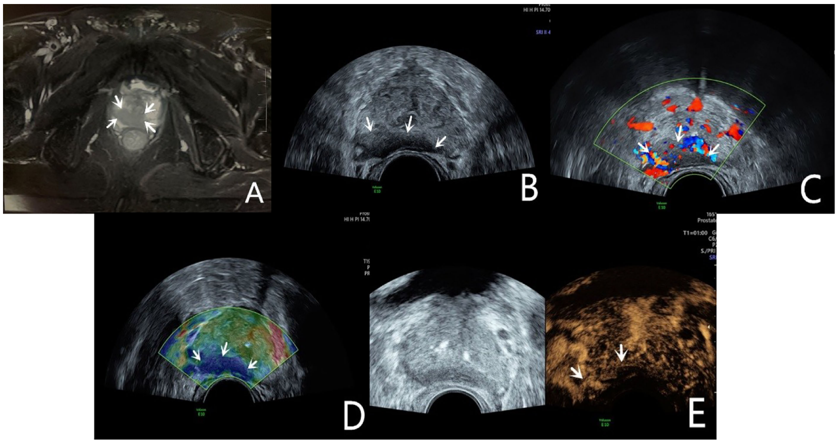


**FIGURE 1** | Diagnosis of PCa using mpUS—case 1. A 68-year-old patient has a total serum PSA of 10.4ng/ml. The T2-weighted sequence of MRI (**A**, arrow) shows a slightly low signal shadow in the peripheral zone of the prostate, suggesting PCa in the diagnosis. The lesion showed a slightly hypoechoic area on the B-mode (**B**, arrow), and it's not clearly demarcated from the seminal vesicle gland. CDU shows an abnormally increased blood flow in the lesion (**C**, arrow). TRTE shows that the slightly hypoechoic area of the prostate's peripheral zone is highly stiff (**D**, arrow). CEUS shows early high enhancement within the lesion (**E**, arrow). TRUS-guided systematic biopsy confirmed that the peripheral zone of the prostate was a Gleason 4 + 4 PCa.

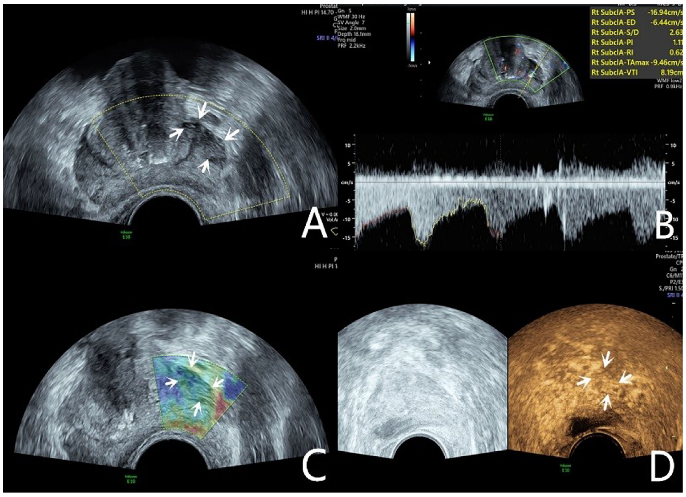


**FIGURE 2** | Diagnosis of PCa using mpUS—case 2. A 76-year-old patient has a total PSA of 14.3ng/ml. Multiparameter US starts from conventional transrectal ultrasound, and the lesion is a hypoechoic nodule at the junction of the inner and outer glands in the prostate's left lobe (**A**, arrow). The lesion appears on the CDU as a rich blood flow of the arterial spectrum (**B**, arrow). The operator uses the endocavitary transducer to alternately compress and decompress the lesion, which appears mostly blue on the TRTE (**C**, arrow). Hypoechoic nodule appears on CEUS as hypervascular nodule with enhanced "fast forward and fast exit" (**D**, arrow). Histopathology shows that the prostate lesions were clinically significant with a Gleason 4 + 3 PCa.

59.4% (112). Brock et al. performed TRTE and CEUS examinations on 100 patients before RP. Compared with pathological biopsy, the examinations showed a SE and SP in detecting PCa by TRTE of 49% and 74%, respectively. Compared with the combination of TRTE and CEUS, the false positive value of TRTE alone was reduced from 34.9% to 10.3%, and the PPV of cancer detection was increased from 65% to 90% (64). Among 153 prostate nodules, the SE, accuracy and NPV of the combination of TRTE and CEUS in diagnosing PCa were 92.1%, 86.2%, and 84.6%, respectively. Multiple ultrasound imaging modes combined with targeted-guided prostate biopsy can not only increase the detection rate of malignant lesions but also reduce the number of tissue punctures.

Mp-US and mp-MRI provide complementary information in the diagnosis of PCa. Zhang et al. performed mp-US and mp-MRI examinations on 88 patients. The results showed that the SE, NPV, accuracy, and AUC in detecting PCa with mp-US were higher than that mp-MRI (97.4% vs. 94.7%, 96.9% vs. 92.3%, 87.2% vs. 76.9%, 0.874 vs. 0.774, respectively) (113). In 167 patients with primary prostate biopsy, Pat F. Fulgham et al. found that mpUS-targeted biopsy was superior to mpMRI/TRUS fusion-targeted biopsy in terms of the positive rate of PCa and the ability to detect low-malignancy PCa (114). Mp-US has higher diagnostic performance than mp-MRI in diagnosing local PCa.

## CONCLUSION

Due to the poor prognosis of metastatic PCa, early detection of PCa is the most effective strategy to reduce morbidity and mortality. MRI experts from the European Society of Urogenital Radiology (ESUR) developed the PI-RADS scoring system for prostate mpMRI and used Likert-type scales to score the corresponding lesions. At present, mpMRI is still used as the main imaging method for diagnosing PCa in clinical practice, and no multiparameter ultrasound image scoring system has been developed. Ultrasound is also very important in the imaging diagnosis of PCa, due to its low cost, ease of use, real-time functionality, lack of radiation, and the continuing development of more advanced ultrasound techniques. Polymeric NBs targeting PSMA as a new UCA can increase the diagnostic potential of CEUS and may become a popular research topic for targeted ultrasound molecular imaging of PCa. In addition, NBs can be used as drug carriers for PCa-targeted therapy. Ultrasound molecular imaging has become an emerging research field in tumor imaging diagnosis. Our future work will focus on accumulating more patient data, integrating the diagnostic characteristics of PCa under different ultrasound modes, and constructing a complete ultrasound examination scoring system through optimized algorithms. The development of this advanced mpUS scoring system will have important clinical application value in improving PCa diagnosis and follow-up.

## AUTHOR CONTRIBUTIONS

YL analyzed relevant literatures and drafted the article. RX made a critical review of the intellectual content of the article. SZ provided project funding. All authors read and approved the final manuscript.

## FUNDING

State Natural Sciences Foundation of China,Grant/Award Numbers: 81871372. The project leader is SZ.

# REFERENCES

1. Lina S, Li Q, Cheng K, Zhang M, Liu Y. Value of Diffusion-Weighted Imaging in Distinguishing Prostate Cancer From Benign Prostatic Hyperplasia: A Meta-Analysis. *Chin J Med Imaging* (2013) 21(6):464–7.
2. Chiu KF, Ng CF, Semjonow A, Zhu Y, Vincendeau S, Houlgatte A, et al. A Multicentre Evaluation of the Role of the Prostate Health Index (PHI) in Regions With Differing Prevalence of Prostate Cancer: Adjustment of PHI Reference Ranges Is Needed for European and Asian Settings. *Eur Urol* (2019) 75(4):8100–04. doi: 10.1016/j.eururo.2018.10.047
3. Mottet N, Bellmunt J, Bolla M, Briers E, Cumberbatch MG, De Santis M, et al. EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment With Curative Intent. *Eur Urol* (2017) 71 (4):618–29. doi: 10.1016/j.eururo.2016.08.003
4. Presti JC, O'dowd GJ, Miller MC, Mattu R, Veltri RW. Extended Peripheral Zone Biopsy Schemes Increase Cancer Detection Rates and Minimize Variance in Prostate Specific Antigen and Age Related Cancer Rates: Results of a Community Multi-Practice Study. *J Urol* (2003) 169(1):125–9. doi: 10.1016/S0022-5347(05)64051-7
5. Puech P, Rouvi Re O, Renard-Penna R, Villers A, Devos P, Colombel M, et al. Prostate Cancer Diagnosis: Multiparametric MR-Targeted Biopsy With Cognitive and Transrectal US-MR Fusion Guidance Versus Systematic Biopsy–Prospective Multicenter Study. *Radiology* (2013) 268 (2):461–9. doi: 10.1148/radiol.13121501
6. Mian BM, Naya Y, Okihara K, Vakar-Lopez F, Troncoso P, Babaian RJ, et al. Predictors of Cancer in Repeat Extended Multisite Prostate Biopsy in Men With Previous Negative Extended Multisite Biopsy. *J Urol* (2003) 60(5):836–40. doi: 10.1016/S0090-4295(02)01950-7
7. Singh H, Canto EI, Shariat SF, Kadmon D, Miles BJ, Wheeler TM, et al. Predictors of Prostate Cancer After Initial Negative Systematic 12 Core Biopsy. *J Urol* (2004) 171(5):1850–4. doi: 10.1097/01.ju.0000119667.86071.e7
8. Delongchamps NB, Haas GP. Saturation Biopsies for Prostate Cancer: Current Uses and Future Prospects. *Nat Rev Urol* (2009) 6(12):645–52. doi: 10.1038/nrurol.2009.213
9. Giannarini G, Autorino R, Lorenzo GD. Saturation Biopsy of the Prostate: Why Saturation Does Not Saturate. *Eur Urol* (2009) 56(4):619–21. doi: 10.1016/j.eururo.2009.03.044
10. Rosario DJ, Lane JA, Metcalfe C, Donovan JL, Doble A, Goodwin L, et al. Short Term Outcomes of Prostate Biopsy in Men Tested for Cancer by Prostate Specific Antigen: Prospective Evaluation Within ProtecT Study. *Yearbook Urol* (2014) 2012(3):55–6. doi: 10.1016/j.juro.2012.05.070
11. Oberlin DT, Casalino DD, Miller FH, Meeks JJ. Dramatic Increase in the Utilization of Multiparametric Magnetic Resonance Imaging for Detection and Management of Prostate Cancer. *Abdom Radiol* (2016) 195(4):1–4. doi: 10.1007/s00261-016-0975-5
12. Pummer K, Rieken M, Augustin H, Gutschi T, Shariat SF. Innovations in Diagnostic Imaging of Localized Prostate Cancer. *World J Urol* (2014) 32 (4):881–90. doi: 10.1007/s00345-013-1172-6
13. Ellis JH, Tempany C, Sarin MS, Gatsonis C, Mcneil BJ. MR Imaging and Sonography of Early Prostatic Cancer: Pathologic and Imaging Features That Influence Identification and Diagnosis. *Am J Roentgenol* (1994) 162 (4):865–72. doi: 10.2214/ajr.162.4.8141009
14. Heijmink S, Tterer J F, Strum SS, Oyen WJG, Frauscher F, Witjes JA, et al. State-Of-the-Art Uroradiologic Imaging in the Diagnosis of Prostate Cancer. *Acta Oncol* (2011) 50 Suppl 1:25–38. doi: 10.3109/0284186X.2010.578369
15. Eisenberg ML, Cowan JE, Carroll PR, Shinohara K. The Adjunctive Use of Power Doppler Imaging in the Preoperative Assessment of Prostate Cancer. *BJU Int* (2010) 105(9):1237–41. doi: 10.1111/j.1464-410X.2009.08958.x
16. Brock M, Bodman CV, Sommerer F, Ppenberg BL, Klein T, Deix T, et al. Comparison of Real-Time Elastography With Grey-Scale Ultrasonography for Detection of Organ-Confined Prostate Cancer and Extra Capsular Extension : A Prospective Analysis Using Whole Mount Sections After Radical Prostatectomy. *BJU Int* (2011) 108(8b):E217–E22. doi: 10.1111/j.1464-410X.2011.10209.x
17. Nelson ED, Slotoroff CB, Gomella LG, Halpern EJ. Targeted Biopsy of the Prostate: The Impact of Color Doppler Imaging and Elastography on Prostate Cancer Detection and Gleason Score. *Urology* (2007) 70(6):1136–40. doi: 10.1016/j.urology.2007.07.067
18. Taverna G, Morandi G, Seveso M, Giusti G, Graziotti P. Colour Doppler and Microbubble Contrast Agent Ultrasonography Do Not Improve Cancer Detection Rate in Transrectal Systematic Prostate Biopsy Sampling. *BJU Int* (2011) 108(11):1723–7. doi: 10.1111/j.1464-410X.2011.10199.x
19. Klein J, De Górski A, Benamran D, Vallee JP, De Perrot T, Wirth GJ, et al. Transrectal Ultrasound-Guided Prostate Biopsy for Cancer Detection: Performance of 2D-, 3D- and 3D-MRI Fusion Targeted Techniques. *Urol Int* (2017) 98(1):7–14. doi: 10.1159/000452250
20. Hwang ST, Lee HJ. The Future Perspectives in Transrectal Prostate Ultrasound Guided Biopsy. *Prostate Int* (2014) 2(4):153–60. doi: 10.12954/PI.14062
21. Ahmed HU, El-Shater Bosaily A, Brown LC, Gabe R, Kaplan R, Parmar MK, et al. Diagnostic Accuracy of Multi-Parametric MRI and TRUS Biopsy in Prostate Cancer (PROMIS): A Paired Validating Confirmatory Study. *Lancet* (2017) 389(10071):815–22. doi: 10.1016/S0140-6736(16)32401-1
22. Heidenreich A, Abrahamsson PA, Artibani W, Catto J, Montorsi F, Van Poppel H, et al. Early Detection of Prostate Cancer: European Association of Urology Recommendation. *Eur Urol* (2013) 64(3):347–54. doi: 10.1016/j.eururo.2013.06.051
23. Long JA, Daanen V, Moreau-Gaudry A, Troccaz J, Rambeaud JJ, Descotes JL, et al. Prostate Biopsies Guided by Three-Dimensional Real-Time (4-D) Transrectal Ultrasonography on a Phantom: Comparative Study Versus Two-Dimensional Transrectal Ultrasound-Guided Biopsies. *Eur Urol* (2007) 52(4):1097–105. doi: 10.1016/j.eururo.2006.11.034
24. Zhao HX, Zhu Q, Wang ZC. Detection of Prostate Cancer With Three-Dimensional Transrectal Ultrasound: Correlation With Biopsy Results. *Br J Radiol* (2012) 85(1014):714–9. doi: 10.1259/bjr/68418881
25. Guo Y, Werahera PN, Narayanan R, Li L, Kumar D, Crawford ED, et al. Image Registration Accuracy of a 3-Dimensional Transrectal Ultrasound-Guided Prostate Biopsy System. *J Ultrasound Med* (2009) 28(11):1561–8. doi: 10.7863/jum.2009.28.11.1561
26. Zhang M, Wang R, Wu Y, Jing J, Chen M. Micro-Ultrasound Imaging for Accuracy of Diagnosis in Clinically Significant Prostate Cancer: A Meta-Analysis. *Front Oncol* (2019) 9:1368. doi: 10.3389/fonc.2019.01368
27. Lughezzani G, Saita A, Lazzeri M, Paciotti M, Maffei D, Lista G, et al. Comparison of the Diagnostic Accuracy of Micro-Ultrasound and Magnetic Resonance Imaging/Ultrasound Fusion Targeted Biopsies for the Diagnosis of Clinically Significant Prostate Cancer. *Eur Urol* (2018) 10):329–32. doi: 10.1016/j.euo.2018.10.001
28. Avolio PP, Paciotti M, Maffei D, Uleri A, Frego N, Fasuolo V, et al. The Use of 29 MHz Transrectal Micro-Ultrasound to Stratify the Prostate Cancer Risk in Patients With PI-RADS III Lesions at Multiparametric MRI: A Single Institutional Analysis. *Urol Oncol* (2021) 32(10):S50. doi: 10.1016/S2666-1683(21)00792-8
29. Sountoulides P, Pyrgidis N, Polyzos SA, Mykoniatis I, Asouhidou E, Papatsoris A, et al. Micro-Ultrasound-Guided vs Multiparametric Magnetic Resonance Imaging-Targeted Biopsy in the Detection of Prostate Cancer: A Systematic Review and Meta-Analysis. *J Urol* (2021) 205(5):1254–62. doi: 10.1097/JU.0000000000001639
30. Wiemer L, Hollenbach M, Heckmann R, Kittner B, Cash H. Evolution of Targeted Prostate Biopsy by Adding Microultrasound to the Magnetic Resonance Imaging Pathway. *Eur Urol* (2021) 7(11):1292–9. doi: 10.1016/j.euf.2020.06.022
31. Pepe P, Garufi A, Priolo G, Pennisi M. Transperineal Versus Transrectal MRI/TRUS Fusion Targeted Biopsy: Detection Rate of Clinically Significant Prostate Cancer. *Clin Genitourinary Cancer* (2017) 15(1):e33–6. doi: 10.1016/j.clgc.2016.07.007
32. Kongnyuy M, George AK, Rastinehad AR, Pinto PA. Magnetic Resonance Imaging Ultrasound Fusion-Guided Prostate Biopsy: Review of Technology, Techniques and Outcomes. *Curr Urol Rep* (2016) 17(4):17:32. doi: 10.1007/s11934-016-0589-z
33. Pepe P, Aragona F. Morbidity After Transperineal Prostate Biopsy in 3000 Patients Undergoing 12 vs 18 vs More Than 24 Needle Cores. *Urology* (2013) 81(6):1142–6. doi: 10.1016/j.urology.2013.02.019
34. Rastinehad AR, Abboud SF, George AK, Frye TP, Pinto PA. Reproducibility of Multiparametric MRI and Fusion-Guided Prostate Biopsy: Multi-

Institutional External Validation by a Propensity Score Matched Cohort. *J Urol* (2016) 195(6):1737–43. doi: 10.1016/j.juro.2015.12.102

35. Roethke MC, Kuru TH, Schultze S, Tichy D, Hadaschik BA. Evaluation of the ESUR PI-RADS Scoring System for Multiparametric MRI of the Prostate With Targeted MR/TRUS Fusion-Guided Biopsy at 3.0 Tesla. *Eur Radiol* (2014) 24(2):344–52. doi: 10.1007/s00330-013-3017-5
36. Grabski B, Baeurle L, Loch A, Wefer B, Paul U, Loch T. Computerized Transrectal Ultrasound of the Prostate in a Multicenter Setup (C-TRUS-MS): Detection of Cancer After Multiple Negative Systematic Random and in Primary Biopsies. *World J Urol* (2011) 29(5):573–9. doi: 10.1007/s00345-011-0713-0
37. Loch T. Computerized Transrectal Ultrasound (C-TRUS) of the Prostate: Detection of Cancer in Patients With Multiple Negative Systematic Random Biopsies. *World J Urol* (2007) 25(4):375–80. doi: 10.1007/s00345-007-0181-8
38. Walz J, Thomassin-Piana J, Poizat F, Vallier C, Van Hove A, Brunelle S, et al. 2033 External Validation of the ANNA/C-TRUS System Regarding the Correct Identification of Prostate Cancer Lesions in the Diagnosis of Prostate Cancer. *Eur Urol Suppl* (2012) 187(1):e1050. doi: 10.1016/j.juro.2012.02.2197
39. Loch T, Loch A, Grabski B, Baeurle L, Paul U, Tokas T. Computerized Transrectal Ultrasound Targeted Biopsies: Preoperative Prediction of the Radical Prostatectomy Gleason Grade. *J Urol* (2014) 191(4):e470–1. doi: 10.1016/j.juro.2014.02.1184
40. Strunk T, Decker G, Willinek W, Mueller SC, Rogenhofer S. Combination of C-TRUS With Multiparametric MRI: Potential for Improving Detection of Prostate Cancer. *World J Urol* (2014) 32(2):335–9. doi: 10.1007/s00345-012-0924-z
41. Nunez-Mora C, Garcia-Mediero JM, Patiño P, Orellana C, Garrido A, Rojo A, et al. Utility of Histoscanning (TM) Prior to Prostate Biopsy for the Diagnosis of Prostate Adenocarcinoma. *Actas Urol* (2013) 37(6):342–6. doi: 10.1016/j.acuroe.2013.01.003
42. Braeckman J, Autier P, Soviany C, Nir R, Nir D, Michielsen D, et al. The Accuracy of Transrectal Ultrasonography Supplemented With Computer-Aided Ultrasonography for Detecting Small Prostate Cancers. *BJU Int* (2008) 102(11):1560–5. doi: 10.1111/j.1464-410X.2008.07878.x
43. Macek P, Barret E, Sanchez-Salas R, Galiano M, Rozet F, Ahallal Y, et al. Prostate Histoscanning in Clinically Localized Biopsy Proven Prostate Cancer: An Accuracy Study. *J Endourol* (2014) 28(3):371–6. doi: 10.1089/end.2013.0419
44. Sivaraman A, Sanchez-Salas R, Barret E, Macek P, Validire P, Galiano M, et al. Prostate Histoscanning True Targeting Guided Prostate Biopsy: Initial Clinical Experience. *World J Urol* (2015) 33(10):1475–9. doi: 10.1007/s00345-014-1434-y
45. Braeckman J, Autier P, Garbar C, Marichal MP, Emberton M. Computer-Aided Ultrasonography (HistoScanning): A Novel Technology for Locating and Characterizing Prostate Cancer. *BJU Int* (2008) 101(3):293–8. doi: 10.1111/j.1464-410X.2007.07232.x
46. Simmons LAM, Autier P, Zát'ura F, Braeckman J, Emberton M. Detection, Localisation and Characterisation of Prostate Cancer by Prostate HistoScanning? *BJU Int* (2012) 110(1):28–35. doi: 10.1111/j.1464-410X.2011.10734.x
47. Schiffmann J, Fischer J, Tennstedt P, Beyer B, Salomon G. Comparison of Prostate Cancer Volume Measured by HistoScanning and Final Histopathological Results. *World J Urol* (2013) 32(4):939–44. doi: 10.1007/s00345-013-1211-3
48. Javed S, Chadwick E, Edwards AA, Beveridge S, Laing R, Bott S, et al. Does Prostate HistoScanning$^{TM}$ Play a Role in Detecting Prostate Cancer in Routine Clinical Practice? Results From Three Independent Studies. *BJU Int* (2015) 114(4):541–8. doi: 10.1111/bju.12568
49. Nougaret S, Robertson N, Golia Pernicka J, Molinari N, Hötker AM, Ehdaie B, et al. The Performance of PI-RADSv2 and Quantitative Apparent Diffusion Coefficient for Predicting Confirmatory Prostate Biopsy Findings in Patients Considered for Active Surveillance of Prostate Cancer. *Abdom Radiol* (2017) 42(7):1968–74. doi: 10.1007/s00261-017-1086-7
50. Faria R, Soares MO, Spackman E, Ahmed HU, Brown LC, Kaplan R, et al. Optimising the Diagnosis of Prostate Cancer in the Era of Multiparametric Magnetic Resonance Imaging: A Cost-Effectiveness Analysis Based on the Prostate MR Imaging Study (PROMIS) - Sciencedirect. *Eur Urol* (2018) 73 (1):23–30. doi: 10.1016/j.eururo.2017.08.018
51. Russo G, Mischi M, Scheepens W, Rosette JJDL, Wijkstra H. Angiogenesis in Prostate Cancer: Onset, Progression and Imaging. *BJU Int* (2013) 110(11C): e794–808. doi: 10.1111/j.1464-410X.2012.11444.x
52. Okihara K, Kojima M, Nakanouchi T, Okada K, Miki T. Transrectal Power Doppler Imaging in the Detection of Prostate Cancer. *BJU Int* (2015) 85 (9):1053–7. doi: 10.1046/j.1464-410x.2000.00663.x
53. Sauvain JL, Sauvain E, Rohmer P, Louis D, Nader N, Papavero R, et al. Value of Transrectal Power Doppler Sonography in the Detection of Low-Risk Prostate Cancers. *Diagn Intervent Imaging* (2012) 94(1):60–7. doi: 10.1016/j.diii.2012.09.003
54. O'Kane P, Gomella LG, Frauscher F, Halpern EJ, Forsberg F, Nazarian LN, et al. High-Frequency Doppler US of the Prostate: Effect of Patient Position. *Radiology* (2002) 222(3):634–9. doi: 10.1148/radiol.2223010946
55. Brock M, Von Bodman C, Palisaar RJ, Löppenberg B, Sommerer F, Deix T, et al. The Impact of Real-Time Elastography Guiding a Systematic Prostate Biopsy to Improve Cancer Detection Rate: A Prospective Study of 353 Patients. *J Urol* (2012) 187(6):2039–43. doi: 10.1016/j.juro.2012.01.063
56. Good DW, Stewart GD, Hammer S, Scanlan P, Shu W, Phipps S, et al. Elasticity as a Biomarker for Prostate Cancer: A Systematic Review. *BJU Int* (2014) 113(4):523–34. doi: 10.1111/bju.12236
57. Salomon G. Transrectal Sonoelastography in the Detection of Prostate Cancers: A Meta-Analysis. *BJU Int* (2013) 110(11B):e621–1. doi: 10.1111/j.1464-410X.2012.11349.x
58. Zhang B, Ma X, Zhan W, Zhu F, Li M, Huang J, et al. Real-Time Elastography in the Diagnosis of Patients Suspected of Having Prostate Cancer: A Meta-Analysis. *Ultrasound Med Biol* (2014) 40(7):1400–7. doi: 10.1016/j.ultrasmedbio.2014.02.020
59. Miyanaga N, Akaza H, Yamakawa M, Oikawa T, Shiina T. Tissue Elasticity Imaging for Diagnosis of Prostate Cancer: A Preliminary Report. *Int J Urol* (2010) 13(12):1514–8. doi: 10.1111/j.1442-2042.2006.01612.x
60. Aigner F, Pallwein L, Junker D, SchFer G, Mikuz G, Pedross F, et al. Value of Real-Time Elastography Targeted Biopsy for Prostate Cancer Detection in Men With Prostate Specific Antigen 1.25 Ng/Ml or Greater and 4.00 Ng/Ml or Less. *J Urol* (2010) 184(3):913–7. doi: 10.1016/j.juro.2010.05.026
61. Aigner F, Pallwein L, Schocke M, Lebovici A, Frauscher F. Comparison of Real-Time Sonoelastography With T2-Weighted Endorectal Magnetic Resonance Imaging for Prostate Cancer Detection. *J Ultrasound Med* (2011) 30(5):643–9. doi: 10.7863/jum.2011.30.5.643
62. Pelzer AE, Heinzelbecker J, Weiß C, Frühbauer D, Weidner AM, Kirchner M, et al. Real-Time Sonoelastography Compared to Magnetic Resonance Imaging Using Four Different Modalities at 3.0 T in the Detection of Prostate Cancer: Strength and Weaknesses. *Eur J Radiol* (2013) 82(5):814–21. doi: 10.1016/j.ejrad.2012.11.035
63. Brock M, Roghmann F, Sonntag C, Sommerer F, Tian Z, LoPpenberg B, et al. Fusion of Magnetic Resonance Imaging and Real-Time Elastography to Visualize Prostate Cancer: A Prospective Analysis Using Whole Mount Sections After Radical Prostatectomy. *Ultraschall der Med (Stuttgart Germany: 1980)* (2015) 36(4):355–61. doi: 10.1055/s-0034-1366563
64. Brock M, Eggert T, Palisaar RJ, Roghmann F, Braun K, LoPpenberg B, et al. Multiparametric Ultrasound of the Prostate: Adding Contrast Enhanced Ultrasound to Real-Time Elastography to Detect Histopathologically Confirmed Cancer. *J Urol* (2013) 189(1):93–8. doi: 10.1016/j.juro.2012.08.183
65. Barr RG, Cosgrove D, Brock M, Cantisani V, Dietrich CF. WFUMB Guidelines and Recommendations on the Clinical Use of Ultrasound Elastography: Part 5.Prostate. *Ultrasound Med Biol* (2017) 43(1):27–48. doi: 10.1016/j.ultrasmedbio.2016.06.020
66. Miyagawa T, Tsutsumi M, Matsumura T, Kawazoe N, Ishikawa S, Shimokama T, et al. Real-Time Elastography for the Diagnosis of Prostate Cancer: Evaluation of Elastographic Moving Images. *Jpn J Clin Oncol* (2009) 39(6):394–8. doi: 10.1093/jjco/hyp026
67. Tsutsumi M, Miyagawa TT, Endo T, Kandori S, Shimokama T, Ishikawa S. Real-Time Balloon Inflation Elastography for Prostate Cancer Detection and Initial Evaluation of Clinicopathologic Analysis. *AJR Am J Roentgenol* (2010) 194(6):W471–w476. doi: 10.2214/AJR.09.3301

68. Woo S, Kim SY, Lee MS, Cho JY, Kim SH. Shear Wave Elastography Assessment in the Prostate: An Intraobserver Reproducibility Study. *Clin Imaging* (2015) 39(3):484–7. doi: 10.1016/j.clinimag.2014.11.013
69. Gennisson JL, Deffieux T, Macé E, Montaldo G, Tanter M. Viscoelastic and Anisotropic Mechanical Properties of *In Vivo* Muscle Tissue Assessed by Supersonic Shear Imaging. *Ultrasound Med Biol* (2010) 36(5):789–801. doi: 10.1016/j.ultrasmedbio.2010.02.013
70. Boehm K, Salomon G, Beyer B, Schiffmann J, Tennstedt P, Budaeus L. Shear Wave Elastography for Localization of Prostate Cancer Lesions and Assessment of Elasticity Thresholds: Implications for Targeted Biopsies and Active Surveillance Protocols. *J Urol* (2014) 191(4):794–800. doi: 10.1016/j.juro.2014.09.100
71. Kaplan SA. Re: Shear Wave Ultrasound Elastography of the Prostate: Initial Results. *J Urol* (2013) 189(1):229. doi: 10.1016/j.juro.2012.09.140
72. Sang L, Wang XM, Xu DY, Cai YF. Accuracy of Shear Wave Elastography for the Diagnosis of Prostate Cancer: A Meta-Analysis. *Sci Rep* (2017) 7(1):1949. doi: 10.1038/s41598-017-02187-0
73. Woo S, Kim SY, Cho JY, Kim SH. Shear Wave Elastography for Detection of Prostate Cancer: A Preliminary Study. *Korean J Radiol* (2014) 15(3):346. doi: 10.3348/kjr.2014.15.3.346
74. Correas JM, Tissier AM, Khairoune A, Khoury G, Hélénon O. Ultrasound Elastography of the Prostate: State of the Art. *Diagn Intervent Imaging* (2013) 94(5):551–60. doi: 10.1016/j.diii.2013.01.017
75. Ahn BM, Kim J, Ian L, Rha KH, Kim HJ. Mechanical Property Characterization of Prostate Cancer Using a Minimally Motorized Indenter in an *Ex Vivo* Indentation Experiment. *Urology* (2010) 76 (4):1007–11. doi: 10.1016/j.urology.2010.02.025
76. Carson WC, Gerling GJ, Krupski TL, Kowalik CG, Harper JC, Moskaluk CA, et al. Material Characterization of *Ex Vivo* Prostate Tissue *via* Spherical Indentation in the Clinic. *Med Eng Phys* (2011) 33(3):302–9. doi: 10.1016/j.medengphy.2010.10.013
77. Ahmad S, Rui C, Varghese T, Bidaut L, Nabi G. Transrectal Quantitative Shear Wave Elastography in the Detection and Characterisation of Prostate Cancer. *Surg Endoscopy* (2013) 27(9):3280–7. doi: 10.1007/s00464-013-2906-7
78. Correas JM, Tissier AM, Khairoune A, Vassiliu V, Méjean A, Hélénon O, et al. Prostate Cancer: Diagnostic Performance of Real-Time Shear-Wave Elastography. *Radiology* (2015) 275(1):280–9. doi: 10.1148/radiol.14140567
79. Shoji S, Hashimoto A, Nakamura T, Hiraiwa S, Sato H, Sato Y, et al. Novel Application of Three-Dimensional Shear Wave Elastography in the Detection of Clinically Significant Prostate Cancer. *Biomed Rep* (2018) 8 (4):373–7. doi: 10.3892/br.2018.1059
80. Su R, Xu G, Xiang L, Ding S, Wu R. A Novel Scoring System for Prediction of Prostate Cancer Based on Shear Wave Elastography and Clinical Parameters. *Urology* (2018) 121(11):112–7. doi: 10.1016/j.urology.2018.08.026
81. Zhai L, Madden J, Foo WC, Rosenzweig S, Palmeri ML, Mouraviev V, et al. Acoustic Radiation Force Impulse Imaging of Human Prostates *Ex Vivo*. *Ultrasound Med Biol* (2010) 36(1):576–88. doi: 10.1016/j.ultrasmedbio.2009.12.006
82. Wang TY. Imaging Feedback of Histotripsy Treatments Using Ultrasound Shear Wave Elastography. *IEEE Trans Ultrason Ferroelectr Freq Control* (2012) 59(6):1167–81. doi: 10.1109/TUFFC.2012.2307
83. Chen M, Zhang KQ, Xu YF, Zhang S, Cao Y, Sun W. Shear Wave Elastography and Contrast-Enhanced Ultrasonography in the Diagnosis of Thyroid Malignant Nodules. *Mol Clin Oncol* (2016) 5(6):724–30. doi: 10.3892/mco.2016.1053
84. Sidhu P, Cantisani V, Dietrich C, Gilja O, Saftoiu A, Bartels E, et al. The EFSUMB Guidelines and Recommendations for the Clinical Practice of Contrast-Enhanced Ultrasound (CEUS) in non-Hepatic Applications: Update 2017 (Long Version). *Ultraschall der Med* (2018) 39(2):154–80. doi: 10.1055/a-0586-1107
85. Andreas M, Thomas F, Julia S, Alexander B, Carsten S, Robert P, et al. Contrast-Enhanced Ultrasound (CEUS) and Quantitative Perfusion Analysis in Patients With Suspicion for Prostate Cancer. *Ultraschall der Med* (2018) 40(3):340–8. doi: 10.1055/a-0594-2093
86. Zhao HX, Xia CX, Yin HX, Guo N, Zhu Q. The Value and Limitations of Contrast-Enhanced Transrectal Ultrasonography for the Detection of Prostate Cancer. *Eur J Radiol* (2013) 82(11):e641–7. doi: 10.1016/j.ejrad.2013.07.004
87. Li Y, Tang J, Fei X, Gao Y. Diagnostic Performance of Contrast Enhanced Ultrasound in Patients With Prostate Cancer: A Meta-Analysis. *Acad Radiol* (2013) 20(2):156–64. doi: 10.1016/j.acra.2012.09.018
88. Sedelaar JPM, Van Leenders GJLH, Hulsbergen-Van dKCA, Van dPHG, van der Laak JA, Jeroen AWM, et al. Microvessel Density: Correlation Between Contrast Ultrasonography and Histology of Prostate Cancer. *Eur Urol* (2001) 40(3):285–93. doi: 10.1159/000049788
89. Sedelaar JPM, Leenders GJLHV, Goossen TEB, Kaa HVD, de la Rosette JJMCH. Value of Contrast Ultrasonography in the Detection of Significant Prostate Cancer: Correlation With Radical Prostatectomy Specimens. *Prostate* (2010) 53 (3):246–53. doi: 10.1002/pros.10145
90. Mitterberger MJ, Aigner F, Horninger W, Ulmer H, Cavuto S, Halpern EJ, et al. Comparative Efficiency of Contrast-Enhanced Colour Doppler Ultrasound Targeted Versus Systematic Biopsy for Prostate Cancer Detection. *Eur Radiol* (2010) 20(12):2791–6. doi: 10.1007/s00330-010-1860-1
91. Strazdina A, Krumina G, Sperga M. The Value and Limitations of Contrast-Enhanced Ultrasound in Detection of Prostate Cancer. *Anticancer Res* (2011) 31(4):1421–6. doi: 10.1023/B:JNMR.0000032550.18424.aa
92. Halpern EJ, Gomella LG, Forsberg F, Mccue PA, Trabulsi EJ. Contrast Enhanced Transrectal Ultrasound for the Detection of Prostate Cancer: A Randomized, Double-Blind Trial of Dutasteride Pretreatment. *J Urol* (2012) 188(5):1739–45. doi: 10.1016/j.juro.2012.07.021
93. Hong WZ, Luo JH, Hui XX, Wang DH, Ying RL, Chen MN, et al. The Value of Contrast-Enhanced Transrectal Ultrasound in Predicting the Nature of Prostate Diseases and the Gleason Score of Prostate Cancer by a Subjective Blood Flow Grading Scale. *Urol Int* (2011) 87(2):165–70. doi: 10.1159/000327988
94. Mike F. Contrast-Enhanced Ultrasound may Aid Prostate Cancer Detection. *J Natl Cancer Inst* (2013) 105(7):444–6. doi: 10.1093/jnci/djt079
95. Numao N, Kawakami S, Sakura M, Yoshida S, Kihara K. Characteristics and Clinical Significance of Prostate Cancers Missed by Initial Transrectal 12-Core Biopsy. *BJU Int* (2012) 109(5):665–71. doi: 10.1111/j.1464-410X.2011.10427.x
96. Yang JC, Tang J, Li J, Luo Y, Li Y, Shi H. Contrast-Enhanced Gray-Scale Transrectal Ultrasound-Guided Prostate Biopsy in Men With Elevated Serum Prostate-Specific Antigen Levels. *Acad Radiol* (2008) 15(10):1291–7. doi: 10.1016/j.acra.2008.03.022
97. Halpern EJ, Mccue PA, Aksnes AK, Hagen EK, Frauscher F, Gomella LG. Contrast-Enhanced US of the Prostate With Sonazoid: Comparison With Whole-Mount Prostatectomy Specimens in 12 Patients. *Radiology* (2002) 222(2):361–6. doi: 10.1148/radiol.2222010582
98. Matsumoto K, Nakagawa K, Hashiguchi A, Kono H, Kikuchi E, Nagata H, et al. Contrast-Enhanced Ultrasonography of the Prostate With Sonazoid. *Jpn J Clin Oncol* (2010) 40(11):1099–104. doi: 10.1093/jjco/hyq102
99. Jung EM, Wiggermann P, Greis C, Ehrich J, Ganzer R. First Results of Endocavity Evaluation of the Microvascularization of Malignant Prostate Tumors Using Contrast Enhanced Ultrasound (CEUS) Including Perfusion Analysis: First Results. *Clin Hemorheol Microcirculation* (2012) 52(2-4):167–77. doi: 10.3233/CH-2012-1594
100. Drudi FM, Cantisani V, Angelini F, Ciccariello M, Scialpi M. Multiparametric MRI Versus Multiparametric US in the Detection of Prostate Cancer. *Anticancer Res* (2019) 39(6):3101–10. doi: 10.21873/anticanres.13446
101. Schalk SG, Huang J, Li J, Libertario D, Hessel W, Pintong H, et al. 3-D Quantitative Dynamic Contrast Ultrasound for Prostate Cancer Localization. *Ultrasound Med Biol* (2018) 44(4):807–14. doi: 10.1016/j.ultrasmedbio.2017.12.005
102. Zhu Y, Sun Y, Liu W, Guan W, Chen Y. Magnetic Polymeric Nanobubbles With Optimized Core Size for MRI/ultrasound Bimodal Molecular Imaging of Prostate Cancer. *Nanomedicine* (2020) 15(1):2901–16. doi: 10.2217/nnm-2020-0188
103. Sanna V, Pintus G, Bandiera P, Anedda R, Punzoni S, Sanna B, et al. Development of Polymeric Microbubbles Targeted to Prostate-Specific Membrane Antigen as Prototype of Novel Ultrasound Contrast Agents. *Mol Pharm* (2011) 8(3):748–57. doi: 10.1021/mp100360g

104. Ding Y, Cao Q, Qian S, Chen X, Xu Y, Chen J, et al. Optimized Anti–Prostate-Specific Membrane Antigen Single-Chain Variable Fragment–Loaded Nanobubbles as a Novel Targeted Ultrasound Contrast Agent for the Diagnosis of Prostate Cancer. *Wiley-Blackwell Online Open* (2019) 39 (4):1–13. doi: 10.1002/jum.15155
105. Wang L, Li L, Guo Y, Tong H, Fan X, Ding J, et al. Construction and *In Vitro*/ *In Vivo* Targeting of PSMA-Targeted Nanoscale Microbubbles in Prostate Cancer. *Prostate* (2013) 73(11):1147–58. doi: 10.1002/pros.22663
106. Fan X, Guo Y, Wang L, Xiong X, Zhu L, Fang K. Diagnosis of Prostate Cancer Using Anti-PSMA Aptamer A10-3.2-Oriented Lipid Nanobubbles. *Int J Nanomed* (2016) 11:3939–50. doi: 10.2147/IJN.S112951
107. Brock M, Penberg B L, Roghmann F, Pelzer A, Dickmann M, Becker W, et al. Impact of Real-Time Elastography on Magnetic Resonance Imaging/ Ultrasound Fusion Guided Biopsy in Patients With Prior Negative Prostate Biopsies. *J Urol* (2015) 193(4):1191–7. doi: 10.1016/j.juro.2014. 10.106
108. Baco E, Ukimura O, Rud E, Vlatkovic L, Svindland A, Aron M, et al. Magnetic Resonance Imaging-Transectal Ultrasound Image-Fusion Biopsies Accurately Characterize the Index Tumor: Correlation With Step-Sectioned Radical Prostatectomy Specimens in 135 Patients. *Eur Urol* (2015) 67 (4):787–94. doi: 10.1016/j.eururo.2014.08.077
109. Gershman B, Karnes RJ. Comparison of MR/Ultrasound Fusion-Guided Biopsy With Ultrasound-Guided Biopsy for the Diagnosis of Prostate Cancer. *JAMA* (2015) 68(3):536–7. doi: 10.1016/j.eururo.2015.05.026
110. Tewes S, Hueper K, Hartung D, Imkamp F, Herrmann TRW, Weidemann J, et al. Targeted MRI/TRUS Fusion-Guided Biopsy in Men With Previous Prostate Biopsies Using a Novel Registration Software and Multiparametric MRI PI-RADS Scores: First Results. *World J Urol* (2015) 33(11):1707–14. doi: 10.1007/s00345-015-1525-4
111. Shao WX, Hong LL, Jing D, Jian GX, Yi FG, Mei X, et al. Contrast-Enhanced Ultrasonography With Contrast-Tuned Imaging Technology for the Detection of Prostate Cancer: Comparison With Conventional Ultrasonography. *BJU Int* (2012) 109(11):1620–6. doi: 10.1111/j.1464-410x.2011.10577.x
112. Aigner F, Schäfer G, Steiner E, Jaschke W, Horninger W, Herrmann TRW, et al. Value of Enhanced Transrectal Ultrasound Targeted Biopsy for Prostate Cancer Diagnosis: A Retrospective Data Analysis. *World J Urol* (2012) 30(3):341–6. doi: 10.1007/s00345-011-0809-6
113. Zhang M, Tang J, Luo Y, Wang Y, Wu M, Memmott B, et al. Diagnostic Performance of Multiparametric Transrectal Ultrasound in Localized Prostate Cancer: A Comparative Study With Magnetic Resonance Imaging. *J Ultrasound Med* (2019) 38(7):1823–30. doi: 10.1002/jum.14878
114. Fulgham PF. Multiparametric Ultrasound-Targeted Biopsy Compares Favorably To Multiparametric MRI Transrectal Ultrasound Fusion-Targeted Biopsy on Initial Biopsy of Men at Risk for Prostate Cancer. *World J Urol* (2018) 36(5):713–8. doi: 10.1007/s00345-018-2187-9

**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.

## GLOSSARY

| | |
|---|---|
| PCa | Prostate cancer |
| DRE | Digital rectal examination |
| PSA | Prostate specific antigen |
| TRUS | Transrectal ultrasound |
| SE | Sensitivity |
| SP | Specificity |
| MRI | Magnetic resonance imaging |
| US | Ultrasound |
| mp-MRI | Multiparameter MRI |
| T2WI | T2- weighted imaging |
| DWI | Diffusion weighted imaging |
| DCE-MRI | Dynamic contrast-enhanced MRI |
| mp-US | multiparametric ultrasound |
| EAU | European Urology Association |
| 3D-TUS | Transrectal three-dimensional ultrasound |
| DOR | Diagnostic odds ratio |
| AUC | Area under curve |
| csPCa | Clinically significant prostate cancer |
| ANNA/C-TRUS | Artificial neural network analysis/computerized-transrectal ultrasound |
| RP | Radical prostatectomy |
| NPV | Negative predictive value |
| PPV | Positive predictive value |
| HS | Histoscanning |
| PHS-TT | Prostate tissue scanning targeting |
| TTB | Transperineal template prostate biopsy |
| GSU | Grayscale ultrasound |
| CDU | Color Doppler ultrasound |
| PDU | Power Doppler ultrasound |
| MVD | Microvessel density |
| UE | Ultrasound elastography |
| TRTE | Transrectal real-time tissue elastography |
| SWE | Shear-wave elastography |
| ROI | Region of interest |
| RBIE | Real-time balloon inflatable elastography |
| SD | Strain index |
| ARFI | Acoustic radiation force pulse imaging |
| CEUS | Contrast-enhanced ultrasound |
| UCA | Ultrasound contrast agent |
| MBs | Microbubbles |
| TIC | Time intensity curve |
| 3DCE-PDU | Three-dimensional contrast-enhanced Doppler ultrasound |
| CUDI | Contrast-enhanced ultrasound diffusion imaging |
| EIR | Early irrigation rate |
| MTT | Mean transit time |
| RT | Rise time |
| WIT | Wash-in time |
| PSMA | Prostate-specific membrane antigen |
| NBs | Nanobubbles |
| ICG | Indocyanine green |
| scFv | Single-chain variable fragment |
| ESUR | European Society of Urogenital Radiology |

frontiers | Frontiers in Oncology

REVIEW
published: 04 July 2022
doi: 10.3389/fonc.2022.879697

# Progresses in Fluorescence Imaging Guidance for Bone and Soft Tissue Sarcoma Surgery

*Zhenyi Chen, Huayi Huang, Siyuan He, Yi Wang, Lin Cai and Yuanlong Xie**

*Department of Orthopedics, Zhongnan Hospital, Wuhan University, Wuhan, China*

R0 surgical resection is the preferred treatment for bone and soft tissue sarcoma. However, there is still a lack of precise technology that can visualize bone and soft tissue sarcoma during surgery to assist the surgeon in judging the tumor surgical boundary. Fluorescence imaging technology has been used in the diagnosis of cancer. It is a simple and essentially safe technique that takes no additional time during the operation. Intraoperative fluorescence imaging has potential application prospects in assisting the surgeons in judging the tumor boundary and improving the accuracy of surgical resection. This review mainly starts with clinical studies, animal experimentation, and newly designed probes of intraoperative fluorescence imaging of bone and soft tissue sarcoma, to appraise the application prospects of fluorescence imaging technology in bone and soft tissue sarcoma.

**Keywords: fluorescence imaging, bone sarcoma, soft tissue sarcoma, fluorescent probes, cancer imaging**

**Edited by:**
*Haibin Shi,*
*Soochow University, China*

**Reviewed by:**
*Yong Shi,*
*Karolinska Institutet, Sweden*
*Yao Sun,*
*Central China Normal University, China*

***Correspondence:**
*Yuanlong Xie*
*yuanlongxie@whu.edu.cn*

**Specialty section:**
*This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology*

**Received:** *20 February 2022*
**Accepted:** *14 June 2022*
**Published:** *04 July 2022*

**Citation:**
*Chen Z, Huang H, He S, Wang Y, Cai L and Xie Y (2022) Progresses in Fluorescence Imaging Guidance for Bone and Soft Tissue Sarcoma Surgery. Front. Oncol. 12:879697. doi: 10.3389/fonc.2022.879697*

## INTRODUCTION

Traditionally, surgeons mainly use preoperative CT(computed tomography) and MRI(magnetic resonance imaging) to assess the tumor boundary within the surgeons' naked eyes to select the scope of resection during the surgery (1). CT and MRI cannot be used in real-time and have limited tumor specificity (2). Assessments of the resection boundary based on the surgeons' naked eyes are inaccurate and rely on surgeons' experience. The intraoperative frozen section reduces surgical efficiency because of the unavoidable extension of the surgery period (3). It is urgent to find an auxiliary examination during the operation to judge the boundary between the tumor tissue and the normal tissue, which can improve the accuracy of the operation together with assistance in finding the tumor satellite foci (4).

The essence of intraoperative tumor fluorescence imaging is to allow fluorescent dyes to accumulate in tumor tissue during the operation so that the surgeons can find the boundary of the tumor. Fluorescent probes have different principles, such as EPR effects (enhanced permeability and retention) and antigen-antibody reactions. To date, fluorescence imaging has exhibited

---

**Abbreviations:** CT, computed tomography; MRI, magnetic resonance imaging; SPECT, single photon emission computed tomography; TBR, tumor-to-background ratio; ICG, indocyanine green; NIR, near-infrared; IHC, immunohistochemistry; VEGF, vascular endothelial growth factor; EGFR, epidermal growth factor receptor; VEGFR, vascular endothelial growth factor receptor; 5-ALA, 5-aminolevulinic acid; FR-α, folate receptor-α; insulin-like growth factor, IGF; EPR, enhanced permeability and retention; CME, clathrin-mediated endocytosis; glutathione, GSH.

promising advantages in various tumors, such as brain tumors (5), breast cancer (6), and gastric cancer (7).

In recent years, intraoperative fluorescence imaging of bone and soft tissue sarcoma has been explored in clinical studies, animal experimentation, and these studies have led to the development of new probes.

# PRINCIPLE OF FLUORESCENCE IMAGING

## Non-Specific Fluorescent Probes

Based on the principle of fluorescent probes, we divide the current fluorescent probes into four types and summarize them in **Figure 1**.

Most non-specific probes rely on a passive targeting strategy, which preferentially accumulates molecules in tumors. The passive targeting strategy attributes to the tumor microenvironment such as accumulation of acidity, anoxic habitat, and necrotic tissue (8). It is not specific so that burns, wounded, and other tissues can also retain more fluorescent dye than surrounding tissues.

The principle fluorescent component of tetracycline is achelate formed upon combination with calcium ions ontrabecular bone (9). OWEN et al. (10) studied the fluorescence of tetracycline medicines in bone cancers and normal bone in 1961. Normal bone tissue has strong fluorescence, while necrotic bone has no or low fluorescence. Tetracycline is nowadays used only infrequently for intraoperative imaging of bone and soft tissue sarcoma, not only because tetracycline drugs are toxic and have a high rate of adverse reactions, but also because the wavelength of tetracycline excitation light is 450-490nm, which is in the visible light range. This wavelength overlaps with normal tissues and is heavily absorbed in tissues such as hemoglobin and myoglobin (11).

Compared with fluorescent probes for fluorescence imaging in the visible region, the research direction in recent years has focused more on fluorescent probes with excitation wavelength in the near-infrared region NIR-1 (700-900nm). Near-infrared fluorescence with a wavelength of 700-900nm is rarely absorbed in tissues (12).

For example, the most commonly used and clinically approved fluorescent probe is indocyanine green (ICG). ICG has absorption and fluorescence spectra in the near-infrared (NIR) region. The excitation wavelength is 780nm, and it emits fluorescence in the range of 700-850nm. The red light is visible to the naked eye, but most of the light is not (13). Most researchers believe that the accumulation of ICG within solid tumors attributes to the EPR (enhanced permeability and retention) effect (14). Due to the presence of defective endothelial cells and wide fenestrations (600 to 800 nm) in nascent blood vessels, small molecules such as ICG are injected systemically and passively accumulate in tumors (15). However, Pandit et al. (16) pointed out that in addition to the EPR effect, transcytosis is the principle of molecular accumulation in tumors. It is the same as the research on ICG in Colorectal Cancer. Cancer cells have a high endocytic rate (17). ICG was preferentially taken up by cancer cells *via* clathrin-mediated endocytosis (CME) (18). Indocyanine green is a safe, basically non-toxic drug, which rarely reacts with other drugs (19). However, Indocyanine green accumulates in bone tumors, inflammation, and bone deformities.

Many factors can influence the EPR effect, including tumor type, size, and vascular mediators. As a result, the intensity of the ICG signal is unpredictable (20). If the patient has a fracture or

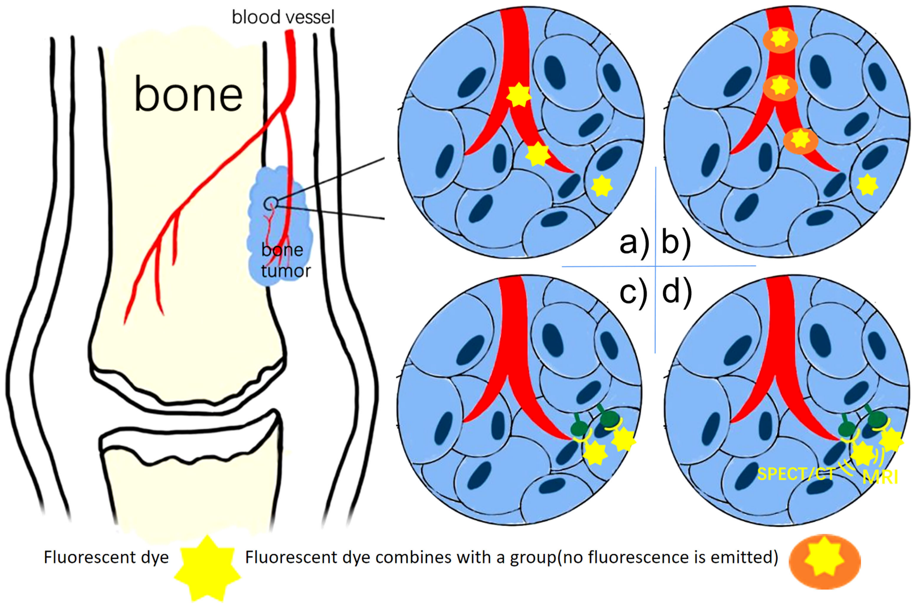


**FIGURE 1** | **(A)** Non-Specific Fluorescent Probes (ICG): the fluorescent dye remains in areas where the vasculature is highly disordered. **(B)** Activatable Fluorescence Probes: the dye fluorescence only when the group breaks down under tumor circumstances. **(C)** Specific fluorescent Probes: the fluorescent dye combines antibodies bind to biomarkers. **(D)** Multimodality Fluorescence Imaging Probes: the fluorescent dye used for both NIR imaging and SPECT/CT, MRI.

ischemia at the surgical site during surgery, it will cause false-positive results and affect the judgment. According to research, encapsulation of ICG improves its targeting abilities and circulation time (21, 22).

## Activatable Fluorescence Probes

Some scientists have also designed activatable fluorescent probes that emit fluorescence only in tumor tissues. This probe contains a chemical group, which can be broken down *via* some enzymes in the tumor and microenvironment, therefore this probe is activated.

Many activatable fluorescent probes have applications in other types of tumors, and could theoretically be used for intraoperative imaging of bone and soft tissue tumors. The activatable fluorescent probes for the tumor microenvironment are mainly activated by extracellular enzymes specifically emitted in the tumor microenvironment. In addition to cathepsin-activated fluorescent probes that have been used in soft tissue sarcoma animal experimentation, there are also fluorescent probes activated by matrix metalloproteinases (23). The activatable fluorescent probes for tumor cells consist of two parts. One is the activation of intracellular enzymes, such as β-galactosidase bioactivation (24) and glutathione (GSH) bioactivation (25). And the other one is the activation of fluorescence by the tumor cell hypoxia environment (26). Besides, the pH of the tumor microenvironment is generally between 6.7-7.1, the pH of tumor cells is between 5.9-6.2, and the pH of advanced tumor cells can even reach 5.0-5.5, which is an acidic environment compared with normal tissues. Some probes are sensitive to pH, and their fluorescence is activated in an acidic environment allowing fluorescence imaging of tumors and tumor microenvironments (27).

Activatable fluorescence probes reduce the fluorescence intensity of normal tissues and further increase the tumor-to-background ratio (TBR). But at the same time, the chemical synthesis of probes is complicated, and there is still a lack of further research on the adverse reactions of these probes.

## Specific Fluorescent Probes

Unlike ICG with the EPR effect, specific probes do not rely on the tumor microenvironment but instead rely on a targeting moiety conjugated to a contrast agent with a high binding affinity. These probes have higher targeting properties than indocyanine green (28).

The original design method is to combine monoclonal antibodies with fluorescent dyes to create fluorescent probes. Previously, monoclonal antibodies were used as targeted drugs to treat tumors. For example, Bevacizumab is a monoclonal antibody that binds to vascular endothelial growth factor-A (VEGF-A)which is highly expressed in tumor cells (29) and plays a direct role in vascular endothelial production (30). Combine bevacizumab with the fluorescent dye IRDye800CW to synthesize a fluorescent probe that can specifically bind to tumors. Scientists designed Panitumumab-800CW (31) and Cetuximab-800CW (32) based on the principle of similars. Panitumumab is a monoclonal IgG2 antibody that binds to the Epidermal Growth Factor Receptor (EGFR) with high specificity (33). EGFR is highly expressed in bone and soft tissue sarcoma and is involved in osteolytic metastases of bone tumors. Cetuximab is also an anti-EGFR monoclonal antibody.

In recent years, with the development of chemical synthesis technology, moieties for active targeting have become available, such as nanoparticle scaffolds, peptides, ligands, and aptamers. Compared with antibodies, the moieties have similar binding characteristics but show better tumor penetration and more rapid clearance from non-targeted tissues (34). For example, ABY-029 is an EGFR-targeted affibody molecule labeled with IRDye 800CW (35). While performing intraoperative tumor fluorescence imaging, ABY-029 can be injected on the same day. Besides, compared with bevacizumab, panitumumab, and cetuximab, ABY-029 retains high EGFR specificity (36) with low immunogenicity and low toxicity (37).

Specific fluorescent probes are based on active targeting, their synthesis is complicated. Tumors are heterogeneous, so we can't find a tumor marker expressed in each tumor tissue. The majority of specific probes are still in the pre-clinical stage. It requires more feasibility and toxicity studies, particularly for small molecule probes before clinical trials.

## Multimodality Fluorescence Imaging Probes

SPECT/CT, MRI, and NIR combined multimodal imaging technology have gained significant popularity. Scientists have designed fluorescent probes with SPECT/CT, MRI sensitive groups, and fluorescent dyes (38) (**Figure 2**). The contrast of preoperative SPECT/CT, MRI tumor imaging is improved by preoperative injection of multimodality fluorescent probes. The fluorescent sign of the tumor can also be collected during the operation. This combination of imaging and fluorescence imaging can significantly increase the detection rate of tumors and obtain more accurate tumor boundaries. This probe is used for preoperative tumor imaging, surgical planning, and intraoperative tumor fluorescence imaging.

Schematically, this kind of probe has much potential. For example, if we expand our scope to treatment, scientists have designed probes that combine Photodynamic therapy with fluorescence imaging (39, 40).

All imaging techniques have their limitations, e.g., MRI has problems with relatively low sensibility, and optical imaging has issues with low spatial resolution and small penetration depth (41). Multiple imaging techniques aid in early diagnosis and treatment planning. However, it is worth exploring whether it is necessary to enhance preoperative MRI and SPECT/CT tumor signal intensity in clinical practice (42).

In recent years, some researchers have focused on fluorescent probes in the NIR-2 range (1000-1700nm) and have produced several fluorescent probes for tumor imaging in the NIR-2 range (43). According to some researchers, fluorescence with a wavelength of 1000-1700nm, can reduce scattering when passing through the skin and is less affected by normal tissue autofluorescence. Compared with NIR-1 imaging, it can penetrate deeper tissues (44).

# PRE-CLINICAL RESEARCH

## Non-Specific Fluorescent Probes

Presently the widely used non-specific fluorescent probe in tumor surgery is indocyanine green. Overall, these cases in the

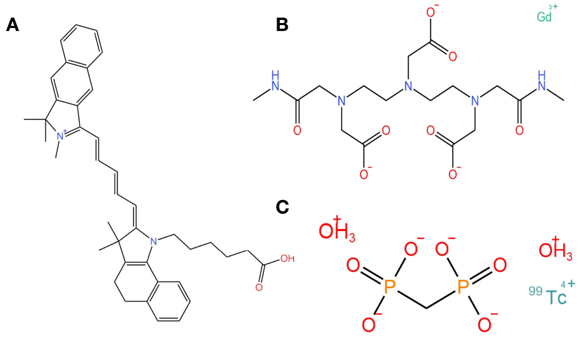


**FIGURE 2** | The chemical structure of fluorescence, magnetic, and SPECT nanoparticles that can compose multimodality probes. **(A)** A fluorescent dye: Cy5.5 carboxylic acid. **(B)** A magnetic nanoparticle as molecular imaging agent: gadodiamide. **(C)** A SPECT-CT tumor imaging agent: technetium Methylenediphosphonate (99mTc-MDT).

past five years support that bone and tissue sarcoma can fluoresce after injection, and the tumor boundary is consistent with the pathological section control.

For example, Fourman (2018) (45) injected osteosarcoma cells into the left hind limb of BALB/c Mice. Researchers used pathological sections to confirm that the fluorescent part of the hind limb was a bone sarcoma. Mice with fluorescent lung tissue developed lung metastases from osteosarcoma. Interestingly, the researchers discovered that the higher the fluorescence intensity of the primary bone tumor, the greater the possibility of lung metastases. This finding suggests that we can early predict the probability of lung metastases in osteosarcoma patients through intraoperative bone tumor fluorescence intensity.

Another example of what is meant by Mahjoub (46), is that they injected ICG into 11 osteosarcoma mice 12 hours before surgery for fluorescence-guided osteosarcoma surgery. The recurrence rate of mice with osteosarcoma resection guided by ICG was much lower than that of mice with conventional resection.

In addition to indocyanine green, non-specific fluorescent probes such as Alizarin Red and Tetracycline had been expected to have great potential for intraoperative imaging of bone and soft tissue sarcoma. However, the fluorescence excitation wavelengths are 465nm and 490nm, which are both in the visible light range, and the fluorescence area overlaps with normal tissues. There have been few related studies in recent years.

## Activatable Fluorescence Probes

According to the different characteristics of tumor cells and tumor microenvironments from normal tissues, scientists have designed activatable optical probes. In 2016, Bartholf Dewitt S (47) used the cathepsin-activated fluorescent probe LUM015 in dogs with soft tissue sarcoma for intraoperative fluorescence imaging. According to previous studies, cathepsin is overexpressed in soft tissue sarcoma and other tumors while rarely expressed in normal tissues. All the dogs' soft tissue sarcoma fluorescence when imaging. The pathologist took 33 parts of the excised tissue for biopsy, all of which were tumor tissues. The cathepsin-activated fluorescent probe is further exemplified in studies by Prince et al. (48). The researchers compare the TBR and effect of prosense750EX (another cathepsin-activated fluorescent probe) with multiple fluorescent probes for fluorescence imaging of soft tissue sarcoma. Prosense750EX, like the other probes in the study, can identify tumor beds with a diameter of less than 1mm intraoperatively. Therefore, the Prosense750EX can be used as a probe for fluorescence imaging.

## Specific Fluorescent Probes

Some specific probes have been designed, and animal experimentation has proved their specificity and sensitivity. Most specific probes are created by combining fluorescent dyes with antibodies or ligands that precisely bind to tumor cells (**Table 1**).

Based on this probe design idea, our team designed a specific fluorescent probe CS2-N-E9R for Ewing's sarcoma-specific fusion protein EWS-FLI1 (E/F) in 2021 (49). Our specific probe can make Ewing's sarcoma fluorescence imaging in cell experimentation and animal experimentation. Besides, it does not show fluorescence for E/F-negative osteosarcoma cells.

For example, Li et al. (50) combined the non-antibody binding protein of CD105 with fluorescein isothiocyanate (FITC) to obtain a fluorescent probe targeting osteosarcoma. This fluorescent probe causes the osteosarcoma cells, dissected osteosarcoma tissues, and osteosarcoma in mice to emit fluorescence, proving that it can label osteosarcoma.

Another example of what is designed by Zhou (51) is CH1055-PEG-PT and CH1055-PEG-Affibody. These probes combined Small molecule protein binding to 143b osteosarcoma cells with Fluorescent dyes in the NIR-2 region. Both of these probes can image fluorescence in osteosarcoma. The surgeons used fluorescence guidance for tumor resection. Pathology specialists sectioned the tumor and adjacent tissues and stained them for microscopic examination after the surgery. The results revealed that the fluorescence intensity of the tumor was higher than that of adjacent tissues. The researcher suggests that, compared with CT, the fluorescent probe can image tumors smaller than 1 cm in diameter and has a clear fluorescence boundary.

This technology is further exemplified in animal experimentation using indocyanine green and ABY-029 combined fluorescence imaging in soft tissue sarcoma surgery.

**TABLE 1** | Specific fluorescent probes for bone and soft tissue tumor imaging.

| Target | Probe | Type of tumor | Name | Observation from postinjection | Year | author |
|---|---|---|---|---|---|---|
| CD105 | Non-antibody-binding proteins | Osteosarcoma | A novel peptide targets CD105 | 1h | 2018 | Xiaolong Li (50) |
| EGFR | Affibody molecule | Synovial sarcoma | ABY-029 | 4h | 2021 | Hira Shahzad Sardar (52) |
|  | Antibody | Fibrosarcoma | Cetuximab -IRDye800CW | 9d | 2018 | Andrew C. Prince BSc (48) |
| EWS-FLI1 | Peptide | Ewing sarcoma | CS2-N-E9R | 6h | 2021 | Yu Wang (49) |
| integrin αvβ3 | Small molecule | Fibrosarcoma | IntegriSense750 | 9d | 2018 | Andrew C. Prince BSc (48) |
|  | Small molecule | Osteosarcoma | 68Ga-CHS2 | 4h | 2018 | Yao Sun (73) |
| osteocalcin property receptor | Peptide | the lung metastases of osteosarcoma | CH1055-PEG -Affibody | 12-36h | 2020 | Hui Zhou (51) |
|  | Peptide | Osteosarcoma | CH1055-PEG-PT | 12-36h | 2020 | Hui Zhou (51) |
| VEGFR2 | Antibody | Fibrosarcoma | DC101 -IRDye800CW | 7d | 2018 | Andrew C. Prince BSc (48) |

Sardar et al. (52) discovered that fluorescence imaging with ICG and ABY-029 is superior to ICG or ABY-029 alone. Among them, ABY-029 is more concentrated in the high-cell living tissue area, while ICG is more concentrated in the low-cell area. The article did not explore the reasons further. A possible explanation for these results may be related to the difference in imaging principles between the two probes. ICG accumulates more in new blood vessels, whereas ABY-029 binds to cancer cells specifically. It suggests that combining two fluorescent probes with different localization areas and fluorescence imaging principles could improve the specificity and sensitivity of fluorescence imaging in bone and soft tissue tumor surgery.

Xu's experimentation study explored the feasibility of specific fluorescent probes used for intraoperative imaging after radiotherapy and chemotherapy (36). Xu designed a mouse model of soft tissue sarcoma after chemotherapy and radiotherapy and injected ABY-029 intraoperative fluorescence imaging into the mice 4-8 hours before surgery. It might be possible to estimate whether most patients with soft tissue sarcoma undergoing preoperative radiotherapy and chemotherapy can use ABY-029 Intraoperative fluorescence imaging. The results confirmed the feasibility of fluorescence imaging of soft tissue sarcoma in mice after radiotherapy and chemotherapy. This outcome is contrary to that of Nicoli et al. (53) who found indocyanine green could not fluorescently label osteosarcoma after radiotherapy. This result demonstrates the superiority of specific fluorescent probes compared to fluorescence imaging in indocyanine green.

Another research compared several fluorescent probes on soft tissue sarcoma mice (48). The researchers compare the intraoperative tissue fluorescence range with HE stained sections, and immunohistochemistry(IHC) to quantitatively compare TBR. Compared to DC101(binding to VEGFR-2) TBR 3.7, IntegriSense750(A small-molecule probe binding to integrin αvβ3) TBR 7.0, and ProSense750EX (activated by locally expressed cathepsin)TBR 5.8, the TBR of cetuximab-IRDye800CW was 16.8, which was significantly higher than other fluorescent probes.

In addition to the fluorescent probes that have been assessed on bone and soft tissue sarcoma, many newly designed fluorescent probes may have the potential to be used in intraoperative imaging of bone and soft tissue sarcoma. Mahalingam et al. (54) designed the Centyrin-Based Near-Infrared Probe, a fluorescent probe that images EGFR-positive tumors. Reviews show that osteosarcoma and soft tissue sarcoma can overexpress EGFR (55). In the future, we can build mouse models and conduct further animal experimentation to explore whether this probe is used for bone and soft tissue tumor imaging.

For a ligand or antibody that specifically binds to bone and soft tissue sarcoma, the ideal is to find a target not expressed in other tissues and expressed in all bone and soft tissue sarcoma, especially tissue cells surrounding the tumor. There are many studies on tumor-specific markers of bone and soft tissue tumor cells. CxCR4 (Cys-X-Cys receptor 4), PDGFR-β(Platelet-derived growth factor receptor-β), TEM1 (Tumor Endothelial Marker 1), VEGFR-1, EGFR, VEGFR-2, IGF-1R, IGF-2R, CD40, et al. are high specific tumor markers (56–58). Scientists use these tumor-specific markers to create antibodies or ligands and combine antibodies and ligands with fluorescent dyes to make specific fluorescent probes. According to animal experiments in the past five years, ligands and small-molecule peptides spread faster than antibodies and are more likely to accumulate in tumor tissue. There are numerous fluorescent dyes on the market currently, most of them are classified as rhodamines, oxazines, fluoresceins, cyanines, and carbopyronines in structure (59). The commonly used near-infrared fluorescent dyes such as IRDye800CW still have high development prospects.

In the case of specific fluorescent probes, future research could focus on developing new probes specifically binding to bone and soft tissue sarcoma, determining whether existing fluorescent probes can be used for bone and soft tissue sarcoma, and evaluating the advantages, disadvantages, and effectiveness of the probes.

## Multimodality Fluorescence Imaging Probes

Probes for multimodal visualization in MRI, SPECT/CT, and Near-Infrared Optical Imaging have gotten attention in the past five years. These probes have the potential for preoperative tumor imaging, surgical planning, and intraoperative tumor fluorescence imaging.

It is exemplified in the animal experimentation undertaken by Xu with 99mTc-Gd@OVA-Cy nanoprobe (60). Researchers performed preoperative NIR fluorescence imaging, MRI, and SPECT/CT of osteosarcoma with nanoprobe. After 15 minutes of intravenous injection of the fluorescent probe, the images of all three modes showed enhanced signals of osteosarcoma. In MRI, SPECT/CT, and NIR imaging, researchers can observe a clear boundary of osteosarcoma, and the tumor boundary is consistent with the results of HE staining sections. Surprisingly, the researchers also found that the fluorescent probe can show lymph drainage and sentinel lymph nodes. Therefore Xu considered that this probe might be used for osteosarcoma to improve lymph node resection and preoperative planning.

Scientists designed many fluorescent probes for multimodal imaging in the past five years. But there are few animal experiments on whether these fluorescent probes can be applied to bone and soft tissue sarcoma. Lee et al. (61) designed an Nd3+-UCNPs nanoprobe specifically binding to CD44. The nanoprobe is injected into the hepatocellular carcinoma of patients, used for preoperative MRI detection and intraoperative NIR tumor imaging. Related literature shows that bone and soft tissue sarcoma can express CD44 (62). Therefore, this multimodal probe may be significant in intraoperative and preoperative tumor imaging for CD44-positive bone and soft tissue sarcoma.

Researchers also focus on probes for multimodal visualization in SPECT/CT and intraoperative near-infrared optical imaging. A notable example is the folate-ECG-ROX targeted folate receptor in the tumor (63). Another example designed by Manca is the ICG-99mTc probe, which facilitates visualization of lymph drainage and assesses the sentinel lymph node (64).

## Clinical Trials

Reports about intraoperative fluorescence imaging of bone and soft tissue sarcoma are limited (**Table 2**). In 2019, Samkoe et al. (65) reported a case of using ABY-029 intraoperative fluorescence imaging for soft tissue sarcoma. The intraoperative fluorescence intensity ratio of soft tissue sarcoma to normal tissue/background is 2.0/3.4, which is sufficient to distinguish tumor from normal tissue by fluorescence during operation. The tumor was stained with hematoxylin-eosin staining and IHC postoperatively, and the fluorescent tissue was confirmed to be soft tissue sarcoma, and the fluorescence signal was highly associated with the expression of EGFR.

In a similar case in the UK, 11 patients with bone and soft tissue sarcoma were admitted for ICG intraoperative fluorescence imaging (53). ICG was injected intravenously 16-24 hours before the operation, and the Stryker Spy Phi near-infrared device collected the fluorescence signal during the operation. Surgeons believe that in three of the 11 cases, they removed more tissue during the operation due to fluorescence. Nine of the 11 instances revealed tumor fluorescence during surgery. Two instances exhibited no fluorescence during surgery, one was grade 1 myxofibrosarcoma, and the other was osteosarcoma with more than 90% necrosis after chemotherapy. The failure could be because ICG fluorescence imaging is better suited to tumors with a higher degree of malignancy, no treatment, and fewer necrotic areas.

This technology is further demonstrated in studies using Bevacizumab-IRDye800cw fluorescence imaging in 15 patients with soft tissue sarcoma during surgery (66). Researchers found fluorescence in soft tissue sarcoma during and after the operation in all 15 cases and no adverse reactions. Furthermore, the researcher discovered that the necrotic area of soft tissue sarcoma treated by neoadjuvant chemotherapy had no fluorescence. Auspiciously, we noticed in clinical practice that the necrotic area is more inside the tumor and has few effects on the fluorescence of the tumor border.

Furthermore, bone and soft tissue tumor metastasis are frequent. Fluorescence imaging can detect tumor metastasis in bone and soft tissue. These clinical trials reveal the need for fluorescence imaging among metastases. Patients subjected toI CG injection were assessed after 24 hours (not overlapping with the optimal time for ICG to show bone and soft tissue sarcoma). Among 44 patients with soft tissue sarcoma lung metastases, 40 lung metastases showed fluorescence during Video-assisted thoracoscopic surgery(VATS). Among 40 cases of osteosarcoma lung metastases, 36 cases had fluorescence. The depth of all lung metastases without fluorescence imaging was more than 2 cm. According to Predina, fluorescence imaging during ICG surgery is better for detecting tumor metastasis with a depth smaller than 2 cm and a diameter greater than 5 mm (67).

Scheichel (68) performed a clinical trial using 5-aminolevulinic acid (5-ALA) intraoperative fluorescence imaging in fifty patients with bone and soft tissue infiltrating meningiomas. All bone fluorescence shows tumor invasion into bone tissue. Three patients showed additional fluorescence in the periosteum and temporal muscles, and histopathological examination confirmed tumor infiltration (68).

Predina and colleagues studied a patient with osteosarcoma lung metastases undergoing surgery and showed that fluorescence imaging with OTL38 enabled the detection of Lung metastases. According to previous studies, FR-$\alpha$ is overexpressed in 80% of primary osteosarcoma. The lung metastases had strong fluorescence after intravenous injection of 0.025mg/kg OTL38. However, the researchers did not specify whether fluorescence was observed in the primary osteosarcoma (69).

## Future Perspectives

At present, intraoperative fluorescence imaging does not use quantitative norms to determine whether it is tumor tissue. There is no standard for how high the fluorescence contrast should be to indicate a tumor in intraoperative fluorescence imaging technology. To determine the standard, it is important to conduct clinical trials including large sample size and compare with pathological results. Futhermore, a technique combining biophysics-inspired modeling and artificial intelligence (AI) was envisioned to monitor intraoperative changes in NIR intensities over time in different tissue and provide clinically significant lesion identification (70). In addition, mixed reality(MR) techniques that combine fluorescence imaging with CT have been used in liver resection (71). We can embed an augmented reality (AR)-based navigation system in the fluorescence imaging devices (72), and evaluate the usefulness of the system in the experimental study.

**TABLE 2 |** Clinical Trials in fluorescence imaging for bone and soft tissue sarcoma surgery.

| | Author | Year | Study type | Studypopulation | Fluorescentpopulation | Type of probe | Target | Type of tumor | Interval betweeninjection andsurgery |
|---|---|---|---|---|---|---|---|---|---|
| ICG | Nicoli, Fabio MD (53) | 2021 | Prospective study | 10 | 8 | Non-specific | . | Osteosarcoma(0/1)<br>Chondrosarcoma(1/1)<br>Myxofibrosarcoma(4/5)<br>Pleomorphic sarcoma(1/1)<br>Leiomyosarcoma(1/1)<br>Myxofibrosarcoma(1/1) | 16-24h |
| ICG | Jarrod D Predina (67) | 2019 | Phase 1 clinical trial | 30 | 30 | Non-specific | . | Lung metastases of bone and soft tissue sarcoma | 24h |
| 5-ALA | Florian Scheichel (68) | 2020 | Retrospective study | 11 | 11 | Non-specific | . | Bone-infiltrating meningiomas | 3h |
| OTL38 | Jarrod D. Predina (69) | 2018 | Case report | 1 | 1 | Specific | FR-α | Lung metastases of osteosarcoma | 4h |
| Bevacizumab-800CW | Pieter J Steinkamp (66) | 2021 | Phase 1 clinical trial | 15 | 15 | Specific | VEGF-A | Myxofibrosarcoma(7/7)<br>Liposarcoma(3/3) Synovial sarcoma(2/2)<br>Leiomyosarcoma(1/1)<br>Angiosarcoma(1/1)<br>Undifferentiated pleomorphic sarcoma(1/1) | 3d |
| ABY-029 | Kimberley S. Samkoe (65) | 2019 | Case report | 1 | 1 | Specific | EGFR affibody | Pleomorphic sarcoma(1/1) | 1-3h |

The manufacture and use of fluorescent probes for intraoperative fluorescence imaging of bone and soft tissue sarcoma have a potential future. Non-specific probes may additionally fluoresce in non-tumor areas, which can cause surgeons to misjudge. The main direction of new fluorescent probes will be specific fluorescent probes with high specificity to label tumors. With the further investigation of the mechanism of bone and soft tissue sarcoma, scientists will discover more specific tumor-expressed molecules. We can accordingly design specific fluorescent probes with high specificity and sensitivity.

Simultaneously, we noticed that tumors are heterogeneous, and it is difficult for a probe to image all tumors of the same type. Experiments are currently underway to combine two fluorescent probes with different principles to increase accuracy and lower the negative rate. In the future, we can design fluorescent probes with multiple responses to tumors and the microenvironment to further reduce the false-negative rate of fluorescence during tumor surgery.

## AUTHOR CONTRIBUTIONS

ZC: writing-original draft. HH: visualization, formal analysis. SH: investigation. YW: supervision. LC: conceptualization. YX: writing—review and editing. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by National Natural Science Foundation of China (NO. 82103285), the Improvement Project for Theranostic ability on Difficulty miscellaneous disease(Tumor) (ZLYNXM202005), the Research Fund from Medical Sci-Tech Innovation Platform of Zhongnan Hospital, Wuhan University (PTXM2021003).

## REFERENCES

1. Reijnders K, Coppes MH, van Hulzen ALJ, Gravendeel JP, van Ginkel RJ, Hoekstra HJ. Image Guided Surgery: New Technology for Surgery of Soft Tissue and Bone Sarcomas. *Eur J Surg Oncol (EJSO)* (2007) 33:390–8. doi: 10.1016/j.ejso.2006.10.030
2. Israel O, Pellet O, Biassoni L, De Palma D, Estrada-Lobato E, Gnanasegaran G, et al. Two Decades of SPECT/CT - the Coming of Age of a Technology: An Updated Review of Literature Evidence. *Eur J Nucl Med Mol Imaging* (2019) 46:1990–2012. doi: 10.1007/s00259-019-04404-6
3. Yu Y, Tan Y, Xie C, Hu Q, Ouyang J, Chen Y, et al. Development and Validation of a Preoperative Magnetic Resonance Imaging Radiomics-Based Signature to Predict Axillary Lymph Node Metastasis and Disease-Free Survival in Patients With Early-Stage Breast Cancer. *JAMA Netw Open* (2020) 3:e2028086. doi: 10.1001/jamanetworkopen.2020.28086
4. Hernot S, van Manen L, Debie P, Mieog J, Vahrmeijer AL. Latest Developments in Molecular Tracers for Fluorescence Image-Guided Cancer Surgery. *Lancet Oncol* (2019) 20:e354–67. doi: 10.1016/S1470-2045(19)30317-1
5. Zhang DY, Singhal S, Lee J. Optical Principles of Fluorescence-Guided Brain Tumor Surgery: A Practical Primer for the Neurosurgeon. *Neurosurgery* (2019) 85:312–24. doi: 10.1093/neuros/nyy315
6. Lee J, Kim B, Park B, Won Y, Kim SY, Lee S. Real-Time Cancer Diagnosis of Breast Cancer Using Fluorescence Lifetime Endoscopy Based on the Ph. *Sci Rep* (2021) 11:16864. doi: 10.1038/s41598-021-96531-0

7. Loshchenov M, Levkin V, Kalyagina N, Linkov K, Kharnas S, Efendiev K, et al. Laser-Induced Fluorescence Diagnosis of Stomach Tumor. *Lasers Med Sci* (2020) 35:1721–8. doi: 10.1007/s10103-020-02963-x
8. Ali I, Alfarouk KO, Reshkin SJ, Ibrahim ME. Doxycycline as Potential Anti-Cancer Agent. *Anticancer Agents Med Chem* (2017) 17:1617–23. doi: 10.2174/1871520617666170213111951
9. Frost HM, Villanueva AR, Roth H, Stanisavljevic S. Tetracycline Bone Labeling. *J New Drugs* (1961) 1:206–16. doi: 10.1177/009127006100100503
10. Owen LN. Fluorescence of Tetracyclines in Bone Tumours, Normal Bone and Teeth. *Nature* (1961) 190:500–2. doi: 10.1038/190500a0
11. Ristow O, Pautke C. Auto-Fluorescence of the Bone and its Use for Delineation of Bone Necrosis. *Int J Oral Max Surg* (2014) 43:1391–3. doi: 10.1016/j.ijom.2014.07.017
12. Nagaya T, Nakamura YA, Choyke PL, Kobayashi H. Fluorescence-Guided Surgery. *Front Oncol* (2017) 7:314. doi: 10.3389/fonc.2017.00314
13. Matsuura Y, Ichinose J, Nakao M, Okumura S, Mun M. Recent Fluorescence Imaging Technology Applications of Indocyanine Green in General Thoracic Surgery. *Surg Today* (2020) 50:1332–42. doi: 10.1007/s00595-019-01906-6
14. Maeda H, Nakamura H, Fang J. The EPR Effect for Macromolecular Drug Delivery to Solid Tumors: Improvement of Tumor Uptake, Lowering of Systemic Toxicity, and Distinct Tumor Imaging *In Vivo*. *Adv Drug Delivery Rev* (2013) 65:71–9. doi: 10.1016/j.addr.2012.10.002
15. Matsumura Y, Maeda H. A New Concept for Macromolecular Therapeutics in Cancer Chemotherapy: Mechanism of Tumoritropic Accumulation of Proteins and the Antitumor Agent Smancs. *Cancer Res* (1986) 46:6387–92.
16. Pandit S, Dutta D, Nie S. Active Transcytosis and New Opportunities for Cancer Nanomedicine. *Nat Mater* (2020) 19:478–80. doi: 10.1038/s41563-020-0672-1
17. Onda N, Mizutani-Morita R, Yamashita S, Nagahara R, Matsumoto S, Yoshida T, et al. Fluorescence Contrast-Enhanced Proliferative Lesion Imaging by Enema Administration of Indocyanine Green in a Rat Model of Colon Carcinogenesis. *Oncotarget* (2017) 8:90278–90. doi: 10.18632/oncotarget.21744
18. Onda N, Kimura M, Yoshida T, Shibutani M. Preferential Tumor Cellular Uptake and Retention of Indocyanine Green for *In Vivo* Tumor Imaging. *Int J Cancer* (2016) 139:673–82. doi: 10.1002/ijc.30102
19. Reinhart MB, Huntington CR, Blair LJ, Heniford BT, Augenstein VA. Indocyanine Green. *Surg Innov* (2016) 23:166–75. doi: 10.1177/1553350615604053
20. Rijs Z, Jeremiasse B, Shifai N, Gelderblom H, Sier C, Vahrmeijer AL, et al. Introducing Fluorescence-Guided Surgery for Pediatric Ewing, Osteo-, and Rhabdomyosarcomas: A Literature Review. *Biomedicines* (2021) 9. doi: 10.3390/biomedicines9101388
21. Hill TK, Abdulahad A, Kelkar SS, Marini FC, Long TE, Provenzale JM, et al. Indocyanine Green-Loaded Nanoparticles for Image-Guided Tumor Surgery. *Bioconjug Chem* (2015) 26:294–303. doi: 10.1021/bc5005679
22. Chen Q, Xu L, Liang C, Wang C, Peng R, Liu Z. Photothermal Therapy With Immune-Adjuvant Nanoparticles Together With Checkpoint Blockade for Effective Cancer Immunotherapy. *Nat Commun* (2016) 7:13193. doi: 10.1038/ncomms13193
23. Ren L, Wang Y, Zhu L, Shen L, Zhang J, Wang J, et al. Optimization of a MT1-MMP-Targeting Peptide and its Application in Near-Infrared Fluorescence Tumor Imaging. *Sci Rep-Uk* (2018) 8:10334. doi: 10.1038/s41598-018-28493-9
24. Huang J, Jiang Y, Li J, Huang J, Pu K. Molecular Chemiluminescent Probes With a Very Long Near-Infrared Emission Wavelength for *In Vivo* Imaging. *Angew Chemie Int Edition* (2021) 60:3999–4003. doi: 10.1002/anie.202013531
25. Guo R, Huang F, Zhang B, Yan Y, Che J, Jin Y, et al. GSH Activated Biotin-Tagged Near-Infrared Probe for Efficient Cancer Imaging. *Theranostics* (2019) 9:3515–25. doi: 10.7150/thno.32742
26. Zhang Z, Wang R, Huang X, Zhu W, He Y, Liu W, et al. A Simple Aggregation-Induced Emission Nanoprobe With Deep Tumor Penetration for Hypoxia Detection and Imaging-Guided Surgery*In Vivo*. *Anal Chem* (2021) 93:1627–35. doi: 10.1021/acs.analchem.0c04101
27. Voskuil FJ, Steinkamp PJ, Zhao T, van der Vegt B, Koller M, Doff JJ, et al. Exploiting Metabolic Acidosis in Solid Cancers Using a Tumor-Agnostic pH-Activatable Nanoprobe for Fluorescence-Guided Surgery. *Nat Commun* (2020) 11:3257. doi: 10.1038/s41467-020-16814-4
28. Srinivasarao M, Galliford CV, Low PS. Principles in the Design of Ligand-Targeted Cancer Therapeutics and Imaging Agents. *Nat Rev Drug Discovery* (2015) 14:203–19. doi: 10.1038/nrd4519
29. Zhao H, Wu Y, Chen Y, Liu H. Clinical Significance of Hypoxia-Inducible Factor 1 and VEGF-A in Osteosarcoma. *Int J Clin Oncol* (2015) 20:1233–43. doi: 10.1007/s10147-015-0848-x
30. Papadopoulos N, Martin J, Ruan Q, Rafique A, Rosconi MP, Shi E, et al. Binding and Neutralization of Vascular Endothelial Growth Factor (VEGF) and Related Ligands by VEGF Trap, Ranibizumab and Bevacizumab. *Angiogenesis* (2012) 15:171–85. doi: 10.1007/s10456-011-9249-6
31. Nishio N, van den Berg NS, van Keulen S, Martin BA, Fakurnejad S, Zhou Q, et al. Optimal Dosing Strategy for Fluorescence-Guided Surgery With Panitumumab-IRDye800CW in Head and Neck Cancer. *Mol Imaging Biol* (2020) 22:156–64. doi: 10.1007/s11307-019-01358-x
32. Gao RW, Teraphongphom N, de Boer E, Berg NSVD, Divi V, Kaplan MJ, et al. Safety of Panitumumab-IRDye800CW and Cetuximab-IRDye800CW for Fluorescence-Guided Surgical Navigation in Head and Neck Cancers. *Theranostics* (2018) 8:2488–95. doi: 10.7150/thno.24487
33. Napier TS, Udayakumar N, Jani AH, Hartman YE, Houson HA, Moore L, et al. Comparison of Panitumumab-IRDye800CW and 5-Aminolevulinic Acid to Provide Optical Contrast in a Model of Glioblastoma Multiforme. *Mol Cancer Ther* (2020) 19:1922–9. doi: 10.1158/1535-7163.MCT-19-0819
34. Olson MT, Ly QP, Mohs AM. Fluorescence Guidance in Surgical Oncology: Challenges, Opportunities, and Translation. *Mol Imaging Biol* (2019) 21:200–18. doi: 10.1007/s11307-018-1239-2
35. Samkoe KS, Bates BD, Elliott JT, LaRochelle E, Gunn JR, Marra K, et al. Application of Fluorescence-Guided Surgery to Subsurface Cancers Requiring Wide Local Excision. *Cancer Control* (2018) 25:544015489. doi: 10.1177/1073274817752332
36. Xu X, Samkoe KS, Henderson ER. Effect of Preoperative Cancer Treatment on Epidermal Growth Factor Receptor (EGFR) Receptor Expression Level in ABY-029 Guided Sarcoma Surgery. *Proc SPIE Int Soc Opt Eng* (2020) 11222. doi: 10.1117/12.2546963
37. Samkoe KS, Gunn JR, Marra K, Hull SM, Moodie KL, Feldwisch J, et al. Toxicity and Pharmacokinetic Profile for Single-Dose Injection of ABY-029: A Fluorescent Anti-EGFR Synthetic Affibody Molecule for Human Use. *Mol Imaging Biol* (2017) 19:512–21. doi: 10.1007/s11307-016-1033-y
38. Zhao M, Ding J, Mao Q, Zhang Y, Gao Y, Ye S, et al. A Novel Alphavbeta3 Integrin-Targeted NIR-II Nanoprobe for Multimodal Imaging-Guided Photothermal Therapy of Tumors *In Vivo*. *Nanoscale* (2020) 12:6953–8. doi: 10.1039/c9nr10720g
39. Xu Y, Tuo W, Yang L, Sun Y, Li C, Chen X, et al. Design of a Metallacycle-Based Supramolecular Photosensitizer for *In Vivo* Image-Guided Photodynamic Inactivation of Bacteria. *Angew Chem Int Ed Engl* (2022) 61: e202110048. doi: 10.1002/anie.202110048
40. Jiang Y, Pang X, Liu R, Xiao Q, Wang P, Leung AW, et al. Design of an Amphiphilic iRGD Peptide and Self-Assembling Nanovesicles for Improving Tumor Accumulation and Penetration and the Photodynamic Efficacy of the Photosensitizer. *ACS Appl Mater Interfaces* (2018) 10:31674–85. doi: 10.1021/acsami.8b11699
41. Luengo MY, Ovejero PK, Lozano CL, Marciello M, Filice M. Recent Advances in Multimodal Molecular Imaging of Cancer Mediated by Hybrid Magnetic Nanoparticles. *Polymers (Basel)* (2021) 13. doi: 10.3390/polym13172989
42. Liu M, Anderson RC, Lan X, Conti PS, Chen K. Recent Advances in the Development of Nanoparticles for Multimodality Imaging and Therapy of Cancer. *Med Res Rev* (2020) 40:909–30. doi: 10.1002/med.21642
43. Shou K, Qu C, Sun Y, Chen H, Chen S, Zhang L, et al. Multifunctional Biomedical Imaging in Physiological and Pathological Conditions Using a NIR-II Probe. *Adv Funct Mater* (2017) 27:1700995. doi: 10.1002/adfm.201700995
44. Zhang X, Wang H, Antaris AL, Li L, Diao S, Ma R, et al. Traumatic Brain Injury Imaging in the Second Near-Infrared Window With a Molecular Fluorophore. *Adv Mater* (2016) 28:6872–9. doi: 10.1002/adma.201600706
45. Fourman MS, Mahjoub A, Mandell JB, Yu S, Tebbets JC, Crasto JA, et al. Quantitative Primary Tumor Indocyanine Green Measurements Predict Osteosarcoma Metastatic Lung Burden in a Mouse Model. *Clin Orthop Relat Res* (2018) 476:479–87. doi: 10.1007/s11999.0000000000000003
46. Mahjoub A, Morales-Restrepo A, Fourman MS, Mandell JB, Feiqi L, Hankins ML, et al. Tumor Resection Guided by Intraoperative Indocyanine Green Dye

Fluorescence Angiography Results in Negative Surgical Margins and Decreased Local Recurrence in an Orthotopic Mouse Model of Osteosarcoma. *Ann Surg Oncol* (2019) 26:894–8. doi: 10.1245/s10434-018-07114-9

47. Bartholf DeWitt S, Eward WC, Eward CA, Lazarides AL, Whitley MJ, Ferrer JM, et al. A Novel Imaging System Distinguishes Neoplastic From Normal Tissue During Resection of Soft Tissue Sarcomas and Mast Cell Tumors in Dogs. *Vet Surg* (2016) 45:715–22. doi: 10.1111/vsu.12487
48. Prince AC, McGee AS, Siegel H, Rosenthal EL, Behnke NK, Warram JM. Evaluation of Fluorescence-Guided Surgery Agents in a Murine Model of Soft Tissue Fibrosarcoma. *J Surg Oncol* (2018) 117:1179–87. doi: 10.1002/jso.24950
49. Wang Y, Mai H, Yuan Y, Chen H, Wu S, Hu X, et al. EWS–FLI1-Targeting Peptide Identifies Ewing Sarcoma Tumor Boundaries and Lymph Node Metastasis *via* Near-Infrared Imaging. *Mol Oncol* (2021) 15:3706–20. doi: 10.1002/1878-0261.13081
50. Li X, Huang X, Zhang J, Huang H, Zhao L, Yu M, et al. A Novel Peptide Targets CD105 for Tumour Imaging *In Vivo*. *Oncol Rep* (2018) 40:2935–43. doi: 10.3892/or.2018.6643
51. Zhou H, Yi W, Li A, Wang B, Ding Q, Xue L, et al. Specific Small-Molecule NIR-II Fluorescence Imaging of Osteosarcoma and Lung Metastasis. *Adv Healthc Mater* (2019) 9:1901224. doi: 10.1002/adhm.201901224
52. Sardar HS, Zai Q, Xu X, Gunn JR, Pogue BW, Paulsen KD, et al. Dual-Agent Fluorescent Labeling of Soft-Tissue Sarcomas Improves the Contrast Based Upon Targeting Both Interstitial and Cellular Components of the Tumor Milieu. *J Surg Oncol* (2020) 122:1711–20. doi: 10.1002/jso.26190
53. Nicoli F, Saleh DB, Baljer B, Chan CD, Beckingsale T, Ghosh KM, et al. Intraoperative Near-Infrared Fluorescence (NIR) Imaging With Indocyanine Green (ICG) can Identify Bone and Soft Tissue Sarcomas Which may Provide Guidance for Oncological Resection. *Ann Surg* (2021) 273:e63–8. doi: 10.1097/SLA.0000000000003857
54. Mahalingam SM, Dudkin VY, Goldberg S, Klein D, Yi F, Singhal S, et al. Evaluation of a Centyrin-Based Near-Infrared Probe for Fluorescence-Guided Surgery of Epidermal Growth Factor Receptor Positive Tumors. *Bioconjugate Chem* (2017) 28:2865–73. doi: 10.1021/acs.bioconjchem.7b00566
55. Dobashi Y, Suzuki S, Sato E, Hamada Y, Yanagawa T, Ooi A. EGFR-Dependent and Independent Activation of Akt/mTOR Cascade in Bone and Soft Tissue Tumors. *Modern Pathol* (2009) 22:1328–40. doi: 10.1038/modpathol.2009.104
56. Bosma SE, van Driel PB, Hogendoorn PC, Dijkstra PS, Sier CF. Introducing Fluorescence Guided Surgery Into Orthopedic Oncology: A Systematic Review of Candidate Protein Targets for Ewing Sarcoma. *J Surg Oncol* (2018) 118:906–14. doi: 10.1002/jso.25224
57. Hassan SE, Bekarev M, Kim MY, Lin J, Piperdi S, Gorlick R, et al. Cell Surface Receptor Expression Patterns in Osteosarcoma. *Cancer-Am Cancer Soc* (2012) 118:740–9. doi: 10.1002/cncr.26339
58. Yang J, Zhang W. New Molecular Insights Into Osteosarcoma Targeted Therapy. *Curr Opin Oncol* (2013) 25:398–406. doi: 10.1097/CCO.0b013e3283622c1b
59. Dempsey GT, Vaughan JC, Chen KH, Bates M, Zhuang X. Evaluation of Fluorophores for Optimal Performance in Localization-Based Super-Resolution Imaging. *Nat Methods* (2011) 8:1027–36. doi: 10.1038/nmeth.1768
60. Xu Z, Wang Y, Han J, Xu Q, Ren J, Xu J, et al. Noninvasive Multimodal Imaging of Osteosarcoma and Lymph Nodes Using A99m Tc-Labeled Biomineralization Nanoprobe. *Anal Chem* (2018) 90:4529–34. doi: 10.1021/acs.analchem.7b04925
61. Lee J, Gordon AC, Kim H, Park W, Cho S, Lee B, et al. Targeted Multimodal Nano-Reporters for Pre-Procedural MRI and Intra-Operative Image-Guidance. *Biomaterials* (2016) 109:69–77. doi: 10.1016/j.biomaterials.2016.09.013
62. Kim C, Oh S, Kim S, Leem S, Heo J, Chung S. Correlation of IGF1R Expression With ABCG2 and CD44 Expressions in Human Osteosarcoma. *Genes Genom* (2018) 40:381–8. doi: 10.1007/s13258-017-0639-z
63. Kim MH, Kim S, Kim D. A Novel Dual-Modality Imaging Agent Targeting Folate Receptor of Tumor for Molecular Imaging and Fluorescence-Guided Surgery. *Ann Nucl Med* (2019) 33:606–16. doi: 10.1007/s12149-019-01369-2
64. Manca G, Garau LM, Mazzarri S, Mazzuca L, Muccioli S, Ghilli M, et al. Novel Experience in Hybrid Tracers. *Clin Nucl Med* (2021) 46:e181–7. doi: 10.1097/RLU.0000000000003478
65. Samkoe KS, Sardar HS, Gunn J, Feldwisch J, Linos K, Henderson E, et al. Measuring Microdose ABY-029 Fluorescence Signal in a Primary Human Soft-Tissue Sarcoma Resection. *Proc SPIE Int Soc Opt Eng* (2019) 10862. doi: 10.1117/12.2510935
66. Steinkamp PJ, Pranger BK, Li MF, Linssen MD, Voskuil FJ, Been LB, et al. Fluorescence-Guided Visualization of Soft-Tissue Sarcomas by Targeting Vascular Endothelial Growth Factor a: A Phase 1 Single-Center Clinical Trial. *J Nucl Med* (2021) 62:342–7. doi: 10.2967/jnumed.120.245696
67. Predina JD, Newton AD, Corbett C, Shin M, Sulfyok LF, Okusanya OT, et al. Near-Infrared Intraoperative Imaging for Minimally Invasive Pulmonary Metastasectomy for Sarcomas. *J Thorac Cardiovasc Surg* (2019) 157:2061–9. doi: 10.1016/j.jtcvs.2018.10.169
68. Scheichel F, Popadic B, Kitzwoegerer M, Ungersboeck K, Marhold F. Fluorescence-Guided Resection in Bone and Soft Tissue Infiltrating Meningiomas. *Acta Neurochir (Wien)* (2020) 162:605–11. doi: 10.1007/s00701-019-04179-7
69. Predina JD, Newton A, Deshpande C, Low P, Singhal S. Utilization of Targeted Near-Infrared Molecular Imaging to Improve Pulmonary Metastasectomy of Osteosarcomas. *J BioMed Opt* (2018) 23:1–4. doi: 10.1117/1.JBO.23.1.016005
70. Cahill RA. Ways of Seeing - It's All in the Image. *Colorectal Dis* (2018) 20:467–8. doi: 10.1111/codi.14265
71. Aoki T, Koizumi T, Sugimoto M, Murakami M. Holography-Guided Percutaneous Puncture Technique for Selective Near-Infrared Fluorescence-Guided Laparoscopic Liver Resection Using Mixed-Reality Wearable Spatial Computer. *Surg Oncol* (2020) 35:476–7. doi: 10.1016/j.suronc.2020.10.013
72. Cho HS, Park MS, Gupta S, Han I, Kim HS, Choi H, et al. Can Augmented Reality be Helpful in Pelvic Bone Cancer Surgery? An *In Vitro* Study. *Clin Orthop Relat Res* (2018) 476:1719–25. doi: 10.1007/s11999.0000000000000233
73. Sun Y, Zeng X, Xiao Y, Liu C, Zhu H, Zhou H, et al. Novel Dual-Function Near-Infrared II Fluorescence and PET Probe for Tumor Delineation and Image-Guided Surgery. *Chem Sci* (2018) 9:2092–7. doi: 10.1039/c7sc04774f

**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.

frontiers | Frontiers in Oncology

SYSTEMATIC REVIEW
published: 04 July 2022
doi: 10.3389/fonc.2022.929174

# Clinical and CT Features of Subsolid Pulmonary Nodules With Interval Growth: A Systematic Review and Meta-Analysis

*Xin Liang[1†], Mengwen Liu[2†], Meng Li[2] and Li Zhang[2*]*

*[1] Medical Statistics Office, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, [2] Department of Diagnostic Radiology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China*

**Background:** Establishing risk-based follow-up management strategies is crucial to the surveillance of subsolid pulmonary nodules (SSNs). However, the risk factors for SSN growth are not currently clear. This study aimed to perform a systematic review and meta-analysis to identify clinical and CT features correlated with SSN growth.

**Methods:** Relevant studies were retrieved from Web of Science, PubMed, Cochrane Library, and EMBASE. The correlations of clinical and CT features with SSN growth were pooled using a random-effects model or fixed-effects model depending on heterogeneity, which was examined by the $Q$ test and $I^2$ test. Pooled odds ratio (OR) or pooled standardized mean differences (SMD) based on univariate analyses were calculated to assess the correlation of clinical and CT features with SSN growth. Pooled ORs based on multivariate analyses were calculated to find out independent risk factors to SSN growth. Subgroup meta-analysis was performed based on nodule consistency (pure ground-glass nodule (pGGN) and part-solid nodule (PSN). Publication bias was examined using funnel plots.

**Results:** Nineteen original studies were included, consisting of 2444 patients and 3012 SSNs. The median/mean follow-up duration of these studies ranged from 24.2 months to 112 months. Significant correlations were observed between SSN growth and eighteen features. Male sex, history of lung cancer, nodule size $> 10$ mm, nodule consistency, and age $> 65$ years were identified as independent risk factors for SSN growth based on multivariate analyses results. Eight features, including male sex, smoking history, nodule size $> 10$ mm, larger nodule size, air bronchogram, higher mean CT attenuation, well-defined border, and lobulated margin were detected to be significantly correlated with pGGNs growth. Smoking history showed no significant correlation with pGGN growth based on the multivariate analysis results.

**Edited by:**
*Marco Rengo,*
*Sapienza University of Rome, Italy*

**Reviewed by:**
*David Naidich,*
*New York University, United States*
*Mark Hammer,*
*Brigham and Women's Hospital and Harvard Medical School, United States*

***Correspondence:**
*Li Zhang*
*zhangli_cicams@163.com*

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

**Specialty section:**
*This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology*

***Received:*** *26 April 2022*
***Accepted:*** *09 June 2022*
***Published:*** *04 July 2022*

**Citation:**
*Liang X, Liu M, Li M and Zhang L (2022) Clinical and CT Features of Subsolid Pulmonary Nodules With Interval Growth: A Systematic Review and Meta-Analysis.*
*Front. Oncol. 12:929174.*
*doi: 10.3389/fonc.2022.929174*

**Conclusions:** Eighteen clinical and CT features were identified to be correlated with SSN growth, among which male sex, history of lung cancer, nodule size > 10 mm, nodule consistency and age > 65 years were independent risk factors while history of lung cancer was not correlated with pGGN growth. These factors should be considered when making risk-based follow-up plans for SSN patients.

**Keywords: subsolid nodule, clinical features, CT features, interval growth, meta-analysis**

## INTRODUCTION

Subsolid pulmonary nodules (SSNs) refer to both part-solid nodules (PSNs) and pure ground-glass nodules (pGGNs) (1), and they are defined as nodules that contain components higher than normal lung tissue but less opaque than consolidated bronchovascular margins (2). The widespread availability of high-resolution computed tomography (CT) and the promotion of low-dose chest CT (LDCT) screening programs have increased the detection rate of SSN. Especially because of the COVID-19 epidemic, people are actively undergoing CT scans, so the probability of finding SSNs in the lungs has greatly increased.

The majority (60%-90%) of persistent SSNs have a more indolent clinical course than solid nodules during 5 to 10 years of observation (3–6), and these nodules often represent precursors of invasive adenocarcinoma. A prospective study suggested that SSN growth often indicated a higher risk of invasive adenocarcinoma (7). For pGGNs, the transition to mixed GGNs (solid component within the ground-glass nodule by thin-section CT at a lung window setting) indicates more rapid growth (8). Moreover, it was reported in two studies that 2% and 13% of SSNs showed growth after 5 years of stability (6, 9). The complex growth characteristics and potential malignant properties of SSNs lead to challenges in clinical management. The current guidelines for SSNs take nodule growth as the basis to adjust the follow-up plan and recommend definitive therapy (10). Although there is no consensus on the duration and frequency of SSN follow-up in the guidelines, the 2017 Fleischner Society guidelines, the American College of Chest Physicians guideline (ACCP) and the National Comprehensive Cancer Network (NCCN) all recommend further evaluation and/or consideration of resection if solid component(s) or growth develops in SSNs (11, 12). Considering the risk of a missed diagnosis of lung cancer and worse prognosis, many patients with persistent SSN(s) may switch to more frequent CT surveillance or definitive treatment. This leads to more overexamination and overtreatment in clinical practice. If we can predict whether a nodule will grow, we can adopt different follow-up schemes for different patients to ease their anxiety and solve these problems.

CT surveillance is the sole effective approach for evaluating SSN growth at present. Although studies have revealed that long-term surveillance of SSN(s) with LDCT is a safe strategy, repeated CT scans over several years have nonnegligible consequences, such as anxiety, radiation exposure, false-positive results and unnecessary costs (13). Therefore, risk-based follow-up management for these patients is greatly desired. Several studies have shown that nodule size and history of lung cancer are important risk factors for SSN growth (14, 15). Other studies have shown that lobular margins and a bubble-like appearance are correlated with the growth of SSN (16). Due to the lack of large-sample data comparisons and analyses of clinical and CT features of SSN growth, we retrieved relevant studies up to December 2021 and carried out a meta-analysis, which aimed to clarify the risk factors correlated with SSN growth and provide information for establishing risk-based follow-up strategy for SSN(s) patients.

## METHODS

This meta-analysis was carried out in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines (17, 18). The primary procedures are outlined in the following sections.

### Literature Search

We performed a systematic literature search of Web of Science, PubMed, Cochrane Library and EMBASE up to December 31, 2021. The search terms "non-solid nodule", "part-solid nodule", "subsolid nodule" and their synonyms combined with "growth" or "follow-up" were used without language restriction, and medical subject headings (MeSH) were applied if available. The reference lists of the retrieved articles and review articles were manually searched for other relevant studies. Two authors (L.Z. and M.W.L.) independently performed the search and reviewed all identified publications for inclusion using predetermined criteria.

### Inclusion Criteria

Studies were included when they met the following criteria: (a) studies published in English or Chinese; (b) the cases included in the studies were of SSNs; (c) clinical or CT features were analyzed in the studies; and (d) nodule growth was defined as the whole nodule growing by > 2 mm in diameter, the emergence of a solid component in a pGGN or the solid area growing by > 2 mm in diameter in a PSN. Reports of lectures, conference papers, and reviews were excluded.

### Data Extraction and Quality Assessment

For each eligible study, two authors independently extracted the following data: (a) general information of the studies, (b) mean value and standard deviation of numerical clinical and CT features included in univariate analysis, (c) number of negative and positive cases for categorical clinical and CT features included in univariate analysis, and (d) odds ratio (OR) value with 95% confidence interval of clinical and CT features in multiple logistic regression model. The Newcastle–Ottawa Scale (NOS) was used to assess the methodological quality of the included studies (19).

## Data Analysis

In the univariate analyses, pooled ORs and pooled standardized mean differences (SMDs) were used to detect the strength of each correlation between binary and continuous features and SSN growth, respectively. To facilitate the analyses, we converted ordered categorical variables in some studies into binary variables. Then, we used Pearson's chi-square test or Fisher's exact test to detect if there was a significant difference in the fourfold table for each study and pooled the corresponding ORs. We also estimated the means and standard deviations based on the corresponding medians, ranges and sample sizes for continuous variables whose means and standard deviations were absent (20). Then, we used Student's t test to detect if there was a significant difference for each study and pooled the corresponding SMDs. In the multivariate analyses, pooled adjusted ORs obtained from multiple logistic regression models were used to assess the strength of each correlation of a CT or clinical feature with nodule growth. *P*-values < 0.05 was considered statistically significant. Statistical heterogeneity was determined using the *Q* test and $I^2$ test (21). If $P < 0.1$ or $I^2 > 50\%$, the random-effects model (DerSimonian–Laird model) was used. Otherwise, the fixed-effects model (Mantel–Haenszel model/inverse variance model) was used. Subgroup meta-analysis was performed based on nodule consistency (pGGN and PSN). Publication bias was evaluated by Begg's funnel plot. *P*-values ≥ 0.05 was considered to indicate that no publication bias existed (22). Statistical analyses were performed with R version 4.0.5 and the Meta package.

# RESULTS

## Study Selection

**Figure 1** provides an overview of the literature search and study selection process. Nineteen original studies (5, 6, 8, 9, 14–16, 23–34) were retrieved from 827 potential publications that assessed the relationship between SSN growth and CT or clinical features.

## Study Characteristics and Quality Assessment

All included studies were case–control studies. The group with nodule growth was the case group, and the group without nodule growth was the control group. Among the 19 included studies, 2 analyzed SSNs after 5 years of stability, 2 analyzed SSNs after 3 years of stability, and the remaining analyzed SSNs at baseline. The characteristics of the included studies are shown in **Table 1**. In total, 2444 patients with 3012 SSNs were included. The frequency of SSN growth ranged from 2.13% to 51.61% on a per-nodule basis.

According to the NOS, 14 studies (74%) were high quality (more than five stars), and the other 5 (26%) were low quality (**Supplementary Appendix, Part 1, Table S1**).

## Categorization of Clinical and CT Features

Eighty-six descriptions were used to describe CT or clinical features in the 19 studies. One CT feature (emergence of a solid component) was removed because it is one of the criteria for SSN growth. After merging and subsuming similar descriptions that referred to the same CT findings as a single CT characteristic, 74 features remained. Among them, 11 features were investigated in both one single study for multivariate analysis and more than one study for univariate analysis, 47 features were only investigated in one single study, and 16 features were only investigated in more than one study. Finally, 27 clinical and CT features were included in the meta-

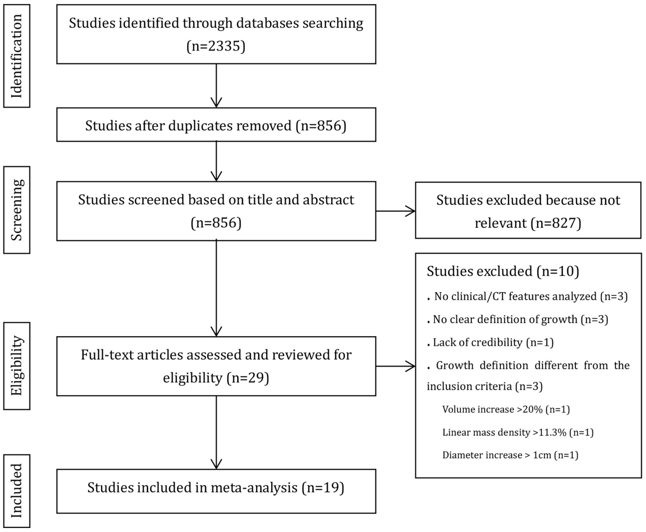


**FIGURE 1** | Flow chart shows summary of the literature review process.

**TABLE 1** | Study characteristics.

| First author | Year | Country/ Region | Data period | Follow-up duration (months) | Total number (patients/ nodules) | Nodule consistency | Number of nodules (pGGN/ PSN) | Analyzed by nodule consistentcy | Number of CT Detector Rows | CT Scan DOSE | Slice thickness (mm) | Definition of nodule growth$^\dagger$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Haruhisa Matsuguma et al. (15) | 2013 | Japan | Jan 2000 - Jun 2008 | 29 (1-136) | 174/174 | SSN | 98/76 | Yes | 4, 64 | Standard-dose | 1 or 0.5 | a, b, c |
| Takashi Eguchi et al. (24) | 2014 | Japan | Sep 1998 - Sep 2013 | 57 (24.1-113.6) | 124/124 | pGGN | 124/0 | No | NA | Low-dose or standard-dose | 1.25 | a, b |
| Shotaro Takahashi et al. (16) | 2012 | Japan | Apr 1999 - Jun 2010 | 66 ± 25 | 111/150 | pGGN | 150/0 | No | 4, 64 | Standard-dose | 2 | a |
| Hyun Woo Lee et al. (9) | 2019 | South Korea | Jan 2003 - Dec 2017 | 136 (120-179) | 160/208 | SSN | 162/46 | No | 16, 256 | Low-dose | 1 or 2 | a |
| Jaeyoung Cho et al. (28) | 2016 | South Korea | May 2003 - Jun 2015 | 77.5 (38.1-117.1) | 218/453 | SSN | 438/15 | No | 64, 256 | NA | 1 to 3 | a, b, c |
| Yuki Sato et al. (31) | 2017 | Japan | Apr 2008 - Dec 2014 | 44 (24.1-87.0) | 187/187 | SSN | 134/53 | No | NA | NA | 0.625 to 2 | a, b, c |
| Boksoon Chang et al. (14) | 2013 | South Korea | Jun 1997 - Sep 2006 | 59 (25-140) | 89/122 | pGGN | 122/0 | No | 64 | Low-dose | 1 or 5 | a |
| Jong Hyuk Lee et al. (29) | 2016 | South Korea | May 2005 - Feb 2013 | days:849 (90-2900) | 213/213 | SSN | 136/77 | Yes | 16, 64 | Low-dose or standard-dose | ≤ 1.25 | a, b, c |
| Masaya Tamura et al. (26) | 2014 | Japan | Oct 2008 - Oct 2012 | 26.1 ± 4.6 | 53/63 | pGGN | 63/0 | No | NA | NA | 2 | a, b |
| Miyako Hiramatsu et al. (23) | 2008 | Japan | 1999-2006 | days: 1048 (177-3269) | 125/125 | SSN | 95/30 | No | NA | Standard-dose | 1.25 or 2 | a, b, c |
| So Hyeon Bak et al. (27) | 2016 | South Korea | Jan 2004 - Jan 2014 | 24.2 ± 16.9 (2.2-64.9) | 49/54 | pGGN | 54/0 | No | 64 | Standard-dose | 2 to 2.5 | a, b |
| Zhe Shi et al. (34) | 2019 | China | Jan 2011 - Dec 2012 | 52 (32-69) | 59/101 | pGGN | 101/0 | No | 64 | Standard-dose | 1 | a, b |
| Xianqun Xu et al. (33) | 2017 | China | Jan 2010 - May 2016 | NA | 69/69 | SSN | NA | No | 16 | Standard-dose | 1 | a |
| Yoshihisa Kobayashi et al. (25) | 2014 | Japan | Jan 1999 - Feb 2013 | NA | 67/120 | SSN | NA | No | NA | NA | NA | a |
| Sei Won Lee et al. (8) | 2013 | South Korea | Apr 2004 - Jul 2011 | 48 (24-99) | 114/175 | SSN | 143/32 | No | 64, 256 | Standard-dose | 3 or 1 | a |
| Wu Fang et al. (30) | 2016 | China | Jun 2008 - | NA | 100/108 | pGGN | 108/0 | No | 256 | Standard-dose | 1.5 | a, b |

*(Continued)*

**TABLE 1** | Continued

| First author | Year | Country/ Region | Data period | Follow-up duration (months) | Total number (patients/ nodules) | Nodule consistency | Number of nodules (pGGN/ PSN) | Analyzed by nodule consistentcy | Number of CT Detector Rows | CT Scan DOSE | Slice thickness (mm) | Definition of nodule growth$^{\dagger}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Apr 2015 | | | | | | | | | |
| En-Kuei Tang et al. (5) | 2019 | Taiwan | Jan 2002 - Aug 2016 | 42.84 ± 35.16 | 128/128 | SSN | 93/35 | No | 16, 64, 256 | NA | 1 to 2.5 | a, b, c |
| Bixiong Wang et al. (32) | 2017 | China | Feb 2009-2016 | 37 (24-81) | 169/203 | SSN | 189/14 | No | 40 | Standard-dose | 5 and 1 | a |
| Jong Hyuk Lee et al. (6) | 2020 | Korea | Jan 2002 - Dec2018 | 112 (84-208) | 235/235 | SSN | 212/24 | No | 16, 64 | Standard-dose | ≤1.5 | a, b, c |

*$^{\dagger}$ a, the whole nodule grew by ≥ 2 mm in diameter; b, emergence of a new solid component; c, the solid area grew by > 2 mm in diameter in part-solid nodules. pGGN, pure ground-glass nodule; PSN, part-solid nodule; SSN, subsolid nodule.*

analysis, and 58 clinical and CT features which investigated in only one study were extracted from their original studies (5, 8, 9, 16, 23, 24, 27, 28, 32–36) and summarized in **Supplementary Appendix** (**Part 1, Table S2**). Among the 27 features included in the meta-analysis, age was analyzed in three ways: as a continuous variable, a binary variables with a threshold of 65 years, and a binary variables with a threshold of 60 years; nodule size was analyzed in two ways: as a continuous variable, and a binary variable with a threshold of 10 mm. The process of categorizing the clinical and CT features is shown in **Supplementary Appendix** (**Part 1, Table S3**).

## Features Correlated With SSN Growth

Twenty-seven clinical and CT features were included in the meta-analysis. Eighteen features, including male sex, history of lung cancer, smoking history, nodule size > 10 mm, larger nodule size, older age, nodule consistency, bubble-like appearance, air bronchogram, spiculated margin, higher mean CT attenuation, well-defined border, lesion below major fissure, larger volume, larger solid component, lobulated margin, higher STD CT attenuation and higher max CT attenuation, were detected to be significantly correlated with SSN growth, while 9 features, including multiple nodules, longer follow-up duration, age > 60 years, age > 65 years, emphysema, nodule shape, peripheral distribution, pleural/fissure retraction and larger mass, showed no significant correlation with SSN growth. Among the 27 CT and clinical features, 13 features, including male sex, number of nodules, history of lung cancer, smoking history, nodule > 10 mm, nodule size, age (years), follow-up duration, nodule consistency, bubble-like appearance, air bronchogram, spiculated margin and mean CT attenuation, were investigated in five or more studies. The pooled OR/SMD of these features are summarized in **Table 2** and forest plots are shown in **Supplementary Appendix** (**Part 1, Figure S1**).

## Independent Risk Factors for SSN Growth

Pooled ORs of six features, including sex, history of lung cancer, smoking history, nodule size > 10 mm, nodule consistency and age > 65 years, were calculated based on the multivariate analysis results. Forest plots of these six features are shown in **Figure 2**. Sex, history of lung cancer, nodule size > 10 mm, nodule consistency and age > 65 years were proven to be independent risk factors for SSN growth. Male patients showed a 2.351-fold higher probability of SSN growth (pooled OR 2.351, 95% CI 1.370-4.032, $P$ = 0.002). Patients with a history of lung cancer had a 3.030-fold higher probability of SSN growth (pooled OR 3.030, 95% CI 1.933-4.749, $P$ < 0.001). Patients with a nodule size > 10 mm had a 4.236-fold higher probability of SSN growth (pooled OR 4.236, 95% CI 1.488-12.059, $P$ = 0.002). PSNs (nodule consistency) had a 2.951-fold higher probability of SSN growth (pooled OR 2.951, 95% CI 1.821-4.782, $P$ < 0.001). Patients aged > 65 years had a 2.260-fold higher probability of SSN growth (pooled OR 2.260, 95% CI 1.308-3.903, $P$ = 0.003). Smoking history showed no significant correlation with SSN growth based on the multivariate analysis results (pooled OR 1.941, 95% CI 0.935-4.029, $P$ = 0.075).

## Subgroup Analyses on Nodule Consistency

The nodule consistency of the 19 original studies we included was pGGN in 7 studies and SSN in 12 studies. Among the 12 studies taking SSN as research objects, 2 studies analyzed the correlation between features and nodule growth by nodule consistency. In total, 9 studies analyzed the correlation between features and pGGN growth, and 2 studies analyzed the correlation between features and PSN. The number of studies on PSN is too few to perform a meta-analysis. Therefore, we just performed the meta-analysis on pGGN including a total of 16 features based on univariate analysis and 1 feature based on multivariate analysis.

Eight features, including male sex, smoking history, nodule size > 10 mm, larger nodule size, air bronchogram, higher mean CT attenuation, well-defined border, and lobulated margin were detected to be significantly correlated with pGGN growth ($P$ = 0.03, 0.003, 0.047, 0.009, 0.001, 0.005, 0.044 and 0.001, respectively), while 8 features, including multiple nodules,

**TABLE 2 |** Clinical and CT features included in the meta-analysis in SSN.

| Features | Studies (patients/nodules) | Test of Correlation | | | Test of Heterogeneity | |
|---|---|---|---|---|---|---|
| | | Pooled OR or SMD | 95% CI | *P* Value | $I^2$ (%) | *P* Value |
| Sex (Male) | 16 (1846/2400) † | 1.469 | 1.066-2.026 | 0.019 | 42.50 | 0.037 |
| No. of nodules (Multiple) | 16 (1877/2439) † | 0.976 | 0.768-1.241 | 0.843 | 0.00 | 0.693 |
| History of lung cancer (Yes) | 12 (1501/1980) † | 1.738 | 1.098-2.750 | 0.018 | 54.70 | 0.012 |
| Smoking history (Yes) | 12 (1424/1878) † | 1.692 | 1.137-2.520 | 0.010 | 39.40 | 0.078 |
| Nodule size (> 10 mm) | 12 (1466/1978) † | 6.386 | 3.514-11.605 | < 0.001 | 65.60 | 0.001 |
| Nodule size, mm | 10 (1107/1517) | 0.678 | 0.310-1.046 | < 0.001 | 83.40 | < 0.001 |
| Age, years | 8 (958/1355) | 0.305 | 0.089-0.521 | 0.006 | 42.20 | 0.097 |
| Follow-up duration, months | 8 (784/936) | 0.077 | -0.613-0.767 | 0.827 | 94.20 | < 0.001 |
| Nodule consistency (PSN) | 8 (1167/1597) † | 3.682 | 2.655-5.107 | < 0.001 | 22.80 | 0.248 |
| Bubble like appearance (Yes) | 6 (792/1216) | 3.938 | 1.214-12.772 | 0.022 | 72.20 | 0.003 |
| Air bronchogram (Yes) | 5 (824/1154) | 4.858 | 2.593-9.101 | < 0.001 | 20.30 | 0.285 |
| Spiculated margin (Yes) | 5 (824/1154) | 10.786 | 1.006-115.624 | 0.049 | 63.70 | 0.041 |
| Mean of CT attenuation, HU | 5 (405/465) | 1.952 | 0.780-3.125 | 0.001 | 95.40 | < 0.001 |
| Age (> 60 years) | 4 (299/385) † | 1.578 | 0.969-2.570 | 0.067 | 37.30 | 0.188 |
| Well-defined border (Yes) | 4 (353/443) | 0.544 | 0.301-0.983 | 0.044 | 0.00 | 0.652 |
| Age (> 65 years) | 3 (354/425) | 1.738 | 0.792-3.812 | 0.168 | 62.70 | 0.068 |
| Emphysema (Yes) | 3 (506/593) | 0.607 | 0.207-1.774 | 0.361 | 40.90 | 0.184 |
| Lesion location (Below major fissure) | 3 (374/455) | 0.448 | 0.242-0.832 | 0.011 | 0.00 | 0.761 |
| Nodule shape (Round) | 3 (264/321) | 0.559 | 0.263-1.187 | 0.130 | 42.70 | 0.175 |
| Volume, mm3 | 3 (177/224) | 0.988 | 0.041-1.936 | 0.041 | 88.90 | < 0.001 |
| Solid part size, mm | 2 (288/336) | 0.429 | 0.164-0.695 | 0.002 | 0.00 | 0.575 |
| Lobulated margin (Yes) | 2 (200/272) | 15.081 | 3.050-74.575 | 0.001 | 0.00 | 0.650 |
| Peripheral distribution (Yes) | 2 (324/357) | 3.342 | 0.432-25.874 | 0.248 | 0.00 | 0.904 |
| Pleural/fissure retraction (Yes) | 2 (453/688) | 1.963 | 0.244-15.785 | 0.526 | 0.00 | 0.824 |
| STD of CT attenuation, HU | 2 (128/170) | 1.067 | 0.693-1.440 | < 0.001 | 0.00 | 0.957 |
| Max of CT attenuation, HU | 2 (128/170) | 1.299 | 0.574-2.024 | < 0.001 | 72.00 | 0.059 |
| Mass, mg | 2 (108/155) | 0.715 | -0.986-2.417 | 0.410 | 94.50 | < 0.001 |

*†Total number of patients in Yoshihisa Kobayashi's study, and Bixiong Wang's study were not reported; PSN, part solid nodule; OR, odds ratio; SMD, standardized mean difference; CI, confidence interval.*

history of lung cancer, age, longer follow-up duration, bubble-like appearance, nodule shape, volume and larger mass, showed no significant correlation with pGGN growth (*P* = 0.675, 0.366, 0.071, 0.796, 0.234, 0.130, 0.176 and 0.410, respectively) based on univariate analysis (**Table 3**; **Supplementary Appendix, Part 1, Figure S2**). Smoking history showed no significant correlation with pGGN growth based on the multivariate analysis results (*P* = 0.071, **Figure 3**).

## Publication Bias

The funnel plots did not show significant publication bias for the CT and clinical features analyzed based on both the univariate analysis results and multivariate analysis results (*P* > 0.05). (**Supplementary Appendix, Part 1, Figure S3** and **Figure S4**)

## DISCUSSION

In this systematic review and meta-analysis, 18 clinical and CT features were found to be significantly correlated with SSN growth, and 5 features including male sex, history of lung cancer, nodule size > 10 mm, nodule consistency and age > 65 years were identified to be independent risk factors for SSN growth. Eight features, including male sex, smoking history, nodule size > 10 mm, larger nodule size, air bronchogram, higher mean CT attenuation, well-defined border, and lobulated margin were associated with pGGN growth. Among the features associated with SSN growth, 11 features, including sex, history of lung cancer, smoking history, nodule size (> 10 mm), nodule size (mm), age (years), nodule consistency, bubble-like appearance, air bronchogram, spiculated margin and mean CT attenuation, were investigated in 5 or more studies, while the other 7 features were only investigated in 2 to 4 studies. Among the features associated with pGGN growth, 5 features including male sex, number of nodules, smoking history, nodule size (mm) and follow-up duration, were investigated in 5 studies or more studies, while the other 11 features were only investigated in 2 to 4 studies. The real clinical significance of the features investigated in less than 5 studies needs to be further studied due to the small number of studies included.

Studies have proven that PSNs are more aggressive than pGGNs (37, 38). In this meta-analysis, PSNs had a 2.95-fold higher probability of SSN growth than pGGNs, which is consistent with the previous studies. When we analyzed the characteristics of pGGN, we found that history of lung cancer, age, bubble like appearance, and volume, which are significantly correlated SSN growth, are not associated with pGGN growth. Although the including studies for these four features are very limited (only 2 to 4 studies), the results may imply that the features correlated with the growth of the pGGN and the PSN are different. Therefore, to predict the growth of SSN based on risk features, it should be identified whether the nodule is PSN or pGGN at first.

A history of lung cancer showed high correlation with SSN growth, and it forecasted a 3.498-fold higher probability of SSN

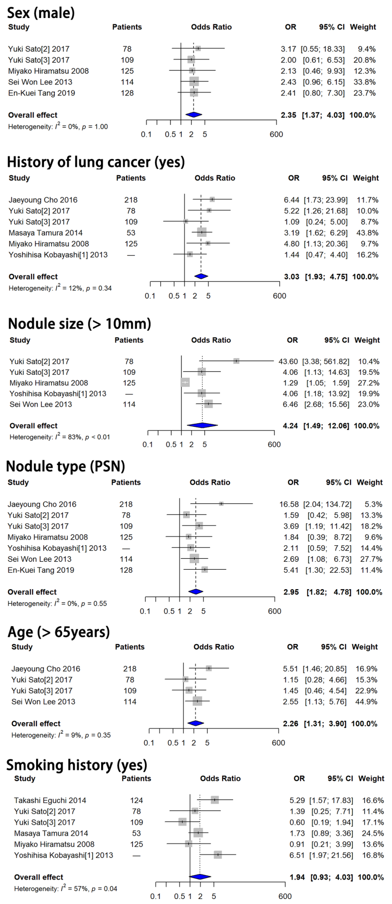


**FIGURE 2** | Forest plots showed that male sex, history of lung cancer, nodule size > 10 mm, PSN and age > 65 years were independent risk factors for SSN growth. Smoking history showed no significant correlation with SSN growth. OR, odds ratio; CI, confidence interval; PSN, part-solid nodule.

growth. However, history of lung cancer was not associated with pGGN growth. Multiple studies suggest that lung cancer history is associated with SSN growth (15, 23, 28, 39), but a few studies found that lung cancer history was not associated with pGGN growth (24, 29). A prospective multicenter study concluded that history of lung cancer was not a factor for SSN growth, and 85.1% of nodules included in the study were pGGNs (7). The relationship between the history of lung cancer and the growth of pGGN and PSN needs further study. Shewale et al. (40) retrospectively reviewed 210 patients with GGNs and a history of lung cancer and demonstrated that patients with a lung adenocarcinoma history had a 6.85-fold higher likelihood for SSN growth than patients with other lung cancer subtypes. As we know, among non-small cell lung cancer, adenocarcinoma is a histological subtype prone to epidermal growth factor receptor (EGFR) mutation, and previous studies have shown that the growth of SSN(s) is closely related to gene mutation status. EGFR mutations have been found to be a promoter of GGN growth in both mice and humans (41, 42). This may be one of the reasons why SSNs are more likely to grow in patients with lung cancer, especially those with adenocarcinoma. The correlation between other tumor types and SSN growth needs to be elucidated.

Among the 5 independent risk factors for SSN growth, a nodule size > 10 mm showed the highest correlation with SSN growth, leading to a 4.236-fold higher probability of SSN growth than nodules ≤ 10 mm in size. Nodule size is an important factor in both SSN follow-up guidelines and pulmonary nodule malignancy prediction models. In the Fleischner society, ACCP and NCCN guidelines, the management of SSN(s) differs based on nodule size (11, 12, 43). Nodule size is also a risk factor used in the Mayo model, PanCan model and Vancouver model to estimate the malignancy risk of pulmonary nodules (11, 44). Several studies also use volume to describe the size of nodules. Han et al. (45) found that in lung cancer screening, semiautomatic volume measurements showed higher accuracy than diameter measurements. In this meta-analysis, three original studies were included, and a significant correlation between volume and nodule growth was found (pooled SMD 0.988, 95% CI 0.041-1.936, $P$ = 0.041). However, considering the difficulties in identifying the SSN boundary for computer-aided measurements, using diameter to describe the size of the SSN is still a better choice.

In this meta-analysis, we also found another interesting result. The follow-up duration showed no significant correlation with either SSN growth or pGGN growth ($P$ = 0.827 and 0.796 respectively). Kobayashi et al. (4) found the tendency to grow was clear within the first 3 years for SSNs. Lee et al. (6) followed SSNs that had been stable for initial 5 years and found that only 5 (2.1%) of nodules grew. The frequency of SSNs increasing in size after prolonged stability is quite small, which is consistent to our results. Based on the above, we speculate that the frequency of CT examination can be reduced for SSN followed up for more than 5 years.

Studies have reported that approximately 37%-70% of SSNs detected on CT screening are transient and resolve spontaneously or with antibiotic therapy within 3 months of the initial examination (46–48). Features associated with transient SSNs include younger age, male sex, peripheral eosinophilia, multiplicity, ill-defined margins, nonspiculated margins, and large solid components (47, 49). Some features, such as ill-defined margins, male sex and large solid components coincide with the features of SSNs growth which may lead to the determination of a transient nodule as a growing nodule. Therefore, we suggest that the prediction of SSN growth should be performed at least 3 months after the initial examination. SSN growth prediction is important not only to make follow-up plan but also to optimize surgical timing. Based on the included studies, we found that only 49.23% stable SSNs were invasive adenocarcinoma (IAC) while 81.01% growth SSNs were IAC, and growth SSNs showed a 4.32-fold higher probability of invasive adenocarcinoma than stable SSNs (**Supplementary Appendix, Part 2**). These results suggest that surgical resection after the growth of SSNs may be more appropriate than upon detection, which needs further study.

Our study had several limitations. First, the study subjects in the included studies were not completely homogenous. Among the 19 included studies, 2 analyzed SSNs after 5 years of stability,

**TABLE 3** | Clinical and CT features included in the meta-analysis based on univariate analyses in pGGN.

| Features | Studies (patients/nodules) | Test of Correlation | | | Test of Heterogeneity | |
|---|---|---|---|---|---|---|
| | | Pooled OR or SMD | 95% CI | *P* Value | $I^2$ (%) | *P* Value |
| Sex (Male) | 6 (534/658) | 1.615 | 1.049-2.488 | 0.030 | 0.00% | 0.614 |
| Number of nodules (Multiple) | 7 (634/766) | 1.092 | 0.724-1.648 | 0.675 | 0.00% | 0.474 |
| History of lung cancer (Yes) | 4 (386/435) | 1.634 | 0.564-4.738 | 0.366 | 74.90% | 0.008 |
| Smoking history (Yes) | 5 (423/508) | 2.143 | 1.292-3.554 | 0.003 | 0.00% | 0.508 |
| Nodule size (> 10 mm) | 4 (351/433) | 4.975 | 1.024-24.164 | 0.047 | 81.80% | 0.001 |
| Nodule size, mm | 6 (532/659) | 0.847 | 0.209-1.485 | 0.009 | 89.20% | 0.000 |
| Age, years | 4 (383/497) | 0.217 | -0.019-0.452 | 0.071 | 0.00% | 0.905 |
| Follow-up duration, months | 6 (496/600) | 0.095 | -0.623-0.812 | 0.796 | 91.60% | 0.000 |
| Bubble like appearance (Yes) | 3 (300/380) | 3.005 | 0.491-18.379 | 0.234 | 80.20% | 0.006 |
| Air bronchogram (Yes) | 2 (211/258) | 4.374 | 1.764-10.845 | 0.001 | 0.00% | 0.492 |
| Mean of CT attenuation, HU | 4 (336/396) | 2.228 | 0.659-3.798 | 0.005 | 96.50% | 0.000 |
| Well-defined border (Yes) | 4 (353/443) | 0.544 | 0.301-0.983 | 0.044 | 0.00% | 0.652 |
| Nodule shape (Round) | 3 (264/321) | 0.559 | 0.263-1.187 | 0.130 | 42.70% | 0.175 |
| Volume, mm3 | 2 (108/155) | 1.149 | -0.517-2.815 | 0.176 | 93.90% | 0.000 |
| Lobulated margin (Yes) | 2 (200/272) | 15.081 | 3.050-74.575 | 0.001 | 0.00% | 0.650 |
| Mass, mg | 2 (108/155) | 0.715 | -0.986-2.417 | 0.410 | 94.50% | 0.000 |

*pGGN, pure ground glass nodule; OR, odds ratio; SMD, standardized mean difference; CI, confidence interval.*

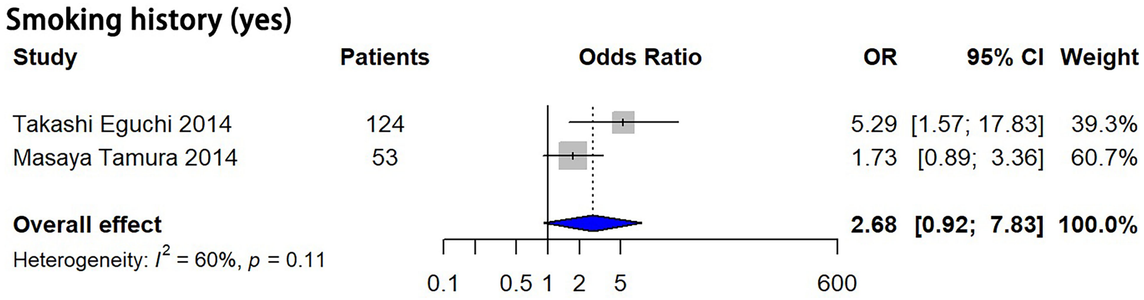


**FIGURE 3** | Forest plots showed no significant correlation between smoking history and pGGN growth based on multivariate analysis. OR, odds ratio; CI, confidence interval; pGGN, pure ground glass nodule.

2 analyzed SSNs after 3 years of stability, and the remaining analyzed SSNs at baseline. Fortunately, no significant differences in the included features were detected in publication bias tests. Second, most of the features had no multivariate analysis results in the original studies, and their specific role in SSN growth needs to be further studied. Third, compared with a large number of clinical and CT features, the number of studies included is so small that the meta-analysis of some features may lack credibility. Fourth, the number of studies on PSN is too few to perform a meta-analysis. Nevertheless, to the best of our knowledge, our meta-analysis is the only study to investigate the correlation of clinical and CT features with SSN growth, and this analysis included all available literature.

In conclusion, in this meta-analysis, eighteen clinical and CT features were detected to be significantly correlated with SSN growth, and 5 features including male sex, history of lung cancer, nodule size > 10 mm, nodule consistency and age > 65 years were identified to be independent risk factors for SSN growth. For pGGN, history of lung cancer, older age, bubble-like appearance, and larger volume were not risk factors for growth, although these factors were associated with SSN growth. A risk-based SSN follow-up strategy should consider these factors and nodule consistency, and separate strategies should be planned for each single nodule in patients with multiple SSNs.

## 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

(I) Conception and design: All authors; (II) Administrative support: None; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: LZ, MWL, ML; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors

## FUNDING

This work was supported by National Natural Science Foundation of China (81701692); and Beijing Municipal Natural Science Foundation (7184238).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.929174/full#supplementary-material

## REFERENCES

1. Godoy MCB, Naidich DP. Subsolid Pulmonary Nodules and the Spectrum of Peripheral Adenocarcinomas of the Lung: Recommended Interim Guidelines for Assessment and Management. *Radiology* (2009) 253(3):606–22. doi: 10.1148/radiol.2533090179
2. Hansell DM, Bankier AA, MacMahon H, McLoud TC, Müller NL, Remy J. Fleischner Society: Glossary of Terms for Thoracic Imaging. *Radiology* (2008) 246(3):697–722. doi: 10.1148/radiol.2462070712
3. Hasegawa M, Sone S, Takashima S, Li F, Yang ZG, Maruyama Y, et al. Growth Rate of Small Lung Cancers Detected on Mass Ct Screening. *Br J Radiol* (2000) 73(876):1252–9. doi: 10.1259/bjr.73.876.11205667
4. Kobayashi Y, Fukui T, Ito S, Usami N, Hatooka S, Yatabe Y, et al. How Long Should Small Lung Lesions of Ground-Glass Opacity Be Followed? *J Thorac Oncol* (2013) 8(3):309–14. doi: 10.1097/JTO.0b013e31827e2435
5. Tang EK, Chen CS, Wu CC, Wu MT, Yang TL, Liang HL, et al. Natural History of Persistent Pulmonary Subsolid Nodules: Long-Term Observation of Different Interval Growth. *Heart Lung Circ* (2019) 28(11):1747–54. doi: 10.1016/j.hlc.2018.08.015
6. Lee JH, Lim WH, Hong JH, Nam JG, Hwang EJ, Kim H, et al. Growth and Clinical Impact of 6-Mm or Larger Subsolid Nodules After 5 Years of Stability at Chest Ct. *Radiology* (2020) 295(2):448–55. doi: 10.1148/radiol.2020191921
7. Kakinuma R, Noguchi M, Ashizawa K, Kuriyama K, Maeshima AM, Koizumi N, et al. Natural History of Pulmonary Subsolid Nodules: A Prospective Multicenter Study. *J Thorac Oncol* (2016) 11(7):1012–28. doi: 10.1016/j.jtho.2016.04.006
8. Lee SW, Leem CS, Kim TJ, Lee KW, Chung JH, Jheon S, et al. The Long-Term Course of Ground-Glass Opacities Detected on Thin-Section Computed Tomography. *Respir Med* (2013) 107(6):904–10. doi: 10.1016/j.rmed.2013.02.014
9. Lee HW, Jin KN, Lee JK, Kim DK, Chung HS, Heo EY, et al. Long-Term Follow-Up of Ground-Glass Nodules After 5 Years of Stability. *J Thorac Oncol* (2019) 14(8):1370–7. doi: 10.1016/j.jtho.2019.05.005
10. Loverdos K, Fotiadis A, Kontogianni C, Iliopoulou M, Gaga M. Lung Nodules: A Comprehensive Review on Current Approach and Management. *Ann Thorac Med* (2019) 14(4):226–38. doi: 10.4103/atm.ATM_110_19
11. MacMahon H, Naidich DP, Goo JM, Lee KS, Leung ANC, Mayo JR, et al. Guidelines for Management of Incidental Pulmonary Nodules Detected on Ct Images: From the Fleischner Society 2017. *Radiology* (2017) 284(1):228–43. doi: 10.1148/radiol.2017161659
12. Gould MK, Donington J, Lynch WR, Mazzone PJ, Midthun DE, Naidich DP, et al. Evaluation of Individuals With Pulmonary Nodules: When Is It Lung Cancer? Diagnosis and Management of Lung Cancer, 3rd Ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest* (2013) 143(5 Suppl):e93S–e120S. doi: 10.1378/chest.12-2351
13. Silva M, Prokop M, Jacobs C, Capretti G, Sverzellati N, Ciompi F, et al. Long-Term Active Surveillance of Screening Detected Subsolid Nodules Is a Safe Strategy to Reduce Overtreatment. *J Thorac Oncol* (2018) 13(10):1454–63. doi: 10.1016/j.jtho.2018.06.013
14. Chang B, Hwang JH, Choi YH, Chung MP, Kim H, Kwon OJ, et al. Natural History of Pure Ground-Glass Opacity Lung Nodules Detected by Low-Dose Ct Scan. *Chest* (2013) 143(1):172–8. doi: 10.1378/chest.11-2501
15. Matsuguma H, Mori K, Nakahara R, Suzuki H, Kasai T, Kamiyama Y, et al. Characteristics of Subsolid Pulmonary Nodules Showing Growth During Follow-Up With Ct Scanning. *Chest* (2013) 143(2):436–43. doi: 10.1378/chest.11-3306
16. Takahashi S, Tanaka N, Okimoto T, Tanaka T, Ueda K, Matsumoto T, et al. Long Term Follow-Up for Small Pure Ground-Glass Nodules: Implications of Determining an Optimum Follow-Up Period and High-Resolution Ct Findings to Predict the Growth of Nodules. *Jpn J Radiol* (2012) 30(3):206–17. doi: 10.1007/s11604-011-0033-8
17. Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, et al. The Prisma Extension Statement for Reporting of Systematic Reviews Incorporating Network Meta-Analyses of Health Care Interventions: Checklist and Explanations. *Ann Internal Med* (2015) 162(11):777–84. doi: 10.7326/M14-2385
18. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (Prisma-P) 2015 Statement. *Systematic Rev* (2015) 4:1. doi: 10.1186/2046-4053-4-1
19. Wells GA, Shea B, O'Connell D, Peterson J, Welch V. *The Newcastle–Ottawa Scale (Nos) for Assessing the Quality of Nonrandomized Studies in Meta-Analysis (2014)* (2014). Available at: http://www.ohrica/programs/clinical_epidemiology/oxfordasp.
20. Hozo SP, Djulbegovic B, Hozo I. Estimating the Mean and Variance From the Median, Range, and the Size of a Sample. *BMC Med Res Methodol* (2005) 5(1 Supplement):13. doi: 10.1186/1471-2288-5-13
21. Higgins JP, Thompson SG. Quantifying Heterogeneity in a Meta-Analysis. *Stat Med* (2002) 21(11):1539–58. doi: 10.1002/sim.1186
22. Lau J, Ioannidis JPA, Terrin N, Schmid CH, Olkin I. The Case of the Misleading Funnel Plot. *BMJ (online)* (2006) 333(7568):597–600. doi: 10.1136/bmj.333.7568.597
23. Hiramatsu M, Inagaki T, Inagaki T, Matsui Y, Satoh Y, Okumura S, et al. Pulmonary Ground-Glass Opacity (Ggo) Lesions-Large Size and a History of Lung Cancer Are Risk Factors for Growth. *J Thorac Oncol* (2008) 3(11):1245–50. doi: 10.1097/JTO.0b013e318189f526
24. Eguchi T, Kondo R, Kawakami S, Matsushita M, Yoshizawa A, Hara D, et al. Computed Tomography Attenuation Predicts the Growth of Pure Ground-Glass Nodules. *Lung Cancer (Amsterdam Netherlands)* (2014) 84(3):242–7. doi: 10.1016/j.lungcan.2014.03.009
25. Kobayashi Y, Sakao Y, Deshpande GA, Fukui T, Mizuno T, Kuroda H, et al. The Association Between Baseline Clinical-Radiological Characteristics and Growth of Pulmonary Nodules With Ground-Glass Opacity. *Lung Cancer* (2014) 83(1):61–6. doi: 10.1016/j.lungcan.2013.10.017
26. Tamura M, Shimizu Y, Yamamoto T, Yoshikawa J, Hashizume Y. 15-Predictive Value of One-Dimensional Mean Computed Tomography Value of Ground-Glass Opacity on High-Resolution Images for the Possibility of Future Change. *J Thorac Oncol* (2014) 9(4):469–72. doi: 10.1097/JTO.0000000000000117
27. Bak SH, Lee HY, Kim JH, Um SW, Kwon OJ, Han J, et al. Quantitative Ct Scanning Analysis of Pure Ground-Glass Opacity Nodules Predicts Further Ct Scanning Change. *Chest* (2016) 149(1):180–91. doi: 10.1378/chest.15-0034
28. Cho J, Kim ES, Kim SJ, Lee YJ, Park JS, Cho YJ, et al. Long-Term Follow-Up of Small Pulmonary Ground-Glass Nodules Stable for 3 Years: Implications of the Proper Follow-Up Period and Risk Factors for Subsequent Growth. *J Thorac Oncol* (2016) 11(9):1453–9. doi: 10.1016/j.jtho.2016.05.026
29. Lee JH, Park CM, Lee SM, Kim H, McAdams HP, Goo JM. Persistent Pulmonary Subsolid Nodules With Solid Portions of 5 Mm or Smaller: Their Natural Course and Predictors of Interval Growth. *Eur Radiol* (2016) 26(6):1529–37. doi: 10.1007/s00330-015-4017-4
30. Wu F, Cai ZL, Tian SP, Jin X, Jing R, Yang YQ, et al. Value of Baseline Clinical and Ct Characteristics for Predicting the Progression of Persistent Pure Ground-Glass Nodule 10 Mm or Less in Diameter. *Zhongguo Yi Xue Ke Xue Yuan Xue Bao Acta Acad Med Sinicae* (2016) 38(4):371–7. doi: 10.3881/j.issn.1000-503X.2016.04.001
31. Sato Y, Fujimoto D, Morimoto T, Uehara K, Nagata K, Sakanoue I, et al. 9-Natural History and Clinical Characteristics of Multiple Pulmonary Nodules With Ground Glass Opacity. *Respirology* (2017) 22(8):1615–21. doi: 10.1111/resp.13089

32. Wang B, You X, Sun X. The Associations Among Radiological Characteristics, Growth, and Pathological Results of Pulmonary Subsolid Nodules. *Biomed Res* (2017) 28(15):6829–33.
33. Xu X, Wu K, Zhao Y, Mei L. Stage I Lung Adenocarcinoma: The Value of Quantitative Ct in Differentiating Pathological Subtypes and Predicting Growth of Subsolid Nodules. *Med (Baltimore)* (2017) 96(16):e6595. doi: 10.1097/MD.0000000000006595
34. Shi Z, Deng J, She Y, Zhang L, Ren Y, Sun W, et al. Quantitative Features Can Predict Further Growth of Persistent Pure Ground-Glass Nodule. *Quant Imaging Med Surg* (2019) 9(2):283–91. doi: 10.21037/qims.2019.01.04
35. Sato Y, Fujimoto D, Morimoto T, Uehara K, Nagata K, Sakanoue I, et al. Natural History and Clinical Characteristics of Multiple Pulmonary Nodules With Ground Glass Opacity. *Respirology* (2017) 22(8):1615–21. doi: 10.1111/resp.13089
36. Tamura M, Shimizu Y, Yamamoto T, Yoshikawa J, Hashizume Y. Predictive Value of One-Dimensional Mean Computed Tomography Value of Ground-Glass Opacity on High-Resolution Images for the Possibility of Future Change. *J Thorac Oncol* (2014) 9(4):469–72. doi: 10.1097/JTO.0000000000000117
37. Oda S, Awai K, Murao K, Ozawa A, Utsunomiya D, Yanaga Y, et al. Volume-Doubling Time of Pulmonary Nodules With Ground Glass Opacity at Multidetector Ct: Assessment With Computer-Aided Three-Dimensional Volumetry. *Acad Radiol* (2011) 18(1):63–9. doi: 10.1016/j.acra.2010.08.022
38. Song YS, Park CM, Park SJ, Lee SM, Jeon YK, Goo JM. Volume and Mass Doubling Times of Persistent Pulmonary Subsolid Nodules Detected in Patients Without Known Malignancy. *Radiology* (2014) 273(1):276–84. doi: 10.1148/radiol.14132324
39. de Margerie-Mellon C, Ngo LH, Gill RR, Monteiro Filho AC, Heidinger BH, Onken A, et al. The Growth Rate of Subsolid Lung Adenocarcinoma Nodules at Chest Ct. *Radiology* (2020) 297(1):189–98. doi: 10.1148/radiol.2020192322
40. Shewale JB, Nelson DB, Rice DC, Sepesi B, Hofstetter WL, Mehran RJ, et al. Natural History of Ground-Glass Lesions Among Patients With Previous Lung Cancer. *Ann Thorac Surg* (2018) 105(6):1671–7. doi: 10.1016/j.athoracsur.2018.01.031
41. Aoki T, Hanamiya M, Uramoto H, Hisaoka M, Yamashita Y, Korogi Y. Adenocarcinomas With Predominant Ground-Glass Opacity: Correlation of Morphology and Molecular Biomarkers. *Radiology* (2012) 264(2):590–6. doi: 10.1148/radiol.12111337
42. Kobayashi Y, Mitsudomi T, Sakao Y, Yatabe Y. Genetic Features of Pulmonary Adenocarcinoma Presenting With Ground-Glass Nodules: The Differences Between Nodules With and Without Growth. *Ann Oncol* (2015) 26(1):156–61. doi: 10.1093/annonc/mdu505
43. Ettinger DS, Wood DE, Aisner DL, Akerley W, Bauman JR, Bharat A, et al. Nccn Guidelines Insights: Non-Small Cell Lung Cancer, Version 2.2021. *J Natl Compr Canc Netw* (2021) 19(3):254–66. doi: 10.6004/jnccn.2021.0013
44. Al-Ameri A, Malhotra P, Thygesen H, Plant PK, Vaidyanathan S, Karthik S, et al. Risk of Malignancy in Pulmonary Nodules: A Validation Study of Four Prediction Models. *Lung Cancer* (2015) 89(1):27–30. doi: 10.1016/j.lungcan.2015.03.018
45. Han D, Heuvelmans MA, Oudkerk M. Volume Versus Diameter Assessment of Small Pulmonary Nodules in Ct Lung Cancer Screening. *Transl Lung Cancer Res* (2017) 6(1):52–61. doi: 10.21037/tlcr.2017.01.05
46. Oh JY, Kwon SY, Yoon HI, Lee SM, Yim JJ, Lee JH, et al. Clinical Significance of a Solitary Ground-Glass Opacity (Ggo) Lesion of the Lung Detected by Chest Ct. *Lung Cancer* (2007) 55(1):67–73. doi: 10.1016/j.lungcan.2006.09.009
47. Lee SM, Park CM, Goo JM, Lee CH, Lee HJ, Kim KG, et al. Transient Part-Solid Nodules Detected at Screening Thin-Section Ct for Lung Cancer: Comparison With Persistent Part-Solid Nodules. *Radiology* (2010) 255(1):242–51. doi: 10.1148/radiol.09090547
48. Felix L, Serra-Tosio G, Lantuejoul S, Timsit JF, Moro-Sibilot D, Brambilla C, et al. Ct Characteristics of Resolving Ground-Glass Opacities in a Lung Cancer Screening Programme. *Eur J Radiol* (2011) 77(3):410–6. doi: 10.1016/j.ejrad.2009.09.008
49. Choi WS, Park CM, Song YS, Lee SM, Wi JY, Goo JM. Transient Subsolid Nodules in Patients With Extrapulmonary Malignancies: Their Frequency and Differential Features. *Acta Radiol* (2015) 56(4):428–37. doi: 10.1177/0284185114528325

**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.

frontiers | Frontiers in Oncology

TYPE Review
PUBLISHED 14 July 2022
DOI 10.3389/fonc.2022.960984

EDITED BY
Humberto Rocha,
University of Coimbra, Portugal

REVIEWED BY
Niccolò Marini,
HES-SO Valais-Wallis, Switzerland
Alireza Sadeghian,
Toronto Metropolitan University,
Canada
Wenbin Chen,
Southern Medical University, China

*CORRESPONDENCE
Jan-Niklas Eckardt
jan-niklas.eckardt@uniklinikum-dresden.de

SPECIALTY SECTION
This article was submitted to
Cancer Imaging and
Image-directed Interventions,
a section of the journal
Frontiers in Oncology

RECEIVED 03 June 2022
ACCEPTED 24 June 2022
PUBLISHED 14 July 2022

CITATION
Eckardt J-N, Bornhäuser M, Wendt K and Middeke JM (2022) Semi-supervised learning in cancer diagnostics.
*Front. Oncol.* 12:960984.
doi: 10.3389/fonc.2022.960984

# Semi-supervised learning in cancer diagnostics

Jan-Niklas Eckardt[1,2]*, Martin Bornhäuser[1,3,4], Karsten Wendt[2,5] and Jan Moritz Middeke[1,2]

[1]Department of Internal Medicine I, University Hospital Carl Gustav Carus, Dresden, Germany, [2]Else Kröner Fresenius Center for Digital Health, Technical University Dresden, Dresden, Germany, [3]German Consortium for Translational Cancer Research, Heidelberg, Germany, [4]National Center for Tumor Disease (NCT), Dresden, Germany, [5]Institute of Software and Multimedia Technology, Technical University Dresden, Dresden, Germany

In cancer diagnostics, a considerable amount of data is acquired during routine work-up. Recently, machine learning has been used to build classifiers that are tasked with cancer detection and aid in clinical decision-making. Most of these classifiers are based on supervised learning (SL) that needs time- and cost-intensive manual labeling of samples by medical experts for model training. Semi-supervised learning (SSL), however, works with only a fraction of labeled data by including unlabeled samples for information abstraction and thus can utilize the vast discrepancy between available labeled data and overall available data in cancer diagnostics. In this review, we provide a comprehensive overview of essential functionalities and assumptions of SSL and survey key studies with regard to cancer care differentiating between image-based and non-image-based applications. We highlight current state-of-the-art models in histopathology, radiology and radiotherapy, as well as genomics. Further, we discuss potential pitfalls in SSL study design such as discrepancies in data distributions and comparison to baseline SL models, and point out future directions for SSL in oncology. We believe well-designed SSL models to strongly contribute to computer-guided diagnostics in malignant disease by overcoming current hinderances in the form of sparse labeled and abundant unlabeled data.

KEYWORDS
semi-supervised learning, cancer, diagnostics, artificial intelligence, machine learning

## Introduction

In the daily routine of cancer diagnostics, an abundance of medical data in the form of images, health records and genetic assays are gathered. Potentially, these data can serve as training input for supervised machine learning classifiers, however, the availability of large-scale labeled datasets represents a substantial bottleneck that limits the advancement of supervised learning (SL) techniques for diagnostic purposes. As the currently most popular technique in ML-guided diagnostics, SL requires data with high-quality labels to train a classifier that is subsequently tested on previously unseen data and evaluated based on its hit-rate to accurately predict labels in a test set that is withheld from training. The major obstacle in this setting is the disparity between overall available data and available data with labels. The latter is the essential prerequisite for supervised learning, however, obtaining a sufficiently large set of labeled data is time- and cost-intensive, especially in highly specialized domains as cancer diagnostics. The discrepancy between an increasing number of cancer patients in an aging society and the receding physician workforce as well as the correspondingly ever-growing workload of radiologists, pathologists and oncologists poses a further constraint on the labeling process as their experience and knowledge is needed to provide high-quality labels. Still, time and resources for the generation of such large-scale labeled data sets is often missing (1, 2). Therefore, strategies are needed that leverage the overall amount of available data while imposing manageable needs for labeling.

Conceptually, Semi-Supervised Learning (SSL) can be positioned at midway between Unsupervised Learning (UL), where no labels are provided and algorithms deconstruct patterns from unlabeled data e. g. for cluster analysis, and SL, where a classifier is trained on labeled data to correctly map labels to unseen data from the same distribution (3). Hence, SSL offers the opportunity to leverage the vast amounts of unlabeled medical data that are acquired in clinical routine to boost classification performance in a diagnostic setting without the need for fully-labeled extensive data sets. Nevertheless, there are critical assumptions for SSL to function properly and models have to be conceptualized and developed with diligence in order to actually provide a performance boost compared to SL models.

In this review, we aim to provide medical professionals with an outline of key concepts of SSL and how to apply it to medical data with a focus on oncology. First, we introduce main functionalities of SSL and delineate it from SL and UL. Subsequently, we provide an overview of SSL techniques applied to cancer diagnostics and care differentiating between image-based and non-image-based use-cases. Finally, we discuss pitfalls in SSL research design for medical applications and provide an outlook on future prospects.

## What is semi-supervised learning?

The key concept to delineate SL, SSL and UL is the labeling process as well as whether at all and if so, how labeled data is being processed. Labeling refers to the process of attaching meaningful information for classification to raw data. One way to do this is to have experts, e. g. medical doctors, evaluate the raw data, e. g. medical images (4). For example, whole-slide images (WSI) of tumor tissue can be labeled by pathologists or chest CAT scans for potentially malignant lesions can be labeled by radiologists. Alternatively in SSL, a limited number of labels can be used to self-train an algorithm iteratively to attach labels to unlabeled raw data and subsequently train a classifier on these self-labeled data (5). Conceptually, these labeled data provide the basis for training SL algorithms (training stage) that are subsequently supposed to apply previously learned patterns to unseen data and assign correct labels (testing stage, **Figure 1A**) (6). UL on the other hand does not use any labeled data at all. In UL, unlabeled data is sorted according to inherent patterns that delineate different clusters (7), e. g. UL can identify patient clusters with co-occurring genetic variants (**Figure 1B**). SSL uses both labeled and unlabeled data in the sense that labeled data are used to train a classifier for a given use-case and the addition of unlabeled data is intended to leverage information gain and thus boost classification performance (**Figure 1C**) (8). It is therefore advantageous when a large dataset is available for which only a limited number of labels can be obtained, i. e. due to time or cost constraints as is usually the case for medical data.

While the addition of unlabeled data can be advantageous, it can also cause issues with model performance leading to stagnation or even degradation if crucial assumptions of SSL design are not met (9). For SSL models to work robustly, it is necessary that the unlabeled data should contain information that is relevant for label prediction. Therefore, it is crucial that both labeled and unlabeled data follow the same distribution (10). For example, if a classifier is trained on labeled histopathological images of colorectal cancer, the unlabeled data should ideally encompass the same tumor entity, same staining procedure and same magnification. Hence, the algorithm can infer that two samples that are close to each other at the input level (according to their features) should also be close to each other at the output level, i. e. should receive the same labels (smoothness assumption) (8). If these high-dimensional data points at the input level are mapped to a lower dimension in Euclidean space, they are usually clustered along low-dimensional structures, so-called manifolds. Data points that lie on the same manifold should therefore be of the same class (8). If both previous assumptions – inputs with similar feature vectors will be close to each other in an *n*-dimensional feature space and be located on the same manifold if mapped to a lower dimensional space - are true, the decision boundary for a classifier should then lie in an area with low density, i. e. where data points are separate and of different classes (8). Thus, the inclusion of unlabeled data (as long as it is from the same distribution as labeled data) can improve the designation of the decision boundary and therefore boost classification performance (**Figure 2**).

As is the case for most machine learning applications, there is no 'one-size fits all' approach and different methods and algorithms have to be evaluated for any given use-case. What further complicates model selection in SSL is a non-standardized taxonomy of methodologies which makes it harder to reproduce techniques proposed in the literature. Van Engelen et al. (3)

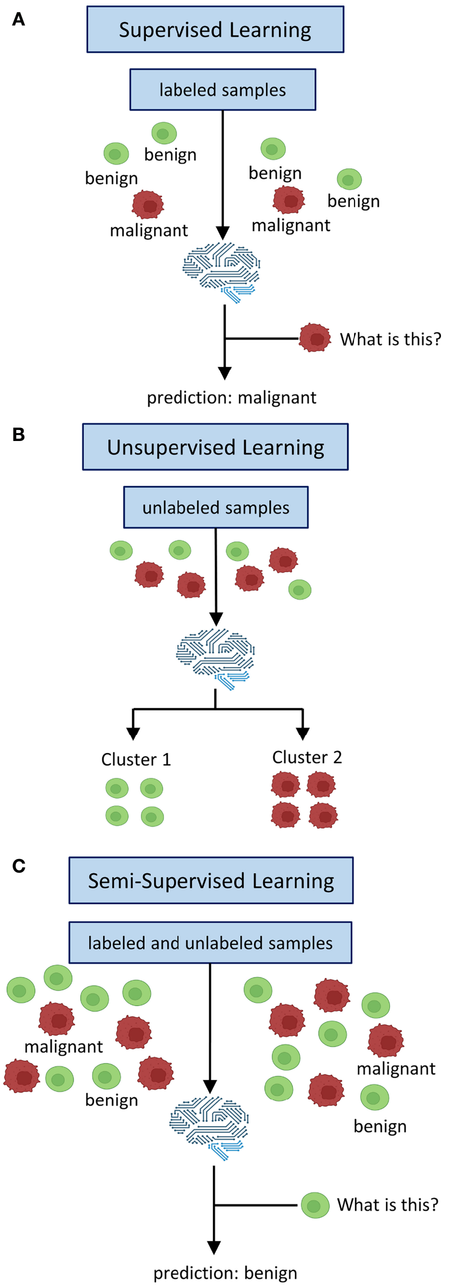


FIGURE 1
Inputs and Outputs of supervised, unsupervised and semi-supervised learning. In supervised learning **(A)** all data is labeled. Labels are used to train a classifier to map learned labels to previously unseen data. Unsupervised learning **(B)** does not use labels. Data is being clustered into groups based on inherent patterns. Semi-supervised learning **(C)** uses both labeled and unlabeled data. Labels are used to train a classifier which is augmented by unlabeled data of the same distribution to derive additional information in order to boost performance.

recently proposed a taxonomy based on the distinction of inductive or transductive methods. The former encompass methods such as clustering with subsequent label assignment, pseudo-labeling or self- and co-training, i. e. methods that assign labels to unseen data and thus can potentially generalize, and the latter include graph-based methods that transfer information along connections of dataset-specific graphs only including data points in a given sample which then cannot be generalized to other data outside the specific sample (3). As for medical applications, the development of robust generalizable

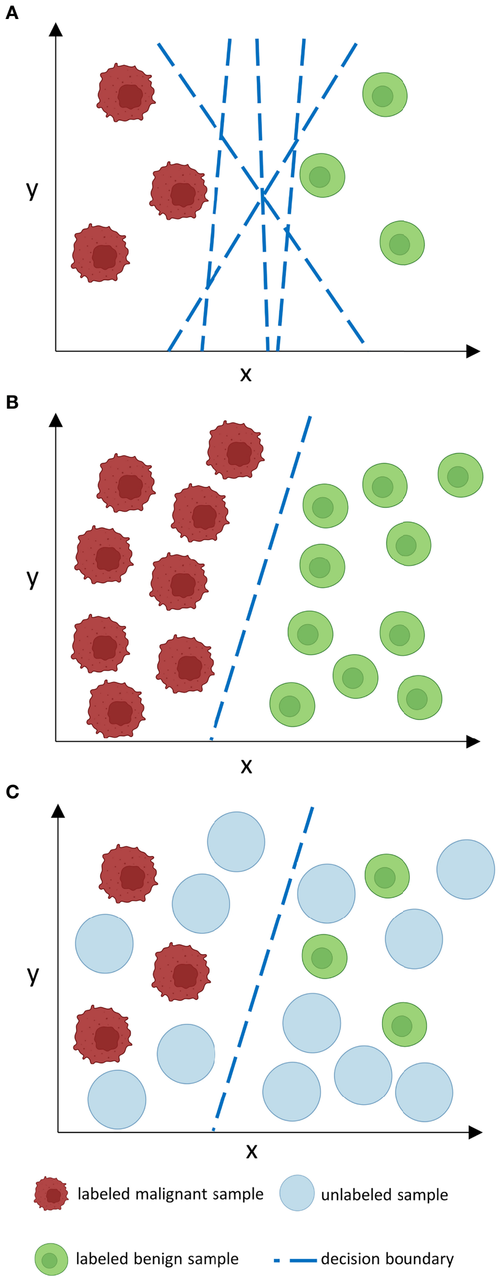


FIGURE 2
How does unlabeled data boost classification performance? Consider a number of features *n* at the input level which corresponds to an *n*-dimensional feature space. In such an n-dimensional coordinate system, every input is located according to its feature vector given by its *n* features and can thus be sorted by similarities and differences in relation to other inputs which is represented by proximity or distance points in the feature space. For clarity reasons, we only consider two features (x, y) in a two-dimensional feature space. When labeled data is sparse **(A)**, as is often the case in medical data sets, the decision boundary of a classifier is less constraint. This may lead to inaccuracies and poor generalization on external data. If many labels are given, the decision boundary is more constraint and thus a more accurate classifier is given that can potentially generalize better. However, manual labeling of such large data sets is often time- and cost-ineffective. Unlabeled data is often available in abundance **(C)** and can be used to constrain the decision boundary of a classifier in a way as large labeled data sets could do, however, without the need for excessive labeling. The decision boundary then lies in an area with low density. Nevertheless, as can be derived from **(B)** and **(C)**, the performance gap between supervised and semi-supervised learning shrinks as the amount of labeled data grows if no further unlabeled samples are provided.

algorithms is desirable for utilization in clinical practice and hence most applied techniques in cancer diagnostics should be developed as inductive methods.

# Studies on semi-supervised learning in cancer diagnostics

Research efforts in applying SSL for diagnostics and care in oncology can broadly be divided data-wise by usage of images or non-image data for model development. Naturally, image-based use-cases most frequently stem from the fields of histopathology, radiology and radiotherapy, while non-image-based applications most frequently include genetic data.

## Image-based semi-supervised learning for cancer detection

### Histopathology

In histopathology as a use-case, classification tasks using computer vision have to be divided into patch- or image-level diagnosis, i. e. whether areas with suspected malignancies should be distinguished from normal surrounding tissue or whether the sample as a whole should be labeled 'malignant' if any sign of neoplastic tissue is present. Importantly for model building, patch-level classification requires image segmentation *a priori* to classification, i. e. different areas of the sample have to be discriminated according to e. g. shapes, patterns and colors. Using a multi-center dataset of > 13.000 colorectal cancer WSI, Yu et al. (11) developed a mean teacher model to detect malignant patches that achieves a comparable area under the curve (AUC) compared to a multi-pathologist benchmark. They report a substantial improvement of SSL over SL when only a limited number of labels is available also validating their model on lung cancer and lymph node samples, but add that with a fully labeled set (with well above 10.000 labels) no difference between SSL and SL was detected. Similarly, Shaw et al. (12) deploy a student-teacher chain model where an iterative process of training a student model that subsequently becomes the teacher model for the following student and so on allows to utilize only 0.5% labeled data to detect colorectal adenocarcinoma from WSI. Wenger et al. (13) utilized consistency regularization and self-ensembling in order to detect and grade bladder cancer samples and report a 19% higher accuracy over baseline SL using only 3% labeled data. Jaiswal et al. (14) compared pre-trained models in detecting neoplastic infiltration of lymph node WSI and reported a high risk of overfitting after short training epochs which was tackled using ensemble learning. Addressing the challenge of variation within classes and similarities between classes, Su et al. (15) propose association cycle consistency loss and maximal conditional association to optimize the loss function reporting improved performance over learning by association on breast cancer histopathological images. Comparing SL and SSL, Al Azzam et al. (16) report similar accuracies for SSL when using only half the number of labels needed for SL in breast cancer prediction from fine needle aspirates. To grade breast cancer samples, Das et al. (17) employ a Generative Adversarial Network (GAN) where the discriminator uses an unsupervised model that is stacked over a supervised model with shared parameters to utilize both labeled and unlabeled samples. An Auxiliary Classifier GAN that divides lung cancer samples into malignant and benign patches which allows for subsequent pixel-based PD-L1 scoring is reported by Kapil et al. (18) for non-small cell lung cancer tissue needle aspirates. Both Marini et al. (19) and Li et al. (20) address the challenge of Gleason scoring prostate cancer samples. The former use a teacher-student approach with different combinations of a pseudo-labeling teacher training a student model utilizing both SSL and semi-weakly supervised learning that are compared to a student-only baseline (19). The latter use a pixel-based approach on prostate WSI with expectation maximization by a fully convolutional encoder-decoder net incorporating both internally annotated and external weakly annotated image data compared to a model trained on a fully labeled dataset alone (20). Both report performance improvements for the SSL methods using additional un- or weakly-labeled data. Lastly, to detect melanoma, Masood et al. (21) train deep belief networks in parallel to support vector machines that are supposed to counteract misclassified data with adjusted weights and finally compare their model to several SL-based models and report superior performance for their SSL-based approach. **Table 1** provides an overview of recent studies that use SSL in histopathology.

### Radiology and radiotherapy

The detection of lung nodules in computer-assisted tomography (CAT) scans is a common theme in SSL-based research in radiology. Khosravan et al. (22) use a multi-tasking CNN to concomitantly learn nodule segmentation and false positive nodule reduction on chest CAT scans incorporating SSL to accommodate for unlabeled data in the segmentation process and report high accuracies compared to baseline. Xie et al. (23) address the task of differentiating between benign and malignant nodules using a semi-supervised adversarial model with an autoencoder unsupervised reconstruction net, learnable transition layers, and a supervised classification net and report high accuracies on a benchmark dataset for lung nodule classification. Using a similarity metric function to iteratively include unlabeled samples *via* SSL, Shi et al. (24) use a transfer learning approach with a pre-trained network that differentiates between nodules and nodule-like tissue to identify lung nodules and report high accuracies in their initial dataset, but

TABLE 1 Overview of Studies on Semi-Supervised Learning in Histopathology.

| Authors and Reference | Entity | Objective | Technique | Publicly Available Code |
|---|---|---|---|---|
| Yu et al. (11) | colorectal and lung cancer as well as lymph nodes | detecting malignant patches in WSI | mean teacher | yes |
| Shaw et al. (12) | colorectal cancer | detecting malignant patches in WSI | student-teacher-chain | no |
| Wenger et al. (13) | bladder cancer | detection and grading | consistency regularization and self-ensembling | no |
| Jaiswal et al. (14) | metastasized tumors | detecting metastases in lymph node WSI | pseudo-labeling | no |
| Su et al. (15) | breast cancer | detecting malignant samples in WSI | combination of association cycle consistency loss and maximal conditional association loss | no |
| Das et al. (17) | breast cancer | grading samples | stacked semi-supervised GAN | no |
| Al Azzam et al. (16) | breast cancer | cancer detection from nuclei morphologies | comparison of 9 SL and SSL classifiers | no |
| Kapil et al. (18) | lung cancer | PD-L1 scoring | auxiliary classifier GAN and pixel-based quantification | no |
| Marini et al. (19) | prostate cancer | Gleason scoring | teacher-student chain and pseudo-labeling | yes |
| Li et al. (20) | prostate cancer | Gleason scoring | expectation maximization-based fully convolutional encoder-decoder network | no |
| Masood et al. (21) | melanoma | detecting malignant samples | Co-training of Deep Belief Network and advised SVM | no |

GAN, generative adversarial networks; SL, supervised learning; SLL, semi-supervised learning; SVM, support vector machines; WSI, whole-slide-images.

acknowledge performance drops in an independent validation set. For breast cancer detection in mammogram images, both Sun et al. (25) and Azary et al. (26) use a co-training approach. In the former study, a three-step method of adjusting weights, selecting features and co-training-based labeling is proposed and a 7.4% performance gain for the combination of labeled and unlabeled data compared to labeled data only is reported (25). The latter study incorporates SSL in pixel-based tumor segmentation and proposes co-training with support vector machines and Bayesian classifiers (26). Using breast ultrasound images for tumor detection in a joint dataset of many weakly and few strongly annotated images, Shin et al. (27) propose a self-training method and report similar accuracies for only ten strongly annotated images joined by a large number of weakly annotated ones compared to 800 strongly annotated images only. Wodzinski et al. (28) aim to identify target volumes for postoperative tumor bed irradiation in breast cancer using a semi-supervised volume penalty *via* a multi-level encoder decoder architecture and report a decrease in target registration error and tumor volume ratio. For brain tumor detection, Ge et al. (29), Chen et al. (30), and Meier et al. (31) investigate brain magnetic resonance imaging (MRI) scans. Ge et al. (29) utilize a graph-based approach to create pseudo-labels and accommodate for moderate-sized data sets by generating additional images with GANs. They use their model for glioma grading and IDH-mutation status prediction (29). In a step-wise approach, Chen et al. (30) deploy a student-teacher-based model and extract hierarchical features using an adversarial network to detect lesions in brain MRI scans that correspond to either multiple sclerosis, ischemic stroke or tumor tissue. In a pre- and postoperative comparative setting, Meier et al. (31) investigate residual tumor tissue in brain MRI scans of ten high-grade glioma patients with semi-supervised decision forest and report improved performance and computation time compared to conventional segmentation methods. Lastly, Turk et al. (32) address thyroid cancer detection in ultrasound texture data with linked clinical scoring systems as additional features using an autoencoder-based model and report a high sensitivity despite their imbalanced dataset by using synthetic minority oversampling. **Table 2** provides an overview of studies using SSL in radiology or radiotherapy.

## Non-image-based semi-supervised learning for cancer management

While the aforementioned image-based studies primarily focus on detection of cancer, research efforts of SSL in oncology that do not use images predominantly address the task of estimating survival, predicting relapse and identifying genetic subtypes. Examining gene expression data from patients with breast, lung, gastric and liver cancer, Chai et al. (33) use a semi-supervised self-paced learning framework with Cox proportional hazard and accelerated failure time models to classify cancer patients and predict censored data thereby reporting improved separation of survival curves for their model compared to baseline supervised models. Also using gene expression data but in the context of colorectal and breast cancer, Shi et al. (34) predict recurrence *via* low density separation. They report increasing accuracies for SSL over

TABLE 2 Overview of Studies on Semi-Supervised Learning in Radiology and Radiotherapy.

| Authors and Reference | Entity | Objective | Technique | Publicly Available Code |
|---|---|---|---|---|
| Khosravan et al. (22) | lung cancer | detecting malignant nodules in chest CAT scans | SSL-based multi-task network | no |
| Xie et al. (23) | lung cancer | detecting malignant nodules in chest CAT scans | semi-supervised adversarial autoencoders, learnable transition layers, and supervised classification | no |
| Shi et al. (24) | lung cancer | detecting malignant nodules in chest CAT scans | transfer learning and semi-supervised feature matching | no |
| Sun et al. (25) | breast cancer | detecting breast cancer in mammogram images | co-training | no |
| Azary et al. (26) | breast cancer | detecting breast cancer in mammogram images | co-training | no |
| Shin et al. (27) | breast cancer | detecting breast cancer in ultrasound images | joint weakly- and strongly-supervised framework and self-training | yes |
| Wodzinski et al. (28) | breast cancer | identifying target volumes for radiotherapy | semi-supervised multilevel encoder-decoder | yes |
| Ge et al. (29) | brain tumor | glioma grading and IDH-mutation prediction in MRI scans | GAN-augmented networks in a graph-based framework | no |
| Chen et al. (30) | brain tumor, multiple sclerosis, ischemic stroke | detecting pathological samples in MRI scans | student-teacher chain combined with adversarial learning | yes |
| Meier et al. (31) | brain tumor | detecting residual tumor tissue in postoperative brain MRI | semi-supervised decision forest | no |
| Turk et al. (32) | thyroid cancer | detecting thyroid cancer from ultrasound textures and clinical scoring systems | autoencoders and synthetic minority oversampling | no |

CAT, computer-assisted tomography; GAN, generative adversarial networks; MRI, magnetic resonance imaging.

baseline SL classifiers with increasing numbers of unlabeled data. Addressing the same task in the same tumor entities, Park et al. (35) resort to a semi-supervised graph regularization algorithm to identify functionally similar gene pairs and thereby predict recurrence in breast and colorectal cancer gene expression data including labeled and unlabeled nodes. Hassanzadeh et al. (36) designed an ensemble model based on decision trees and boosting to predict survival for patients harboring kidney, ovarian, or pancreatic cancer for whom only incomplete clinical data was available and report improved accuracy for SSL compared to SL baselines. Cristovao et al. (37) compared SL and SSL in subtyping breast cancer using multi-omic data, however, did not find any performance improvements when comparing SSL to baseline logistic regression. Also investigating multi-omics data, Ma et al. (38) developed affinity fusion networks to cluster patients based on their specific omics profile into lung, kidney, uterus or adrenal gland cancer groups. The authors report a high predictive accuracy with training on less than one percent of labeled data. Sherafat et al. (39) developed a positive-unlabeled learning model using auto machine learning to predict tumor-rejection mediation neoepitopes from exome sequencing data in ovarian cancer. The authors report improved performance over model-based classifiers for somatic variant calling and peptide identification. Both Camargo et al. (40) and Livieris et al. (41) propose novel active learning models that are tested on either data of acute myeloid leukemia, E. coli, and plant leaves, or breast and lung cancer, respectively. In both studies, the authors report higher accuracies for their respective models, root distance boundary sampling (40) and improved CST voting (41), compared to both SSL and SL classifiers. **Table 3** summarizes non-image-based applications of SSL with relevance to cancer detection and management.

## Discussion

SSL represents a viable approach to the dilemma of big data in cancer medicine, especially in the context of image data which is usually acquired in abundance during clinical routine work-ups, but adequate labeling by medical experts is often time consuming and thus cost-ineffective. The main goal of SSL in this context is to achieve classification performances that surpass those of SL alone when labeled data is limited and at the same time abundant unlabeled data is available. Crucially, SSL models have to satisfy the above-mentioned assumptions: i) both labeled and unlabeled data have to be drawn from the same distribution, ii) similarity of data on the input level results in similarity of data at the output level (smoothness), iii) hence data points on the same low-dimensional structures (manifolds) receive the same labels and thus, iv) the decision boundary runs through an area of low density, i.e. where data points are separated and of

TABLE 3 Overview of Studies on Semi-Supervised Learning using non-image-based data.

| Authors and Reference | Entity | Objective | Technique | Publicly Available Code |
|---|---|---|---|---|
| Chai et al. (33) | breast, lung, gastric and liver cancer | predicting survival | self-paced learning with Cox proportional hazard and accelerated failure time models | no |
| Shi et al. (34) | colorectal and breast cancer | predicting relapse | low density separation | no |
| Park et al. (35) | colorectal and breast cancer | predicting relapse | graph-based regularization | no |
| Hassanzadeh et al. (36) | kidney, ovarian and pancreatic cancer | predicting survival | ensemble learning with robust boost and decision trees | no |
| Cristovao et al. (37) | breast cancer | subtyping, model comparison | comparison of different SL and SSL algorithms | no |
| Ma et al. (38) | lung, kidney, uterus and adrenal gland cancer | predicting primary tumor site | Affinity Network Fusion | yes |
| Sherafat et al. (39) | ovarian cancer | predicting tumor-rejection mediating neoepitopes | Positive-unlabeled Learning using Auto-ML | no |
| Camargo et al. (40) | acute myeloid leukemia, E. coli, plant leaves | model comparison | root distance boundary sampling | yes |
| Livieris et al. (41) | breast and lung cancer | model comparison | self- and co-training with ensemble learning | no |

different classes. Divergence from these key assumptions can not only lead to performance stagnation, but also degradation as unlabeled data is handled as noise that blurs information abstraction of the classifier (42). Importantly, this is what delineates SSL from transfer learning, where a classifier is first trained on one use-case and subsequently transferred to another similar use-case where it is supposed to perform a similar task (43), e. g. a classifier trained by identifying alteration A in immunohistochemistry on WSI in a supervised setting could potentially be transferred to also identify alteration B if staining is similar. Therefore, the most important question before conducting SSL experiments is whether labeled and unlabeled data are actually from the same distribution and if so whether an inclusion of the unlabeled samples might lead to a performance gain over baseline SL.

Several of the above-mentioned studies reported substantial performance gains for SSL as long as the model was short on labeled data, however, when the amount of labeled data was increased or only labeled data was used the gap between SSL and SL performance shrunk. However, the frequent lack of a comparison between baseline SL and SSL classifiers further complicates the evaluation of such studies and only few studies do report baseline comparisons (11, 13, 19, 22, 33, 37) and still even fewer report equal tuning of hyperparameters (11, 19) for SSL and SL classifiers to make results comparable. When it comes to model design, it is essential to note that different algorithms may perform differently with regard to different tasks (9). While this sounds obvious, it is still the case that often only the use of a single algorithm is reported which either may be due to a lack of comparative testing or due to publication bias as only the successful algorithm is selected for a given manuscript. However, to evaluate suitable model designs for different tasks, we advocate for a full report on tested algorithms ideally including a comparison between different SSL model set-ups, their SL baseline, adequate hyperparameter tuning for both SSL and SL, and the models' individual performance in comparison. Further, varying the amount of labeled and unlabeled data for both training and testing sets seems warranted to find the equilibrium of optimal performance for different tasks in future studies of SSL in oncology. The lack of reproducibility in research on artificial intelligence in general (44) is also likely to be a future issue in biomedical use-cases of SSL as unfortunately only a minority of studies provide publicly accessible code to support their results (11, 19, 27, 28, 30, 38, 40). As is evident from previous studies on SSL in oncology, use cases mainly include tumor entities with high prevalence such as breast (15–17, 25–28, 33–35, 37, 41), lung (18, 22, 23, 33, 34, 38, 41), and colorectal cancer (11, 12, 34, 35) where single centers can amass sufficiently sized data sets to conduct SSL experiments. This is also reflected in the overwhelming absence of studies on SSL in hematology with only one single study (40) including any hematological neoplasm at all. Therefore, data-sharing is crucial in order to expand use-cases to rare tumor entities. Slight differences between centers in how training data is handled – e.g. differences in imaging devices used and thus consecutive differences in image format, shape, contrast, resolution and brightness – may also influence individual models. A model trained solely on single center image data may therefore significantly drop in performance if it is introduced to data of another source. Hence, pooling heterogenous data of different sources for initial model training is useful in order to obtain classifiers that can be widely generalized beyond in-house use for single institutions. Not only may the crowd-sourcing of research in biomedical SSL vastly enlarge the pool of unlabeled (and possibly labeled) data, but it may also help identify and modify promising models for multi-center prospective validation. The latter is another key shortcoming of previous studies that were often confined to single centers and retrospective evaluation. Thus, publicly available code, data-sharing for both labeled and unlabeled data and prospective collaborative research efforts will be key to evaluate models for future clinical applicability. Shared data and models may then also enable the evaluation of a variety of tumor entities in the same diagnostic modality, i. e. differential diagnosis of tumor entities in histopathological WSI.

This, however, leads to a frequent problem of artificial intelligence in general that is even more pronounced in the sensitive context of oncology where diagnostic accuracy is essential to provide high quality care to patients with life-threatening diseases: explainability of ML models. ML and especially deep learning has often been referred to as a 'black box' (45) and the path of decision making within a model is hard to interpret. While this is already a key issue in SL, SSL adds to the confusion as information is also derived from unlabeled samples. The apparent lack of interpretability when it comes to clinical validation of model outputs stresses the urgent need to incorporate mechanisms of explainability into SSL models that make outputs or even intermediate steps such as label assignment on unlabeled samples traceable for clinical experts. The virtual lack thereof in previous studies signals a discrepancy between what is technologically possible and what is clinically acceptable for routine use as 'black box' models will likely have it harder to be included in routine clinical workflows due to a lack of acceptance in diagnostic specialties and ethical concerns in cancer management (46). Still, given large unlabeled data sets that often are routinely acquired in cancer diagnostics combined with the trend of a shrinking physician workforce that is occupied with complex tasks that have to be performed in increasingly shorter periods of time (1), SSL provides a low-cost and potentially high-benefit solution to develop clinically meaningful ML models for diagnostic tasks in oncology.

## Conclusion

While SSL provides a possible solution to the vast discrepancy between available labeled and unlabeled data in cancer diagnostics, it should not be considered a silver bullet in the development of accurate classifiers for cancer detection. Adequate selection of labeled and unlabeled data of the same distribution as well as comparisons to baseline SL, among others, are crucial to build robust SSL models. While previous research efforts of SSL in oncology have mainly comprised retrospective single-center studies, future research is warranted in multi-center prospective model evaluation to design robust and explainable classifiers for implementation in the clinical routine of cancer diagnostics.

## Author contributions

J-NE performed the literature search and wrote the initial draft. All authors provided critical scientific insights, reviewed and edited the draft and approved its final version for submission. All authors agree to be accountable on the contents of the work. All authors contributed to the article and approved the submitted version.

## Funding

J-NE is grateful for a research scholarship from the Mildred-Scheel-Nachwuchszentrum Dresden (German Cancer Aid). The funder had no role in the design and conduct of the study, analysis, and interpretation of the data, preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

## 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

## References

1. Zhang X, Lin D, Pforsich H, Lin VW. Physician workforce in the united states of America: forecasting nationwide shortages. *Hum Resour Health* (2020) 18:8. doi: 10.1186/s12960-020-0448-3

2. Metter DM, Colgan TJ, Leung ST, Timmons CF, Park JY. Trends in the US and Canadian pathologist workforces from 2007 to 2017. *JAMA Netw Open* (2019) 2:e194337. doi: 10.1001/jamanetworkopen.2019.4337

3. van Engelen JE, Hoos HH. A survey on semi-supervised learning. *Mach Learn* (2020) 109:373–440. doi: 10.1007/s10994-019-05855-6

4. Willemink MJ, Koszek WA, Hardell C, Wu J, Fleischmann D, Harvey H, et al. Preparing medical imaging data for machine learning. *Radiology* (2020) 295:4–15. doi: 10.1148/radiol.2020192224

5. Triguero I, García S, Herrera F. Self-labeled techniques for semi-supervised learning: taxonomy, software and empirical study. *Knowl Inf Syst* (2015) 42:245–84. doi: 10.1007/s10115-013-0706-y

6. Cunningham P, Cord M, Delany SJ. Supervised learning. In: M Cord and P Cunningham, editors. *Machine learning techniques for multimedia: case studies on organization and retrieval. cognitive technologies*. Berlin, heidelberg: springer (2008). p. P.21–49. doi: 10.1007/978-3-540-75171-7_2

7. Barlow HB. Unsupervised learning. *Neural Comput* (1989) 1:295–311. doi: 10.1162/neco.1989.1.3.295

8. O Chapelle, B Schölkopf and A Zien eds. *Semi-supervised learning.* Cambridge, MA, USA: MIT Press (2006). 528 p.

9. Oliver A, Odena A, Raffel C, Cubuk ED, Goodfellow IJ. Realistic evaluation of deep semi-supervised learning algorithms (2019) (Accessed March 9, 2022).

10. Zhu X, Goldberg AB. Introduction to semi-supervised learning. *Synthesis Lectures Artif Intell Mach Learn* (2009) 3:1–130. doi: 10.2200/S00196ED1V01Y200906AIM006

11. Yu G, Sun K, Xu C, Shi X-H, Wu C, Xie T, et al. Accurate recognition of colorectal cancer with semi-supervised deep learning on pathological images. *Nat Commun* (2021) 12:6311. doi: 10.1038/s41467-021-26643-8

12. Shaw S, Pajak M, Lisowska A, Tsaftaris SA, O'Neil AQ. Teacher-student chain for efficient semi-supervised histology image classification (Accessed February 22, 2022).

13. Wenger K, Tirdad K, Dela Cruz A, Mari A, Basheer M, Kuk C, et al. A semi-supervised learning approach for bladder cancer grading. *Mach Learn Appl* (2022) 9:100347. doi: 10.1016/j.mlwa.2022.100347

14. Jaiswal AK, Panshin I, Shulkin D, Aneja N, Abramov S. Semi-supervised learning for cancer detection of lymph node metastases (2019) (Accessed February 22, 2022).

15. Su L, Liu Y, Wang M, Li A. Semi-HIC: A novel semi-supervised deep learning method for histopathological image classification. *Comput Biol Med* (2021) 137:104788. doi: 10.1016/j.compbiomed.2021.104788

16. Al-Azzam N, Shatnawi I. Comparing supervised and semi-supervised machine learning models on diagnosing breast cancer. *Ann Med Surg* (2021) 62:53–64. doi: 10.1016/j.amsu.2020.12.043

17. Das A, Mishra S, Mishra DK, Gopalan SS. Machine learning to predict 5-year survival among pediatric acute myeloid leukemia patients and development of OSPAM-c online survival prediction tool. *medRxiv* (2020) 2020:4. doi: 10.1101/2020.04.16.20068221

18. Kapil A, Meier A, Zuraw A, Steele KE, Rebelatto MC, Schmidt G, et al. Deep semi supervised generative learning for automated tumor proportion scoring on nsclc tissue needle biopsies. *Sci Rep* (2018) 8:17343. doi: 10.1038/s41598-018-35501-5

19. Marini N, Otálora S, Müller H, Atzori M. Semi-supervised training of deep convolutional neural networks with heterogeneous data and few local annotations: an experiment on prostate histopathology image classification. *Med Image Anal* (2021) 73:102165. doi: 10.1016/j.media.2021.102165

20. Li J, Speier W, Ho KC, Sarma KV, Gertych A, Knudsen BS, et al. An EM-based semi-supervised deep learning approach for semantic segmentation of histopathological images from radical prostatectomies. *Comput Med Imaging Graph* (2018) 69:125–33. doi: 10.1016/j.compmedimag.2018.08.003

21. Masood A, Al-Jumaily A. Semi-advised learning model for skin cancer diagnosis based on histopathalogical images. *Annu Int Conf IEEE Eng Med Biol Soc* (2016) 2016:631–4. doi: 10.1109/EMBC.2016.7590781

22. Khosravan N, Bagci U. Semi-supervised multi-task learning for lung cancer diagnosis. *Annu Int Conf IEEE Eng Med Biol Soc* (2018) 2018:710–3. doi: 10.1109/EMBC.2018.8512294

23. Xie Y, Zhang J, Xia Y. Semi-supervised adversarial model for benign-malignant lung nodule classification on chest CT. *Med Image Anal* (2019) 57:237–48. doi: 10.1016/j.media.2019.07.004

24. Shi F, Chen B, Cao Q, Wei Y, Zhou Q, Zhang R, et al. Semi-supervised deep transfer learning for benign-malignant diagnosis of pulmonary nodules in chest ct images. *IEEE Trans Med Imaging* (2021) 41(4):771–81. doi: 10.1109/TMI.2021.3123572

25. Sun W, Tseng T-LB, Zhang J, Qian W. Computerized breast cancer analysis system using three stage semi-supervised learning method. *Comput Methods Programs BioMed* (2016) 135:77–88. doi: 10.1016/j.cmpb.2016.07.017

26. Azary H, Abdoos M. A semi-supervised method for tumor segmentation in mammogram images. *J Med Signals Sens* (2020) 10:12–8. doi: 10.4103/jmss.JMSS_62_18

27. Shin YS, Lee S, Yun IlD, Kim SM, Lee KM. Joint weakly and semi-supervised deep learning for localization and classification of masses in breast ultrasound images. *IEEE Trans Med Imaging* (2019) 38:762–74. doi: 10.1109/TMI.2018.2872031

28. Wodzinski M, Ciepiela I, Kuszewski T, Kedzierawski P, Skalski A. Semi-supervised deep learning-based image registration method with volume penalty for real-time breast tumor bed localization. *Sensors (Basel)* (2021) 21:4085. doi: 10.3390/s21124085

29. Ge C, Gu IY-H, Jakola AS, Yang J. Deep semi-supervised learning for brain tumor classification. *BMC Med Imaging* (2020) 20:87. doi: 10.1186/s12880-020-00485-0

30. Chen G, Ru J, Zhou Y, Rekik I, Pan Z, Liu X, et al. MTANS: Multi-scale mean teacher combined adversarial network with shape-aware embedding for semi-supervised brain lesion segmentation. *Neuroimage* (2021) 244:118568. doi: 10.1016/j.neuroimage.2021.118568

31. Meier R, Bauer S, Slotboom J, Wiest R, Reyes M. Patient-specific semi-supervised learning for postoperative brain tumor segmentation. *Med Image Comput Comput Assist Interv* (2014) 17:714–21. doi: 10.1007/978-3-319-10404-1_89

32. Turk G, Ozdemir M, Zeydan R, Turk Y, Bilgin Z, Zeydan E. On the identification of thyroid nodules using semi-supervised deep learning. *Int J Numer Method BioMed Eng* (2021) 37:e3433. doi: 10.1002/cnm.3433

33. Chai H, Li Z, Meng D, Xia L, Liang Y. A new semi-supervised learning model combined with cox and sp-aft models in cancer survival analysis. *Sci Rep* (2017) 7:13053. doi: 10.1038/s41598-017-13133-5

34. Shi M, Zhang B. Semi-supervised learning improves gene expression-based prediction of cancer recurrence. *Bioinformatics* (2011) 27:3017–23. doi: 10.1093/bioinformatics/btr502

35. Park C, Ahn J, Kim H, Park S. Integrative gene network construction to analyze cancer recurrence using semi-supervised learning. *PLoS One* (2014) 9: e86309. doi: 10.1371/journal.pone.0086309

36. Hassanzadeh HR, Phan JH, Wang MD. A semi-supervised method for predicting cancer survival using incomplete clinical data. *Annu Int Conf IEEE Eng Med Biol Soc* (2015) 2015:210–3. doi: 10.1109/EMBC.2015.7318337

37. Cristovao F, Cascianelli S, Canakoglu A, Carman M, Nanni L, Pinoli P, et al. Investigating deep learning based breast cancer subtyping using pan-cancer and multi-omic data. *IEEE/ACM Trans Comput Biol Bioinform* (2022) 19:121–34. doi: 10.1109/TCBB.2020.3042309

38. Ma T, Zhang A. Affinity network fusion and semi-supervised learning for cancer patient clustering. *Methods* (2018) 145:16–24. doi: 10.1016/j.ymeth.2018.05.020

39. Sherafat E, Force J, Măndoiu II. Semi-supervised learning for somatic variant calling and peptide identification in personalized cancer immunotherapy. *BMC Bioinf* (2020) 21:498. doi: 10.1186/s12859-020-03813-x

40. Camargo G, Bugatti PH, Saito PTM. Active semi-supervised learning for biological data classification. *PLoS One* (2020) 15:e0237428. doi: 10.1371/journal.pone.0237428

41. Livieris I, Pintelas E, Kanavos A, Pintelas P. An improved self-labeled algorithm for cancer prediction. *Adv Exp Med Biol* (2020) 1194:331–42. doi: 10.1007/978-3-030-32622-7_31

42. Cozman F, Cohen I. Risks of semi-supervised learning: How unlabeled data can degrade performance of generative classifiers. *Semi-Supervised Learning MIT Press* (2006):57–71. doi: 10.7551/mitpress/9780262033589.003.0004

43. Weiss K, Khoshgoftaar TM, Wang D. A survey of transfer learning. *J Big Data* (2016) 3:9. doi: 10.1186/s40537-016-0043-6

44. Hutson M. Artificial intelligence faces reproducibility crisis. *Science* (2018) 359(6377):725–6. doi: 10.1126/science.359.6377.725

45. Castelvecchi D. Can we open the black box of AI? *Nature* (2016) 538:20–3. doi: 10.1038/538020a

46. Grote T, Berens P. On the ethics of algorithmic decision-making in healthcare. *J Med Ethics* (2020) 46:205–11. doi: 10.1136/medethics-2019-105586

frontiers | Frontiers in Oncology

TYPE Review
PUBLISHED 25 July 2022
DOI 10.3389/fonc.2022.953934

OPEN ACCESS

EDITED BY
Oliver Diaz,
University of Barcelona, Spain

REVIEWED BY
Zhaohua Miao,
Hefei University of Technology, China
Bingxia Zhao,
Southern Medical University, China
Shanhui Fan,
Hangzhou Dianzi University, China

*CORRESPONDENCE
Hengyu Zhao
xzzx_paper@163.com
Qingliang Zhao
zhaoql@xmu.edu.cn
Jingsong Mao
maojingsong163@163.com

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

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 26 May 2022
ACCEPTED 29 June 2022
PUBLISHED 25 July 2022

CITATION
Yang L, Chen Y, Ling S, Wang J, Wang G, Zhang B, Zhao H, Zhao Q and Mao J (2022) Research progress on the application of optical coherence tomography in the field of oncology.
*Front. Oncol.* 12:953934.
doi: 10.3389/fonc.2022.953934

# Research progress on the application of optical coherence tomography in the field of oncology

Linhai Yang[1†], Yulun Chen[2†], Shuting Ling[1], Jing Wang[3], Guangxing Wang[1], Bei Zhang[1], Hengyu Zhao[3*], Qingliang Zhao[1*] and Jingsong Mao[1,4*]

[1]State Key Laboratory of Molecular Vaccinology and Molecular Diagnostics, Center for Molecular Imaging and Translational Medicine, Department of Laboratory Medicine, School of Public Health, Shenzhen Research Institute of Xiamen University, Xiamen University, Xiamen, China, [2]School of Medicine, Xiamen University, Xiamen, China, [3]Department of Imaging, School of Medicine, Xiamen Cardiovascular Hospital of Xiamen University, Xiamen University, Xiamen, China, [4]Department of Radiology, Xiamen Key Laboratory of Endocrine-Related Cancer Precision Medicine, Xiang'an Hospital of Xiamen University, Xiamen, China

Optical coherence tomography (OCT) is a non-invasive imaging technique which has become the “gold standard” for diagnosis in the field of ophthalmology. However, in contrast to the eye, nontransparent tissues exhibit a high degree of optical scattering and absorption, resulting in a limited OCT imaging depth. And the progress made in the past decade in OCT technology have made it possible to image nontransparent tissues with high spatial resolution at large (up to 2mm) imaging depth. On the one hand, OCT can be used in a rapid, noninvasive way to detect diseased tissues, organs, blood vessels or glands. On the other hand, it can also identify the optical characteristics of suspicious parts in the early stage of the disease, which is of great significance for the early diagnosis of tumor diseases. Furthermore, OCT imaging has been explored for imaging tumor cells and their dynamics, and for the monitoring of tumor responses to treatments. This review summarizes the recent advances in the OCT area, which application in oncological diagnosis and treatment in different types: (1) superficial tumors:OCT could detect microscopic information on the skin's surface at high resolution and has been demonstrated to help diagnose common skin cancers; (2) gastrointestinal tumors: OCT can be integrated into small probes and catheters to image the structure of the stomach wall, enabling the diagnosis and differentiation of gastrointestinal tumors and inflammation; (3) deep tumors: with the rapid development of OCT imaging technology, it has shown great potential in the diagnosis of deep tumors such in brain tumors, breast cancer, bladder cancer, and lung cancer.

KEYWORDS

**optical coherence tomography, cancer imaging, oncological diseases, imaging technique, tumor diagnoses**

## Introduction

OCT is a noninvasive optical imaging technique, that can capture high-resolution and three-dimensional (3D) images on biological tissues with label-free. Huang et al., in 1991 (1), first proposed a concept of OCT. Through the weak coherent light interferometer theory, images of biological tissues with excellent axial resolution (<10 μm) can be obtained in real-time utilizing near-infrared (NIR) light waves reflected by microstructures within the tissue (2). Based on the above advantages, OCT is widely used in a variety of biomedical fields, including ophthalmology, dentistry, dermatology, oncological and cardiovascular, among others.

In biomedical fields, optical imaging technology is extensively used such as Laser scanning confocal imaging (LSCI), Two-Photon imaging, Fluorescence imaging (FI), Laser speckle imaging (LSI), Laser doppler imaging (LDI) and OCT, etc. Although LSCI and two-photon imaging techniques could provide high spatial resolution images of biological tissues, imaging under aqueous or oily objectives requires contact with tissues, while the relatively small imaging view field and low penetration depth make them impractical for clinical application (3, 4). In terms of LSI, it is also difficult to perform depth-resolved *in vivo* 3D imaging due to depth limitation, even though it can offer high-resolution, non-contact imaging (5). A millimeter- resolution LDI can only monitor microcirculatory vessels (6). FI involves the use of fluorescent materials as imaging labels (7), which may cause adverse effects like allergies. In contract, OCT offers a non-invasive method for imaging tumor tissue at multiple scales, with high contrast and resolution *in vivo*, as well as displaying high endogenous contrast in biological tissues (8). Moreover, OCT allows for deep penetration of tissue up to 2 mm and can rapidly produce 3D images with high temporal resolution (9, 10).

After the development of OCT, it was widely used in ocular imaging for glaucoma (11–13), macular degeneration (14–16), retinal vein obstruction (17–19), diabetic retinal microaneurysm (20, 21), uveitis (22–24), etc. The advances in OCT imaging applications for ophthalmology were detailed in relevant reviews in 2018 (25), 2019 (22), 2020 (26), and 2022 (27). OCT has also contributed to the fields other than ophthalmology due to its miniaturization and integration with catheters and endoscopes (**Figure 1**). Hence, an overview of OCT imaging technology will be given in this article, along with its recent developments in oncological diseases: (1) superficial tumors: OCT could detect microscopic information on the skin's surface at high resolution and has been demonstrated to help diagnose common skin cancers; (2) gastrointestinal tumors: OCT can be integrated into small probes and catheters to image the structure of the stomach wall, enabling the diagnosis and differentiation of gastrointestinal tumors and inflammation; (3) deep tumors: with the rapid development of OCT imaging technology, it has shown great potential in the diagnosis of deep tumors such in brain tumors, breast cancer, bladder cancer and lung cancer. In addition, the possible future development direction of OCT is prospected.

## Development of the OCT

Low coherence interference of light is the basis for OCT, which is similar to ultrasound imaging in principle. Based on the Michelson interferometer, the OCT imaging equipment architecture obtains tissue reflection signals with depth resolution characteristics by detecting an interference signal formed between the reflected light of the reference lens and the backscattered light of the sample objective. By changing the relative position of the reference mirror, the intensity of the backscattered light of tissue can be detected at different depths. In the axial direction, echo sequences with different depths form an intensity distribution (A-scan). Multi-points A-Scan reconstruction results in a two-dimensional (2D) cross-sectional image of the tissue, called B-Scan. And then the 3D structure of the tissue can be obtained by reconstructing the B-Scan at different locations (28, 29).

OCT imaging technology has gone through three generations, namely time-domain OCT (TD-OCT), spectral-domain OCT (SD-OCT), and swept OCT (SS-OCT), with the advancement of laser and computer technology and the optimization of imaging algorithms (**Figure 2**). The first-generation OCT system based on time-domain detection relied on time delay measurement of the reflected signal from tissue relative to the reflected signal from the reference mirror. The optical signal reflected from the target tissue superimposes and interferes with the optical signal reflected by the reference mirror, resulting in the formation of the OCT. This procedure requires mechanically shifting the reference mirror, thereby changing the depth of the tissue being scanned (31, 32). However, with the advancement of technology and technology and for different needs, TD-OCT has emerged many variants, such as line-field confocal OCT (LC-OCT) (33, 34), full-field OCT (FF-OCT) (35), polarization-sensitive OCT (PS-OCT) (36), etc., to achieve more efficient and wide applications in the clinic.

Unlike TD-OCT, the reference mirror of the reference arm is fixed in the second-generation SD-OCT structure. The interference of the optical signal can be achieved by varying the frequency of the light source, and the data acquisition rate can be raised by 45-100 times (37). Simultaneously, SD-OCT can measure the spatial and structural information on all echo delays (axial pixels) by evaluating the interference spectrum between the light signal from the rest-reference mirror and the light signal reflected from biological tissue (38).

Though SD-OCT and SS-OCT use Fourier domain techniques, spectrometers and high-speed line scan cameras are utilized to measure the interferometer spectra of the SD-OCT. On the other hand, SS-OCT detects OCT signals with sweep light sources and photodetectors. Moreover, the SS-OCT

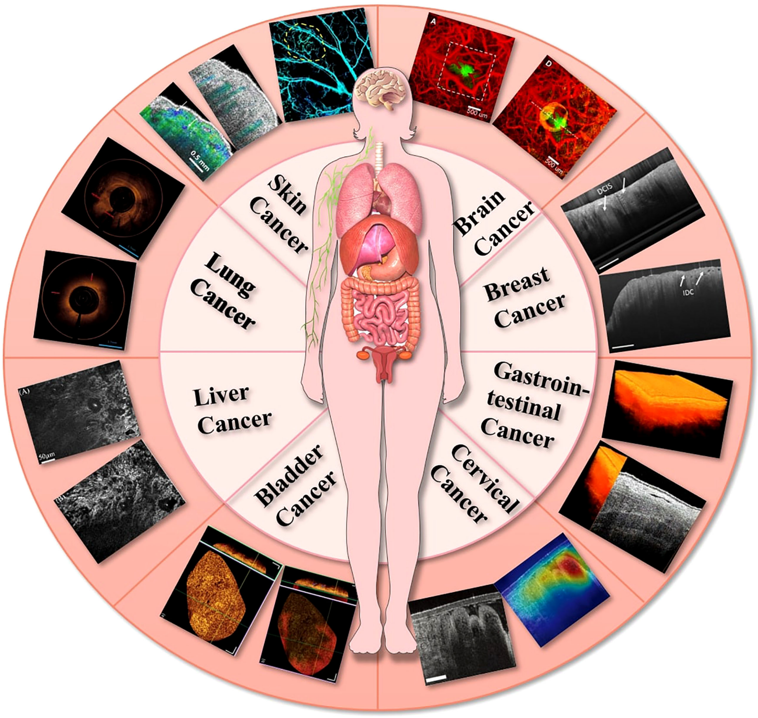


FIGURE 1
Application of OCT in the field of oncology.

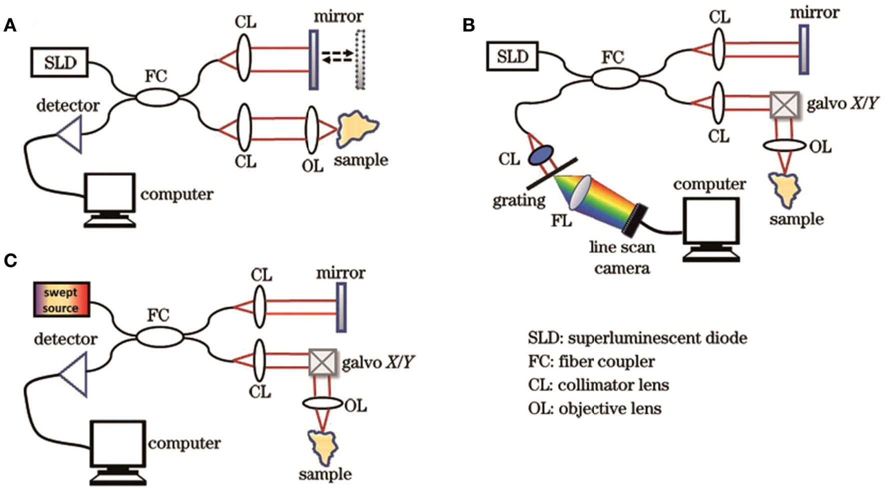


FIGURE 2
Structural diagrams of three generations of OCT systems. **(A)** TD-OCT; **(B)** SD-OCT; **(C)** SS-OCT (30). Copyright 2022, www.opticsjournal.net.

has a high-speed scan rate and a tunable scanning laser, resulting longer wavelengths than traditional spectral-domain OCT (20, 39). Therefore, the third-generation SS-OCT has faster scan speeds, higher scan densities, less deeply dependent signal-to-noise ratios, and higher resolutions. These properties enable them to reduce the impact of motion artifacts and better visualize tissues such as blood vessels while imaging larger areas, improving the quality of OCT *in vivo* imaging, and expanding its usage in biomedical research. Toward Pi Company has recently developed an SS-OCT system with 400,000 cycles per second that can rapidly reach an imaging depth of 6 mm. Simultaneously, the scanning length of a single image reaches 24 mm, and the axial resolution reaches as high as 3.8 μm. The images shine in the commercial field of ophthalmic OCT because of its excellent imaging parameters in both local and foreign markets (40).

Optical coherence tomography angiography (OCTA) is a kind of vascular imaging technology, belonging to the SS-OCT, which can visualize and quantify the morphological information of blood vessels by detecting the red blood cell (RBC) movement of the intravascular dynamic scattering signal (41). It has become the "gold standard" in the field of ophthalmic diagnostics. Currently, applications in the field of oncology are also widely studied.

## Application of OCT in oncology

Cancer is a leading cause of death worldwide and most patients are in the middle and advanced stages of treatment due to the subtle onset of early cancer and the inconspicuous symptoms. Traditional medical imaging methods focus on morphological tumor diagnosis, however, association of the imaging characteristics with early cancer is not apparent. Therefore, achieving multi-angle, all-round imaging and diagnosis of early cancer occurrence and progression from structural and functional levels, and providing timely radical treatment, is a significant component in the long-term survival of cancer patients. OCT has become a novel approach to early cancer diagnosis due to its rapid development.

### Application of OCT in superficial tumors

The skin is not only the largest and most accessible organ of the human body, but it also has relatively clear layered structures. Therefore, the microstructural information of the skin surface can be easily visualized using the OCT of near-infrared light. OCT not only generates micron-level images of living skin with a depth of 2 mm, but also is convenient, real-time, dynamic, great repeatable, and inexpensive. Hence, it is widely used in the diagnosis of superficial tumors. It was first used in 1997 to diagnose skin lesions as an additional tool for diagnosing and monitoring skin lesions (42). The high-resolution OCT detects the epidermis, dermis, appendages, and blood vessels of the skin, as well as evaluates the response to treatment of some diseases. OCT has already been demonstrated to help diagnose common skin cancers.

OCT is considered to be an advantageous diagnostic method for non-melanoma skin cancer, offering potential for diagnosis in the early stages of the disease. Non-melanoma skin cancers are generally classified as basal cell carcinoma (BCC) and squamous cell carcinoma (SCC) (43).

In 2021, an international consensus statement on Basal cell carcinoma (BCC) OCT, including BCC term sets for different subtypes was proposed. The publication of this statement helps implement OCT imaging of basal cell carcinoma in clinical and research settings (44). Adan et al. used the established diagnostic value of OCT features in 99 patients to determine whether OCT features could accurately distinguish BCC from non-BCC and BCC subtypes. The results showed that a limited number of OCT features were able to distinguish superficial BCC from non-superficial BCC and non-BCC lesions. The diagnostic method was able to detect 97.8% of BCC lesions, 84.3% of superficial BCC lesions and 98.8% of non-superficial BCC lesions (45). The LC-OCT technique, which combines reflex confocal microscopy with OCT technology, explains the basal cell carcinoma characteristics under LC-OCT examination and offers a theoretical basis for the diagnosis, classification, and treatment of later basal cell carcinoma (33, 34).

Cutaneous squamous cell carcinoma (SCC) is another common non-melanoma skin cancer that, unlike BCC, has the potential to metastasize. Early recognition and treatment are critical to reducing this risk, and actinic keratosis (AK) is considered a precursor lesion in SCC (46). Zhou et al., used an SD-OCT to image AK lesions of varying degrees in mice, which showed that the irregular wavy dermal-epidermal junction (DEJ) and persistent thickening of the epidermis are useful diagnostic parameters for AK. It demonstrates the great potential of OCT for non-invasive diagnosis of precancerous lesions (47). Cinotti et al. imaged 158 patients preoperatively using LC-OCT devices and performed histological examinations postoperatively. Conclusions show that LC-OCT is a new non-invasive imaging technique that can identify the main features of AK and SCC, which can help clinicians detect cellular and structural changes in keratinocyte skin tumors in real-time (48). Ho et al. based on a convolutional neural network (CNN) developed a mouse skin SCC classification model that integrates a FF-OCT device. This model provides a rapid, non-invasive, and accurate SCC classification, achieving 87.12% and 90.10% classification accuracy at the image level and tomography image level, respectively (35).

Conventional OCT is considered to be less sensitive for detecting early-stage melanoma, but it has the highest sensitivity compared to other techniques such as confocal microscopy, ultrasonography, and multispectral imaging. The imaging results of high-definition OCT (HD-OCT) and speckle variance OCT (SV-OCT) for melanoma are more positive than conventional OCT. However, due to the limited data available, more reports are needed to draw conclusions about their effects (49).

OCT technology offer changes in tumor microvascular before and after treatment to assess tumor microvascular response to nano therapy. This creates the theoretical and technical base for developing new tumor-specific diagnostic and treatment approaches (50). Welzel et al. observed blood vessels in skin and malignant melanoma using Dynamic optical coherence tomography (D-OCT) D-OCT and proposed that increased blood vessel density and irregular vascular patterns were more common in melanoma and more common in higher-stage melanoma (51). OCT can visually exhibit microscopic characteristics within tissues and distinguish lymph node tissue and surrounding adipose tissue, revealing changes in nodular microstructure during metastatic tumor invasion (**Figure 3**). Si et al. generated "flow-gated" spectral OCT images using a dual-band signal processing algorithm that demonstrates lymphatic drainage pathways for melanoma blood vessels and peritumoral tissue at micron-scale resolution (**Figure 4**) (52). **Table 1** summarizes the imaging capabilities of OCT in superficial tumors of the skin.

The introduction of non-invasive, efficient, and cost-effective screening tools will enhance the early detection of oral cancer and hence, the patient's lifespan. A Local Residual Adaptation Network (LRAN) model based on deep learning technology was developed for qualitative and quantitative analysis of oral cancer OCT image datasets with high accuracy and sensitivity (58). Furthermore, a 3D technique of SD-OCT was developed for evaluating the structural changes in oral epithelial cells, which

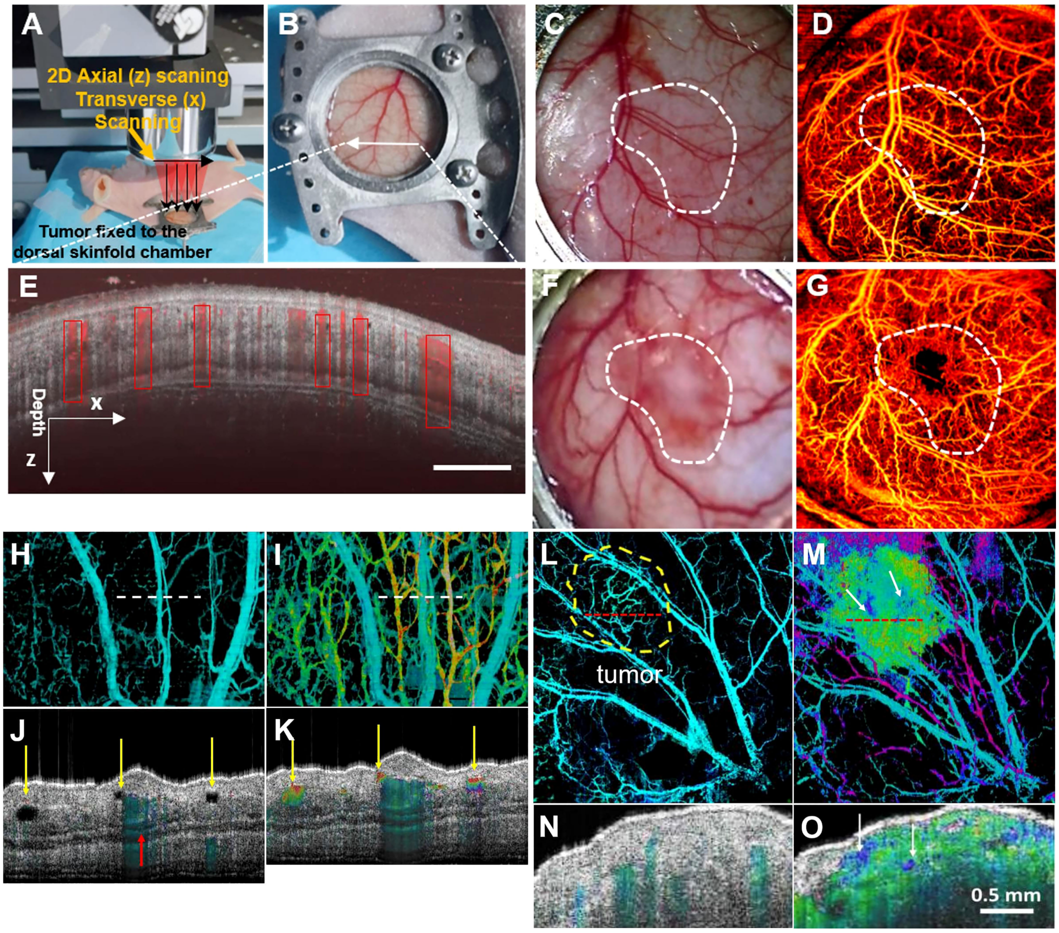


FIGURE 3
**(A)** Schematic of subcutaneous tumor-bearing nude mice dorsal window imaging; **(B)** Subcutaneous tumors of nude mice with tumors enlarge the skin window chamber images; **(C)** The zoom-in skin window chamber image in the healthy nude mice; **(D)** The corresponding enface microvascular image *in vivo* is shown in **(C)**; **(E)** Representative tissue cross sectional structural image (gray) and blood flow image (red border); **(F)** The zoom-in skin window chamber image in the subcutaneous tumor-bearing nude mice; **(G)** The corresponding enface microvascular image *in vivo* is shown in **(F)** (50); Copyright 2021, Wiley. **(H)** Normal vascular OCT images; **(I)** OCT images of angiogram and lymphangiography. The dotted line indicates the position of the cross-section image in A-B. **(J, K)** normal angiography and lymphangiography OCT en-face images. The red arrow indicates a large blood vessel, and the yellow arrow indicates the lymphatic vessel. **(L)** Melanoma vascular OCT image. **(M)** OCT images of melanoma angiography and lymphangiography. The dotted line indicates the position of the cross-section image in **(E–F)**. **(N, O)** Melanoma angiography and lymphangiography OCT en-face image (52). Copyright 2020, American chemistry society.

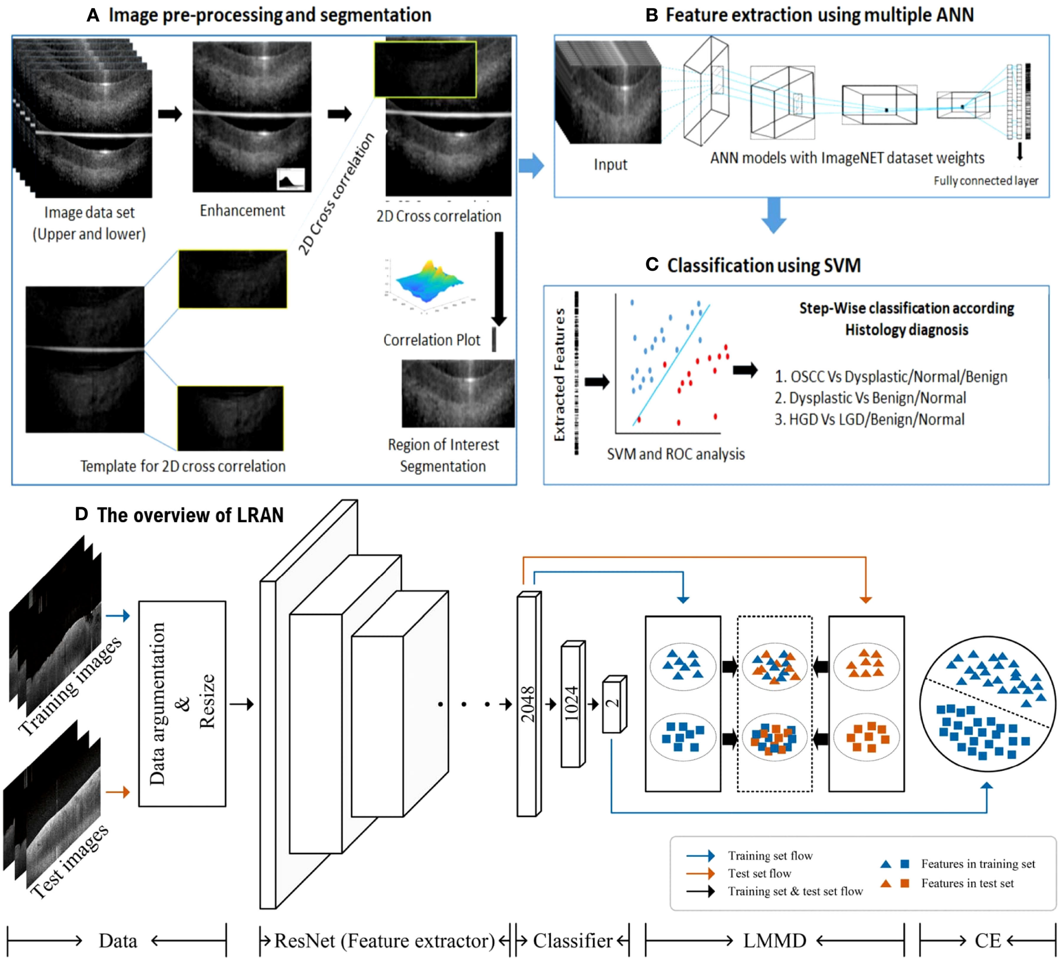


FIGURE 4
**(A-C)** ANN-SVM based image analysis pipeline (53); Copyright 2021, MDPI. **(D)** The overview of LRAN (58) Copyright 2022, Springer.

TABLE 1 The imaging capabilities of OCT in superficial tumors.

| Cancer | Authors | Main Findings |
|---|---|---|
| Cutaneous melanoma | Von Knorring et al. (2022) (54); Schuh et al. (2022) (55) | OCT can distinguish benign and malignant pigmented cutaneous tumors |
| Cutaneous basal cell carcinoma | Gust et al. (2022) (56); Suppa M et al. (2021) (33); Verzì et al. (2021) (57) | LC-OCT can describe the characteristics of basal cell carcinoma at the bedside for differential diagnosis of basal cell carcinoma and for typing of basal cell carcinoma and assessment of response to treatment of BCC. |
| Squamous cell carcinoma of the skin | Ho et al. (2021) (35) | FF-OCT can provide fast, non-invasive, and accurate SCC classification with high accuracy. |
| Oral cancer | Yuan et al. (2022) (58); Trebing et al. (2021) (59); Ilhan et al. (2020) (60); Chen et al. (2020) (61) | OCT can assess structural changes in oral epithelial cells, distinguish oral cancer from precancerous lesion tissue, and conduct noninvasive screening, detection, evaluation of differentiation, and staging oral dysplasia and early cancer. |

improved the time efficiency and quality of diagnosing epithelial lesions (59). Automatic image processing algorithms in OCT images can differentiate between heterotypic oral potentially malignant lesions (OPML) and malignant lesions, resulting in high sensitivity. Evidence is provided by using reliable and low-cost OCT instruments as point-of-care devices in resource-constrained settings and potential clinical applications in oral cancer screening and surveillance (53).

## Application of OCT in gastrointestinal tumors

OCT has shown significant potential in cavity organ tumors using techniques, such as endoscopy, catheterization, and laparoscopy. OCT imaging can distinguish between the four layers of the stomach wall structure, namely the glandular epithelium, mucosal muscle layer, submucosal layer, and muscle layer, where the submucosal layer is visible to the blood vessels. Jansen et al. used prospective research to investigate 26 patients with esophageal cancer. Calculate the contrast of plaques in an M-mode scan to distinguish between blood flow areas and resting tissues. This study is the first to confirm the OCT imaging of gastric tissue and blood flow detection *in vivo* during surgery in esophageal cancer patients, reducing the occurrence of anastomotic leakage after operation and improving surgical outcomes for patients (62).

OCT imaging of the esophageal and gastrointestinal parietal structures can detect various digestive tract diseases because tumors and normal tissues exhibit different light scattering patterns on OCT images. Lee et al. used volumetric OCTA imaging and corresponding histological diagnosis of 52 dysplasia patients who received Barrett's esophagus (BE) monitoring or endoscopic eradication of dysplasia, which can differentiate between low-grade dysplasia, low-grade dysplasia (LGD), and high-grade dysplasia (HGD) with the ability to visualize LGD/HGD-associated microvascular features with high accuracy (63) (**Figure 5**). Rodriguez et al. reviewed 14 studies, including endoluminal laser microscopy and OCT imaging of Barrett's esophagus, and found that endoscopic imaging of Barrett's esophagus with OCT and laser intraluminal microscopy could perform targeted biopsies and improve the probability of early detection of esophageal tumors (64).

OCT's ability to image the layers of the gastrointestinal wall can be used to diagnose cross-wall inflammation in Crohn's disease (CD) and differentiate it from ulcerative colitis (UC). Shen et al. used colonoscopic OCT to express the lamellar structures of colon wall disintegration. They discovered that

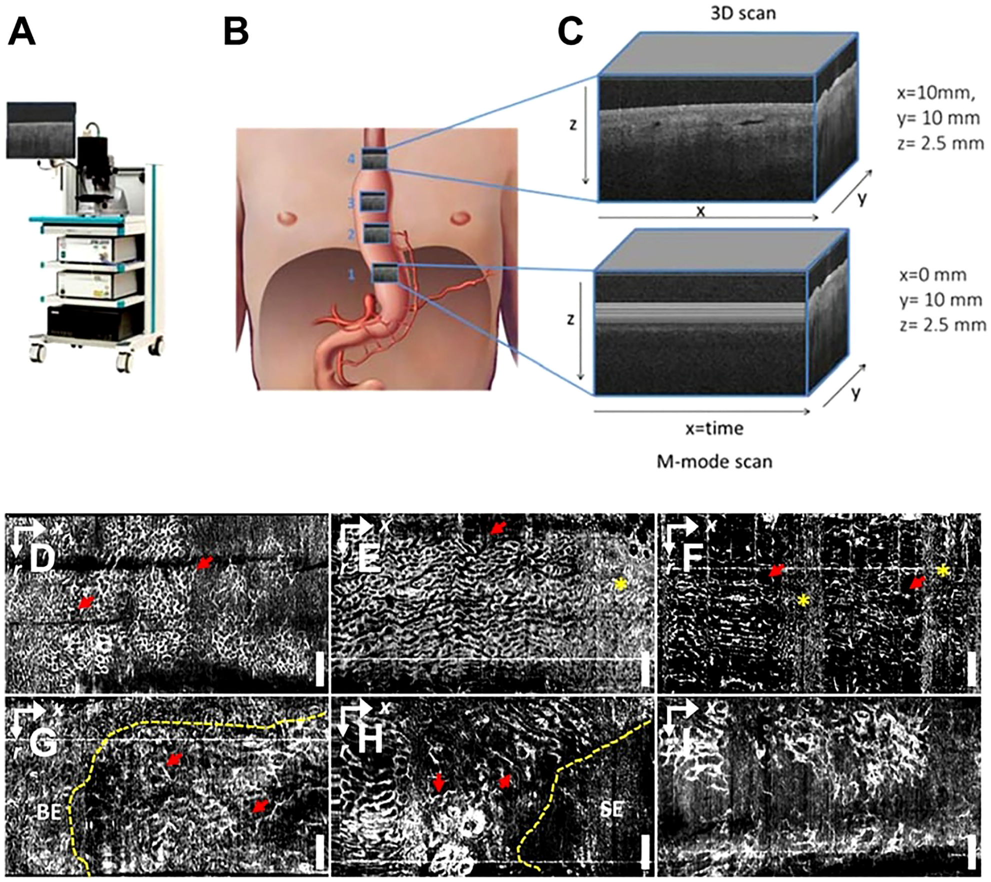


FIGURE 5
**(A)** Schematic diagram of a commercial OCT; **(B)** Gastric tube perfusion areas and **(C)** ROI region OCT grayscale image, cross-sectional OCT image showing vessels shadow (62); Copyright 2018, MDPI. OCTA vascular imaging of **(D-F)** non-dysplasia and **(G-I)** dysplasia BE (63). Nanoparticles targeting the hypoxic tumor microenvironment. Copyright 2017, ELSEVIER.

the destruction of colon wall layers on OCT is a reliable indicator of CD transmural inflammation (65).

OCT has a critical role in achieving qualitative real-time analysis and targeted biopsy. Ding et al. used OCT images to differentiate dysplasia and cancer from normal colonic tissue (66). For the first time, Hariri et al. used endoscopic OCT and laser-induced fluorescence (LIF) OCT-LIF to demonstrate repetitive, minimally invasive, cross-sectional colon imaging in mice, observing the development of adenoma with atypical hyperplasia of the colonic mucosal epithelium, mucosal thickening. The mucosal/submucosal barrier can be uplifted and disrupted by mucosal cancer tissue (67). Harpel et al. used OCT to track the onset and development of colorectal cancer in mice. They discovered that OCT could be used to allow for the monitoring of morphological changes in the distal colon due to tumor development and the presence of lymphoid aggregates. In addition, the role of inflammation on tumor development and the immune system can be elucidated. So, they could be used as novel therapeutic agents to prevent disease progression and increase the efficacy of anti-cancer agents. OCT can also be useful for initiating early therapy and assessing the benefit of combination therapy targeting inflammation (68).

Overall, OCT imaging is useful in the early differential detection of gastrointestinal tumors. The intraluminal optical tomography scanner (62, 69)could become a helpful reference for rapid, low-cost, non-invasive light biopsy, early differential diagnosis, and treatment of gastrointestinal cancers (**Table 2**).

## Application of OCT in deep tumors

Many research groups (71) have developed new OCT technologies to perform extensive studies in deep tumors with the rapid growth of lasers and computers.

OCT enables fast, wide-field, and label-free imaging of the living brain. In 2019, Katta et al. (72) used OCT to coagulate blood vessels and performed laser ablation of brain tumors (**Figure 6**). Yecies et al. published a new *in vivo* imaging approach using speckle-modulating OCT (SM-OCT) for label-free *in vivo* nerve and tumor edge identification in the same year. SM-OCT was used to show the white matter bundle and cortical layer structure in the brains of live mice. They identified the edges of glioblast tumors *in situ* in a mouse brain at an imaging of 10 μm (10).

Rapid and accurate evaluation of the intraoperative margin is vital for minimizing the resection rate in breast cancer. Using OCT images, Singla et al. used an active reverse-learning pre-trained inception-v3 CNN to distinguish between healthy and cancerous breast tissue. The method is highly sensitive, specific, and accurate (73). Likewise, Mojahed concluded that using CNN-based algorithms, it is possible to accurately identify malignant spots in OCT images (74).

Kansal et al. also developed a high-resolution automated full-field polarization-sensitive optical coherence tomography (FF-PS-OCT) system that was utilized to visualize 12 breast tissue samples, including four healthy tissues and eight malignant (cancerous) tissues. They used 106 OCT pictures to extract various phase features. This system can detect breast cancer models with up to 90.90% sensitivity and 85.0% specificity (75). Yang et al. analyzed the resection of normal breast tissue, breast cancer tissue, benign breast lesions, and axillary lymph nodes using FF-OCT and dynamic cell imaging (DCI). The findings reveal that FF-OCT and DCI have high accuracy in diagnosing breast cancer and have good diagnostic potential in breast surgery (76). Many research organizations have developed deep learning technology to improve qualitative leaps in image recognition and diagnostic characterization. More recently, Chen et al. created a computer-aided diagnosis (CADx) method that designs a contrastive texture learning strategy, with a sensitivity of 91.17% ± 4.99% for OCT image plaques. The specificity was 93.96% ± 4.72%, providing better interpretability based on texture features, which could lead to immediate clinical treatment (77).

Bladder cancer patients may benefit from OCT in addition to endoscopy for staging and grading. A prospective multicenter phase II trial revealed that OCT-assisted cystoscopy is a real-time, noninvasive, and maneuverable facility that increases the accuracy of bladder cancer staging and tumor invasion prediction (78). Xu et al. used intracellular motion (IM) as a dynamic contrast agent to track the distribution of urinary celiac cells. This contrast could provide a novel mechanism for OCT to accurately depict urothelial cancer cells' the depth and kind of invasion to stage bladder cancer (79). Wurster et al. offered a

TABLE 2 The imaging capabilities of OCT in gastrointestinal tumors.

| Cancer | Authors | Main Findings |
|---|---|---|
| Esophageal adenocarcinoma | Rodriguez et al. (2019) (64); Lee et al. (2017) (63) | OCTA can distinguish between LGD and HGD while showing LGD/HGD-related microvascular features with high accuracy. Endoscopic imaging using OCT and laser intraluminal microscopy allows for targeted biopsy to improve the probability of early detection of esophageal tumors. |
| Colon cancer | Kendall et al. (2022) (70) | OCT can show the "texture" of tissue well and is an excellent way to assess the mucosal thickness and the number of layers for quick identification and classification of tissues. |
| Gastric cancer | Jansen et al. (2018) (62) | OCT can perform real-time visual blood flow detection imaging in a surgical setting to evaluate the efficacy of surgery or drugs by using Vivo microcirculatory perfusion destined data. |

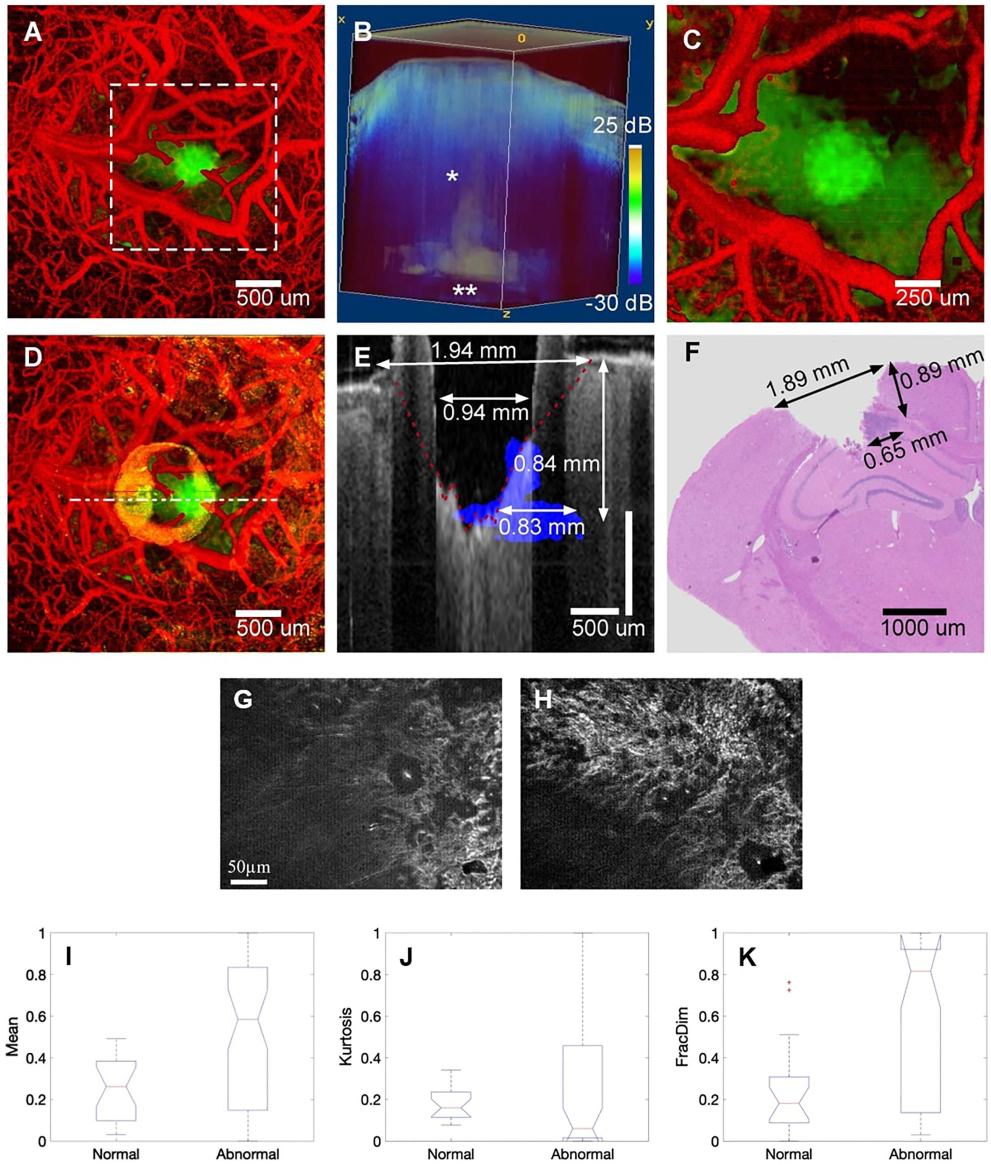


FIGURE 6
**(A)** Pretreatment of cerebral surface vascular construction (red) and tumor area (green) in mice. **(B)** Renders an image in 3D with an attenuation rate threshold mask superimposed on the OCT intensity (blue). **(C)** Maximum intensity projection after coagulation. **(D)** Maximum intensity projection after ablation. **(E)** After overlapping the tumor margins (blue) before ablation, stain the corresponding area with the post-ablation b-scan (gray) **(F)** H&E stained of the corresponding region (72); Copyright 2019, Theranostics. **(G)** FF-OCT image of normal and **(H)** cancerous hepatic cell. **(I-K)** Boxplot of selected features for the Mean, Kurtosis, and FracDim. mean and kurtosis are not sensitive in distinguishing between normal and cancerous hepatocytes, the mean and kurtosis are not sensitive in distinguishing between normal and cancerous hepatocytes, when the liver becomes cancerous, the value of the fractal parameter increases (88). Copyright 2020, Wiley.

piezoelectric fiber-optic scanner-based forward-imaging endoscope for OCTA. The instrument combines morphological tissue comparison with qualitative dynamic blood flow information to improve the early diagnosis of diseases like bladder cancer (80).

For the time being, OCT is primarily used as a bronchoscopic auxiliary tool to display the microstructure of each layer of the bronchial wall and achieve a similar histopathological diagnosis without tissue biopsy, which helps to reduce the invasive examination and improve the early diagnosis rate of lung cancer, which is critical for the diagnosis and treatment of lung cancer (81). According to Shostak et al., ultra-high resolution images provide essential microstructural information that effectively distinguishes lymph nodes from adjacent airway walls through the characteristics of these microstructures and reveals lymphoid follicles, adipose tissue, pigment-laden histiocytes, and blood vessels information based on needle-based OCT-guided lymph node sampling for lung cancer staging (82).

Hohert et al. used a combination of OCT and autofluorescence imaging (OCT-AFI) to improve diagnostic rates for areas of the lungs not accessible by more extensive imaging methods (83). Furthermore, OCT can aid in determining the tumor's depth of invasion (84). In malignant lung disease, discriminate between normal and malignant sections of the central airway, lung parenchyma, lymph nodes, and pleura by visualizing illness-related anatomical partitions of the lungs in real-time (85). The results of an *in vitro* scan of 64 specimens of lung nodules suggest that PS-OCT may be able to distinguish between tumors and fibrosis and can be used to guide intraoperative tissue sampling *in vivo* or to assess sufficiency for rapid biopsy *in vitro* (36). Nandy et al. came to a similar conclusion (86).

Presently, the use of OCT in liver cancer is under-reported. In 2015, Zhu et al. performed rapid and high-resolution tomography of human liver specimens using an FF-OCT scanner (87). Nuclear atypia and thicker fibrous bands of hepatocellular carcinoma can be observed on en-face tomography images of FF-OCT. They proposed the support vector machine (SVM) for classifying normal liver tissue and cancerous liver tissue using en-face tomography images. They used the label-free human liver tomography stack to extract seven quantitative parameters, including mean, variance, skewness, kurtosis, energy, entropy, and fractal dimension (FD). The value of FD grows as the liver becomes cancerous, signifying that a divided-dimensional classifier can be utilized for label-free quantitative tumor detection (88).

These encouraging research results suggest that OCT technology will become an important imaging method for deep tumor clinical applications. **Table 3** summarizes the imaging capabilities of OCT in the deep tumor.

## Conclusions and perspectives

As a mature imaging method being used in new fields, OCT has its unique advantages. Firstly, OCT can provide non-invasive, high-quality detailed images. Through optical, electrical and image processing, OCT can provide micrometer-resolution images of tissues, as well as high-resolution 3D imaging, which can be used for early diagnosis and treatment of diseases. The high soft tissue contrast that OCT can provide facilitates detailed analysis of soft tissue anatomy, which is of great significance for early diagnosis of cancer. In addition, OCT can be integrated into small probes and catheters, making it suitable for entering internal organs for cancer imaging and diagnosis. Due to the limited penetration depth and visual field of OCT itself, it is difficult for an oncologist to diagnose from an image of small area tissue. Besides, the imaging depth is limited in evaluating intraoperative tumor margins, usually within 2mm, which greatly limits its application in surgery. Moreover, it is difficult to fix the probe or imaging module well when imaging with small probes and catheters, so it is difficult to obtain clear images. Therefore, the first step in the future development of OCT is to improve the imaging depth, and combining artificial intelligence algorithms and a variety of imaging methods, so that it can perform imaging at a relatively large depth and field of view. Secondly, under the advantages of ultra-high sampling speed and high resolution of OCT itself, the multi-frame synthesis technology is used to improve the stability of sampling. Finally, OCT will integrate with other disciplines and technologies in the future, such as artificial intelligence, medical image analysis, intelligent machinery manufacturing, safe and

TABLE 3 The imaging capabilities of OCT in deep tumors.

| Cancer | Authors | Main Findings |
|---|---|---|
| Brain cancer | NyúlTóth et al. (2021) (89); Hartmann et al. (2020) (90); Yecies et al. (2019) (10); Tsai et al. (2018) (91) | OCT can not only identify and quantify cerebrovascular morphology and degree of relaxation *in vivo* but also conduct long-term monitoring of cerebrovascular dynamics in dilated. It can also show hidden brain microanatomy to identify brain tumor margins, improving intraoperative safety. |
| Breast cancer | Mojahed et al. (2020) (74); Kansal et al. (2020) (75); Yang et al. (2020) (76) | FF-OCT has good diagnostic potential in breast surgery and enables real-time assessment of intraoperative margins. |
| Bladder cancer | Sung et al. (2021) (78); Xu et al. (2021) (79) | OCT can show the depth and type of invasion of urothelial cancer cells, accurately grading and staging bladder cancer. Assist in intraoperative decision-making through real-time disease staging for more accurate diagnosis, resection, and reduced recurrence rates. |
| Cervical cancer | Chen et al. (2022) (77); Ren et al. (2021) (92); Placzek et al. (2020) (93); Ma et al. (2019) (94); Zeng et al. (2018) (95) | OCT can identify cervical morphological features and lesions noninvasively in real-time. |
| Lung cancer | Ding et al. (2021) (81) | Endobronchial OCT (EB-OCT) combined with machine learning algorithms can identify malignant lung nodules at a low cost. |
| Hepatocellular carcinoma | Zhu et al. (2020) (88) | FF-OCT can quantitatively detect hepatocellular carcinoma without markers. |

environmentally friendly new material processes, et. It can be used not only in early disease diagnosis and facilitating scientific research to provide a more objective and precise imaging measurement basis, but also for the routine detection of diseases to provide safer, faster, and inexpensive technology solutions.

## Author contributions

LY, JM, HZ, and QZ conceived the article. LY and YC wrote the manuscript. SL, JW, GW, and BZ helped to improve the manuscript. All authors have read and agreed to the published version of the manuscript.

## Funding

The work was supported by the Guangdong Basic and Applied Basic Research Foundation (2021A1515011654), Fundamental Research Funds for the Central Universities of China (20720210117), Joint Funds for the Innovation of Science and Technology of Fujian province (2019Y9128), Xiamen Key Laboratory of Endocrine-Related Cancer Precision Medicine (XKLEC2021KF03, XKLEC2020KF05). Key Laboratory of OptoElectronic Science and Technology for Medicine of Ministry of Education, Fujian Provincial Key Laboratory of Photonics Technology (JYG2105). XMU Undergraduate Innovation and Entrepreneurship Training Programs (202210384051, S202210384404, 2021X1119, 2021Y1119, S202110384391), and Shenzhen Bay Laboratory (SZBL2019062801005). Key Laboratory of Nanomedical Technology (Education Department of Fujian Province), School of Pharmacy, Nano Medical Technology Research Institute, Fujian Medical University (2022KLNT201). The Science and Technology Project of Xiamen municipal Bureau of Science and Technology (3502Z20194044).

## Conflict of interest

## Publisher's note

## References

1. Huang D, Swanson EA, Lin CP, Schuman JS, Stinson WG, Chang W, et al. Optical coherence tomography. *Science* (1991) 254:1178–81. doi: 10.1126/science.1957169

2. Hsieh YS, Ho YC, Lee SY, Chuang CC, Tsai JC, Lin KF, et al. Dental optical coherence tomography. *Sensors (Basel)* (2013) 13:8928–49. doi: 10.3390/s130708928

3. Huang H, Lai JY, Do J, Liu D, Li L, Del Rosario J, et al. Specifically targeting angiopoietin-2 inhibits angiogenesis, Tie2-expressing monocyte infiltration, and tumor growth. *Clin Cancer Res* (2011) 17:1001–11. doi: 10.1158/1078-0432.CCR-10-2317

4. Zhang SS, Ni YH, Zhao CR, Qiao Z, Yu HX, Wang LY, et al. Capsaicin enhances the antitumor activity of sorafenib in hepatocellular carcinoma cells and mouse xenograft tumors through increased ERK signaling. *Acta Pharmacol Sin* (2018) 39:438–48. doi: 10.1038/aps.2017.156

5. Zhang C, Feng W, Zhao Y, Yu T, Li P, Xu T, et al. A large, switchable optical clearing skull window for cerebrovascular imaging. *Theranostics* (2018) 8:2696–708. doi: 10.7150/thno.23686

6. Briers JD. Laser Doppler, speckle and related techniques for blood perfusion mapping and imaging. *Physiol Meas* (2001) 22:R35–66. doi: 10.1088/0967-3334/22/4/201

7. Stummer W, Reulen HJ, Meinel T, Pichlmeier U, Schumacher W, Tonn JC, et al. Extent of resection and survival in glioblastoma multiforme: identification of and adjustment for bias. *Neurosurgery* (2008) 62:564–76. doi: 10.1227/01.neu.0000317304.31579.17

8. Yubin L, Xu M, Da Y. NIR-II dual-modal optical coherence tomography and photoacoustic imaging-guided dose-control cancer chemotherapy. *ACS Appl Polymer Mater* (2020) 2:9. doi: 10.1021/acsapm.0c00155

9. Wojtkowski M. High-speed optical coherence tomography: basics and applications. *Appl Opt* (2010) 49:D30–61. doi: 10.1364/AO.49.000D30

10. Yecies D, Liba O, Sorelle ED, Dutta R, Yuan E, Vogel H, et al. Speckle modulation enables high-resolution wide-field human brain tumor margin detection and *in vivo* murine neuroimaging. *Sci Rep* (2019) 9:10388. doi: 10.1038/s41598-019-45902-9

11. Wang X, Jiang C, Ko T, Kong X, Yu X, Min W, et al. Correlation between optic disc perfusion and glaucomatous severity in patients with open-angle glaucoma: an optical coherence tomography angiography study. *Graefes Arch Clin Exp Ophthalmol* (2015) 253:1557–64. doi: 10.1007/s00417-015-3095-y

12. Brusini P. OCT glaucoma staging system: a new method for retinal nerve fiber layer damage classification using spectral-domain OCT. *Eye (Lond)* (2018) 32:113–9. doi: 10.1038/eye.2017.159

13. Moon Y, Lim HT. Relationship between peripapillary atrophy and myopia progression in the eyes of young school children. *Eye (Lond)* (2021) 35:665–71. doi: 10.1038/s41433-020-0945-6

14. Spaide RF. Optical coherence tomography angiography signs of vascular abnormalization with antiangiogenic therapy for choroidal neovascularization. *Am J Ophthalmol* (2015) 160:6–16. doi: 10.1016/j.ajo.2015.04.012

15. Wu Z, Luu CD, Ayton LN, Goh JK, Lucci LM, Hubbard WC, et al. Fundus autofluorescence characteristics of nascent geographic atrophy in age-related macular degeneration. *Invest Ophthalmol Vis Sci* (2015) 56:1546–52. doi: 10.1167/iovs.14-16211

16. Schaal KB, Gregori G, Rosenfeld PJ. En face optical coherence tomography imaging for the detection of nascent geographic atrophy. *Am J Ophthalmol* (2017) 174:145–54. doi: 10.1016/j.ajo.2016.11.002

17. Adhi M, Filho MA, Louzada RN, Kuehlewein L, De Carlo TE, Baumal CR, et al. Retinal capillary network and foveal avascular zone in eyes with vein occlusion and fellow eyes analyzed with optical coherence tomography angiography. *Invest Ophthalmol Vis Sci* (2016) 57:Oct486–494. doi: 10.1167/iovs.15-18907

18. Nobre Cardoso J, Keane PA, Sim DA, Bradley P, Agrawal R, Addison PK, et al. Systematic evaluation of optical coherence tomography angiography in retinal vein occlusion. *Am J Ophthalmol* (2016) 163:93–107.e106. doi: 10.1016/j.ajo.2015.11.025

19. Chung CY, Li KKW. Optical coherence tomography angiography wide-field montage in branch retinal vein occlusion before and after anti-vascular endothelial-derived growth factor injection. *Int Ophthalmol* (2018) 38:1305–7. doi: 10.1007/s10792-017-0568-5

20. Jia Y, Bailey ST, Hwang TS, Mcclintic SM, Gao SS, Pennesi ME, et al. Quantitative optical coherence tomography angiography of vascular abnormalities in the living human eye. *Proc Natl Acad Sci U.S.A.* (2015) 112:E2395–2402. doi: 10.1073/pnas.1500185112

21. Hamada M, Ohkoshi K, Inagaki K, Ebihara N, Murakami A. Visualization of microaneurysms using optical coherence tomography angiography: comparison of OCTA en face, OCT b-scan, OCT en face, FA, and IA images. *Jpn J Ophthalmol* (2018) 62:168–75. doi: 10.1007/s10384-018-0570-0

22. Tian M, Tappeiner C, Zinkernagel MS, Huf W, Wolf S, Munk MR. Evaluation of vascular changes in intermediate uveitis and retinal vasculitis using swept-source wide-field optical coherence tomography angiography. *Br J Ophthalmol* (2019) 103:1289–95. doi: 10.1136/bjophthalmol-2018-313078

23. Cao JL, Srivastava SK, Venkat A, Lowder CY, Sharma S. Ultra-widefield fluorescein angiography and OCT findings in tubulointerstitial nephritis and uveitis syndrome. *Ophthalmol Retina* (2020) 4:189–97. doi: 10.1016/j.oret.2019.08.012

24. Pichi F, Dolz-Marco R, Francis JH, Au A, Davis JL, Fawzi A, et al. Advanced OCT analysis of biopsy-proven vitreoretinal lymphoma. *Am J Ophthalmol* (2021) 238:16–26. doi: 10.1016/j.ajo.2021.11.023

25. Ang M, Baskaran M, Werkmeister RM, Chua J, Schmidl D, Aranha Dos Santos V, et al. Anterior segment optical coherence tomography. *Prog Retin Eye Res* (2018) 66:132–56. doi: 10.1016/j.preteyeres.2018.04.002

26. Rao HL, Pradhan ZS, Suh MH, Moghimi S, Mansouri K, Weinreb RN. Optical coherence tomography angiography in glaucoma. *J Glaucoma* (2020) 29:312–21. doi: 10.1097/IJG.0000000000001463

27. Sampson DM, Dubis AM, Chen FK, Zawadzki RJ, Sampson DD. Towards standardizing retinal optical coherence tomography angiography: a review. *Light Sci Appl* (2022) 11:63. doi: 10.1038/s41377-022-00740-9

28. Marschall S, Sander B, Mogensen M, Jørgensen TM, Andersen PE. Optical coherence tomography-current technology and applications in clinical and biomedical research. *Anal Bioanal Chem* (2011) 400:2699–720. doi: 10.1007/s00216-011-5008-1

29. Tsang SH, Sharma T. Optical coherence tomography. *Adv Exp Med Biol* (2018) 1085:11–3. doi: 10.1007/978-3-319-95046-4_3

30. Wei Yingzhao YX, Lan G, Huang Y, Qin J, Lin AN, Tan H, et al. Research progress and application of cardiovascular optical CoherenceTomography. *Laser Optoelectron Prog* (2021) 58:30–44. doi: 10.3788/LOP202158.2400002

31. Massatsch P, Charrière F, Cuche E, Marquet P, Depeursinge CD. Time-domain optical coherence tomography with digital holographic microscopy. *Appl Opt* (2005) 44:1806–12. doi: 10.1364/AO.44.001806

32. Podoleanu AG. Optical coherence tomography. *J Microsc* (2012) 247:209–19. doi: 10.1111/j.1365-2818.2012.03619.x

33. Suppa M, Fontaine M, Dejonckheere G, Cinotti E, Yélamos O, Diet G, et al. Line-field confocal optical coherence tomography of basal cell carcinoma: a descriptive study. *J Eur Acad Dermatol Venereol* (2021) 35:1099–110. doi: 10.1111/jdv.17078

34. Monnier J, De Carvalho N, Harris U, Garfinkel J, Saud A, Navarrete-Dechent C, et al. Combined reflectance confocal microscopy and optical coherence tomography to improve the diagnosis of equivocal lesions for basal cell carcinoma. *J Am Acad Dermatol* (2022) 86:934–6. doi: 10.1016/j.jaad.2021.03.066

35. Ho CJ, Calderon-Delgado M, Lin MY, Tjiu JW, Huang SL, Chen HH. Classification of squamous cell carcinoma from FF-OCT images: Data selection and progressive model construction. *Comput Med Imaging Graph* (2021) 93:101992. doi: 10.1016/j.compmedimag.2021.101992

36. Hariri LP, Adams DC, Applegate MB, Miller AJ, Roop BW, Villiger M, et al. Distinguishing tumor from associated fibrosis to increase diagnostic biopsy yield with polarization-sensitive optical coherence tomography. *Clin Cancer Res* (2019) 25:5242–9. doi: 10.1158/1078-0432.CCR-19-0566

37. Titiyal JS, Kaur M, Nair S, Sharma N. Intraoperative optical coherence tomography in anterior segment surgery. *Surv Ophthalmol* (2021) 66:308–26. doi: 10.1016/j.survophthal.2020.07.001

38. Yaqoob Z, Wu J, Yang C. Spectral domain optical coherence tomography: a better OCT imaging strategy. *Biotechniques* (2005) 39:S6–13. doi: 10.2144/000112090

39. Huber R, Adler DC, Fujimoto JG. Buffered fourier domain mode locking: unidirectional swept laser sources for optical coherence tomography imaging at 370,000 lines/s. *Opt Lett* (2006) 31:2975–7. doi: 10.1364/OL.31.002975

40. *Toward pi medical received another 200 million yuan of b round financing this year, leading the rise of domestic ophthalmic equipment* . Available at: https://www.shangyexinzhi.com/article/4440235.html.

41. Si P, Yuan E, Liba O, Winetraub Y, Yousefi S, Sorelle ED, et al. Gold nanoprisms as optical coherence tomography contrast agents in the second near-infrared window for enhanced angiography in live animals. *ACS Nano* (2018) 12:11986–94. doi: 10.1021/acsnano.8b03862

42. Welzel J, Lankenau E, Birngruber R, Engelhardt R. Optical coherence tomography of the human skin. *J Am Acad Dermatol* (1997) 37:958–63. doi: 10.1016/S0190-9622(97)70072-0

43. Sattler E, Kastle R, Welzel J. Optical coherence tomography in dermatology. *J BioMed Opt* (2013) 18:061224. doi: 10.1117/1.JBO.18.6.061224

44. Fuchs CSK, Ortner VK, Mogensen M, Rossi AM, Pellacani G, Welzel J, et al. 2021 international consensus statement on optical coherence tomography for basal cell carcinoma: image characteristics, terminology and educational needs. *J Eur Acad Dermatol Venereol* (2022) 36(6):772–8. doi: 10.1111/jdv.17969

45. Adan F, Mosterd K, Kelleners-Smeets NWJ, Nelemans PJ. Diagnostic value of optical coherence tomography image features for diagnosis of basal cell carcinoma. *Acta Derm Venereol* (2021) 101:adv00607. doi: 10.2340/actadv.v101.421

46. Dreyfuss I, Kamath P, Frech F, Hernandez L, Nouri K. Squamous cell carcinoma: 2021 updated review of treatment. *Dermatol Ther* (2022) 35:e15308. doi: 10.1111/dth.15308

47. Zhou C, Zhang H, Wang P, Shi L, Wen L, Chen Q, et al. Optical coherence tomography-based non-invasive evaluation of premalignant lesions in SKH-1 mice. *J Biophoton* (2021) 14:e202000490. doi: 10.1002/jbio.202000490

48. Cinotti E, Tognetti L, Cartocci A, Lamberti A, Gherbassi S, Orte Cano C, et al. Line-field confocal optical coherence tomography for actinic keratosis and squamous cell carcinoma: a descriptive study. *Clin Exp Dermatol* (2021) 46:1530–41. doi: 10.1111/ced.14801

49. Rajabi-Estarabadi A, Bittar JM, Zheng C, Nascimento V, Camacho I, Feun LG, et al. Optical coherence tomography imaging of melanoma skin cancer. *Lasers Med Sci* (2019) 34:411–20. doi: 10.1007/s10103-018-2696-1

50. Huang D, Wang G, Mao J, Liu C, Fan Z, Zhang Y, et al. Intravital whole-process monitoring thermo-chemotherapy *via* 2d silicon nanoplatform: a macro guidance and long-term microscopic precise imaging strategy. *Adv Sci (Weinh)* (2021) 8:e2101242. doi: 10.1002/advs.202101242

51. Welzel J, Schuh S, De Carvalho N, Themstrup L, Ulrich M, Jemec GBE, et al. Dynamic optical coherence tomography shows characteristic alterations of blood vessels in malignant melanoma. *J Eur Acad Dermatol Venereol* (2021) 35:1087–93. doi: 10.1111/jdv.17080

52. Si P, Shevidi S, Yuan E, Yuan K, Lautman Z, Jeffrey SS, et al. Gold nanobipyramids as second near infrared optical coherence tomography contrast agents for in vivo multiplexing studies. *Nano Lett* (2020) 20:101–8. doi: 10.1021/acs.nanolett.9b03344

53. James BL, Sunny SP, Heidari AE, Ramanjinappa RD, Lam T, Tran AV, et al. Validation of a point-of-care optical coherence tomography device with machine learning algorithm for detection of oral potentially malignant and malignant lesions. *Cancers (Basel)* (2021) 13(14):3583. doi: 10.3390/cancers13143583

54. Von Knorring T, Israelsen NM, Ung V, Formann JL, Jensen M, Hædersdal M, et al. Differentiation between benign and malignant pigmented skin tumours using bedside diagnostic imaging technologies: a pilot study. *Acta Derm Venereol* (2022) 102:adv00634. doi: 10.2340/actadv.v101.571

55. Schuh S, Ruini C, Perwein MKE, Daxenberger F, Gust C, Sattler EC, et al. Line-field confocal optical coherence tomography: a new tool for the differentiation between nevi and melanomas? *Cancers (Basel)* (2022) 14(5):1140. doi: 10.3390/cancers14051140

56. Gust C, Schuh S, Welzel J, Daxenberger F, Hartmann D, French LE, et al. Line-field confocal optical coherence tomography increases the diagnostic accuracy and confidence for basal cell carcinoma in equivocal lesions: a prospective study. *Cancers (Basel)* (2022) 14(4):1082 doi: 10.3390/cancers14041082.

57. Verzì AE, Micali G, Lacarrubba F. Line-field confocal optical coherence tomography may enhance monitoring of superficial basal cell carcinoma treated with imiquimod 5% cream: a pilot study. *Cancers (Basel)* (2021) 13(19):4913. doi: 10.3390/cancers13194913

58. Yuan W, Cheng L, Yang J, Yin B, Fan X, Yang J, et al. Noninvasive oral cancer screening based on local residual adaptation network using optical coherence tomography. *Med Biol Eng Comput* (2022) 60:1363–75. doi: 10.1007/s11517-022-02535-x

59. Trebing CT, Sen S, Rues S, Herpel C, Schöllhorn M, Lux CJ, et al. Non-invasive three-dimensional thickness analysis of oral epithelium based on optical coherence tomography-development and diagnostic performance. *Heliyon* (2021) 7:e06645. doi: 10.1016/j.heliyon.2021.e06645

60. Ilhan B, Lin K, Guneri P, Wilder-Smith P. Improving oral cancer outcomes with imaging and artificial intelligence. *J Dent Res* (2020) 99:241–8. doi: 10.1177/002203452090212

61. Chen PH, Lee HY, Chen YF, Yeh YC, Chang KW, Hou MC, et al. Detection of oral dysplastic and early cancerous lesions by polarization-sensitive optical coherence tomography. *Cancers (Basel)* (2020) 12(9):2376. doi: 10.3390/cancers12092376

62. Jansen SM, Almasian M, Wilk LS, De Bruin DM, Van Berge Henegouwen MI, Strackee SD, et al. Feasibility of optical coherence tomography (OCT) for intra-operative detection of blood flow during gastric tube reconstruction. *Sensors (Basel)* (2018) 18(5):1331. doi: 10.3390/s18051331

63. Lee HC, Ahsen OO, Liang K, Wang Z, Figueiredo M, Giacomelli MG, et al. Endoscopic optical coherence tomography angiography microvascular features associated with dysplasia in barrett's esophagus (with video). *Gastrointest Endosc* (2017) 86:476–84.e473. doi: 10.1016/j.gie.2017.01.034

64. Rodriguez MAC, De Moura DTH, Ribeiro IB, Bernardo WM, Morita FHA, Marques SB, et al. Volumetric laser endomicroscopy and optical coherence tomography in barrett's esophagus: a systematic review and meta-analysis. *Endosc Int Open* (2019) 7:E1078–e1091. doi: 10.1055/a-0965-6487

65. Shen B, Zuccaro G Jr., Gramlich TL, Gladkova N, Trolli P, Kareta M, et al. *In vivo* colonoscopic optical coherence tomography for transmural inflammation in inflammatory bowel disease. *Clin Gastroenterol Hepatol* (2004) 2:1080–7. doi: 10.1016/S1542-3565(04)00621-4

66. Ding J, Li Q, Lin J, He S, Chen W, He Q, et al. Optical coherence tomography for the early detection of colorectal dysplasia and cancer: validation in a murine model. *Quant Imaging Med Surg* (2021) 11:371–9. doi: 10.21037/qims-20-13

67. Hariri LP, Tumlinson AR, Besselsen DG, Utzinger U, Gerner EW, Barton JK. Endoscopic optical coherence tomography and laser-induced fluorescence spectroscopy in a murine colon cancer model. *Lasers Surg Med* (2006) 38:305–13. doi: 10.1002/lsm.20305

68. Harpel K, Leung S, Rice PF, Jones M, Barton JK, Bommireddy R. Imaging colon cancer development in mice: IL-6 deficiency prevents adenoma in azoxymethane-treated Smad3 knockouts. *Phys Med Biol* (2016) 61:N60–69. doi: 10.1088/0031-9155/61/3/N60

69. Liu R, Zhang Y, Zheng Y, Liu Y, Zhao Y, Yi L. Automated detection of vulnerable plaque for intravascular optical coherence tomography images. *Cardiovasc Eng Technol* (2019) 10:590–603. doi: 10.1007/s13239-019-00425-2

70. Kendall WY, Bordas J, Mirminachi S, Joseph A, Roper J, Wax A. Spectroscopic optical coherence tomography for classification of colorectal cancer in a mouse model. *J Biophoton* (2022) 15(7):e202100387. doi: 10.1002/jbio.202100387

71. Fan Y, Xia Y, Zhang X, Sun Y, Tang J, Zhang L, et al. Optical coherence tomography for precision brain imaging, neurosurgical guidance and minimally invasive theranostics. *Biosci Trends* (2018) 12:12–23. doi: 10.5582/bst.2017.01258

72. Katta N, Estrada AD, Mcelroy AB, Gruslova A, Oglesby M, Cabe AG, et al. Laser brain cancer surgery in a xenograft model guided by optical coherence tomography. *Theranostics* (2019) 9:3555–64. doi: 10.7150/thno.31811

73. Singla N, Dubey K, Srivastava V. Automated assessment of breast cancer margin in optical coherence tomography images *via* pretrained convolutional neural network. *J Biophoton* (2019) 12:e201800255. doi: 10.1002/jbio.201800255

74. Mojahed D, Ha RS, Chang P, Gan Y, Yao X, Angelini B, et al. Fully automated postlumpectomy breast margin assessment utilizing convolutional neural network based optical coherence tomography image classification method. *Acad Radiol* (2020) 27:e81–6. doi: 10.1016/j.acra.2019.06.018

75. Kansal S, Bhattacharya J, Srivastava V. Automated full-field polarization-sensitive optical coherence tomography diagnostic systems for breast cancer. *Appl Opt* (2020) 59:7688–93. doi: 10.1364/AO.396592

76. Yang H, Zhang S, Liu P, Cheng L, Tong F, Liu H, et al. Use of high-resolution full-field optical coherence tomography and dynamic cell imaging for rapid intraoperative diagnosis during breast cancer surgery. *Cancer* (2020) 126 Suppl 16:3847–56. doi: 10.1002/cncr.32838

77. Chen K, Wang Q, Ma Y. Cervical optical coherence tomography image classification based on contrastive self-supervised texture learning. *Med Phys* (2022) 49(6)3638–53. doi: 10.1002/mp.15630

78. Sung HH, Scherr DS, Slaton J, Liu H, Feeny KL, Lingley-Papadopoulos C, et al. Phase II multi-center trial of optical coherence tomography as an adjunct to white light cystoscopy for intravesical real time imaging and staging of bladder cancer. *Urol Oncol* (2021) 39:434.e423–434.e429. doi: 10.1016/j.urolonc.2021.03.026

79. Xu Z, Zhu H, Wang H. Segmentation of the urothelium in optical coherence tomography images with dynamic contrast. *J BioMed Opt* (2021) 26(8):086002. doi: 10.1117/1.JBO.26.8.086002

80. Wurster LM, Shah RN, Placzek F, Kretschmer S, Niederleithner M, Ginner L, et al. Endoscopic optical coherence tomography angiography using a forward imaging piezo scanner probe. *J Biophoton* (2019) 12:e201800382. doi: 10.1002/jbio.201800382

81. Ding M, Pan SY, Huang J, Yuan C, Zhang Q, Zhu XL, et al. Optical coherence tomography for identification of malignant pulmonary nodules based on random forest machine learning algorithm. *PLoS One* (2021) 16:e0260600. doi: 10.1371/journal.pone.0260600

82. Shostak E, Hariri LP, Cheng GZ, Adams DC, Suter MJ. Needle-based optical coherence tomography to guide transbronchial lymph node biopsy. *J Bronchol Interv Pulmonol* (2018) 25:189–97. doi: 10.1097/LBR.0000000000000491

83. Hohert G, Myers R, Lam S, Vertikov A, Lee A, Lam S, et al. Feasibility of combined optical coherence tomography and autofluorescence imaging for visualization of needle biopsy placement. *J BioMed Opt* (2020) 25(10):106003. doi: 10.1117/1.JBO.25.10.106003

84. Gupta A, Harris K, Dhillon SS. Role of bronchoscopy in management of central squamous cell lung carcinoma in situ. *Ann Transl Med* (2019) 7:354. doi: 10.21037/atm.2019.04.36

85. Goorsenberg A, Kalverda KA, Annema J, Bonta P. Advances in optical coherence tomography and confocal laser endomicroscopy in pulmonary diseases. *Respiration* (2020) 99:190–205. doi: 10.1159/000503261

86. Nandy S, Helland TL, Roop BW, Raphaely RA, Ly A, Lew M, et al. Rapid non-destructive volumetric tumor yield assessment in fresh lung core needle biopsies using polarization sensitive optical coherence tomography. *BioMed Opt Express* (2021) 12:5597–613. doi: 10.1364/BOE.433346

87. Zhu Y, Gao W, Zhou Y, Guo Y, Guo F, He Y. Rapid and high-resolution imaging of human liver specimens by full-field optical coherence tomography. *J BioMed Opt* (2015) 20:116010. doi: 10.1117/1.JBO.20.11.116010

88. Zhu Y, Gao W, Guo Z, Zhou Y, Zhou Y. Liver tissue classification of en face images by fractal dimension-based support vector machine. *J Biophoton* (2020) 13: e201960154. doi: 10.1002/jbio.201960154

89. Nyúl-Tóth Á, Tarantini S, Delfavero J, Yan F, Balasubramanian P, Yabluchanskiy A, et al. Demonstration of age-related blood-brain barrier disruption and cerebromicrovascular rarefaction in mice by longitudinal intravital two-photon microscopy and optical coherence tomography. *Am J Physiol Heart Circ Physiol* (2021) 320:H1370–h1392. doi: 10.1152/ajpheart.00709.2020

90. Hartmann K, Stein KP, Neyazi B, Sandalcioglu IE. Microscope integrated optical coherence tomography of a cerebral arachnoid cyst: A new technique to increase intraoperative security. *J Clin Neurosci* (2020) 82:29–31. doi: 10.1016/j.jocn.2020.10.008

91. Tsai MT, Zhang JW, Wei KC, Yeh CK, Liu HL. Assessment of temporary cerebral effects induced by focused ultrasound with optical coherence tomography angiography. *BioMed Opt Express* (2018) 9:507–17. doi: 10.1364/BOE.9.000507

92. Ren C, Zeng X, Shi Z, Wang C, Wang H, Wang X, et al. Multi-center clinical study using optical coherence tomography for evaluation of cervical lesions in-vivo. *Sci Rep* (2021) 11:7507. doi: 10.1038/s41598-021-86711-3

93. Placzek F, Cordero Bautista E, Kretschmer S, Wurster LM, Knorr F, González-Cerdas G, et al. Morpho-molecular ex vivo detection and grading of non-muscle-invasive bladder cancer using forward imaging probe based multimodal optical coherence tomography and raman spectroscopy. *Analyst* (2020) 145:1445–56. doi: 10.1039/C9AN01911A

94. Ma Y, Xu T, Huang X, Wang X, Li C, Jerwick J, et al. Computer-aided diagnosis of label-free 3-d optical coherence microscopy images of human cervical tissue. *IEEE Trans BioMed Eng* (2019) 66:2447–56. doi: 10.1109/TBME.2018.2890167

95. Zeng X, Zhang X, Li C, Wang X, Jerwick J, Xu T, et al. Ultrahigh-resolution optical coherence microscopy accurately classifies precancerous and cancerous human cervix free of labeling. *Theranostics* (2018) 8:3099–110. doi: 10.7150/thno.24599

frontiers | Frontiers in Oncology

REVIEW
published: 27 July 2022
doi: 10.3389/fonc.2022.892056

# Applications of Artificial Intelligence Based on Medical Imaging in Glioma: Current State and Future Challenges

**Jiaona Xu[1†], Yuting Meng[1†], Kefan Qiu[1†], Win Topatana[2], Shijie Li[2], Chao Wei[3], Tianwen Chen[4], Mingyu Chen[2*], Zhongxiang Ding[5*] and Guozhong Niu[4*]**

*[1] Hangzhou First People's Hospital, Zhejiang University School of Medicine, Hangzhou, China, [2] Department of General Surgery, Sir Run-Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, China, [3] Department of Neurology, Affiliated Ningbo First Hospital, Ningbo, China, [4] Department of Neurology, Affiliated Hangzhou First People's Hospital, Zhejiang University School of Medicine, Hangzhou, China, [5] Department of Radiology, Affiliated Hangzhou First People's Hospital, Zhejiang University School of Medicine, Hangzhou, China*

OPEN ACCESS

**Edited by:**
*Rathan Subramaniam,*
*University of Otago, New Zealand*

**Reviewed by:**
*Radka Stoyanova,*
*University of Miami, United States*
*Jiawen Zhang,*
*Fudan University, China*

***Correspondence:**
*Mingyu Chen*
*mychen@zju.edu.cn*
*Zhongxiang Ding*
*hangzhoudzx73@zju.edu.cn*
*Guozhong Niu*
*ngz001@163.com*

*†These authors have contributed equally to this work*

**Specialty section:**
*This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology*

**Received:** *24 March 2022*
**Accepted:** *22 June 2022*
**Published:** *27 July 2022*

**Citation:**
*Xu J, Meng Y, Qiu K, Topatana W, Li S, Wei C, Chen T, Chen M, Ding Z and Niu G (2022) Applications of Artificial Intelligence Based on Medical Imaging in Glioma: Current State and Future Challenges. Front. Oncol. 12:892056. doi: 10.3389/fonc.2022.892056*

Glioma is one of the most fatal primary brain tumors, and it is well-known for its difficulty in diagnosis and management. Medical imaging techniques such as magnetic resonance imaging (MRI), positron emission tomography (PET), and spectral imaging can efficiently aid physicians in diagnosing, treating, and evaluating patients with gliomas. With the increasing clinical records and digital images, the application of artificial intelligence (AI) based on medical imaging has reduced the burden on physicians treating gliomas even further. This review will classify AI technologies and procedures used in medical imaging analysis. Additionally, we will discuss the applications of AI in glioma, including tumor segmentation and classification, prediction of genetic markers, and prediction of treatment response and prognosis, using MRI, PET, and spectral imaging. Despite the benefits of AI in clinical applications, several issues such as data management, incomprehension, safety, clinical efficacy evaluation, and ethical or legal considerations, remain to be solved. In the future, doctors and researchers should collaborate to solve these issues, with a particular emphasis on interdisciplinary teamwork.

**Keywords: artificial intelligence, medical imaging, neural tumors, glioma, radiomics, machine learning, deep learning**

## INTRODUCTION

Glioma is the most common histological type of primary central nervous system cancer, accounting for 81% of all malignant brain tumors (1). Astrocytomas, oligodendrogliomas, oligoastrocytomas, and ependymomas are all types of gliomas. The World Health Organization (WHO) defines gliomas into four categories; the first two grades and the last two grades are further classified as low-grade glioma (LGG) and high-grade glioma (HGG). The poor 5-year overall survival (OS) rate for WHO grade IV glioma patients are 6.8% (2, 3). Glioblastoma (GBM) is the most aggressive kind of grade IV astrocytoma, accounting for 45% of gliomas and the 5-year OS rate of GBM patients is 5%. Treatment for gliomas generally comprises surgical excision, radiation, and temozolomide chemotherapy. Previous randomized clinical studies indicated that the addition of tumor-treating fields to routine treatment increased life expectancy by 4 months (4, 5).

Glioma diagnosis and treatment mostly involve imaging, segmentation and localization, grading, pathology, gene acquisition, and post-treatment recurrence monitoring (6, 7). Tumor treatment and decision-making are difficult due to the heterogeneity of tumors. Therefore, the rise of artificial intelligence (AI) has significantly alleviated doctors' loads (8, 9). AI enables physicians to examine therapeutically important material that is buried inside massive volumes of data (10). Precision medicine is based on artificial intelligence, a relatively new technique to diagnose and treat a disease that considers various factors such as genetics, environment, and lifestyle. Magnetic resonance imaging (MRI), positron emission tomography (PET), and spectral imaging of the brain all contain a wealth of structural and functional information that can be analyzed by AI algorithms for glioma patient management and decision-making (11). However, neurologists should be aware of its limitations, since the use of algorithms raises concerns regarding transparency, privacy, data encryption, and licensing (12). Additionally, doctors and scientists must bridge gaps in one another's subject expertise (13).

The purpose of this review is to (1) provide an overview of AI technology and its applications in medical imaging analysis; (2) summarize the application and performance of AI-based on MRI, PET, and spectral images in glioma; and (3) discuss future challenges and directions for AI applications in the field of neural tumors.

# 1 ARTIFICIAL INTELLIGENCE

AI broadly refers to the capacity of computers to emulate intelligent tasks, such as explicit rule-based systems and computer algorithms that do not require hard-coded rules (14). It was first proposed by an American computer scientist John McCarthy in 1956 (15). Subsequently, machine learning (ML), which falls under the umbrella of AI, has emerged and been applied in various fields. In the past decade, deep learning (DL), a new model of ML, has shown great potential for applications in a broader range of domains, leading to the third AI boom (16, 17) (**Figure 1**).

## 1.1 Machine Learning

ML is a subfield of AI that integrates algorithms and statistical models trained on sample data, allowing computers to learn from previously-stored "training" data without explicitly programming to anticipate new data points (18). ML can be classified as supervised, unsupervised, semi-supervised, and reinforcement learning. To forecast a regression or classification, supervised learning algorithms must be trained on a labeled dataset (19). The most often used supervised approaches include support vector machine (SVM), linear and logistic regression (LR), random forest (RF), decision trees, and Bayesian networks (20). Unsupervised learning algorithms can discover patterns by grouping unlabeled datasets or reducing data. Gaussian mixture modeling, affinity propagation, mean shift, K-mean clustering, and hierarchical clustering are all frequently used techniques. Semi-supervised learning is a technique that combines labeled and unlabeled data. It is a hybrid of supervised and unsupervised learning. Reinforcement learning is a machine learning-enhanced decision-making technique that develops algorithms for a specific task and learns from future errors and successes to reinforce learning (21).

Since the 1980s, ML has been used to create accurate predictions and classifications based on input data in different disciplines, including military research, life science, and clinical practice. This substantially contributed to the advancement of several fields and allowed AI development to again reach its pinnacle after the 1950s (22). However, the construction of every ML model requires intricate feature engineering, resulting in a convoluted workflow. Besides, the accuracy of ML is not satisfactory. Thus, the breadth and extent of ML applications are restricted, leading to the creation of DL (16, 23).

## 1.2 Deep Learning

Since the 2010s, the advent of DL has fundamentally altered the traditional model, in response to the past two AI booms (16). DL is a subset of ML that derives its technology from the artificial neural network (ANN) (24). In comparison to ML approaches, DL algorithms can identify underlying patterns in data without the requirement to extract individual features. The layer-by-layer updating of DL weights aids in the training of DL systems, while the ML weights are updated concurrently. The primary DL-based networks include a convolutional neural network (CNN), deep neural network (DNN), recurrent neural network (RNN), deep auto-encoder (DA), deep belief network (DBN), and deep

---

**Abbreviations:** WHO, World Health Organization; LGG, low-grade glioma; HGG, high-grade glioma; OS, overall survival; GBM, glioblastoma; AI, artificial intelligence; MRI, magnetic resonance imaging; PET, positron emission tomography; ML, machine learning; SVM, support vector machine; LR, logistic regression; RF, random forest; DL, deep learning; ANN, artificial neural network; CNN, convolutional neural network; DNN, deep neural network; RNN, recurrent neural network; DA, deep auto-encoder; DBN, deep belief network; DBM, deep Boltzmann machine; GAN, generative adversarial network; VAE, variation auto-encoder; VGG, Visual Geometry Group; TL, transfer learning; CT, computed tomography; ROI, region-of-interest; AUC, area under the receiver operating characteristic curve; DSC, dice similarity coefficient; T1, T1-weighted; FLAIR, fluid-attenuated inversion recovery; T2, T2-weighted; DWI, diffusion weighted imaging; DTI, diffusion tensor imaging; DKI, diffusional kurtosis imaging; PWI, perfusion weighted imaging; ASL, arterial spin labeling; DCE, dynamic contrast-enhanced; MRS, magnetic resonance spectroscopy; BraTS, Brain Tumor Segmentation; T1c, T1-weighted contrast-enhanced; XGBoost, eXtreme Gradient Boosting; IDH, isocitrate dehydrogenase; MGMT, methylation of O6-Methylguanine-DNA methyltransferase; EGFR, epidermal growth factor receptor; TERT, telomerase reverse transcriptase promoter; CDKN, cyclin-dependent kinase inhibitor; ATRX, alpha thalassemia/mental retardation syndrome X-linked; TP53, tumor protein 53; LASSO, least absolute shrinkage and selection operator; MPRAGE, Magnetization Prepared Rapid Gradient Echo; RTKII, receptor tyrosine kinase II; VEGF, vascular endothelial growth factor; PsP, pseudoprogression; TTP, true tumor progression; TCIA, imaging archive; IVIM, intravoxel incoherent motion; 18F-FDG, [18F]-fluorodeoxyglucose; 11C-MET, [11C]-methyl-L-methionine; 18F-FET, [18F]-fluoro-ethyl-tyrosine; 18F-FDOPA, [18F]-fluoro-L-phenylalanine; AA-PET, amino acid PET; RANO, response assessment in neuro-oncology; TBR, tumor-brain ratio; TTP, time-to-peak; PCNSL, primary central nervous system lymphoma; LOOCV, leave-one-out cross-validation; IS, infrared spectroscopy; RS, Raman spectroscopy; FS, fluorescence spectroscopy; HI, hyperspectral imaging; MRSI, agnetic resonance spectroscopy imaging.

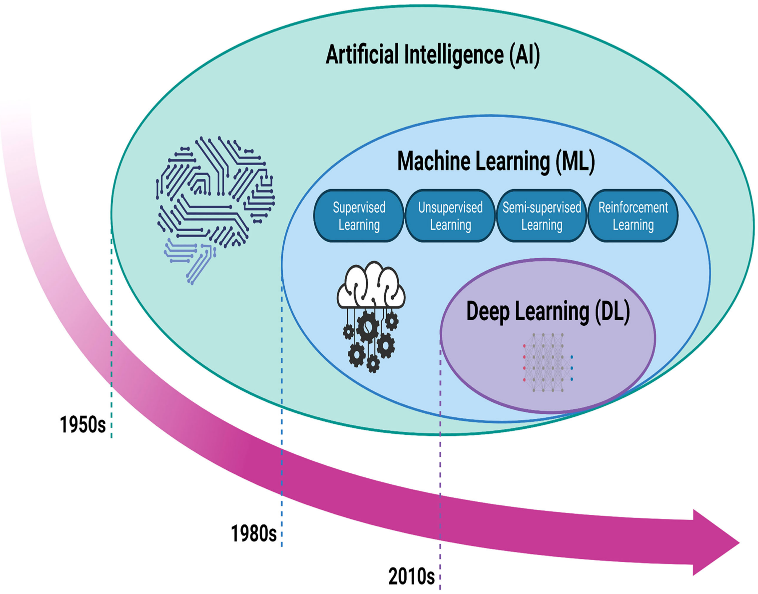


**FIGURE 1** | Artificial intelligence methods and timeline. Machine learning is a form of artificial intelligence that could be classified as supervised learning, unsupervisd learning, semisupervised learning, and reinforcement learning. Deep learning is a form of machine learning. AI: artificial intelligence; ML: machine learning; DL: deep learning.

Boltzmann machine (DBM). Apart from these, generative adversarial network (GAN) and variation auto-encoder (VAE) are two recent approaches for generative and unsupervised learning (25). CNN performs exceptionally well in picture identification; convolutional and pool layers extract obvious information, while fully connected layers conduct final classification. For comparison, CNN's approach encompasses all current ImageNet Classification Challenge winners, with a category mistake rate of 3.6% to date. The development of deep learning models has increased the number of layer designs and the number of model architectures, loss functions, and optimizers available for network construction. Due to the unlimited range of potential computational networks, a significant number of designs have been suggested (for example, AlexNet, ZeNet, Visual Geometry Group (VGG) Net, GoogLeNet, ResNet, DenseNet, Super Resolution CNN, and U-net, among others). Transfer learning (TL) is a subset of DL, and because the weights generated from these networks trained on ImageNet can be applied to different tasks, such as medical pictures, this AI can significantly cut training time (26).

In conclusion, constructing DL models is more time-efficient, simpler, and can achieve greater performance compared to ML. Moreover, DL is readily adaptable to various domains and applications due to TL. Although the DL establishment procedure is straightforward, it requires huge data sets and expensive hardware equipment, therefore ML remains a viable option for smaller data sets (27). Additionally, on a task-specific basis, a tailored image-naive architecture may outperform a DL architecture (16).

## 2 AI IN MEDICAL IMAGING

Over the past few decades, medical imaging techniques including computed tomography (CT), MRI, PET, and ultrasound have been used for early detection, diagnosis, and treatment of diseases (28). In clinical settings, the majority of medical image interpretation has been performed by human specialists such as radiologists and physicians (27). Due to the varying levels of expertise among physicians and the possible exhaustion of human specialists, clinical application of medical imaging has not yielded flawless outcomes. The situation has been improved by the use of AI (29). Following the progression of AI development, ML was initially applied to analyze medical imaging. However, developing ML models necessitates those medical specialists to give well-described regularities or patterns inherent in data, which is a challenge for non-experts in computer science to apply ML to investigate their studies (30). Consequently, DL has been developed and widely used in medical imaging in recent years. Instead of manually extracting features, DL can autonomously find meaningful and useful features in datasets allowing nonexperts in AI to effectively use DL for their research. Besides, with sufficient training data, DL models can achieve greater accuracy (31).

As different forms of AI techniques continue to be applied to medical imaging, radiomics has arisen. Radiomics is the application of computer image processing to transform region-of-interest (ROI) image data into mineable high-dimensional feature data. AI models are constructed based on the extracted feature data to make disease-related diagnoses and predictions

(32). This AI-assisted technology is of great interest to doctors and is widely used in clinical research. Radiomics can elicit measurable objective data that has previously been unavailable and establish its relationship to underlying biological processes (33). Radiomics may be mainly classified into two types: feature-based and deep learning-based radiomics (34). The workflow for feature-based radiomics consists primarily of picture preprocessing, tumor segmentation, feature extraction, and feature selection, followed by the establishment and evaluation of a mathematical model (35). By utilizing various network topologies, deep learning-based radiomics procedures discover and detect classification-related patterns in picture data (36). The feature structure is then merged to form higher-level abstraction features. Finally, the retrieved features can be evaluated further by the network or subjected to a model-building process that is used in feature-based radiomics (**Figure 2**). To assess the AI technique, the model may be tested either internally (through cross-validation or bootstrapping) or externally (by supervised learning). After training and testing the model, it is desirable to apply it to a third dataset, referred known as the external validation dataset. External validation datasets serve as the gold standard for assessing the performance, robustness, and dependability of AI models. Statistical metrics like as accuracy, area under the receiver operating characteristic curve (AUC), sensitivity, specificity, positive/negative predictive values, and dice similarity coefficient (DSC) or dice score can be used to evaluate the effectiveness of AI systems (37).

## 3 APPLICATIONS OF AI-BASED ON MEDICAL IMAGING IN GLIOMA

Neuroimaging techniques, such as contrast-enhanced CT, MRI, PET, and spectral imaging, have been widely applied for the detection, treatment, and prognostic prediction of glioma. However, the numerous amounts of data generated by these techniques and the heterogeneity of tumors are miserable for physicians. AI-based medical imaging could help to release physicians from these large amounts of data by integrating the similarity of these figures and providing directions. This section will mainly demonstrate the strengths and shortages of the application of AI-based MRI, PET, and spectral imaging in glioma.

### 3.1 Magnetic Resonance Imaging

MRI reflects the tumor pathophysiological environment at the voxel level by utilizing geometric, histogram, and texture analysis methods for quantification and prediction of image-based biomarkers *via* radiomics. Compared to biopsy, MRI is a non-invasive method, which could provide relatively comprehensive information on tumors. Whereas MRI can help to get rid of

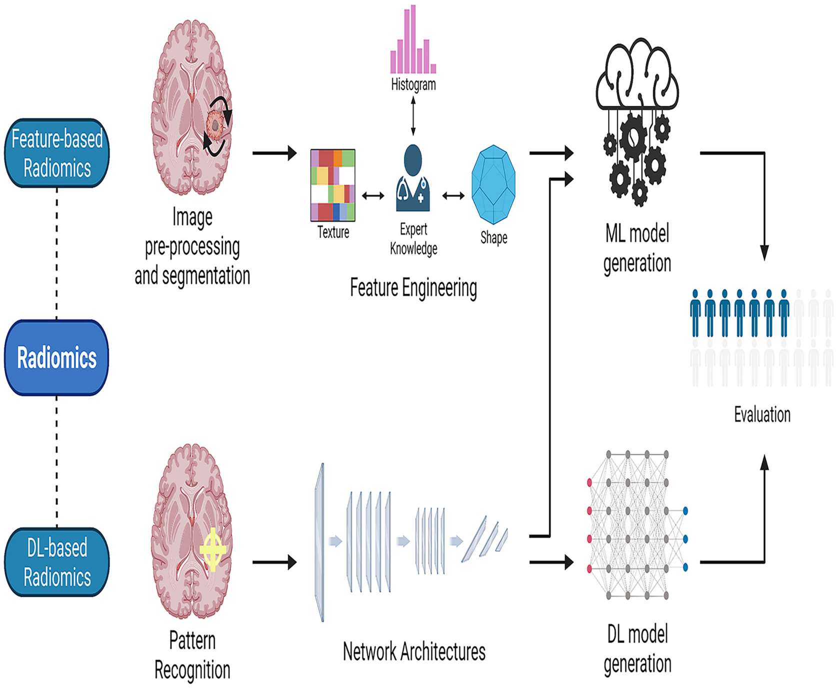


**FIGURE 2** | The workflow of radiomics. Radiomics may be divided into two categories: feature-based radiomics and deep learning-based radiomics. The workflow for feature-based radiomics begins with image preprocessing, tumor segmentation, feature extraction, and feature selection, and concludes with the construction and assessment of a mathematical model. In deep learning-based radiomics, different network architectures are used to find the most relevant features from the input data. Finally, the retrieved features can be processed further by the network for analysis and classification, or they can leave the network and used to generate models in a manner similar to the feature-based radiomics technique by employing different classifiers. ML, machine learning; DL, deep learning.

ionizing radiation and interference from bone artifacts when compared to contrast-enhanced CT. Besides, multiple sequences such as T1-weighted (T1) and T2-weighted (T2) MRI can be applied, which means more information can be obtained through MRI. Among these sequences, T1 images often depict the glioma boundaries, and fluid-attenuated inversion recovery (FLAIR) and T2 images more clearly depict the tumor core (38). The integrity of the blood-brain barrier (BBB) is disrupted in almost all high-grade gliomas, which means that the gadolinium-based contrast agents (GBCA) injected from the vein can successfully enter the extravascular extracellular space of the brain, manifesting as contrast-enhancing hyperintense regions on T1 sequences (39). Along with the T1 and T2 sequences, several other sequences have also been used to comprehensively evaluate the state of glioma (40). In detail, diffusion-weighted imaging (DWI), diffusion tensor imaging (DTI), and diffusional kurtosis imaging (DKI) can indicate changes in cell density, membrane permeability, and tissue microstructure; perfusion-weighted imaging (PWI) can detect changes in the microcirculation and cell proliferation (41); magnetic resonance spectroscopy (MRS) can reveal the metabolic status of malignancies directly, which is most closely connected to gene expression regulation, suggesting the combination of these two methods (42). Herein, we will discuss the application of AI-based MRI in glioma from the following four aspects: tumor segmentation and classification, molecular marker prediction, molecular marker prediction, and tumor cell analysis (**Supplementary Table 1**).

### 3.1.1 Tumor Segmentation and Classification

Glioma is classified into four subtypes: enhanced area, non-enhanced region, necrosis area, and edema area. Several algorithms have been used to segment glioma. Among them, the outstanding performance of CNN has been well known in glioma segmentation, with an accuracy greater than 80-90%. Fu et al. (43) evaluated their multipath denseNet architecture based on 3D CNNs using the Brain Tumor Segmentation (BraTS) 2019 dataset and obtained a DSC of 0.922. Along with the CNN model mentioned above, other AI methodologies have also been applied in glioma segmentation. Another study combined Superpixel fuzzy clustering with the lattice Boltzmann technique can reach a disc of 0.93 (BraTS 2017) (44), demonstrating that the approach is resistant to noise, initialization, and strength inhomogeneity. Besides, Amin et al. (45) proposed a technique merging Local Binary Pattern and Gabor Wavelet Transform features, and generated dices of 0.96 (BraTS 2013), 0.98 (BraTS 2015), and 0.95 (local dataset). In summary, segmentation of glioma is a time-consuming and subjective task through the current manual ways. Through AI-based MRI, these shortages can be largely overcome, and subsequently, radiomics can be performed. Despite this, the large heterogeneity of HGG and the low proliferative state of LGG still bring a huge challenge to this task (71, 72). Besides, the various outcomes in the same datasets caused by different ML methodologies are a major concern for the application of ML in clinical. For instance, the results generated by a two-stage cascaded U-Net (73) and an RDAU-Net (74) using the BraTS 2019 training dataset which comprises 259 cases of HGG and 76 cases of LGG are various.

Additionally, the value of MRI in the grading and categorization of glioma has also been assessed according to its pathophysiology, molecular composition, and transcriptional activity. DL-based MRI, particularly CNN, performed well in a study of glioma classification and grading. For instance, Quon et al. created a modified ResNeXt-50-32x4d architecture to detect and classify gliomas into distinct pathological sub-types using T2 images (46), and this model demonstrated an AUC of 99% for tumor detection and 92% for glioma classification. In 2020, Basha et al. proposed a novel Harris Hawks optimization algorithm for evolving CNN architecture and investigated the classification and grading of brain tumors using two datasets; the former contains 8.000 brain tumors with four grades and 8.000 healthy MRI images, while the latter contains 4.908 MRI images with glioma, pituitary, and meningioma; the accuracy was greater than 95% in all experiments. Luo and colleagues (47) examined the utility of high-throughput network characteristics derived from the 3D U-net for histological and molecular subtype prediction in three cohorts of 655 glioma patients using conventional MRI. For histological diagnosis and molecular subtyping, the novel picture signature-based radiomics model achieved accuracies of 89.8% and 86.1% in the cross-validation cohort and 83.9% and 80.4% in the independent testing cohort. Overall, these studies indicated the high accuracy generated by DL in the grading of glioma.

Besides DL, other AI technologies also performed well in glioma classification. For example, Le et al. (48) identified transcriptome subgroups in GBM patients using conventional MRI in two cohorts of 120 patients. Model generation was performed using an eXtreme Gradient Boosting (XGBoost) machine classifier, and the model was constructed using 13 radiomics features selected from 704 handcrafted radiomics features achieved 70.9%, 73.3%, 88.4%, and 88.4% accuracy in predicting classical, mesenchymal, neural, and proneural subtypes, respectively. Lu and co-workers (49) achieved an accuracy of 81.8% after fivefold cross-validation using an SVM classifier based on radiomics features from multimodal MRI in 456 glioma patients for the classification of five molecular subtypes; this accuracy was increased to 89.2% when combined with histological diagnosis and MR radiomics.

In general, many AI systems can accurately detect and grade gliomas using picture data. However, because various studies use different data and defining criteria, it's impossible to compare them, and it's unclear which algorithm is the most effective.

### 3.1.2 Molecular Marker Prediction

WHO included molecular and histological characteristics in the classification of brain cancers for the first time in 2016, and in 2021, WHO made significant revisions to the categorization of tumors, emphasizing the importance of molecular detection (75). The updated WHO 2016 classification of central nervous system malignancies stresses the prognostic significance of molecular characteristics such as the isocitrate dehydrogenase (IDH) genotype or the 1p/19q chromosomal arm heterozygous deletion (3). 2021 WHO classification approves methylome classifiers for a variety of CNS tumor types and subtypes and promoter methylation of O6-Methylguanine-DNA methyltransferase (MGMT) is related to an improved response to temozolomide therapy and a longer OS. Numerous studies have also demonstrated the predictive abilities of certain molecules. For example, research suggests that grade II or III IDH wild-type astrocytomas may harbor chromosomal +7/-10, epidermal growth factor receptor (EGFR) amplification, and/or telomerase reverse transcriptase promoter (TERT) alterations, with the same prognosis as GBMs (76). The detection of homozygous cyclin-dependent kinase inhibitor (CDKN) 2A/B deletion is critical for properly diagnosing and prognosing patients with diffuse astrocytomas caused by IDH mutations. In diffuse astrocytomas, IDH mutations are related to alpha thalassemia/intellectual disability syndrome X-linked (ATRX) and tumor protein 53 (TP53) functional loss mutations (77). ATRX mutations are mutually exclusive with co-deletion of 1p/19q and are associated with oligodendrocytes (78). TERT and ATRX are telomere maintenance proteins (79, 80).

Recent years have seen a surge in interest in radiogenomics. Radiogenomics needs the establishment of correlations between quantitative or qualitative imaging aspects and genomic data derived from tissue analysis and other clinical data in order to enable the development of imaging alternatives to genetic testing (81, 82). Radiomics can help to distinguish IDH-mutant co-deleted 1p/19q tumors (oligodendrogliomas) from IDH-mutant non-co-deleted 1p/19q tumors (astrocytomas). Researchers reported that the combination of a near-complete or complete hyperintense signal on a T2 sequence and a hypointense signal on a FLAIR (except a potential hyperintense peripheral rim) possesses a 100% predictive value of IDH-mutant astrocytomas, which was termed as T2-FLAIR mismatch (83). Researchers further verified the specificity of this mismatch for anaplastic astrocytomas and diffuse through a retrospective study containing patients with diffuse oligodendroglioma (IDH-mutant 1p/19q co-deleted), diffuse astrocytoma (IDH-mutant), anaplastic oligodendroglioma (IDH-mutant 1p/19q co-deleted), anaplastic astrocytoma (IDH-mutant), and IDH-WT (Glioblastoma-like) (84). It was revealed that the T2-FLAIR mismatch is present in four of five anaplastic astrocytoma tumors, 34 of 70 diffuse astrocytoma tumors, and 0 of 79 other three types of tumors, confirming the 100% specificity differentiating astrocytomas from other LGGs, which has been further verified in other two studies (85, 86). In addition to the T2-FLAIR mismatch, researchers created a model consisting of T1, T2-weighted FLAIR, and an apparent diffusion coefficient (ADC), and reported that the model can differentiate MGMT methylated tumors from non-methylated tumors with an AUC of 0.925 and 0,902 in the training and validation cohort, respectively. This indicated the efficiency of MRI in the prediction of molecular markers. Employing AI-based MRI can help clinicians to clear changes in molecular markers easily (87, 88). In general, the majority of research employed MRI to predict glioma gene mutations with DL (particularly CNN), RF, least absolute shrinkage and selection operator (LASSO), and SVM technologies to obtain strong predictive performance with an accuracy of greater than 80 - 90%. For example, Choi et al. (57) predicted the IDH genotype with an accuracy of 92.8% and 91.7% in the validation and test sets, respectively, using an RNN application based on dynamic susceptibility contrast MRP from 463 patients with gliomas. The H3- -K27M mutation status prediction model based on CNN features and the SVM classifier was tested by Liu et al. in a group of 55 patients with preoperative T1-magnetization prepared rapid gradient echo (MPRAGE) images MRI, and the results indicated an accuracy of 95% upon fivefold cross-validation (60). For the prediction of deletion of Chromosomal Arms 1p/19q, Akkus and co-workers (62) used a multi-scale CNN based on T1c and T2 pictures from 159 LGGs. Using TL and previously trained 3D-dense-UNets on T2 images, Yogananda and colleagues (58) were able to accurately predict the MGMT promoter methylation status in 247 individuals. Similarly, several studies used CNN and/or RF models to predict molecular markers (such as TERT (61), 7/10 aneuploidies, CDKN2 family mutations (66), receptor tyrosine kinase II (RTKII) (67), and tumor proliferation marker (Ki-67) (63) in glioma patients' MRI and reached a high degree of accuracy. Additionally, LASSO regression and/or SVM models based on MRI correctly predicted additional molecular indicators such as ATRX mutation (59), TP53 status (64), and vascular endothelial growth factor (VEGF) expression (65).

### 3.1.3 Response Assessment and Prognosis Prediction

AI has been used in MR imaging sequences to assess response and predict survival in gliomas, excluding the prediction of molecular markers. A significant challenge following chemoradiotherapy is the presence of radiation-induced side effects such as pseudoprogression (PsP), a late benign therapeutic effect that mimics true tumor progression (TTP) at the tumor site or resection margin, which occurs in approximately one-third of GBMs and is usually stable without further treatment (89). Clinicians face significant hurdles because of this discrepancy between PsP and TTP.

SVM has been successfully used to measure response and predict survival in gliomas. Li and co-workers (51) demonstrated a 92% accuracy in differentiating between PsP and TTP after tenfold cross-validation using an SVM classifier based on deep convolutional generative adversarial networks and AlexNet radiomics feature learning from DTI. Conventional MRI data from two institutions, comprising 105 GBMs, was utilized by Ismail and colleagues (52) to distinguish between PsP and TTP. An SVM classifier was utilized to evaluate the test cohort after extracting 30 shape features, and the training and test cohorts had accuracy rates of 91.5% and 90.2%, respectively.

Moreover, some studies have reported the accuracy of AI in predicting glioma prognosis. The cancer imaging archive (TCIA) and local test cohorts were used by Pan et al. to predict the OS using ML techniques with C-indexes of 0.70 and 0.76, respectively, for multiparameter MRI of 152 GBMs (53). When radiomic characteristics were paired with preoperative clinical risk factors (C-index = 0.76 in the TCIA and test cohort), the impact of OS prediction was substantially enhanced. Sanghani and colleagues (56) found that an SVM classifier based on textural characteristics, tumor shape, and volumetric data from conventional MRI was able to accurately predict OS in two- and three-class survival groups following a 5-fold cross-validation. Similarly, Chang and colleagues (55) predicted OS with good accuracy using an RF MRI feature selector and a kernel SVM or neural network classifier. Furthermore, another study demonstrated significant accuracy in identifying survival-relevant high-risk subregions in MRIs from GBMs using the K-means clustering methodology (54).

In summary, ML algorithms are more than 80% accurate in predicting glioma outcomes *via* imaging. One way to improve the efficiency of AI-based MRI in Response assessment and prognosis prediction is to overcome the limitations of MRI. The major disadvantage is that the treatment-related changes can affect MRI results, regardless of the time of evaluation. In this situation, some entities such as radiation necrosis (RN), pseudoprogression (PSP), and pseudoresponse can be introduced (90). Notably, oedema and necrosis caused by postoperative reaction and radio- or chemotherapy could be misinterpreted as disease progression due to the increase in T2/FLAIR signal (91). Therefore, it is suggested to introduce a reliable imaging technique to increase the accuracy of MRI.

### 3.1.4 Tumor Cell Analysis

Non-enhanced aggressive tumors are difficult to detect with MR enhancement but can be aided by assessing a variety of biophysical characteristics. Hu and colleagues (68) trained a TL model using dynamic susceptibility contrast MR imaging and DTI data from 18 GBMs from a single clinical institution and on 82 image-recorded biopsy samples. With a Pearson correlation value of 0.88 and a mean absolute error of 5.66 percent, the tumor cell density could be predicted. In another study, images from High-Resolution Magic Angle Spinning Nuclear MRS of glioma and control samples were analyzed using an RF model with AUCs of 85.6% and 87.1% to differentiate tumor cells and benign samples from controls and malignant samples (69). Similarly, Fathi Kazerooni and colleagues (70) differentiated subregions of brain gliomas in Fifty-one tissue specimens from 10 patients using conventional MRI, DWI, DTI, intravoxel incoherent motion (IVIM), and dynamic susceptibility contrast MRI. An SVM classifier was used to generate models, and a model based on 15 MRI-based parameters had an AUC of greater than 0.90 for identifying the three subregions (active tumor, infiltrative edema, and normal tissue).

Tumor cell analysis enables the direction of postoperative targeted therapy and the assessment of tumor margins intraoperatively. At the moment, artificial intelligence is still in its infancy. Due to financing and data issues, there are still very few relevant studies available now. Future studies can be conducted to improve the use of AI and the verification of cell analysis.

In general, AI has been extensively applied in glioma MRI, including tumor segmentation and classification, molecular marker prediction, and tumor cell analysis. With the rapid advancement of AI, deep learning in image analysis demonstrates both its advantages and limits. AI will eventually assist in the integration of data from disparate sources (clinical examination, other medical imaging, and pathology) to guide therapy and prognosis.

## 3.2 Positron Emission Tomography

As described above, the application of AI-based MRI shows excellent outcomes in glioma. However, MRI may not always be able to answer three essential questions: evaluation of the initial characterization of the brain lesion, monitoring of therapies to clear changes induced by recurrence/progression and treatment, and evaluation of treatment efficacy (92). Furthermore, one of the main advantages of PET is that the radiotracers used for PET are in most cases independent of disruption of the blood-brain barrier (BBB) as opposed to MRI, which is especially useful in LDH (92, 93). Overall, PET provides insights into glioma that exceed MRI and that can be applied for noninvasive grading, differential diagnosis, mapping the extent of tumor involvement, designing surgery and radiotherapy methods, and prognostic prediction.

PET mainly uses [18F]-fluorodeoxyglucose (18F-FDG) and radioactively labeled amino acids as radioactive tracers. Compared with 18F-FDG, the radioactive labeled amino acid, such as [11C]-methyl-L-methionine (11C-MET), [18F]-fluoro-ethyl-tyrosine (18F-FET), 3,4-dihydroxy-6-[18F]-fluoro-L-phenylalanine (18F-FDOPA) show higher contrast in tumor tissues and normal brain tissues (94). Further, the amino acid PET (AA-PET) can provide additional information on the metabolic characteristics of glioma. These two advantages make the United Cooperative produce guidelines encouraging the use of AA-PET for tumor diagnosis and treatment (95, 96), and the response assessment in neuro-oncology (RANO) group made evidence-based recommendations for the use of PET imaging in the planning and monitoring of radiation therapy for glioma patients (97–99). While the tumor-brain ratio (TBR) is currently the gold standard for estimating neoplastic uptake relative to healthy brain tissue in the majority of centers, tracer uptake dynamics, such as slope and time-to-peak, have been shown to increase diagnostic accuracy (100). Dynamic factors were found to be linked with tumor grade, tumor progression, molecular indicators such as IDH gene alterations, and separating patients with actual and false tumor progression in patients with gliomas (3, 101, 102). The following is a summary of recent AI-based PET studies on glioma diagnosis, treatment, and prognosis (**Table 1**).

### 3.2.1 Applications for Diagnosis

Glioma misdiagnosis as another lesion can have a significant impact on patient survival, and although MRI is frequently utilized for the first screening, radiological separation of glioma, primary central nervous system lymphoma (PCNSL), and multiple sclerosis remain challenging. PET is an alternative form of imaging that has been used to assess central nervous system disorders (117). As a result, an increasing number of studies have used AI-based PET to aid in the detection and diagnosis of glioma. For example, 18F-FET-PET imaging may differentiate between multiple sclerosis and WHO grade II-IV glioma with a 91% accuracy by using an SVM classifier,

**TABLE 1 |** Summary of major studies on AI-assisted PET in Glioma.

| Purpose | Ref. | Design ofstudy | Database | Sample size | Performingalgorithm | Modality | Feature | Outcomes (%) | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Accuracy | Sensitivity/ Specificity |
| Detection and segmentation | Blanc-Durand et al, 2018 (134) | Retrospective | Internal | 37 glioma patients | 18F-FET PET | CNN Feature | 3D U-net CNN | Detection: 100;<br>Segmentation: DSC: 82.31 | Detection: 100/100;<br>Segmentation: 88/99 |
| Classification | Kebir et al, 2021 (135) | Retrospective | Internal | 7 multiple sclerosis and 34 glioma patients | 18F-FET PET | TBR | SVM | 91; AUC:94 | 89/100 |
| Classification | Kong et al, 2019 (136) | Retrospective | Internal | 24 lymphoma patients and 53 GBMs | 18F-FDG PET | SUV map | Decision tree | 90.9-97.4; AUC:97.1-99.8 | 90.6-98.1/ 87.5-100 |
| Classification (3-group molecular subtypes) | Matsui et al, 2020 (137) | Retrospective | Internal | 217 LGGs 49/58/100 (IDH-wildtype diffuse astrocytoma/IDH-mutant difuse astrocytoma/ oligodendroglioma) | MRI, PET, and CT | Image and clinical features | residual network | 96.6/68.7 (training/ testing) | NA |
| Discrimination between PsP and TTP | Lohmann et al, 2020 (138) | Retrospective | Internal | 34 glioma patients | 18F-FET-PET | First-order statistics, shape, and texture, Laplacian-of-Gaussian filtered, wavelet-transformed features | RF | Training/ testing: 86/70; AUC:74/74 | 82/90 (training); 100/40 (testing) |
| Discrimination between PsP and TTP | Kebir et al, 2020 (139) | Retrospective | Internal | 44 glioma patients | 18F-FET-PET | TBR and time-to-peak | Linear discriminant analysis | AUC:93 | 100/80 |
| Discrimination between PsP and TTP | Imani et al, 2014 (140) | Retrospective | Internal | 12 grade 2 and 3 gliomas | 18F-FDG PET and MRS | Maximal SUV and multiple 2D maps of choline, creatine | SVM | 92 | 80/100 |
| Discrimination between PsP and TTP | Kebir et al, 2017 (141) | Retrospective | Internal | 14 HGGs | 18F-FET-PET | Textural and conventional features | Clustering based classifier | Positive predictive value: 90 | 90/75 |
| OS prediction | Papp et al, 2018 (142) | Retrospective | Internal | 70 patients with a treatment naive glioma | 11C-MET PET | General and higher-order textural features, *in vivo*, ex vivo, and clinical patient information | K-nearest neighbor classifier | 90; AUC: 91 | 88/95 |
| IDH mutation prediction | Li et all, 2019 (143) | Retrospective | Internal | 127 consecutive gliomas | 18F-FDG PET | Clinical characteristics and the radiomic signature | SVM and multivariate LR | Training/ testing: 79.8/83.7; AUC: 91.1/90 | 78.9/80.4 (training); 92.3/80 (validation) |
| IDH status prediction | Tatekawa et al, 2021 (144) | Retrospective | Internal | 62 treatment-naive glioma patients | Multiparametric MRI and 18F-FDOPA PET | Voxel-wise feature | Two-level clustering and SVM | 76; AUC:81 | NA |
| Classification (HGG and LGG) and IDH status prediction | Kebir et al, 2019 (145) | Retrospective | Internal | 39 gliomas | 11C-MET PET/MRI | TBR | SVM classifier with a linear kernel | Classification: AUC:62;<br>Prediction: AUC:79 | NA |
| MGMT status prediction | Qian et al, 2020 (146) | Prospective | Internal | 86 GBMs | 18F-FDOPA PET | Shape, tumor intensity and tumor texture features | RF | 80 | 100/33 |

*(Continued)*

**TABLE 1** | Continued

| Purpose | Ref. | Design ofstudy | Database | Sample size | Performingalgorithm | Modality | Feature | Outcomes (%) | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Accuracy | Sensitivity/ Specificity |
| Ki-67 prediction | Kong et al, 2019 (147) | Retrospective | Internal | 123 glioma patients<br>82/41 (training/ testing) | 18F-FDG PET | Shape and size, first-order, texture, wavelet, and alternative filtered features | SVM | Training/ validation: 81.7/73.2; AUC:88/76 | 95.6/64.9 (training); 92/43.8 (validation) |

*AI, artificial intelligence; PET, positron emission tomography; Internal, subjects were recruited from insitutional and/or public through media channels; 18F-FET, [18F]-fluoro-ethyl-tyrosine; CNN, convolutional neural network; DSC, dice similarity coefficient; TBR, tumor-brain ratio; SVM, support vector machine; AUC, area under the receiver operating characteristic curve; 18F-FDG, [18F]-fluorodeoxyglucose; SUV, standardized uptake value; IDH, isocitrate dehydrogenase; MRI, magnetic resonance imaging; CT, computed tomography; NA, not available; PsP, pseudoprogression; TTP, true tumor progression; RF, random forest; 2D, two-dimensional; HGG, high-grade glioma; OS, overall survival; 11C-MET, [11C]-methyl-L-methionine; LR, logistic regression; 18F-FDOPA, [18F]-fluoro-L-phenylalanine; LGG, low-grade glioma; MGMT, methylation of O6-Methylguanine-DNA methyltransferase; GBM, glioblastoma.*

according to a study by Kebir et al. (104) In an attempt to identify PCNSL from GBM, Kong et al. (105) used 107 radiomic characteristics from 18F-FDG PET in 77 individuals (24 with lymphoma and 53 with GBM). The decision tree approach algorithm demonstrated great diagnostic performance, according to this study's findings (accuracy 90.9%-97.4%, AUC 97.1%-99.8%). LGG may be classified into three molecular subtypes based on the WHO's 2016 categorization of central nervous system malignancies. The mainstay of care for patients with LGG is surgical excision of the tumor followed by postoperative chemoradiotherapy. Their effectiveness, however, is dependent on the tumor's molecular subtype. Matsui et al. (106) utilized residual networks to predict LGG molecular subtypes using multimodal data from a glioma database, including MRI, PET, and CT, and achieved an overall accuracy of 68.7% for the test dataset.

The above evidence has exhibited the strength of PET in the diagnosis of glioma. However, about 5% of HGG do not show amino-acid tracer uptake (118, 119) and some non-neoplastic lesions, such as vascular malformations, hematomas, inflammatory lesions, and ischemic lesions, can also exhibit unspecific amino-acid uptake (120, 121). Besides, although static 18F-FDG PET has been used for the differentiation of LGG and HGG, overlap can be seen, which may interfere with the judgement (122). Also, static 18F-FDG PET has only a specificity of 56-85% and a sensitivity of 71-80% for the differentiation between LGG and HGG, suggesting the employment of dynamic 18F-FDG PET which can improve the accuracy (95, 123). Therefore, although the application of PET enhances the interpretation of lesions determined by MRI, histological diagnosis and the molecular signature cannot be neglected.

### 3.2.2 Applications for Treatment

Segmentation is a frequently performed operation in medical imaging; automated segmentation significantly reduces the time required for human segmentation. Segmentation objectives such as radiotherapy plans that define the total or biological tumor volume, and surgical plans that quantify the three-dimensional volume of enhancing tumor and surrounding edema are necessary for accurate assessment and monitoring of tumor response and have also demonstrated some independent prognostic value. A 3D U-Net CNN was employed in 37 glioma patients to detect and segment gliomas using 18F-FET PET with 100% detection accuracy and 82.31% DSC (segmentation) (103).

Although PsP is most frequently noticed within the first 12 weeks following the cessation of radiation and chemotherapy (124), it can develop later (125). Detecting PsP in GBMs continues to be an important clinical problem in radiology since it is necessary to avoid continuing ineffective therapy and discontinuation of beneficial treatment. Kebir et al. (108) developed a model for identifying PsP using 18F-FET PET scans from 44 glioma patients and a linear discriminant analysis model with an AUC of 0.93 was utilized. Lohmann et al. (107) used a model for discriminating PsP from TTP by analyzing 18F-FET PET scans from 34 glioma patients. The patient group was separated into a training and a test cohort. The final model used an RF classifier and attained accuracies of 86% and 70% in the training and test data, respectively. In another study, an SVM classifier was developed on twelve post-therapy patients who underwent 18F-FDG PET and MRS to identify brain glioma progression. The classifier's sensitivity and specificity for detecting glioma progression were 80% and 100%, respectively, with an accuracy of 0.92 (109).

### 3.2.3 Applications for Prognosis

PET imaging using radiolabeled amino acid tracers such as 11C-MET and 18F-DOPA is regarded as a potential diagnostic tool for tumor characterization and longitudinal therapy monitoring due to its excellent sensitivity and specificity. Papp et al. (111) assessed the possibility for survival prediction using 11C-MET PET radiomics and clinical patient information in 70 patients with a treatment-naive glioblastoma. The final model incorporated *in vivo*, ex vivo, and clinical patient data and had an AUC of 0.90. Similarly, another study (114) showed a good AUC for IDH status prediction using an SVM classifier while assessing 11C-MET PET scans from glioma patients. Based on 18F-DOPA PET images, RF and SVM models correctly predicted MGMT status (115) and tumor proliferation marker (Ki-67) (116). Additionally, several studies employ a combination of multimodal imaging and machine learning methods to predict tumor genetic markers. For example, Tatekawa et al. (113) performed a radiomics analysis based on multiparametric MRI and 18F-DOPA PET images for the prediction of the IDH status in 62 treatment-naïve glioma patients, a SVM model achieved an AUC of 81% after leave-one-out cross-validation (LOOCV).

Taken together, feature-based PET radiomics has shown promise in the field of neuro-oncology, allowing for the evaluation of more data at a reasonable cost. However, the majority of existing research is retrospective in nature, with insufficient sample sizes and no available database. ML is a technique for fitting statistical models, and its outcomes are sample size-dependent.

## 3.3 Spectral Imaging

Glioma is defined by its proclivity for metastasis and heterogeneity. Due to the lack of specificity of early clinical signs, the majority of glioma patients are frequently overlooked, resulting in patients missing the best treatment window. Histopathology has evolved into the gold standard for classification and diagnosis, whereas molecular pathology has gained increasing attention in the diagnosis and classification of glioma. With the advancement of molecular biology and molecular pathology in 2016, the WHO categorization of recombinant central nervous system malignancies is beneficial for early detection and accurate therapy (3). Spectral imaging is a potential tool for assisting in the histopathological study of cancer samples that contain molecular information. Imaging can be employed for real-time intraoperative evaluation, allowing for earlier detection and more precise intraoperative resection, which is critical for patient survival (126).

### 3.3.1 Infrared spectroscopy

Infrared spectroscopy (IS) is a non-invasive and quick measuring technique used to characterize biological samples and their constituents qualitatively and quantitatively by quantitative detection of molecule internal vibration patterns (127, 128). Several studies have coupled human serum IS with ML methods to identify glioma (129–131). Hands et al. (131) extracted 130 features from Fourier-transform IS pictures of blood samples from 433 individuals with or without glioma. The final SVM classifier model has a sensitivity and specificity of 91.5% and 83.0%, respectively, for detecting glioma. In this test, SVM and RF outperformed other classifiers. Another model was constructed using partial least squares discriminant analysis and synthetic minority over-sampling to classify GBM multiforme and lymphoma from 765 serum samples. The result has a sensitivity of 90.1% and a specificity of 86.3%, respectively (132).

Furthermore, the combination of IS with a microscope enables the spatial distribution of proteins, lipids, nucleic acids, and other compounds in tissue samples to be examined. Peng et al. (133) used Fourier transform infrared microscopy to study 9360 spectra from the tissue of 77 glioma patients. This study employed artificial neural networks to categorize gliomas (HGG and LGG) with higher than 98% accuracy, specificity, and sensitivity. For estimating the secondary structure of proteins, Surowka et al. (134) employed infrared micro-spectroscopy spectral range. ANNs were employed to generate the models, and the accuracy was improved to less than 5%.

### 3.3.2 Raman Spectroscopy

RS is a label-free method that generates spectra by detecting and measuring Raman scattering using narrow-band laser excitation and sensitive spectrometers. For stereotactic brain tumor biopsy, *in vivo* tumor infiltration detection, intra-operative histopathology diagnosis, and molecular categorization, it gives quantitative biochemical information regarding the molecular composition (135–138).

For the creation of a model to grade glioma, Zhou et al. (139) employed label-free visible resonance RS spectra from 125 histologically normal human brain tissues and glioma tissues. The SVM model was able to discriminate normal, LGGs, and HGGs 75.1% of the time. Besides, Pekmezci and coworkers (140) used RS spectral data to differentiate the phenotypes of T-cells and monocytes following incubation with a medium conditioned by GBM stem cells with a variety of genetic backgrounds in three human GBM cell lines. The linear discriminant analysis model was generated using 67% of the dataset (training set) and then verified against 33% of the dataset (test set). The SVM produced sensitivities and specificities of greater than 70% and 67% in the validation and independent test sets, respectively.

### 3.3.3 Fluorescence Spectroscopy

FS offers a comprehensive array of detection tools and procedures for high-grade gliomas that accumulate the endogenous biomarker protoporphyrin IX following exogenous treatment of 5-aminolevulproic acid, boosting tumor tissue fluorescence and directing surgical intervention (141, 142). In ten glioma patients, Valdés and colleagues (143) assessed the possibility of combined FS and reflectance spectroscopy *in vivo* optical data for diagnostic performance during surgery. The SVM model attained an accuracy of 94%. Leclerc et al. (144) used spectral characteristics analysis based on FS to identify healthy tissue from margin tissue in 50 samples from ten patients. A completely automated clustering technique obtained a diagnostic accuracy of 77% in predicting healthy tissues from margin tissues.

### 3.3.4 Hyperspectral Imaging

HI measures the diffuse reflectance of tissue surfaces to generate spectral characteristics that contain both spatial and spectral information (145). Recently, HI has been utilized to identify and diagnose illnesses characterized by alterations in cellular biochemical pathways (146). Urbanos et al. (147) classified tumor tissue in a set of 12 HGGs using thirteen *in-vivo* hyperspectral photos (healthy tissue, tumor, venous blood vessel, arterial blood vessel, and dura mater). Overall accuracies for the three models (RF, SVM, and CNN) ranged from 60% to 95% depending on the training settings. Similarly, Manni and coworkers (148) classified tumor tissue (tumor, healthy tissue, and blood vessels) in 16 tumor patients using 26 *in-vivo* hyperspectral pictures. The hybrid 3D-2D CNN models achieved an overall accuracy of 80%. Ortega et al. (149) employed 527 high-resolution pictures to detect GBM in non-tumor brains and GBM samples from 13 individuals. The CNN models had an average sensitivity and specificity of 88% and 77%, respectively.

In conclusion, these investigations demonstrate that spectral image-based AI is beneficial for diagnosing and applying glioma tissue samples intraoperatively. Due to spectrum imaging's unpopularity, there are few studies and their sample sizes are modest. As a potential intraoperative quick diagnostic method, more research may be directed toward developing applicable AI software.

### 3.3.5 Magnetic Resonance Spectroscopy Imaging

MRSI is a non-invasive technique for evaluating the spatial distribution of metabolic changes in the brain (150, 151). It can provide information on neuron integrity, neurotransmitter levels, and specific biological information like as cell membrane turnover, cell density, and cell proliferation, complementing the structural pictures of conventional MRI. The measured brain MRSI includes complicated signals corresponding to several overlapping peaks of various metabolites, baselines from various macromolecules and lipids, as well as noise and distortion (152). By measuring the concentration levels of specific metabolites, *in vivo* and *in vitro* MRSI studies (153, 154) of the brain can indicate tumor kind, grade, or invasion and distinguish tumor growth from post-radiation necrosis.

In one work, the SVM classifier and the minimum redundancy maximum relevance algorithm were used to predict glioma grade based on twenty-six metabolic characteristics from the preoperative MRSI. This model attained AUCs of 0.825% in the training set and 0.820% in the validation set (50). In another study, the dictionary pair learning approach was designed to evaluate glioma based on 150 spectra; its overall accuracy was 0.9778 (155). For glioma treatment response, in a research of 29 control mice and 34 TMZ-treated mice, the performance of an SVM classifier with a linear kernel over the number of sources picked for the MRS image data was able to identify between treated and untreated mice with GBM with an accuracy of over 80% (156).

## 4 LIMITATIONS AND FUTURE CONSIDERATIONS

Simultaneous advancements in image processing technology (MRI, PET, and spectral imaging) and AI, particularly in machine learning and deep learning, have enabled these data-rich patterns to provide diagnostic and guidance information for glioma patients in a non-destructive manner. The majority of these technologies have demonstrated a moderate to a high degree of accuracy. However, some constraints must be solved before these novel predictive analytics algorithms can become widely used in glioma diagnosis and therapy.

Initially, the use of AI in glioma is still in its infancy, with the majority of research being retrospective with limited sample size. It is difficult to validate the safety and reliability of these models in clinical practice. The present medical scientific environment requires data sharing, data management, data standards, and interoperability. Additionally, as machine learning continues to change the area of healthcare, it has posed a variety of challenging ethical problems. If misdiagnosis happens in the use of AI, issues of moral and legal accountability must be addressed (157). Another difficulty is the "black box" aspect of AI technology, which leaves developers and consumers in the dark about how a computer generates its results and lacks interpretability and transparency (158). Radiomics is an AI-assisted technique that will confront the same challenge in the diagnosis and treatment of glioma, which may restrict clinical application (159). However, there are several directions to modify this problem. Applications of radiomics in glioma belong to more deterministic domains. The AI-calculated result assigns the images a relevant annotation (such as a certain gene mutation, the prognosis of the disease is good or bad). Using other patient samples, medical specialists can objectively validate the tags assigned to the images (17, 160), which could reduce the impact of the "black box". Besides, a new generation of AI which has better reliability, interpretability, accountability, and transparency than black-box AI is worth investing in to overcome the "black box" dilemma. For example, Jia et al. created visualizing surrogate decision trees of convolutional neural networks with python (161).

Notable also is the fact that the research described above consists of analyses of a single type of data, a technique known as single-omics analysis (radiomics). However, single-omics data analysis has limits, and it is not apparent which data types should be used to reflect clinical characteristics. A qualified physician should evaluate not only the type of tumor when diagnosing and treating patients, but also the pathology, genes, medical imaging of the tumor, and clinical aspects of the patients. Along with radiomics, genomics, transcriptomics, and high-throughput proteomics are all examples of "omics" techniques that provide data for the examination of molecular constituents. In a general sense, the multi-omics analysis consists of three components: input data, technique, and output data (162). Multiple omics analysis is crucial in neuro-oncology research with limited sample sizes. While the multi-omics analysis may analyze several types of data in parallel for humans, human-dependent multi-omics analysis is not repeatable or interpretable. This challenge can be tackled by integrating machine learning techniques (163) such as multi-modal learning, multi-task learning, representational learning, semi-supervised learning, and automated feature acquisition. Recently, radiomics and radiogenomics (81, 82, 164) have received considerable interest, as have various studies that analyze both radiographic and histological pictures (165). The multi-omics analysis enables us to acquire a more complete knowledge of the illness to improve clinical applications such as determining therapy efficacy, predicting prognosis, and identifying the optimal treatment (**Figure 3**).

In general, AI will show its superiority and larger-scale research will be carried out. Clinicians need to increase interaction with engineers to complement knowledge gaps in both fields. In the future, multidisciplinary collaboration remains a crucial aspect. Researchers will be able to combine multi-omics data to discover drugs and assess treatment effects, predict prognosis, and discover the best treatment for each patient. Finally, while AI has played a huge role in the medical field, AI still can't replace doctors.

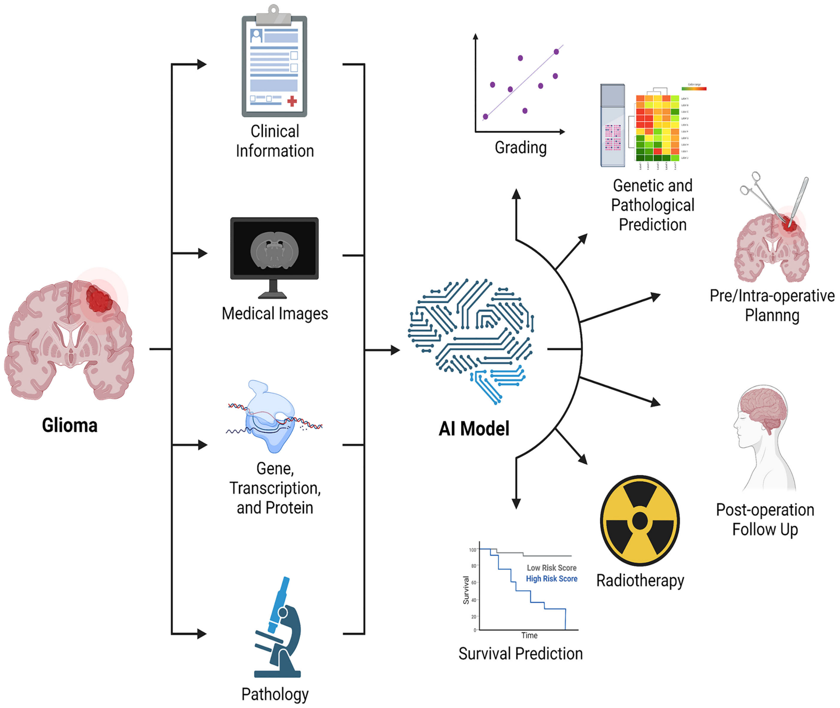


**FIGURE 3** | Combination of multi-omics analysis and artificial intelligence. Artificial intelligence integrates clinical data, medical imaging, genomics, transcriptomics, proteomics, and pathology, among other things, to enable the application of multiple omics in glioma, with the potential to detect and evaluate lesions, promote treatment, and predict treatment response and prognosis. AI: artificial intelligence.

## 5 CONCLUSION

This review retrospectively summarizes some sample studies on the applications of AI in the diagnosis and treatment of glioma using MRI, PET, and spectral imaging. AI is advancing at a breakneck pace and is emerging as a viable tool for medical picture analysis. However, we should be mindful that the implementation of AI in clinical practice is not without flaws. While we are continually working to improve the accuracy of AI, we should not rely excessively on it, as it cannot replace the clinician.

## AUTHOR CONTRIBUTIONS

JX, YM, KQ, MC, ZD and GN wrote this paper. WT, SL, CW, TC, FQ, MC, ZD and GN reviewed this paper. All authors read and approved the final manuscript.

## FUNDING

This work was supported by Natural Science Foundation of Zhejiang Province (NO. LQ22H160003 and NO. Y22H185692).

## ACKNOWLEDGMENTS

The authors would like to thank Dr. Fanghui Qiu for his guidance and contribution in this work. The figures in this article were created using Adobe Illustrator, Adobe Photoshop, BioRender, and Microsoft PowerPoint.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.892056/full#supplementary-material

## REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* (2018) 68:394–424. doi: 10.3322/caac.21492
2. Ostrom QT, Cioffi G, Gittleman H, Patil N, Waite K, Kruchko C, et al. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2012-2016. *Neuro Oncol* (2019) 21:v1–00. doi: 10.1093/noajnl/vdz014.005
3. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 World Health Organization Classification of

Tumors of the Central Nervous System: A Summary. *Acta Neuropathol* (2016) 131:803–20. doi: 10.1007/s00401-016-1545-1

4. Stupp R, Taillibert S, Kanner A, Read W, Steinberg D, Lhermitte B, et al. Effect of Tumor-Treating Fields Plus Maintenance Temozolomide vs Maintenance Temozolomide Alone on Survival in Patients With Glioblastoma: A Randomized Clinical Trial. *JAMA* (2017) 318:2306–16. doi: 10.1001/jama.2017.18718
5. Ballo MT, Urman N, Lavy-Shahaf G, Grewal J, Bomzon Z, Toms S. Correlation of Tumor Treating Fields Dosimetry to Survival Outcomes in Newly Diagnosed Glioblastoma: A Large-Scale Numerical Simulation-Based Analysis of Data From the Phase 3 EF-14 Randomized Trial. *Int J Radiat Oncol Biol Phys* (2019) 104:1106–13. doi: 10.1016/j.ijrobp.2019.04.008
6. Davis ME. Glioblastoma: Overview of Disease and Treatment. *Clin J Oncol Nurs* (2016) 20:S2–8. doi: 10.1188/16.CJON.S1.2-8
7. Sanghvi D. Post-Treatment Imaging of High-Grade Gliomas. *Indian J Radiol Imaging* (2015) 25:102–8. doi: 10.4103/0971-3026.155829
8. Litjens G, Kooi T, Bejnordi BE, Setio AA.A, Ciompi F, Ghafoorian M, et al. A Survey on Deep Learning in Medical Image Analysis. *Med Image Anal* (2017) 42:60–88. doi: 10.1016/j.media.2017.07.005
9. Topol EJ. High-Performance Medicine: The Convergence of Human and Artificial Intelligence. *Nat Med* (2019) 25:44–56. doi: 10.1038/s41591-018-0300-7
10. Patel UK, Anwar A, Saleem S, Malik P, Rasul B, Patel K, et al. Artificial Intelligence as an Emerging Technology in the Current Care of Neurological Disorders. *J Neurol* (2021) 268:1623–42. doi: 10.1007/s00415-019-09518-3
11. Ganapathy K, Abdul SS, Nursetyo AA. Artificial Intelligence in Neurosciences: A Clinician's Perspective. *Neurol India* (2018) 66:934–9. doi: 10.4103/0028-3886.236971
12. Ho CWL, Soon D, Caals K, Kapur J. Governance of Automated Image Analysis and Artificial Intelligence Analytics in Healthcare. *Clin Radiol* (2019) 74:329–37. doi: 10.1016/j.crad.2019.02.005
13. Giger ML. Machine Learning in Medical Imaging. *J Am Coll Radiol JACR* (2018) 15:512–20. doi: 10.1016/j.jacr.2017.12.028
14. Rudie JD, Rauschecker AM, Bryan RN, Davatzikos C, Mohan S. Emerging Applications of Artificial Intelligence in Neuro-Oncology. *Radiology* (2019) 290:607–18. doi: 10.1148/radiol.2018181928
15. Hamet P, Tremblay J. Artificial Intelligence in Medicine. *Metabolism* (2017) 69S:S36–40. doi: 10.1016/j.metabol.2017.01.011
16. Fujita H. AI-Based Computer-Aided Diagnosis (AI-CAD): The Latest Review to Read First. *Radiol Phys Technol* (2020) 13:6–19. doi: 10.1007/s12194-019-00552-4
17. Cao JS, Lu ZY, Chen MY, Zhang B, Juengpanich S, Hu JH, et al. Artificial Intelligence in Gastroenterology and Hepatology: Status and Challenges. *World J Gastroenterol* (2021) 27:1664–90. doi: 10.3748/wjg.v27.i16.1664
18. Van Calster B, Wynants L. Machine Learning in Medicine. *N Engl J Med* (2019) 380:2588. doi: 10.1056/NEJMc1906060
19. McBee MP, Awan OA, Colucci AT, Ghobadi CW, Kadom N, Kansagra AP, et al. Deep Learning in Radiology. *Acad Radiol* (2018) 25:1472–80. doi: 10.1016/j.acra.2018.02.018
20. Kahn CE. Artificial Intelligence in Radiology: Decision Support Systems. *Radiographics* (1994) 14:849–61. doi: 10.1148/radiographics.14.4.7938772
21. Laurent PA. The Emergence of Saliency and Novelty Responses From Reinforcement Learning Principles. *Neural Netw* (2008) 21:1493–9. doi: 10.1016/j.neunet.2008.09.004
22. Handelman GS, Kok HK, Chandra RV, Razavi AH, Lee MJ, Asadi H. Edoctor: Machine Learning and the Future of Medicine. *J Intern Med* (2018) 284:603–19. doi: 10.1111/joim.12822
23. Kohli A, Jha S. Why CAD Failed in Mammography. *J Am Coll Radiol* (2018) 15:535–7. doi: 10.1016/j.jacr.2017.12.029
24. Lee J-G, Jun S, Cho Y-W, Lee H, Kim GB, Seo JB, et al. Deep Learning in Medical Imaging: General Overview. *Korean J Radiol* (2017) 18:570–84. doi: 10.3348/kjr.2017.18.4.570
25. Gautam R, Sharma M. Prevalence and Diagnosis of Neurological Disorders Using Different Deep Learning Techniques: A Meta-Analysis. *J Med Syst* (2020) 44:49. doi: 10.1007/s10916-019-1519-7
26. Cai C, Wang S, Xu Y, Zhang W, Tang K, Ouyang Q, et al. Transfer Learning for Drug Discovery. *J Med Chem* (2020) 63:8683–94. doi: 10.1021/acs.jmedchem.9b02147
27. Chan HP, Samala RK, Hadjiiski LM, Zhou C. Deep Learning in Medical Image Analysis. *Adv Exp Med Biol* (2020) 1213:3–21. doi: 10.1007/978-3-030-33128-3_1
28. Biswas M, Kuppili V, Saba L, Edla DR, Suri HS, Cuadrado-Godia E, et al. State-Of-the-Art Review on Deep Learning in Medical Imaging. *Front Biosci (Landmark Ed)* (2019) 24:392–426. doi: 10.2741/4725
29. Schmidhuber J. Deep Learning in Neural Networks: An Overview. *Neural Netw* (2015) 61:85–117. doi: 10.1016/j.neunet.2014.09.003
30. Pew-Thian Y, Yong Z, Dinggang S. Multi-Tissue Decomposition of Diffusion MRI Signals *via* Sparse-Group Estimation. *IEEE Trans Image Process* (2016) 25:4340–53. doi: 10.1109/TIP.2016.2588328
31. LeCun Y, Bengio Y, Hinton G. Deep Learning. *Nature* (2015) 521:436–44. doi: 10.1038/nature14539
32. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More Than Pictures, They Are Data. *Radiology* (2016) 278:563–77. doi: 10.1148/radiol.2015151169
33. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EE.C, van Timmeren J, et al. Radiomics: The Bridge Between Medical Imaging and Personalized Medicine. *Nat Rev Clin Oncol* (2017) 14:749–62. doi: 10.1038/nrclinonc.2017.141
34. Lohmann P, Galldiks N, Kocher M, Heinzel A, Filss CP, Stegmayr C, et al. Radiomics in Neuro-Oncology: Basics, Workflow, and Applications. *Methods (San Diego Calif)* (2021) 188:112–21. doi: 10.1016/j.ymeth.2020.06.003
35. Rizzo S, Botta F, Raimondi S, Origgi D, Fanciullo C, Morganti AG, et al. Radiomics: The Facts and the Challenges of Image Analysis. *Eur Radiol Exp* (2018) 2:36. doi: 10.1186/s41747-018-0068-z
36. Cha YJ, Jang WI, Kim MS, Yoo HJ, Paik EK, Jeong HK, et al. Prediction of Response to Stereotactic Radiosurgery for Brain Metastases Using Convolutional Neural Networks. *Anticancer Res* (2018) 38:5437–45. doi: 10.21873/anticanres.12875
37. Mayerhoefer ME, Materka A, Langs G, Haggstrom I, Szczypinski P, Gibbs P, et al. Introduction to Radiomics. *J Nucl Med* (2020) 61:488–95. doi: 10.2967/jnumed.118.222893
38. Zhao L, Jia K. Multiscale CNNs for Brain Tumor Segmentation and Diagnosis. *Comput Math Methods Med* (2016) 2016:8356294. doi: 10.1155/2016/8356294
39. Cao Y, Sundgren PC, Tsien CI, Chenevert TT, Junck L. Physiologic and Metabolic Magnetic Resonance Imaging in Gliomas. *J Clin Oncol* (2006) 24:1228–35. doi: 10.1200/JCO.2005.04.7233
40. Mabray MC, Cha S. Advanced MR Imaging Techniques in Daily Practice. *Neuroimaging Clin N Am* (2016) 26:647–66. doi; 10.1016/j.nic.2016.06.010
41. d'Este SH, Nielsen MB, Hansen AE. Visualizing Glioma Infiltration by the Combination of Multimodality Imaging and Artificial Intelligence, a Systematic Review of the Literature. *Diagnost (Basel)* (2021) 11(4):592. doi: 10.3390/diagnostics11040592
42. Menze BH, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby , et al. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Trans Med Imaging* (2015) 34:1993–2024. doi: 10.1109/TMI.2014.2377694
43. Fu J, Singhrao K, Qi XS, Yang Y, Ruan D, Lewis JH. Three-Dimensional Multipath DenseNet for Improving Automatic Segmentation of Glioblastoma on Pre-Operative Multimodal MR Images. *Med Phys* (2021) 48:2859–66. doi: 10.1002/mp.14800
44. Khosravanian A, Rahmanimanesh M, Keshavarzi P, Mozaffari S. Fast Level Set Method for Glioma Brain Tumor Segmentation Based on Superpixel Fuzzy Clustering and Lattice Boltzmann Method. *Comput Methods Programs BioMed* (2021) 198:105809. doi: 10.1016/j.cmpb.2020.105809
45. Amin J, Sharif M, Raza M, Saba T, Anjum MA. Brain Tumor Detection Using Statistical and Machine Learning Method. *Comput Methods Programs BioMed* (2019) 177:69–79. doi: 10.1016/j.cmpb.2019.05.015
46. Quon JL, Bala W, Chen LC, Wright J, Kim LH, Han M, et al. Deep Learning for Pediatric Posterior Fossa Tumor Detection and Classification: A Multi-Institutional Study. *AJNR Am J Neuroradiol* (2020) 41:1718–25. doi: 10.3174/ajnr.A6704
47. Luo H, Zhuang Q, Wang Y, Abudumijiti A, Shi K, Rominger A, et al. A Novel Image Signature-Based Radiomics Method to Achieve Precise Diagnosis and Prognostic Stratification of Gliomas. *Lab Invest* (2021) 101:450–62. doi: 10.1038/s41374-020-0472-x

48. Le NQK, Hung TNK, Do DT, Lam LHT, Dang LH, Huynh T-T. Radiomics-Based Machine Learning Model for Efficiently Classifying Transcriptome Subtypes in Glioblastoma Patients From MRI. *Comput Biol Med* (2021) 132:104320. doi: 10.1016/j.compbiomed.2021.104320
49. Lu C-F, Hsu F-T, Hsieh KL-C, Kao Y-C.J, Cheng S.-J, Hsu JB.-K, et al. Machine Learning-Based Radiomics for Molecular Subtyping of Gliomas. *Clin Cancer Res* (2018) 24:4429–36. doi: 10.1158/1078-0432.CCR-17-3445
50. Qi C, Li Y, Fan X, Jiang Y, Wang R, Yang S, et al. A Quantitative SVM Approach Potentially Improves the Accuracy of Magnetic Resonance Spectroscopy in the Preoperative Evaluation of the Grades of Diffuse Gliomas. *NeuroImage Clin* (2019) 23:101835. doi: 10.1016/j.nicl.2019.101835
51. Li M, Tang H, Chan MD, Zhou X, Qian X. DC-AL GAN: Pseudoprogression and True Tumor Progression of Glioblastoma Multiform Image Classification Based on DCGAN and AlexNet. *Med Phys* (2020) 47:1139–50. doi: 10.1002/mp.14003
52. Ismail M, Hill V, Statsevych V, Huang R, Prasanna P, Correa R, et al. Shape Features of the Lesion Habitat to Differentiate Brain Tumor Progression From Pseudoprogression on Routine Multiparametric MRI: A Multisite Study. *AJNR Am J Neuroradiol* (2018) 39:2187–93. doi: 10.3174/ajnr.A5858
53. Pan Z-Q, Zhang S-J, Wang X-L, Jiao Y-X, Qiu J-J. Machine Learning Based on a Multiparametric and Multiregional Radiomics Signature Predicts Radiotherapeutic Response in Patients With Glioblastoma. *Behav Neurol* (2020) 2020:1712604. doi: 10.1155/2020/1712604
54. Zhang X, Lu D, Gao P, Tian Q, Lu H, Xu X, et al. Survival-Relevant High-Risk Subregion Identification for Glioblastoma Patients: The MRI-Based Multiple Instance Learning Approach. *Eur Radiol* (2020) 30:5602–10. doi: 10.1007/s00330-020-06912-8
55. Chang Y, Lafata K, Sun W, Wang C, Chang Z, Kirkpatrick JP, et al. An Investigation of Machine Learning Methods in Delta-Radiomics Feature Analysis. *PLoS One* (2019) 14:e0226348. doi: 10.1371/journal.pone.0226348
56. Sanghani P, Ang BT, King NKK, Ren H. Overall Survival Prediction in Glioblastoma Multiforme Patients From Volumetric, Shape and Texture Features Using Machine Learning. *Surg Oncol* (2018) 27:709–14. doi: 10.1016/j.suronc.2018.09.002
57. Choi KS, Choi SH, Jeong B. Prediction of IDH Genotype in Gliomas With Dynamic Susceptibility Contrast Perfusion MR Imaging Using an Explainable Recurrent Neural Network. *Neuro Oncol* (2019) 21:1197–209. doi: 10.1093/neuonc/noz095
58. Yogananda CGB, Shah BR, Nalawade SS, Yang X, Wang W, Wan X, et al. MRI-Based Deep-Learning Method for Determining Glioma Promoter Methylation Status. *AJNR Am J Neuroradiol* (2021) 42:845–52. doi: 10.3174/ajnr.A7029
59. Li Y, Liu X, Qian Z, Sun Z, Xu K, Wang K, et al. Genotype Prediction of ATRX Mutation in Lower-Grade Gliomas Using an MRI Radiomics Signature. *Eur Radiol* (2018) 28:2960–8. doi: 10.1007/s00330-017-5267-0
60. Liu J, Chen F, Pan C, Zhu M, Zhang X, Zhang L, et al. A Cascaded Deep Convolutional Neural Network for Joint Segmentation and Genotype Prediction of Brainstem Gliomas. *IEEE Trans BioMed Eng* (2018) 65:1943–52. doi: 10.1109/TBME.2018.2845706
61. Fukuma R, Yanagisawa T, Kinoshita M, Shinozaki T, Arita H, Kawaguchi A, et al. Prediction of IDH and TERT Promoter Mutations in Low-Grade Glioma From Magnetic Resonance Images Using a Convolutional Neural Network. *Sci Rep* (2019) 9:20311. doi: 10.1038/s41598-019-56767-3
62. Akkus Z, Ali I, Sedlář J, Agrawal JP, Parney I.F, Giannini C, et al. Predicting Deletion of Chromosomal Arms 1p/19q in Low-Grade Gliomas From MR Images Using Machine Intelligence. *J Digit Imaging* (2017) 30:469–76. doi: 10.1007/s10278-017-9984-3
63. Gates EDH, Lin JS, Weinberg JS, Hamilton J, Prabhu S.S, Hazle JD, et al. Guiding the First Biopsy in Glioma Patients Using Estimated Ki-67 Maps Derived From MRI: Conventional Versus Advanced Imaging. *Neuro Oncol* (2019) 21:527–36. doi: 10.1093/neuonc/noz004
64. Li Y, Qian Z, Xu K, Wang K, Fan X, Li S, et al. MRI Features Predict P53 Status in Lower-Grade Gliomas *via* a Machine-Learning Approach. *NeuroImage Clin* (2018) 17:306–11. doi: 10.1016/j.nicl.2017.10.030
65. Sun Z, Li Y, Wang Y, Fan X, Xu K, Wang , et al. Radiogenomic Analysis of Vascular Endothelial Growth Factor in Patients With Diffuse Gliomas. *Cancer Imaging* (2019) 19:68. doi: 10.1186/s40644-019-0256-y
66. Calabrese E, Villanueva-Meyer JE, Cha S. A Fully Automated Artificial Intelligence Method for non-Invasive, Imaging-Based Identification of Genetic Alterations in Glioblastomas. *Sci Rep* (2020) 10:11852. doi: 10.1038/s41598-020-68857-8
67. Kickingereder P, Bonekamp D, Nowosielski M, Kratz A, Sill M, Burth S, et al. Radiogenomics of Glioblastoma: Machine Learning-Based Classification of Molecular Characteristics by Using Multiparametric and Multiregional MR Imaging Features. *Radiology* (2016) 281:907–18. doi: 10.1148/radiol.2016161382
68. Hu LS, Yoon H, Eschbacher JM, Baxter LC, Dueck AC, Nespodzany A, et al. Accurate Patient-Specific Machine Learning Models of Glioblastoma Invasion Using Transfer Learning. *AJNR Am J Neuroradiol* (2019) 40:418–25. doi: 10.3174/ajnr.A5981
69. Cakmakci D, Karakaslar EO, Ruhland E, Chenard M-P, Proust F, Piotto M, et al. Machine Learning Assisted Intraoperative Assessment of Brain Tumor Margins Using HRMAS NMR Spectroscopy. *PLoS Comput Biol* (2020) 16: e1008184. doi: 10.1371/journal.pcbi.1008184
70. Fathi Kazerooni A, Nabil M, Zeinali Zadeh M, Firouznia K, Azmoudeh-Ardalan F, Frangi AF, et al. Characterization of Active and Infiltrative Tumorous Subregions From Normal Tissue in Brain Gliomas Using Multiparametric MRI. *J Magn Reson Imaging* (2018) 48:938–50. doi: 10.1002/jmri.25963
71. Durmo F, Latt J, Rydelius A, Engelholm S, Kinhult S, Askaner K, et al. Brain Tumor Characterization Using Multibiometric Evaluation of MRI. *Tomography* (2018) 4:14–25. doi: 10.18383/j.tom.2017.00020
72. De Baene W, Rutten GJM, Sitskoorn MM. The Temporal Pattern of a Lesion Modulates the Functional Network Topology of Remote Brain Regions. *Neural Plast* (2017) 2017:3530723. doi: 10.1155/2017/3530723
73. Jiang Z, Ding C, Liu M, Tao D. Two-Stage Cascaded U-Net: 1st Place Solution to BraTS Challenge 2019 Segmentation Task. In: A Crimi and S Bakas, editors. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*. Cham: Springer International Publishing (2020). p. 231–41.
74. Wang J, Yu Z, Luan Z, Ren J, Zhao Y, Yu G. RDAU-Net: Based on a Residual Convolutional Neural Network With DFP and CBAM for Brain Tumor Segmentation. *Front Oncol* (2022) 12:805263. doi: 10.3389/fonc.2022.805263
75. Gritsch S, Batchelor TT, Gonzalez Castro LN. Diagnostic, Therapeutic, and Prognostic Implications of the 2021 World Health Organization Classification of Tumors of the Central Nervous System. *Cancer* (2021). doi: 10.1002/cncr.33918
76. Tesileanu CMS, Dirven L, Wijnenga MMJ, et al. Survival of Diffuse Astrocytic Glioma, IDH1/2 Wildtype, With Molecular Features of Glioblastoma, WHO Grade IV: A Confirmation of the cIMPACT-NOW Criteria. *Neuro Oncol* (2020) 22:515–23. doi: 10.1093/neuonc/noz200
77. Liu XY, Gerges N, Korshunov A, Sabha N, Khuong-Quang DA, Fontebasso AM, et al. Frequent ATRX Mutations and Loss of Expression in Adult Diffuse Astrocytic Tumors Carrying IDH1/IDH2 and TP53 Mutations. *Acta Neuropathol* (2012) 124:615–25. doi: 10.1007/s00401-012-1031-3
78. Reuss DE, Sahm F, Schrimpf D, Wiestler B, Capper D, Koelsche C, et al. ATRX and IDH1-R132H Immunohistochemistry With Subsequent Copy Number Analysis and IDH Sequencing as a Basis for an "Integrated" Diagnostic Approach for Adult Astrocytoma, Oligodendroglioma and Glioblastoma. *Acta Neuropathol* (2015) 129:133–46. doi: 10.1007/s00401-014-1370-3
79. Abedalthagafi M, Phillips JJ, Kim GE, Mueller S, Haas-Kogen DA, Marshall R.E, et al. The Alternative Lengthening of Telomere Phenotype is Significantly Associated With Loss of ATRX Expression in High-Grade Pediatric and Adult Astrocytomas: A Multi-Institutional Study of 214 Astrocytomas. *Modern Pathol an Off J U States Can Acad Pathol Inc* (2013) 26:1425–32. doi: 10.1038/modpathol.2013.90
80. Koelsche C, Sahm F, Capper D, Reuss D, Sturm D, Jones DT, et al. Distribution of TERT Promoter Mutations in Pediatric and Adult Tumors of the Nervous System. *Acta Neuropathol* (2013) 126:907–15. doi: 10.1007/s00401-013-1195-5
81. Lo Gullo R, Daimiel I, Morris EA, Pinker K. Combining Molecular and Imaging Metrics in Cancer: Radiogenomics. *Insights Imaging* (2020) 11:1. doi: 10.1186/s13244-019-0795-6
82. Bodalal Z, Trebeschi S, Nguyen-Kim TDL, Schats W, Beets-Tan R. Radiogenomics: Bridging Imaging and Genomics. *Abdom Radiol (NY)* (2019) 44:1960–84. doi: 10.1007/s00261-019-02028-w

83. Broen MPG, Smits M, Wijnenga MMJ, Dubbink HJ, Anten M, Schijns O, et al. The T2-FLAIR Mismatch Sign as an Imaging Marker for non-Enhancing IDH-Mutant, 1p/19q-Intact Lower-Grade Glioma: A Validation Study. *Neuro Oncol* (2018) 20:1393–9. doi: 10.1093/neuonc/noy048

84. Lasocki A, Gaillard F, Gorelik A, Gonzales M. MRI Features Can Predict 1p/19q Status in Intracranial Gliomas. *AJNR Am J Neuroradiol* (2018) 39:687–92. doi: 10.3174/ajnr.A5572

85. Tay KL, Tsui A, Phal PM, Drummond KJ, Tress BM. MR Imaging Characteristics of Protoplasmic Astrocytomas. *Neuroradiology* (2011) 53:405–11. doi: 10.1007/s00234-010-0741-2

86. Hanzély Z, Polgár C, Fodor J, Brucher JM, Vitanovics D, Mangel LC, et al. Role of Early Radiotherapy in the Treatment of Supratentorial WHO Grade II Astrocytomas: Long-Term Results of 97 Patients. *J Neurooncol* (2003) 63:305–12. doi: 10.1023/A:1024376719067

87. Wang J, Zheng X, Zhang J, Xue H, Wang L, Jing R, et al. An MRI-Based Radiomics Signature as a Pretreatment Noninvasive Predictor of Overall Survival and Chemotherapeutic Benefits in Lower-Grade Gliomas. *Eur Radiol* (2021) 31:1785–94. doi: 10.1007/s00330-020-07581-3

88. Choi YS, Ahn SS, Chang JH, Kang S.G, Kim EH, Kim S.H, et al. Machine Learning and Radiomic Phenotyping of Lower Grade Gliomas: Improving Survival Prediction. *Eur Radiol* (2020) 30:3834–42. doi: 10.1007/s00330-020-06737-5

89. Taal W, Brandsma D, de Bruin HG, Bromberg JE, Swaak-Kragten AT, Smitt PA, et al. Incidence of Early Pseudo-Progression in a Cohort of Malignant Glioma Patients Treated With Chemoirradiation With Temozolomide. *Cancer* (2008) 113:405–10. doi: 10.1002/cncr.23562

90. Santo G, Laudicella R, Linguanti F, Nappi AG, Abenavoli E, Vergura V, et al. The Utility of Conventional Amino Acid PET Radiotracers in the Evaluation of Glioma Recurrence Also in Comparison With MRI. *Diagnost (Basel)* (2022) 12(4):844. doi: 10.3390/diagnostics12040844

91. Wen PY, Chang SM, Van den Bent MJ, Vogelbaum MA, Macdonald DR, Lee EQ. Response Assessment in Neuro-Oncology Clinical Trials. *J Clin Oncol* (2017) 35:2439–49. doi: 10.1200/JCO.2017.72.7511

92. Verger A, Kas A, Darcourt J, Guedj E. PET Imaging in Neuro-Oncology: An Update and Overview of a Rapidly Growing Area. *Cancers (Basel)* (2022) 14 (5):1103. doi: 10.3390/cancers14051103

93. Verger A, Arbizu J, Law I. Role of Amino-Acid PET in High-Grade Gliomas: Limitations and Perspectives. *Q J Nucl Med Mol Imaging* (2018) 62:254–66. doi: 10.23736/S1824-4785.18.03092-3

94. Holzgreve A, Albert NL, Galldiks N, Suchorska B. Use of PET Imaging in Neuro-Oncological Surgery. *Cancers (Basel)* (2021) 13(9):2093. doi: 10.3390/cancers13092093

95. Law I, Albert NL, Arbizu J, Boellaard R, Drzezga A, Galldiks N, et al. Joint EANM/EANO/RANO Practice Guidelines/SNMMI Procedure Standards for Imaging of Gliomas Using PET With Radiolabelled Amino Acids and [(18)F]FDG: Version 1.0. *Eur J Nucl Med Mol Imaging* (2019) 46:540–57. doi: 10.1007/s00259-018-4207-9

96. Albert NL, Weller M, Suchorska B, Galldiks N, Soffietti R, Kim MM, et al. Response Assessment in Neuro-Oncology Working Group and European Association for Neuro-Oncology Recommendations for the Clinical Use of PET Imaging in Gliomas. *Neuro Oncol* (2016) 18:1199–208. doi: 10.1093/neuonc/now058

97. Galldiks N, Albert NL, Sommerauer M, Grosu AL, Ganswindt U, Law I, et al. PET Imaging in Patients With Meningioma-Report of the RANO/PET Group. *Neuro Oncol* (2017) 19:1576–87. doi: 10.1093/neuonc/nox112

98. Galldiks N, Niyazi M, Grosu AL, Kocher M, Langen KJ, Law I, et al. Contribution of PET Imaging to Radiotherapy Planning and Monitoring in Glioma Patients - a Report of the PET/RANO Group. *Neuro Oncol* (2021) 23:881–93. doi: 10.1093/neuonc/noab013

99. Galldiks N, Langen KJ, Albert NL, Chamberlain M, Soffietti R, Kim MM, et al. PET Imaging in Patients With Brain Metastasis-Report of the RANO/PET Group. *Neuro Oncol* (2019) 21:585–95. doi: 10.1093/neuonc/noz003

100. Suchorska B, Albert NL, Tonn JC. Usefulness of PET Imaging to Guide Treatment Options in Gliomas. *Curr Treat Options Neurol* (2016) 18:4. doi: 10.1007/s11940-015-0384-z

101. Zaragori T, Ginet M, Marie P-Y, Roch V, Grignon R, Gauchotte G, et al. Use of Static and Dynamic [(18)F]-F-DOPA PET Parameters for Detecting Patients With Glioma Recurrence or Progression. *EJNMMI Res* (2020) 10:56. doi: 10.1186/s13550-020-00645-x

102. Ceccon G, Lohmann P, Stoffels G, Judov N, Filss CP, Rapp M, et al. Dynamic O-(2-18F-Fluoroethyl)-L-Tyrosine Positron Emission Tomography Differentiates Brain Metastasis Recurrence From Radiation Injury After Radiotherapy. *Neuro Oncol* (2017) 19:281–8. doi: 10.1093/neuonc/now149

103. Blanc-Durand P, van der Gucht A, Schaefer N, Itti E, Prior JO. Automatic Lesion Detection and Segmentation of 18F-FET PET in Gliomas: A Full 3D U-Net Convolutional Neural Network Study. *PLoS One* (2018) 13:e0195798. doi: 10.1371/journal.pone.0195798

104. Kebir S, Rauschenbach L, Weber M, Lazaridis L, Schmidt T, Keyva , et al. Machine Learning-Based Differentiation Between Multiple Sclerosis and Glioma WHO II°-IV° Using O-(2-[18F] Fluoroethyl)-L-Tyrosine Positron Emission Tomography. *J Neurooncol* (2021) 152:325–32. doi: 10.1007/s11060-021-03701-1

105. Kong Z, Jiang C, Zhu R, Feng S, Wang Y, Li J, et al. F-FDG-PET-Based Radiomics Features to Distinguish Primary Central Nervous System Lymphoma From Glioblastoma. *NeuroImage Clin* (2019) 23:101912. doi: 10.1016/j.nicl.2019.101912

106. Matsui Y, Maruyama T, Nitta M, Saito T, Tsuzuki S, Tamura M, et al. Prediction of Lower-Grade Glioma Molecular Subtypes Using Deep Learning. *J Neurooncol* (2020) 146:321–7. doi: 10.1007/s11060-019-03376-9

107. Lohmann P, Elahmadawy MA, Gutsche R, Werner J-M, Bauer EK, Ceccon G, et al. FET PET Radiomics for Differentiating Pseudoprogression From Early Tumor Progression in Glioma Patients Post-Chemoradiation. *Cancers (Basel)* (2020) 12(12):3835. doi: 10.3390/cancers12123835

108. Kebir S, Schmidt T, Weber M, Lazaridis L, Galldiks N, Langen K-J, et al. A Preliminary Study on Machine Learning-Based Evaluation of Static and Dynamic FET-PET for the Detection of Pseudoprogression in Patients With IDH-Wildtype Glioblastoma. *Cancers (Basel)* (2020) 12(11):3080. doi: 10.3390/cancers12113080

109. Imani F, Boada FE, Lieberman FS, Davis DK, Mountz JM. Molecular and Metabolic Pattern Classification for Detection of Brain Glioma Progression. *Eur J Radiol* (2014) 83:e100-e5. doi: 10.1016/j.ejrad.2013.06.033

110. Kebir S, Khurshid Z, Gaertner FC, Essler M, Hattingen E, Fimmers R, et al. Unsupervised Consensus Cluster Analysis of [18F]-Fluoroethyl-L-Tyrosine Positron Emission Tomography Identified Textural Features for the Diagnosis of Pseudoprogression in High-Grade Glioma. *Oncotarget* (2017) 8:8294–304. doi: 10.18632/oncotarget.14166

111. Papp L, Pötsch N, Grahovac M, Schmidbauer V, Woehrer A, Preusser M, et al. Glioma Survival Prediction With Combined Analysis of *In Vivo* C-MET PET Features, Ex Vivo Features, and Patient Features by Supervised Machine Learning. *J Nucl Med* (2018) 59:892–9. doi: 10.2967/jnumed.117.202267

112. Li L, Mu W, Wang Y, Liu Z, Liu Z, Wang Y, et al. A Non-Invasive Radiomic Method Using F-FDG PET Predicts Isocitrate Dehydrogenase Genotype and Prognosis in Patients With Glioma. *Front Oncol* (2019) 9:1183. doi: 10.3389/fonc.2019.01183

113. Tatekawa H, Hagiwara A, Uetani H, Bahri S, Raymond C, Lai A, et al. Differentiating IDH Status in Human Gliomas Using Machine Learning and Multiparametric MR/PET. *Cancer Imaging* (2021) 21:27. doi: 10.1186/s40644-021-00396-5

114. Kebir S, Weber M, Lazaridis L, Deuschl C, Schmidt T, Mönninghoff C, et al. Hybrid 11c-MET PET/MRI Combined With "Machine Learning" in Glioma Diagnosis According to the Revised Glioma WHO Classification 2016. *Clin Nucl Med* (2019) 44:214–20. doi: 10.1097/RLU.0000000000002398

115. Qian J, Herman MG, Brinkmann DH, Laack NN, Kemp BJ, Hunt CH, et al. Prediction of MGMT Status for Glioblastoma Patients Using Radiomics Feature Extraction From F-DOPA-PET Imaging. *Int J Radiat Oncol Biol Phys* (2020) 108:1339–46. doi: 10.1016/j.ijrobp.2020.06.073

116. Kong Z, Li J, Liu Z, Liu Z, Zhao D, Cheng X, et al. Radiomics Signature Based on FDG-PET Predicts Proliferative Activity in Primary Glioma. *Clin Radiol* (2019) 74:815.e15–.e23. doi: 10.1016/j.crad.2019.06.019

117. Kim MM, Parolia A, Dunphy MP, Venneti S. Non-Invasive Metabolic Imaging of Brain Tumours in the Era of Precision Medicine. *Nat Rev Clin Oncol* (2016) 13:725–39. doi: 10.1038/nrclinonc.2016.108

118. Jansen NL, Suchorska B, Wenter V, Schmid-Tannwald C, Todica A, Eigenbrod S, et al. Prognostic Significance of Dynamic 18F-FET PET in

Newly Diagnosed Astrocytic High-Grade Glioma. *J Nucl Med* (2015) 56:9–15. doi: 10.2967/jnumed.114.144675

119. Unterrainer M, Schweisthal F, Suchorska B, Wenter V, Schmid-Tannwald C, Fendler W.P, et al. Serial 18f-FET PET Imaging of Primarily 18f-FET-Negative Glioma: Does It Make Sense? *J Nucl Med* (2016) 57:1177–82. doi: 10.2967/jnumed.115.171033
120. Hutterer M, Nowosielski M, Putzer D, Jansen NL, Seiz M, Schocke M, et al. [18F]-Fluoro-Ethyl-L-Tyrosine PET: A Valuable Diagnostic Tool in Neuro-Oncology, But Not All That Glitters is Glioma. *Neuro Oncol* (2013) 15:341–51. doi: 10.1093/neuonc/nos300
121. Sala Q, Metellus P, Taieb D, Kaphan E, Figarella-Branger D, Guedj E. 18f-DOPA, a Clinically Available PET Tracer to Study Brain Inflammation? *Clin Nucl Med* (2014) 39:e283-5. doi: 10.1097/RLU.0000000000000383
122. Rapp M, Heinzel A, Galldiks N, Stoffels G, Felsberg J, Ewelt C, et al. Diagnostic Performance of 18F-FET PET in Newly Diagnosed Cerebral Lesions Suggestive of Glioma. *J Nucl Med* (2013) 54:229–35. doi: 10.2967/jnumed.112.109603
123. Pichler R, Dunzinger A, Wurm G, Pichler J, Weis S, Nussbaumer K, et al. Is There a Place for FET PET in the Initial Evaluation of Brain Lesions With Unknown Significance? *Eur J Nucl Med Mol Imaging* (2010) 37:1521–8. doi: 10.1007/s00259-010-1457-6
124. Kebir S, Fimmers R, Galldiks N, Schäfer N, Mack F, Schaub C, et al. Late Pseudoprogression in Glioblastoma: Diagnostic Value of Dynamic O-(2-[18F]Fluoroethyl)-L-Tyrosine PET. *Clin Cancer Res* (2016) 22:2190–6. doi: 10.1158/1078-0432.CCR-15-1334
125. Young RJ, Gupta A, Shah AD, Graber JJ, Zhang Z, Shi W, et al. Potential Utility of Conventional MRI Signs in Diagnosing Pseudoprogression in Glioblastoma. *Neurology* (2011) 76:1918–24. doi: 10.1212/WNL.0b013e31821d74e7
126. Clancy NT, Jones G, Maier-Hein L, Elson DS, Stoyanov D. Surgical Spectral Imaging. *Med Image Anal* (2020) 63:101699. doi: 10.1016/j.media.2020.101699
127. Türker-Kaya S. Huck CW. A Review of Mid-Infrared and Near-Infrared Imaging: Principles, Concepts and Applications in Plant Tissue Analysis. *Molecules (Basel Switzerland)* (2017) 22(1):168. doi: 10.3390/molecules2201016
128. Baker MJ, Trevisan J, Bassan P, Bhargava R, Butler HJ, Dorling KM, et al. Using Fourier Transform IR Spectroscopy to Analyze Biological Materials. *Nat Protoc* (2014) 9:1771–91. doi: 10.1038/nprot.2014.110
129. Chen F, Meng C, Qu H, Cheng C, Chen C, Yang B, et al. Human Serum Mid-Infrared Spectroscopy Combined With Machine Learning Algorithms for Rapid Detection of Gliomas. *Photodiagnosis Photodyn Ther* (2021) 35:102308. doi: 10.1016/j.pdpdt.2021.102308
130. Qu H, Wu W, Chen C, Yan Z, Guo W, Meng C, et al. Application of Serum Mid-Infrared Spectroscopy Combined With an Ensemble Learning Method in Rapid Diagnosis of Gliomas. *Anal Methods* (2021) 13:4642–51. doi: 10.1039/D1AY00802A
131. Hands JR, Clemens G, Stables R, Ashton K, Brodbelt A, Davis C, et al. Brain Tumour Differentiation: Rapid Stratified Serum Diagnostics *via* Attenuated Total Reflection Fourier-Transform Infrared Spectroscopy. *J Neurooncol* (2016) 127:463–72. doi: 10.1007/s11060-016-2060-x
132. Cameron JM, Butler HJ, Smith BR, Hegarty MG, Jenkinson MD, Syed K, et al. Developing Infrared Spectroscopic Detection for Stratifying Brain Tumour Patients: Glioblastoma Multiforme vs. Lymphoma. *Analyst* (2019) 144:6736–50. doi: 10.1039/C9AN01731C
133. Peng W, Chen S, Kong D, Zhou X, Lu X, Chang C. Grade Diagnosis of Human Glioma Using Fourier Transform Infrared Microscopy and Artificial Neural Network. *Spectrochim Acta A Mol Biomol Spectrosc* (2021) 260:119946. doi: 10.1016/j.saa.2021.119946
134. Surowka AD, Adamek D, Szczerbowska-Boruchowska M. The Combination of Artificial Neural Networks and Synchrotron Radiation-Based Infrared Micro-Spectroscopy for a Study on the Protein Composition of Human Glial Tumors. *Analyst* (2015) 140:2428–38. doi: 10.1039/C4AN01867B
135. Kalkanis SN, Kast RE, Rosenblum ML, Mikkelsen T, Yurgelevic SM, Nelson KM, et al. Raman Spectroscopy to Distinguish Grey Matter, Necrosis, and Glioblastoma Multiforme in Frozen Tissue Sections. *J Neurooncol* (2014) 116:477–85. doi: 10.1007/s11060-013-1326-9
136. Jermyn M, Mok K, Mercier J, Desroches J, Pichette J, Saint-Arnaud K, et al. Intraoperative Brain Cancer Detection With Raman Spectroscopy in Humans. *Sci Trans Med* (2015) 7:274ra19. doi: 10.1126/scitranslmed.aaa2384
137. Orringer DA, Pandian B, Niknafs YS, Hollon TC, Boyle J, Lewis S, et al. Rapid Intraoperative Histology of Unprocessed Surgical Specimens *via* Fibre-Laser-Based Stimulated Raman Scattering Microscopy. *Nat Biomed Eng* (2017) 1:0027. doi: 10.1038/s41551-016-0027
138. Uckermann O, Yao W, Juratli TA, Galli R, Leipnitz E, Meinhardt M, et al. IDH1 Mutation in Human Glioma Induces Chemical Alterations That are Amenable to Optical Raman Spectroscopy. *J Neurooncol* (2018) 139:261–8. doi: 10.1007/s11060-018-2883-8
139. Zhou Y, Liu C-H, Wu B, Yu X, Cheng G, Zhu K, et al. Optical Biopsy Identification and Grading of Gliomas Using Label-Free Visible Resonance Raman Spectroscopy. *J BioMed Opt* (2019) 24(9):095001. doi: 10.1117/1.JBO.24.9.095001
140. Robert C, Tsiampali J, Fraser-Miller SJ, Neumann S, Maciaczyk D, Young SL, et al. Molecular Monitoring of Glioblastoma's Immunogenicity Using a Combination of Raman Spectroscopy and Chemometrics. *Spectrochim Acta A Mol Biomol Spectrosc* (2021) 252:119534. doi: 10.1016/j.saa.2021.119534
141. Valdés PA, Kim A, Brantsch M, Niu C, Moses ZB, Tosteson TD, et al. δ-Aminolevulinic Acid-Induced Protoporphyrin IX Concentration Correlates With Histopathologic Markers of Malignancy in Human Gliomas: The Need for Quantitative Fluorescence-Guided Resection to Identify Regions of Increasing Malignancy. *Neuro Oncol* (2011) 13:846–56. doi: 10.1093/neuonc/nor086
142. Stummer W, Pichlmeier U, Meinel T, Wiestler OD, Zanella F, Reulen HJ. Fluorescence-Guided Surgery With 5-Aminolevulinic Acid for Resection of Malignant Glioma: A Randomised Controlled Multicentre Phase III Trial. *Lancet Oncol* (2006) 7:392–401. doi: 10.1016/S1470-2045(06)70665-9
143. Valdés PA, Kim A, Leblond F, Conde OM, Harris BT, Paulsen KD, et al. Combined Fluorescence and Reflectance Spectroscopy for *In Vivo* Quantification of Cancer Biomarkers in Low- and High-Grade Glioma Surgery. *J BioMed Opt* (2011) 16:116007. doi: 10.1117/1.3646916
144. Leclerc P, Ray C, Mahieu-Williame L, Alston L, Frindel C, Brevet P-F, et al. Machine Learning-Based Prediction of Glioma Margin From 5-ALA Induced PpIX Fluorescence Spectroscopy. *Sci Rep* (2020) 10:1462. doi: 10.1038/s41598-020-58299-7
145. Lu G, Fei B. Medical Hyperspectral Imaging: A Review. *J BioMed Opt* (2014) 19:10901. doi: 10.1117/1.JBO.19.1.010901
146. Halicek M, Fabelo H, Ortega S, Callico GM, Fei B. *In-Vivo* and Ex-Vivo Tissue Analysis Through Hyperspectral Imaging Techniques: Revealing the Invisible Features of Cancer. *Cancers (Basel)* (2019) 11(6):756. doi: 10.3390/cancers11060756
147. Urbanos G, Martín A, Vázquez G, Villanueva M, Villa M, Jimenez-Roldan L, et al. Supervised Machine Learning Methods and Hyperspectral Imaging Techniques Jointly Applied for Brain Cancer Classification. *Sensors (Basel)* (2021) 21(11):3827. doi: 10.3390/s21113827
148. Manni F, van der Sommen F, Fabelo H, Zinger S, Shan C, Edström E, et al. Hyperspectral Imaging for Glioblastoma Surgery: Improving Tumor Identification Using a Deep Spectral-Spatial Approach. *Sensors (Basel)* (2020) 20(23):6955. doi: 10.3390/s20236955
149. Ortega S, Halicek M, Fabelo H, Camacho R, Plaza MdlL, Godtliebsen F, et al. Hyperspectral Imaging for the Detection of Glioblastoma Tumor Cells in H&E Slides Using Convolutional Neural Networks. *Sensors (Basel)* (2020) 20 (7):1911. doi: 10.3390/s20071911
150. Law M, Cha S, Knopp EA, Johnson G, Arnett J, Litt AW. High-Grade Gliomas and Solitary Metastases: Differentiation by Using Perfusion and Proton Spectroscopic MR Imaging. *Radiology* (2002) 222:715–21. doi: 10.1148/radiol.2223010558
151. Cordova JS, Shu HK, Liang Z, Gurbani SS, Cooper LA, Holder CA, et al. Whole-Brain Spectroscopic MRI Biomarkers Identify Infiltrating Margins in Glioblastoma Patients. *Neuro Oncol* (2016) 18:1180–9. doi: 10.1093/neuonc/now036
152. Mountford CE, Stanwell P, Lin A, Ramadan S, Ross B. Neurospectroscopy: The Past, Present and Future. *Chem Rev* (2010) 110:3060–86. doi: 10.1021/cr900250y
153. Maxwell RJ, Martinez-Perez I, Cerdan S, Cabanas ME, Arus C, Moreno A, et al. Pattern Recognition Analysis of 1H NMR Spectra From Perchloric Acid

Extracts of Human Brain Tumor Biopsies. *Magn Reson Med* (1998) 39:869–77. doi: 10.1002/mrm.1910390604

154. Somorjai RL, Dolenko B, Nikulin AK, Pizzi N, Scarth G, Zhilkin P, et al. Classification of 1H MR Spectra of Human Brain Neoplasms: The Influence of Preprocessing and Computerized Consensus Diagnosis on Classification Accuracy. *J Magn Reson Imaging* (1996) 6:437–44. doi: 10.1002/jmri.1880060305
155. Adebileje SA, Ghasemi K, Aiyelabegan HT, Saligheh Rad H. Accurate Classification of Brain Gliomas by Discriminate Dictionary Learning Based on Projective Dictionary Pair Learning of Proton Magnetic Resonance Spectra. *Magn Reson Chem* (2017) 55:318–22. doi: 10.1002/mrc.4532
156. Nunez LM, Romero E, Julia-Sape M, et al. Unraveling Response to Temozolomide in Preclinical GL261 Glioblastoma With MRI/MRSI Using Radiomics and Signal Source Extraction. *Sci Rep* (2020) 10:19699. doi: 10.1038/s41598-020-76686-y
157. Stead WW. Clinical Implications and Challenges of Artificial Intelligence and Deep Learning. *JAMA* (2018) 320:1107–8. doi: 10.1001/jama.2018.11029
158. Verma V, Simone CB2nd, Krishnan S, Lin SH, Yang J, Hahn SM. The Rise of Radiomics and Implications for Oncologic Management. *J Natl Cancer Inst* (2017) 109(7). doi: 10.1093/jnci/djx055
159. Limkin EJ, Sun R, Dercle L, Zacharaki EI, Robert C, Reuze S, et al. Promises and Challenges for the Implementation of Computational Medical Imaging (Radiomics) in Oncology. *Ann Oncol* (2017) 28:1191–206. doi: 10.1093/annonc/mdx034
160. Chetan MR, Gleeson FV. Radiomics in Predicting Treatment Response in non-Small-Cell Lung Cancer: Current Status, Challenges and Future Perspectives. *Eur Radiol* (2021) 31:1049–58. doi: 10.1007/s00330-020-07141-9
161. Jia S, Lin P, Li Z, Zhang J, Liu S. Visualizing Surrogate Decision Trees of Convolutional Neural Networks. *J Visualization* (2020) 23:141–56. doi: 10.1007/s12650-019-00607-z
162. Takahashi S, Takahashi M, Tanaka S, Takayanagi S, Takami H, Yamazawa E, et al. A New Era of Neuro-Oncology Research Pioneered by Multi-Omics Analysis and Machine Learning. *Biomolecules* (2021) 11(4):565. doi: 10.3390/biom11040565
163. Hamamoto R, Komatsu M, Takasawa K, Asada K, Kaneko S. Epigenetics Analysis and Integrated Analysis of Multiomics Data, Including Epigenetic Data, Using Artificial Intelligence in the Era of Precision Medicine. *Biomolecules* (2019) 10. doi: 10.3390/biom10010062
164. Chaddad A, Daniel P, Sabri S, Desrosiers C, Abdulkarim B. Integration of Radiomic and Multi-Omic Analyses Predicts Survival of Newly Diagnosed IDH1 Wild-Type Glioblastoma. *Cancers (Basel)* (2019) 11. doi: 10.3390/cancers11081148
165. Zhang Y, Li A, He J, Wang M. A Novel MKL Method for GBM Prognosis Prediction by Integrating Histopathological Image and Multi-Omics Data. *IEEE J Biomed Health Inf* (2020) 24:171–9. doi: 10.1109/JBHI.2019.2898471

**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.

frontiers | Frontiers in Oncology

TYPE Systematic Review
PUBLISHED 15 August 2022
DOI 10.3389/fonc.2022.915481

OPEN ACCESS

EDITED BY
Jakub Nalepa,
Silesian University of Technology,
Poland

REVIEWED BY
Gwang Ha Kim,
Pusan National University Hospital,
South Korea
Hakan Senturk,
Bezmiâlem Vakıf Üniversitesi, Turkey

*CORRESPONDENCE
Xiao-Yu Li
lixiaoyu05@163.com

SPECIALTY SECTION
This article was submitted to
Cancer Imaging and
Image-directed Interventions,
a section of the journal
Frontiers in Oncology

RECEIVED 08 April 2022
ACCEPTED 18 July 2022
PUBLISHED 15 August 2022

CITATION
Liu X-Y, Song W, Mao T, Zhang Q,
Zhang C and Li X-Y (2022) Application
of artificial intelligence in the diagnosis
of subepithelial lesions using
endoscopic ultrasonography: a
systematic review and meta-analysis.
*Front. Oncol.* 12:915481.
doi: 10.3389/fonc.2022.915481

# Application of artificial intelligence in the diagnosis of subepithelial lesions using endoscopic ultrasonography: a systematic review and meta-analysis

Xin-Yuan Liu, Wen Song, Tao Mao, Qi Zhang, Cuiping Zhang
and Xiao-Yu Li*

Department of Gastroenterology, The Affiliated Hospital of Qingdao University, Qingdao, China

Endoscopic ultrasonography (EUS) is the most common method for diagnosing gastrointestinal subepithelial lesions (SELs); however, it usually requires histopathological confirmation using invasive methods. Artificial intelligence (AI) algorithms have made significant progress in medical imaging diagnosis. The purpose of our research was to explore the application of AI in the diagnosis of SELs using EUS and to evaluate the diagnostic performance of AI-assisted EUS. Three databases, PubMed, EMBASE, and the Cochrane Library, were comprehensively searched for relevant literature. RevMan 5.4.1 and Stata 17.0, were used to calculate and analyze the combined sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), diagnostic odds ratio (DOR), and summary receiver-operating characteristic curve (SROC). Eight studies were selected from 380 potentially relevant studies for the meta-analysis of AI-aided EUS diagnosis of SELs. The combined sensitivity, specificity, and DOR of AI-aided EUS were 0.92 (95% CI, 0.85-0.96), 0.80 (95% CI, 0.70-0.87), and 46.27 (95% CI, 19.36-110.59), respectively). The area under the curve (AUC) was 0.92 (95% CI, 0.90-0.94). The AI model in differentiating GIST from leiomyoma had a pooled AUC of 0.95, sensitivity of 0.93, specificity of 0.88, PLR of 8.04, and NLR of 0.08. The combined sensitivity, specificity, and AUC of the AI-aided EUS diagnosis in the convolutional neural network (CNN) model were 0.93, 0.81, and 0.94, respectively. AI-aided EUS diagnosis using conventional brightness mode (B-mode) EUS images had a combined sensitivity of 0.92, specificity of 0.79, and AUC of 0.92. AI-aided EUS diagnosis based on patients had a combined sensitivity, specificity, and AUC of 0.95, 0.83, and 0.96, respectively. Additionally, AI-aided EUS was superior to EUS by experts in terms of sensitivity (0.93 vs. 0.71), specificity (0.81 vs. 0.69), and AUC (0.94 vs. 0.75). In conclusion, AI-assisted EUS is a promising and reliable method for distinguishing SELs, with excellent diagnostic performance. More multicenter cohort and prospective studies are expected to be conducted to further develop AI-assisted real-time diagnostic systems and validate the superiority of AI systems.

**Systematic Review Registration:** PROSPERO (https://www.crd.york.ac.uk/PROSPERO/), identifier CRD42022303990.

KEYWORDS

artificial intelligence, computer-assisted diagnosis, endoscopic ultrasonography, subepithelial lesions, gastrointestinal stromal tumors

## Introduction

Gastrointestinal subepithelial lesions (SELs) are tumors that originate from the muscularis mucosa, submucosa, or muscularis propria (1). According to statistics, one SEL is found in every 300 endoscopy examinations (2). SELs, including gastrointestinal stromal tumors (GIST), leiomyomas, schwannomas, neuroendocrine tumors (NET), lipomas, and ectopic pancreas, are asymptomatic and difficult to distinguish due to their similar morphology in size, shape, surface color, contour, and margin (1). GISTs are the most prevalent SELs, with a prevalence of 14–20 cases per million, and have the potential to evolve into malignancies (3, 4). Approximately 60% of patients with GISTs can be cured by surgery (5). Therefore, it is crucial to differentiate GISTs from other benign tumors.

With the development of endoscopic ultrasonography (EUS), fine-needle aspiration biopsy (FANB), immunohistochemical staining methods, and various new imaging technologies, such as contrast-enhanced harmonic EUS (CH-EUS), the approaches for diagnosing and treating SELs have improved (6). EUS as a useful tool has recently become the conventional inspection method for the discovery and diagnosis of SELs. However, the diagnostic accuracy of EUS is limited and closely related to the professional level and experience of the endoscopists (7). EUS-FNAB can be used to obtain tissue specimens for immunohistochemical staining and is the gold standard for diagnosing SELs. Nevertheless, the diagnostic yield of EUS-FNAB for SELs is not ideal, ranging from 60% to 85% (8–10). FNAB is an invasive and risky operation, and the limited sampling sites are subjectively determined by endoscopists, which may lead to missed diagnoses. Therefore, alternative methods are needed for the accurate diagnosis of SELs to avoid surgical resection of benign lesions as GISTs with malignant potential.

Recently, artificial intelligence (AI) has been extensively used in medical imaging technology, owing to its superior performance. Machine learning (ML) involves the fields of computer science and statistics, generating algorithms to analyze various types of data, and building appropriate descriptive and predictive models (11). Artificial neural networks (ANN), as mathematical models of information processing, are supervised ML models inspired by the structure of brain synaptic connections (11). A convolutional neural network (CNN) is a deep learning algorithm that shows strong performance in image recognition, classification, and processing (12). AI-aided EUS diagnostic tools have been widely applied to differentiate various types of pancreatic diseases, such as pancreatic tumors, chronic pancreatitis, and autoimmune pancreatitis (13–15). In recent years, several studies have explored the value of CNN in distinguishing SELs based on EUS images, mainly in differentiating GIST from benign lesions. In this systematic review and meta-analysis, we aimed to assess the effectiveness and accuracy of AI in diagnosing SELs using EUS images and focused on the performance of computer-aided diagnosis models in differentiating GIST from other benign lesions by comparing AI and EUS experts.

## Methods

### Search strategy

This study followed the preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines (16). The PubMed, Embase, and Cochrane Library databases were systematically and comprehensively searched for studies on the AI-aided diagnostic accuracy of SELs under EUS with or without EUS experts as controls published until February 2022. Search terms in the title, abstract, and keywords are as follows: (“artificial intelligence” OR “AI” OR “machine learning” OR “deep learning” OR “convolutional neural network” OR “computer-assisted” OR “computer-aided” OR “neural network” OR “digital image analysis” OR “digital image processing”) AND (“endoscopic ultrasound” OR “endoscopic ultrasonography” OR “EUS”). To avoid omissions, the SELs were not included in the retrieval strategy. The retrieved articles were screened independently by two investigators (Xin-Y L and WS). Disagreements were discussed and resolved by a third researcher (TM). This protocol was registered with PROSPERO (CRD42022303990).

## Inclusion and exclusion criteria

The inclusion criteria for studies were as follows (1): prospective or retrospective study design; (2) study subjects were adult participants (≥18 years old); (3) all SELs patients were diagnosed based on histopathological diagnosis after surgical or endoscopic resection or EUS-FNAB; (4) AI algorithm was applied to the diagnosis of patients with SELs using EUS images; (5) study results demonstrated the diagnostic performance of computer-aided diagnosis (CAD) algorithms, including area under the curve (AUC), sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), diagnostic odds ratio (DOR), or accuracy, enabling the calculation of true positive (TP), false positive (FP), true negative (TN), and false negative (FN); and (6) the manuscript was written in English. Conference proceedings, case reports, narrative and systematic reviews, meta-analyses, and studies with incomplete data (TP, FP, TN, and FN could not be calculated) were excluded. Studies with failed randomization and significant differences in baseline data between groups were also excluded from this systematic review.

## Data extraction and quality assessment

The number of histologically confirmed SELs that were true-positive (GIST considered to be GIST by AI or experts), true-negative (non-GIST considered to be non-GIST by AI or experts), false-positive (non-GIST considered to be GIST), or false-negative (GIST considered to be non-GIST) were extracted. Additionally, the first author's name; year of publication; country where the study was conducted; study type; number of samples in the training, validation, and test sets; imaging modality; AI model; and video were also retrieved.

The Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool was used to evaluate the quality and potential bias of all included studies in four aspects: patient selection, index test, reference standard, and flow and timing quality (17). Regarding the problem of pre-specified thresholds, we referred to the study by Thaninee et al. and modified the problem as to whether the performance of the AI-assisted diagnostic system was validated in another cohort (18). Two reviewers (Xin-Y L and WS) independently assessed the eight studies, and conflicts were discussed and resolved with a third reviewer (TM).

## Statistical analysis

RevMan 5.4.1 (The Cochrane Collaboration, 2020, London, United Kingdom) and Stata 17.0 (StataCorp, College Station, TX, USA) were used for diagnostic meta-analysis. Published data were extracted, including the reported TP, FP, FN, TN, sensitivity, and specificity of the test datasets. The pooled sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), diagnostic score, and DOR with 95% confidence intervals (CIs) were calculated and analyzed using the bivariate mixed-effects model. A summary receiver-operating characteristic curve (SROC) was also constructed, and the AUC was calculated to assess diagnostic accuracy. A funnel plot and its symmetric distribution were used to evaluate the risk of publication bias. Subgroup and meta-regression analyses were performed to explore the sources of heterogeneity. Heterogeneity among the studies was determined using $I^2$ and Cochran's Q tests. $P < 0.1$ generally suggests significant heterogeneity, and $I^2 > 50\%$ indicates substantial heterogeneity.

# Result

## Literature search and bias assessment

The literature retrieval process and screening results are shown in Figure 1. Initially, 380 potentially relevant studies were retrieved from the three databases, and 98 duplicates were removed. Subsequently, 268 studies were excluded after reviewing the titles and abstracts, as they were irrelevant articles and were not suitable for the research topic or type. After screening the full text of 14 eligible studies, two studies that did not meet the eligibility criteria and four studies related to GIST malignant potential were excluded. Finally, eight studies were selected for the meta-analysis of AI-aided diagnosis of SELs according to the PRISMA flowchart (19–26).

The characteristics of all included studies are summarized in Table 1. A total of eight studies with 339 patients with GIST and 194 patients without GIST were included in the meta-analysis, seven of which were within the last three years. They were all retrospective studies, and one of them used both retrospective and prospective test sets in the stage of testing AI models (22). Three studies were conducted in Japan, two in South Korea, and three in China, Turkey, and the United States. Only one study developed an AI model based on contrast-enhanced harmonic EUS (CH-EUS) images, whereas the others used the conventional brightness mode (B-mode) of EUS. Considering computer-aided models, except for one study that used the ANN model, the remaining studies applied the CNN model. Only one study did not use EUS experts as controls (19). The training, validation, and testing datasets of the included studies are presented in Supplementary Table 1. All the studies trained and developed AI models using a large number of EUS images. One of the studies used videos

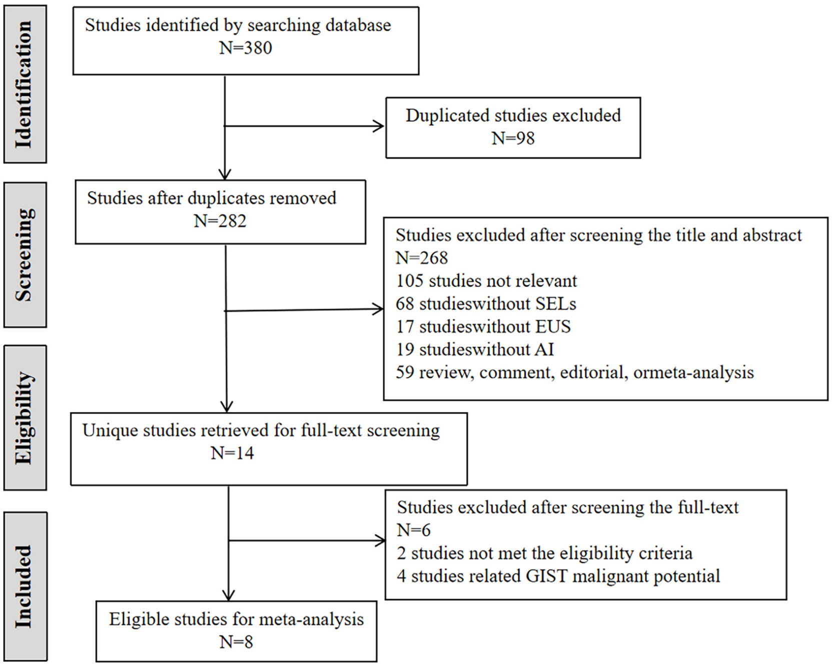


FIGURE 1
Flowchart of literature search.

from each patient divided into 0.1s intervals, yielding images to train the AI model (24).

The quality and risk of bias of the included studies determined using the QUADAS-2 tool are presented in Figure 2. One meta-analysis of AI-aided diagnosis of GIST identified a high-risk bias in patient selection (19).

The slope coefficient of the Deeks' funnel plot was symmetrical ($p = 0.14$) (Figure 3), indicating that publication bias was insignificant.

TABLE 1 Characteristics of included studies.

| Author | Year | Study type | Country | GIST | non-GIST | TP | FP | FN | TN | Reference standard | Imaging modality | AI model | EUS experts as control | Video | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Vien X. Nguyen | 2010 | Retrospective | USA | 124*/ 28 | 217*/ 18 | 100 | 46 | 24 | 171 | Histopathology | B-mode | ANN | N | N | (19) |
| Yosuke Minoda | 2020 | Retrospective | Japan | 47 | 13 | 42 | 4 | 5 | 9 | Histopathology | B-mode | CNN | Y | N | (20) |
| Yoon Ho Kim | 2020 | Retrospective | Korea | 106*/ 32 | 106*/ 37 | 88 | 26 | 18 | 80 | Histopathology | B-mode | CNN | Y | N | (21) |
| Xintian Yang | 2021 | Retrospective & Prospective | China | 30**/ 36 | 54**/ 41 | 27/ 32 | 2/ 14 | 3/4 | 52/ 27 | Histopathology | B-mode | CNN | Y | N | (22) |
| Chang Kyo Oh | 2021 | Retrospective | Korea | 40 | 14 | 40 | 2 | 0 | 12 | Histopathology | B-mode | CNN | Y | N | (23) |
| Keiko Hirai | 2021 | Retrospective | Japan | 85 | 37 | 84 | 12 | 1 | 25 | Histopathology | B-mode | CNN | Y | N | (24) |
| Gulseren Seven | 2021 | Retrospective | Turkey | 35 | 10 | 32 | 4 | 3 | 6 | Histopathology | B-mode | CNN | Y | N | (25) |
| Hidekazu Tanaka | 2022 | Retrospective | Japan | 42 | 11 | 38 | 1 | 4 | 10 | Histopathology | CH-EUS | CNN | Y | Y | (26) |

*ROI, region of interest, not patient.
**Data of retrospective diagnostic test.

FIGURE 2
Quality assessment of included studies using QUADAS-2.

## Diagnostic performance of AI-assisted EUS in GIST

We incorporated data from all retrospective diagnostic test sets and performed a meta-analysis of the eight included studies. The pooled sensitivity of AI-aided EUS diagnosis of GIST was 0.92 (95% CI, 0.85-0.96) (Figure 4A) and specificity was 0.80 (95% CI, 0.70-0.87) (Figure 4B). The pooled PLR and NLR were 4.61 (95% CI, 3.00-7.08) (Figure 4C) and 0.10 (95% CI, 0.05-0.19) (Figure 4D), respectively. The diagnostic score and DOR were 3.83 (95% CI, 2.96-4.71) and 46.27 (95% CI, 19.36-110.59), respectively (Supplementary Figure 1). Figure 5A shows the SROC curve of AI-aided EUS, with an AUC of 0.92 (95% CI, 0.90-0.94).

Subsequently, we expanded the sample size by including the data from a prospective diagnostic test set. The combined results of AI-assisted EUS diagnosis of GIST were shown as follows: AUC of 0.92 (95% CI, 0.89-0.94) (Figure 5B), sensitivity 0.92 (95% CI, 0.85-0.95), specificity 0.78 (95% CI, 0.69-0.86), PLR 4.23 (95% CI, 2.88-6.22), and NLR 0.11 (95% CI, 0.06-0.19) (Supplementary Figure 2). The diagnostic score and DOR were 3.67 (95% CI, 2.90-4.45) and 39.40 (95% CI, 18.20-85.30), respectively (Supplementary Figure 3).

To investigate the clinical application of AI in the diagnosis of GIST, we generated a Fagan diagram (Figure 6). Assuming a 20% prevalence of GIST, the diagram shows a posterior probability of 54% for GIST if the test is positive, and approximately 2% for a negative test.

## Subgroup analysis of AI-assisted EUS

The specific types of SELs in the included studies are shown in Supplementary Table 2. One study involved five SELs, including GIST, leiomyomas, schwannomas, NET, and ectopic pancreas (24). Four studies developed AI only for the differential diagnosis of GIST and leiomyoma (22, 23, 25, 26), and a subgroup analysis of these four studies was conducted to explore the discriminating ability of the two diseases. The AI model had a pooled AUC of 0.95 (95% CI, 0.93-0.97), sensitivity of 0.93 (95% CI, 0.87-0.97), specificity of 0.88 (95% CI, 0.71-0.96), PLR of 8.04 (95% CI,

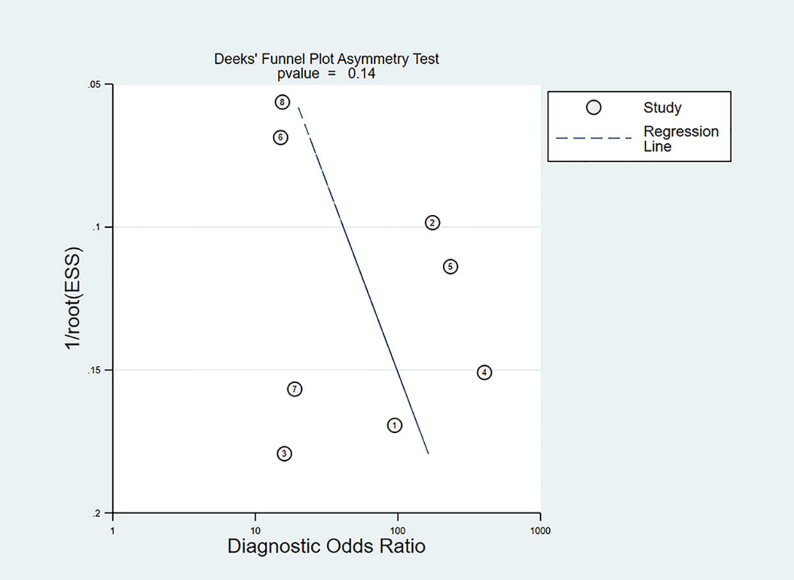


FIGURE 3
Deeks' funnel plot of publication bias.

2.92-22.18), and NLR of 0.08 (95% CI, 0.04-0.15) (Supplementary Figures 4, 5).

We performed a subgroup analysis after excluding Nguyen's study, as the AI model adopted was ANN. The combined sensitivity and specificity of AI-assisted EUS diagnosis of GIST on the CNN model were 0.93 (95% CI, 0.87-0.97) and 0.81 (95% CI, 0.68-0.89) (Supplementary Figures 6A, B), respectively. The pooled PLR was 4.85 (95% CI, 2.81-8.36) and NLR was 0.08 (95% CI, 0.04-0.17) (Supplementary Figures 6C, D). Figure 7A shows the SROC curve of the AI-assisted EUS, with an AUC of 0.94 (95% CI, 0.92-0.96). The $I^2$ was 50.57% for PLR, 74.05% for sensitivity, 71.16% for specificity, and 73.61 for NLR, indicating that significant heterogeneity existed in the pooled sensitivity, specificity, and NLR.

We also conducted a subgroup analysis of seven studies on imaging modalities without CH-EUS. The AI model had a pooled AUC of 0.92 (95% CI, 0.89-0.94) (Figure 7B), sensitivity of 0.92 (95% CI, 0.84-0.97), specificity of 0.79 (95% CI, 0.68-0.87), PLR of 4.39 (95% CI, 2.85-6.78), and NLR of 0.10 (95% CI, 0.04-0.21) (Supplementary Figure 7). However, the heterogeneity within the subgroups was still significantly high.

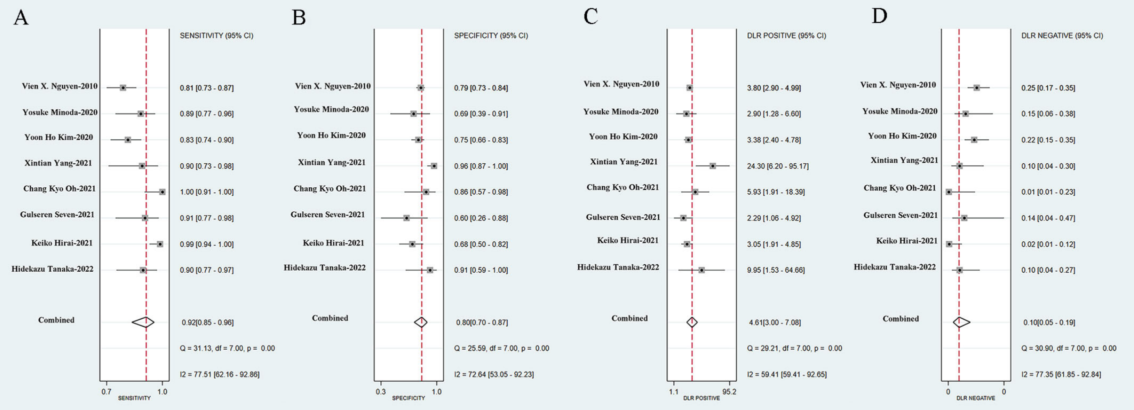


FIGURE 4
Sensitivity **(A)**, specificity **(B)**, positive likelihood ratio **(C)**, negative likelihood ratio **(D)** of AI-assisted EUS diagnosis of GIST.

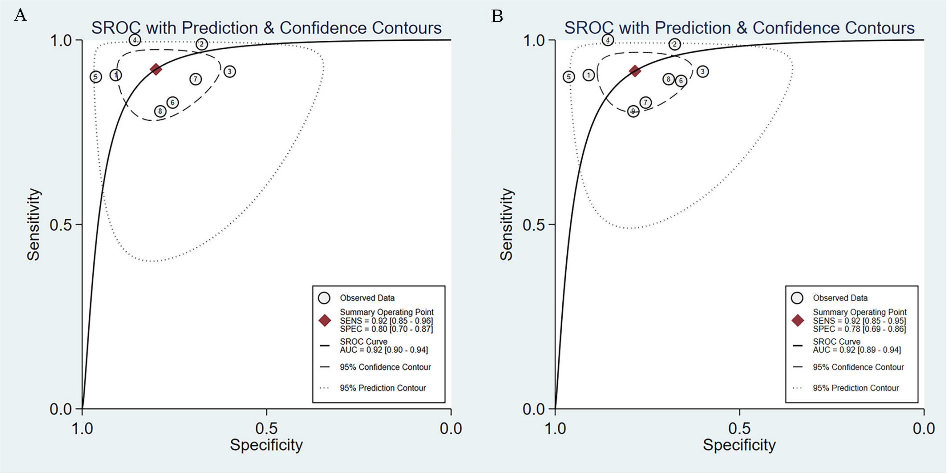


FIGURE 5
SROC curves of AI-assisted EUS diagnosis of GIST. **(A)** The SROC curve of eight studies. **(B)** SROC curve of nine datasets including prospective diagnostic test set.

A subgroup analysis of AI-assisted EUS diagnosis of GIST was performed on the study subjects, namely six studies based on patients and not regions of interest. As shown in Supplementary Figure 8, the combined sensitivity, specificity, PLR, and NLR were 0.95 (95% CI, 0.89-0.97), 0.83 (95% CI, 0.67-0.92), 5.43 (95% CI, 2.75-10.71), 0.07 (95% CI, 0.03-0.13), respectively. The SROC curve, with an AUC of 0.96 (95% CI, 0.94-0.97), is displayed in Figure 7C. $I^2$ was 38.69% for PLR, 51.31% for NLR, 58.46% for sensitivity, and 71.06% for specificity, indicating a low degree of heterogeneity in PLR, whereas there was moderate heterogeneity in NLR, sensitivity, and specificity.

To further explore the source of heterogeneity, we performed meta-regression analysis. The number of samples was a major source of heterogeneity in univariate meta-regression analysis ($p$ <0.001, Figure 8). Study quality ($p$ = 0.03) and study subjects ($p$ = 0.01) were major sources of heterogeneity in the joint meta-regression model (Table 2).

TABLE 2 P-value of parameters in the joint model.

| Parameter | LRTChi2 | P-value | I2 | I2lo | I2hi |
|---|---|---|---|---|---|
| Quality* | 6.73 | 0.03 | 70 | 34 | 100 |
| Number | 4.59 | 0.10 | 56 | 2 | 100 |
| Publish year | 2.11 | 0.35 | 5 | 0 | 100 |
| AI model | 2.11 | 0.35 | 5 | 0 | 100 |
| Imaging form | 0.90 | 0.64 | 0 | 0 | 100 |
| Study subject* | 8.85 | 0.01 | 77 | 51 | 100 |

*p <0.05

## Comparison between AI and EUS experts

Seven studies simultaneously tested the accuracy of EUS experts in the diagnosis of GIST. All EUS experts performed more than 500 EUS examinations or had at least 5-year experience in evaluating gastrointestinal SELs. The SROC curve of the EUS experts, with an AUC of 0.75 (95% CI, 0.71-0.78), is displayed in Figure 7D. The pooled sensitivity of EUS experts in diagnosing GIST was 0.71 (95% CI, 0.63-0.78) (Figure 9A) and specificity was 0.69 (95% CI, 0.61-0.76) (Figure 9B). The combined PLR and NLR are 2.28 (95% CI, 1.85-2.82) (Figure 9C) and 0.42 (95% CI, 0.33-0.54) (Figure 9D), respectively. There was little heterogeneity in the specificity ($p$ = 0.37), PLR ($p$ = 0.69), and NLR ($p$ = 0.12).

For diagnosis of GIST under EUS, AI was superior to EUS experts in terms of sensitivity [0.93 (95% CI, 0.87-0.97) vs. 0.71 (95% CI, 0.63-0.78)], specificity [0.81 (95% CI, 0.68-0.89) vs. 0.69 (95% CI, 0.61-0.76)], and PLR [4.85 (95% CI, 2.81-8.36) vs. 2.28 (95% CI, 1.85-2.82)], and NLR [0.08 (95% CI, 0.04-0.17) vs. 0.42 (95% CI, 0.33-0.54)]. Figure 10 shows the comparison of SROC curves between AI-assisted EUS models and EUS experts with AUC of 0.94 (95% CI, 0.92-0.96) vs. 0.75 (95% CI, 0.71-0.78), suggesting that AI-assisted EUS models have better diagnostic performance.

## Discussion

With the application of artificial intelligence in medical imaging technology, an increasing number of diseases have advanced their diagnosis and treatment methods. In this

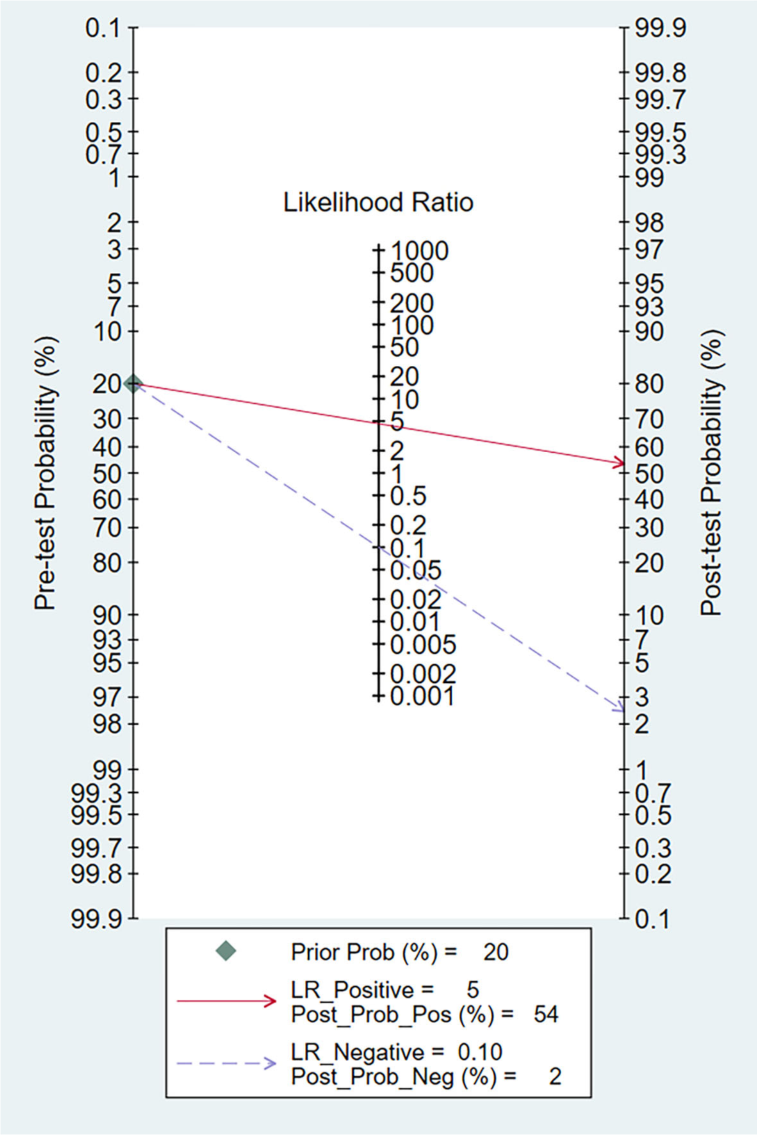


FIGURE 6
Fagan normogram for the prediction of GISTs in EUS images.

systematic review and meta-analysis, we explored the application of computer-aided diagnosis systems in gastrointestinal SELs and found that artificial intelligence algorithm models have excellent diagnostic performance with a sensitivity of 0.92 (95% CI, 0.85-0.96) and specificity of 0.80 (95% CI, 0.70-0.87). EUS is currently the most accurate and prevalent imaging modality for evaluating gastrointestinal SELs because of its ability to penetrate tissue layers and, thus, most likely identify the origin of the lesion (1). A previous study has shown that CH-EUS has better diagnostic performance than B-mode EUS in distinguishing leiomyomas from GIST and discriminating the risk stratification of GIST (27). In addition to improving the equipment performance and imaging technology of EUS, the application of artificial intelligence undoubtedly compensates for the limitations of EUS. With the help of the AI system, it is expected to shorten the diagnostic time, improve diagnostic efficiency, and reduce the misdiagnosis rate of GIST and other benign lesions, thus avoiding unnecessary EUS tests, invasive biopsies, and surgical operations.

In our initial literature search, we found that Kim and Lee used digital image analysis of objective information provided by EUS images to diagnose gastric stromal tumors (28, 29). We excluded these two studies because they were limited to analyzing the features of EUS images and did not develop

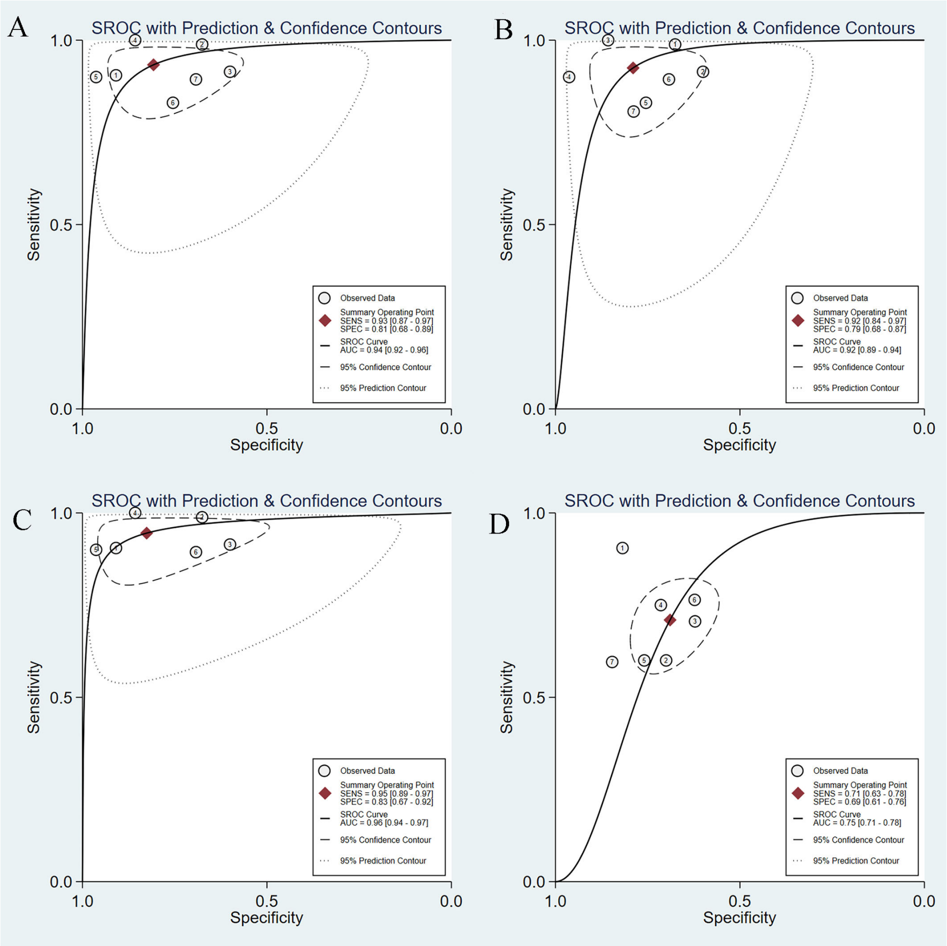


FIGURE 7
SROC curves of AI-assisted EUS and EUS experts in diagnosis of GIST. **(A)** SROC curve of seven studies on CNN AI-models. **(B)** SROC curve of seven studies on imaging modality. **(C)** SROC curve of seven studies based on patients. **(D)** The SROC curve of the EUS experts.

corresponding AI models. We also found four studies that explored the application of AI in the malignancy stratification of GISTs, and the overall accuracy of the AI models in predicting the malignant potential of GISTs was 66.0%-83.4% (30–33). During the literature search, we found that several studies have explored the application of AI in SELs, especially GISTs. Therefore, we systematically and comprehensively summarized the application of AI-assisted EUS for the diagnosis of SELs. Although there are many types of SELs, most studies classified SELs into two categories: GIST and non-GIST, to explore the accuracy of AI-assisted EUS. In four studies, the non-GISTs only referred to leiomyoma, and we performed a subgroup analysis (22, 23, 25, 26). Nguyen et al. developed an ANN with excellent performance for differentiating lipomas (AUC=0.92), carcinoids (AUC=0.86), and GISTs (AUC=0.89) (19). Despite the SELs involved in the Minoda's research, including GIST, leiomyoma, schwannoma, and aberrant pancreas, the results section was still divided into GIST and non-GIST for exploration (20). Kim et al. utilized CNN-CAD to first classify SELs into GIST and non-GIST tumors, and then sub-classified the non-GIST tumors into leiomyomas and schwannomas. Accuracy of the CNN-CAD system in differentiating leiomyomas from schwannomas was 85.0% (95% CI: 81.6-87.7%) (21). In the Hirai's study, accuracy of the AI system for five-category classification was 86.1%, including GIST, leiomyoma, schwannoma, NET, and ectopic pancreas (24).

Nguyen trained, constructed, and internally validated an ANN through unsupervised and supervised learning based on the features extracted through texture analysis (19). In the traditional sense, ANN is a type of machine learning (ML). As a computer application, ML can recognize patterns in training data and generate mathematical models to develop an AI system to realize the recognition and prediction function, similar to the learning behavior of humans (13). Other studies trained CNN

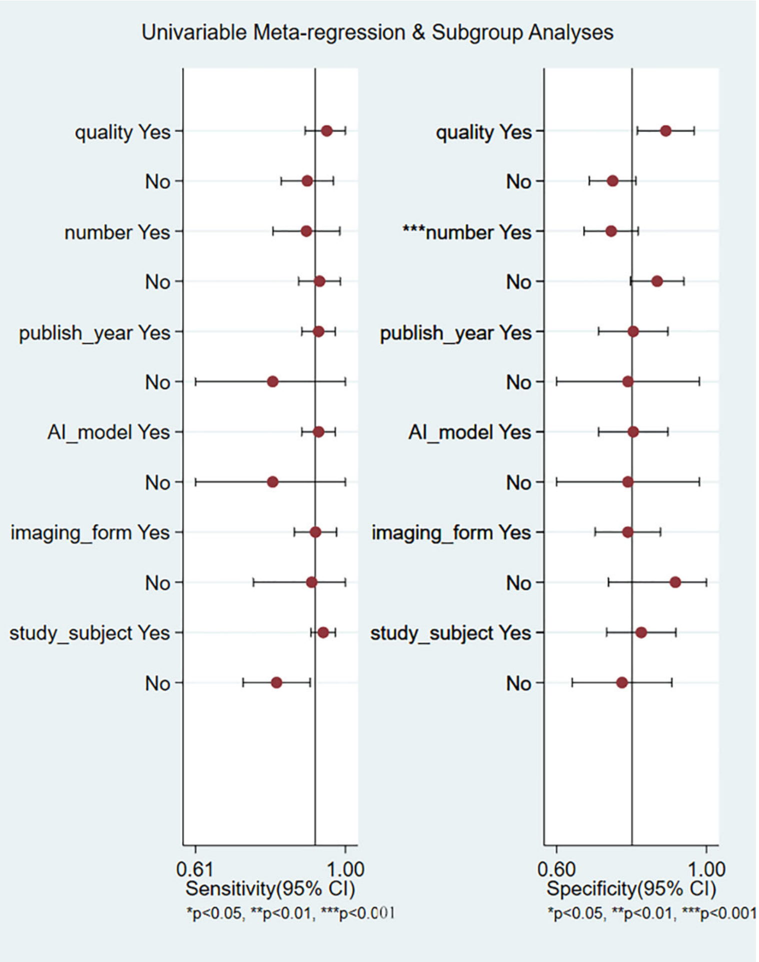


FIGURE 8
Univariate meta-regression for the reason of heterogeneity in sensitivity and specificity.

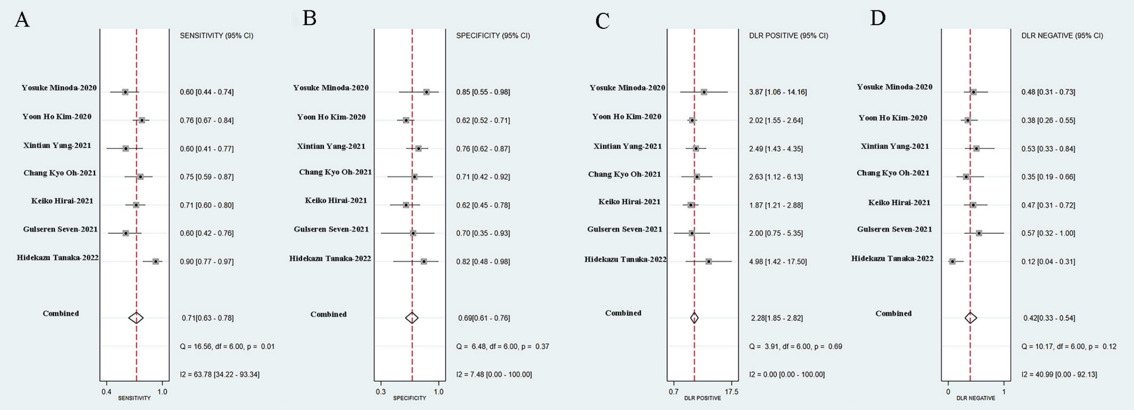


FIGURE 9
Sensitivity **(A)**, specificity **(B)**, positive likelihood ratio **(C)**, negative likelihood ratio **(D)** of EUS experts in diagnosis of GIST.

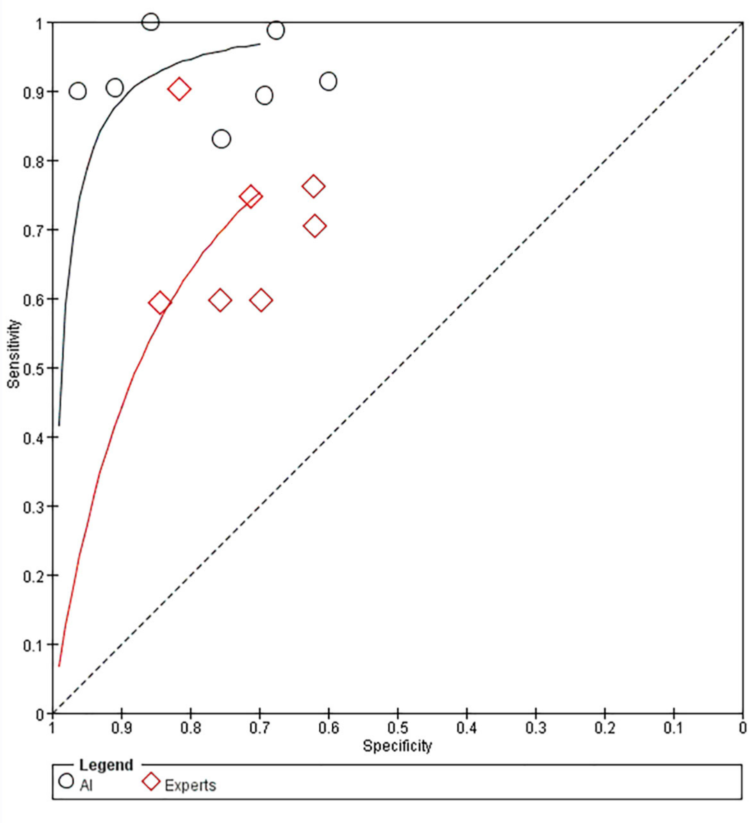


FIGURE 10
Comparison of SROC curves between AI-assisted EUS models and EUS experts.

models using deep-learning algorithms. Deep learning-based analysis does not need to measure characteristic values, as they can be automatically and accurately identified, thereby demonstrating greater diagnostic ability (34). This is consistent with our findings that the combined AUC of CNN model after excluding the ANN model was improved from 0.92 (95% CI, 0.90-0.94) to 0.94 (95% CI, 0.92-0.96).

Heterogeneity is a prominent issue in this meta-analysis. Although we performed subgroup analyses based on the AI models, imaging modalities, and study subjects, the heterogeneity was not completely eliminated. Possible reasons for this are as follows: First, we have to consider the diversity of clinical samples, as most of the included studies were from different countries, and the manufacturers and models of EUS were inconsistent. In addition, the sample size was not sufficiently large. Second, methodological diversity should be considered. The specific algorithms, tools used, and parameter settings were not uniform, despite the fact that seven studies applied the CNN deep-learning model. The EUS expert group had little heterogeneity, probably because all EUS experts were selected on the basis of having performed more than 500 EUS examinations or having at least 5 years of experience in assessing gastrointestinal SELs. Additionally, different trial designs also contributed to the heterogeneity. Only two studies applied training, validation, and test sets (22, 24). Others merely had two datasets: one set to develop the AI model and the other to validate it. Considering the existence of heterogeneity, we avoided directly adopting a fixed-effects model.

In this review and meta-analysis, the diagnostic performance of AI models was superior to EUS experts, with the accuracy of 0.94 (95% CI, 0.92-0.96) vs. 0.75 (95% CI, 0.71-0.78). Additionally, two studies also investigated the diagnostic accuracy of AI-assisted EUS according to the size of SELs, ≥ 20 mm and <20 mm. Minoda et al. found that the accuracy, sensitivity, and specificity of SELs ≥ 20 mm between AI-assisted EUS and EUS experts were 90.0% vs. 53.3%, 91.7% vs. 50.0%, and 83.3% vs. 83.3%, respectively. The diagnostic performance for SELs ≥ 20 mm of AI-assisted EUS was significantly better than that of EUS experts, with an AUC of 0.965 vs. 0.684 ($p$ = 0.007) (20). Tanaka et al. discovered that the diagnostic performance of AI and experts was completely consistent for cases with lesions <20 mm, but the specificity and accuracy of AI in diagnosing GISTs ≥ 20 mm were superior to those of experts (87.5% vs. 75.0% and 88.9% vs. 86.1%, respectively) (26).

Therefore, we need to further develop and improve artificial intelligence algorithms to improve their performance in the diagnosis of small lesions.

This is the first systematic review and meta-analysis of AI-assisted EUS for SEL diagnosis. We summarized recent advances in AI in the diagnosis and differential diagnosis of SELs and evaluated the overall diagnostic performance of AI. Our meta-analysis also has some limitations. Although no publication bias existed, the number of eligible studies was limited (n=8) and most of the included studies were retrospective. Future studies are expected to expand the sample size, supplement videos, add external validation datasets, and conduct prospective real-time clinical studies to further confirm the credibility of AI diagnostic performance. In addition, the issue of heterogeneity among studies is also discussed above.

In conclusion, AI-assisted EUS is a promising and reliable method for differentiating SELs with high accuracy, and may become an important tool to assist endoscopists in diagnosing SELs in the near future.

## 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

X-YLiu, WS, and TM was responsible for literature search and screening, data extraction, analysis and interpretation, and manuscript writing. QZ and CZ was responsible for the statistical analysis and revising the manuscript. X-YLi was responsible for revising the manuscript, financial support, and final approval of the manuscript. All authors have read and approved the final version of the manuscript.

## Funding

The study was supported by the National Natural Science Foundation (No. 81802777), the "Clinical medicine + X" scientific research project of Affiliated Hospital of Qingdao University, and Qingdao Chinese Medicine Technology Project (2021-zyym26).

## Acknowledgments

We would like to thank Editage (www.editage.com) for English language editing. And we thank all the authors for helping with the writing and publication of this article.

## 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

## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.915481/full#supplementary-material

SUPPLEMENTARY FIGURE 1
The diagnostic score and DOR of AI-assisted EUS diagnosis of GIST.

SUPPLEMENTARY FIGURE 2
Sensitivity **(A)**, specificity **(B)**, positive likelihood ratio **(C)**, negative likelihood ratio **(D)** of AI-assisted EUS diagnosis of GIST including prospective test set.

SUPPLEMENTARY FIGURE 3
The diagnostic score and DOR of AI-assisted EUS diagnosis of GIST including prospective diagnostic test set.

SUPPLEMENTARY FIGURE 4
SROC curves of AI-assisted EUS of four studies for the diagnosis of GIST and leiomyoma.

SUPPLEMENTARY FIGURE 5
Sensitivity **(A)**, specificity **(B)**, positive likelihood ratio **(C)**, negative likelihood ratio **(D)** of AI-assisted EUS for the differential diagnosis of GIST and leiomyoma.

SUPPLEMENTARY FIGURE 6
Sensitivity **(A)**, specificity **(B)**, positive likelihood ratio **(C)**, negative likelihood ratio **(D)** of AI-assisted EUS of seven studies on CNN AI-models.

SUPPLEMENTARY FIGURE 7
Sensitivity **(A)**, specificity **(B)**, positive likelihood ratio **(C)**, negative likelihood ratio **(D)** of AI-assisted EUS of seven studies on imaging modality.

SUPPLEMENTARY FIGURE 8
Sensitivity **(A)**, specificity **(B)**, positive likelihood ratio **(C)**, negative likelihood ratio **(D)** of AI-assisted EUS of seven studies based on patients.

SUPPLEMENTARY TABLE 1
The training, validation, and test datasets of included studies.

SUPPLEMENTARY TABLE 2
Types of SELs of included studies.

## References

1. Faulx AL, Kothari S, Acosta RD, Agrawal D, Bruining DH, Chandrasekhara V, et al. The role of endoscopy in subepithelial lesions of the GI tract. *Gastrointest Endosc* (2017) 85:1117–32. doi: 10.1016/j.gie.2017.02.022

2. Papanikolaou IS, Triantafyllou K, Kourikou A, Rösch T. Endoscopic ultrasonography for gastric submucosal lesions. *World J Gastrointest Endosc* (2011) 3:86–94. doi: 10.4253/wjge.v3.i5.86

3. Rubin BP, Heinrich MC, Corless CL. Gastrointestinal stromal tumour. *Lancet* (2007) 369:1731–41. doi: 10.1016/s0140-6736(07)60780-6

4. Park CH, Kim EH, Jung DH, Chung H, Park JC, Shin SK, et al. Impact of periodic endoscopy on incidentally diagnosed gastric gastrointestinal stromal tumors: findings in surgically resected and confirmed lesions. *Ann Surg Oncol* (2015) 22:2933–9. doi: 10.1245/s10434-015-4517-0

5. Joensuu H, Hohenberger P, Corless CL. Gastrointestinal stromal tumour. *Lancet* (2013) 382:973–83. doi: 10.1016/s0140-6736(13)60106-3

6. Alkhatib AA, Faigel DO. Endoscopic ultrasonography-guided diagnosis of subepithelial tumors. *Gastrointest Endosc Clin N Am* (2012) 22:187–205. doi: 10.1016/j.giec.2012.04.006

7. Lim TW, Choi CW, Kang DH, Kim HW, Park SB, Kim SJ. Endoscopic ultrasound without tissue acquisition has poor accuracy for diagnosing gastric subepithelial tumors. *Med (Baltimore)* (2016) 95:e5246. doi: 10.1097/md.0000000000005246

8. Kim GH, Cho YK, Kim EY, Kim HK, Cho JW, Lee TH, et al. Comparison of 22-gauge aspiration needle with 22-gauge biopsy needle in endoscopic ultrasonography-guided subepithelial tumor sampling. *Scand J Gastroenterol* (2014) 49:347–54. doi: 10.3109/00365521.2013.867361

9. Han JP, Lee TH, Hong SJ, Kim HK, Noh HM, Lee YN, et al. EUS-guided FNA and FNB after on-site cytological evaluation in gastric subepithelial tumors. *J Dig Dis* (2016) 17:582–7. doi: 10.1111/1751-2980.12381

10. Zhang XC, Li QL, Yu YF, Yao LQ, Xu MD, Zhang YQ, et al. Diagnostic efficacy of endoscopic ultrasound-guided needle sampling for upper gastrointestinal subepithelial lesions: a meta-analysis. *Surg Endosc* (2016) 30:2431–41. doi: 10.1007/s00464-015-4494-1

11. Le Berre C, Sandborn WJ, Aridhi S, Devignes MD, Fournier L, Smaïl-Tabbone M, et al. Application of artificial intelligence to gastroenterology and hepatology. *Gastroenterology* (2020) 158:76–94.e2. doi: 10.1053/j.gastro.2019.08.058

12. Lee H, Kwon H. Going deeper with contextual CNN for hyperspectral image classification. *IEEE Trans Image Process* (2017) 26:4843–55. doi: 10.1109/tip.2017.2725580

13. Kuwahara T, Hara K, Mizuno N, Haba S, Okuno N, Koda H, et al. Current status of artificial intelligence analysis for endoscopic ultrasonography. *Dig Endosc* (2021) 33:298–305. doi: 10.1111/den.13880

14. Tonozuka R, Itoi T, Nagata N, Kojima H, Sofuni A, Tsuchiya T, et al. Deep learning analysis for the detection of pancreatic cancer on endosonographic images: a pilot study. *J Hepatobiliary Pancreat Sci* (2021) 28:95–104. doi: 10.1002/jhbp.825

15. Marya NB, Powers PD, Chari ST, Gleeson FC, Leggett CL, Abu Dayyeh BK, et al. Utilisation of artificial intelligence for the development of an EUS-convolutional neural network model trained to enhance the diagnosis of autoimmune pancreatitis. *Gut* (2021) 70:1335–44. doi: 10.1136/gutjnl-2020-322821

16. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, Clifford T, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: The PRISMA-DTA statement. *JAMA* (2018) 319:388–96. doi: 10.1001/jama.2017.19163

17. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* (2011) 155:529–36. doi: 10.7326/0003-4819-155-8-201110180-00009

18. Prasoppokakorn T, Tiyarattanachai T, Chaiteerakij R, Decharatanachart P, Mekaroonkamol P, Ridtitid W, et al. Application of artificial intelligence for diagnosis of pancreatic ductal adenocarcinoma by EUS: A systematic review and meta-analysis. *Endosc Ultrasound* (2022) 11:17–26. doi: 10.4103/eus-d-20-00219

19. Nguyen VX, Nguyen CC, Li B, Das A. Digital image analysis is a useful adjunct to endoscopic ultrasonographic diagnosis of subepithelial lesions of the gastrointestinal tract. *J Ultrasound Med* (2010) 29:1345–51. doi: 10.7863/jum.2010.29.9.1345

20. Minoda Y, Ihara E, Komori K, Ogino H, Otsuka Y, Chinen T, et al. Efficacy of endoscopic ultrasound with artificial intelligence for the diagnosis of gastrointestinal stromal tumors. *J Gastroenterol* (2020) 55:1119–26. doi: 10.1007/s00535-020-01725-4

21. Kim YH, Kim GH, Kim KB, Lee MW, Lee BE, Baek DH, et al. Application of a convolutional neural network in the diagnosis of gastric mesenchymal tumors on endoscopic ultrasonography images. *J Clin Med* (2020) 9:3162. doi: 10.3390/jcm9103162

22. Yang X, Wang H, Dong Q, Xu Y, Liu H, Ma X, et al. An artificial intelligence system for distinguishing between gastrointestinal stromal tumors and leiomyomas using endoscopic ultrasonography. *Endoscopy* (2022) 54:251–61. doi: 10.1055/a-1476-8931

23. Oh CK, Kim T, Cho YK, Cheung DY, Lee BI, Cho YS, et al. Convolutional neural network-based object detection model to identify gastrointestinal stromal tumors in endoscopic ultrasound images. *J Gastroenterol Hepatol* (2021) 36:3387–94. doi: 10.1111/jgh.15653

24. Hirai K, Kuwahara T, Furukawa K, Kakushima N, Furune S, Yamamoto H, et al. Artificial intelligence-based diagnosis of upper gastrointestinal subepithelial lesions on endoscopic ultrasonography images. *Gastric Cancer* (2022) 25:382–91. doi: 10.1007/s10120-021-01261-x

25. Seven G, Silahtaroglu G, Seven OO, Senturk H. Differentiating gastrointestinal stromal tumors from leiomyomas using a neural network trained on endoscopic ultrasonography images. *Dig Dis* (2021) 40:427–35. doi: 10.1159/000520032

26. Tanaka H, Kamata K, Ishihara R, Handa H, Otsuka Y, Yoshida A, et al. Value of artificial intelligence with novel tumor tracking technology in the diagnosis of gastric submucosal tumors by contrast-enhanced harmonic endoscopic ultrasonography. *J Gastroenterol Hepatol* (2022) 37:841–46. doi: 10.1111/jgh.15780

27. Lefort C, Gupta V, Lisotti A, Palazzo L, Fusaroli P, Pujol B, et al. Diagnosis of gastric submucosal tumors and estimation of malignant risk of GIST by endoscopic ultrasound. comparison between b mode and contrast-harmonic mode. *Dig Liver Dis* (2021) 53:1486–91. doi: 10.1016/j.dld.2021.06.013

28. Kim GH, Kim KB, Lee SH, Jeon HK, Park DY, Jeon TY, et al. Digital image analysis of endoscopic ultrasonography is helpful in diagnosing gastric mesenchymal tumors. *BMC Gastroenterol* (2014) 14:7. doi: 10.1186/1471-230x-14-7

29. Lee MW, Kim GH, Kim KB, Kim YH, Park DY, Choi CI, et al. Digital image analysis-based scoring system for endoscopic ultrasonography is useful in predicting gastrointestinal stromal tumors. *Gastric Cancer* (2019) 22:980–7. doi: 10.1007/s10120-019-00928-w

30. Seven G, Silahtaroglu G, Kochan K, Ince AT, Arici DS, Senturk H. Use of artificial intelligence in the prediction of malignant potential of gastric gastrointestinal stromal tumors. *Dig Dis Sci* (2022) 67:273–81. doi: 10.1007/s10620-021-06830-9

31. Liu C, Qiao M, Jiang F, Guo Y, Jin Z, Wang Y. TN-USMA net: Triple normalization-based gastrointestinal stromal tumors classification on multicenter EUS images with ultrasound-specific pretraining and meta attention. *Med Phys* (2021) 48:7199–214. doi: 10.1002/mp.15172

32. Li X, Jiang F, Guo Y, Jin Z, Wang Y. Computer-aided diagnosis of gastrointestinal stromal tumors: a radiomics method on endoscopic ultrasound image. *Int J Comput Assist Radiol Surg* (2019) 14:1635–45. doi: 10.1007/s11548-019-01993-3

33. Seven G, Silahtaroglu G, Kochan K, Kiremitci S, Senturk H. Artificial intelligence in the prediction of malignant potential in gastrointestinal stromal tumors of stomach. *United Eur Gastroenterol J* (2020) 8:806. doi: 10.1177/2050640620927345

34. Rawat W, Wang Z. Deep convolutional neural networks for image classification: a comprehensive review. *Neural Comput* (2017) 29:2352–449. doi: 10.1162/NECO_a_00990

frontiers | Frontiers in Oncology

TYPE Review
PUBLISHED 16 August 2022
DOI 10.3389/fonc.2022.972984

OPEN ACCESS

EDITED BY
Min Wu,
Sichuan University, China

REVIEWED BY
Weifeng Lin,
Weizmann Institute of Science, Israel
Hong Yang,
Shanghai Normal University, China
Deteng Zhang,
Qingdao University, China
Fangfang Shen,
Stanford University, United States

*CORRESPONDENCE
Xin Liu
liuxin@njucm.edu.cn
Jiege Huo
huojiege@jsatcm.com

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 19 June 2022
ACCEPTED 27 July 2022
PUBLISHED 16 August 2022

CITATION
Deng S, Chen B, Huo J and Liu X (2022) Therapeutic potential of NR4A1 in cancer: Focus on metabolism. *Front. Oncol.* 12:972984. doi: 10.3389/fonc.2022.972984

# Therapeutic potential of NR4A1 in cancer: Focus on metabolism

Shan Deng[1], Bo Chen[2], Jiege Huo[1]* and Xin Liu[1,3]*

[1]Third School of Clinical Medicine, Nanjing University of Chinese Medicine, Nanjing, China, [2]Materials Science and Devices Institute, Suzhou University of Science and Technology, Suzhou, China, [3]Department of Orthopedics, Nanjing Lishui Hospital of Traditional Chinese Medicine, Nanjing, China

Metabolic reprogramming is a vital hallmark of cancer, and it provides the necessary energy and biological materials to support the continuous proliferation and survival of tumor cells. NR4A1 is belonging to nuclear subfamily 4 (NR4A) receptors. NR4A1 plays diverse roles in many tumors, including melanoma, colorectal cancer, breast cancer, and hepatocellular cancer, to regulate cell growth, apoptosis, metastasis. Recent reports shown that NR4A1 exhibits unique metabolic regulating effects in cancers. This receptor was first found to mediate glycolysis *via* key enzymes glucose transporters (GLUTs), hexokinase 2 (HK2), fructose phosphate kinase (PFK), and pyruvate kinase (PK). Then its functions extended to fatty acid synthesis by modulating CD36, fatty acid-binding proteins (FABPs), sterol regulatory element-binding protein 1 (SREBP1), glutamine by Myc, mammalian target of rapamycin (mTOR), and hypoxia-inducible factors alpha (HIF-1α), respectively. In addition, NR4A1 is involving in amino acid metabolism and tumor immunity by metabolic processes. More and more NR4A1 ligands are found to participate in tumor metabolic reprogramming, suggesting that regulating NR4A1 by novel ligands is a promising approach to alter metabolism signaling pathways in cancer therapy. Basic on this, this review highlighted the diverse metabolic roles of NR4A1 in cancers, which provides vital references for the clinical application.

KEYWORDS

**metabolic reprogramming, NR4A1, cancers, signaling pathways, tumor cells**

## Introduction

Metabolic reprogramming is an important characteristic of tumor cells, which can provide energy and multiple substrates for biosynthesis to support cancer cells' rapid proliferation and survival (1). Furthermore, the malignant transformation, invasion, and metastasis of tumor cells also need metabolic reprogramming. The metabolic pathways include glycolysis, glutamine metabolism, fatty acid metabolism, nucleic acid, and amino acid metabolism (2). In the 1920s, Otto Warburg first reported the Warburg effect, which was once considered the main metabolic pathway in cancer cells. Nowadays, this effect is defined as aerobic glycolysis, which can metabolize glucose into lactate and provides Adenosine triphosphate (ATP) for cancer cell survival. By downregulating aerobic glycolysis via the c-Myc signaling pathway, the proliferation of hepatocellular carcinoma (HCC) was inhibited, and their apoptosis was induced (3). Besides, glutamine has been reported to contribute to tricarboxylic acid (TCA) cycle metabolites. The "glutamine addiction" is a vital metabolic feature to support the rapid proliferation of cancer cells. In breast cancer, some polyphenols, such as catechin, delphinidin, and kaempferol, exhibit an anti-proliferative effect by inhibiting alanine, serine, cysteine transporter 2 (ASCT2) and decreasing total and Na +-dependent 3H-glutamine uptake (4). The fatty acid is the key competitor of the cell membrane and stores energy and acts as the secondary messengers. Thus, fatty acid synthesis (FSAN) is vital for transporting intracellular signal transduction and tumor cell proliferation, differentiation, migration, survival, and apoptosis (5). Fan et al. (6) reported that α-linolenic acid could inhibit osteosarcoma cell proliferation and metastasis by suppressing FASN expression. By blocking fatty acid enzymes hexokinase 2 (HK2) or acyl-CoA synthetase long-chain family member 4 (ACSL4), acetyl-CoA accumulation decreased, leading to a suppressed fatty acid β-oxidation activity. These results effectively inhibit liver cancer growth (7). In order to meet the infinite proliferation in tumor cells, transcription and replication activities are more frequent, so the nucleotides and amino acids are enhanced. Buel et al. (8) reported the crosstalk between amino acid and mTORC1, which can regulate tumor cell fate through the Rag-GTPase pathway. As mentioned above, cancer metabolism is controlled by many factors, including genes, enzymes, and signaling pathways; therefore, exploring novel targets of metabolic reprogramming provides enormous opportunities to regulate tumor cells fate.

In recent years, metabolism has been widely observed during cancer development; NR subfamily 4 (NR4A) receptors are considered the mediators in controlling this metabolic hallmark of tumors. NR4A family receptors are one of 48 human nuclear receptors that act as transcription factors to regulate many cell processes. NR4A nuclear receptors include NR4A1 (NUR77), NR4A2 (NURR1), and NR4A3 (NOR-1), showing similar structures which consist of a DNA-binding domain (DBD), a C-terminal ligand-binding domain (LBD), and an N-terminal transactivation domain (TAD). The TAD contains a ligand-independent activation function 1 (AF-1), responsible for interacting and regulating the activity of transcription factors. The DBD in the middle can specifically interact with DNA sequences known as NGFI-B response element (NBRE) and Nur-responsive element (NurRE); there is over 90% sequence homology in DBD of NR4A receptors (9). The last part, LBD, contains a ligand-dependent activation function 2 (AF-2), which can recognize corresponding ligands to ensure the transcriptional activity (10–12) (Figure 1). Although these receptors share a typical structure, about 60% of the sequences in the LBD region are conserved, while the sequence of the activation domain changes greatly. Because the large hydrophobic residues occupy the binding pocket space, there has no progress in identifying endogenous ligands so far. Recent reports suggest that the NR4A family receptors may bind to unsaturated fatty acids in the LBD to exhibit regulation of metabolism (13). For example, in breast cancer, NR4A1-NR4A3 regulates glycolysis to participate in cell progression (14). The NR4A receptors are also associated with the activation of T cell, which involving in cancer immunotherapy (15). From the literature, increasing evidence proves that among these NR4A receptors, NR4A1 shows more metabolic functions in cancers, such as regulating glycolysis and exhibiting activities in fatty acid synthesis, glutamine, and amino acid metabolism. This paper summarizes the metabolism roles of NR4A1 in the tumor.

## Identification and regulation of NR4A1

Nuclear receptor 4A1 (NR4A1, also called Nur77, NGFIB, TR3) is one of the NR4A subfamily transcription factors, which was firstly identified in mouse fibroblasts in 1988 (16). Next year, Chang (17) isolated NR4A1 from a human prostate lambda gt11 cDNA library. Then it is found in various tissues and cells, including cancer cells. NR4A1 is an immediate gene induced by stress, cytokines, growth factor, glucose, fatty acids, or other stimuli (18–21). NR4A1 plays diverse roles in many physiological and pathological processes, for example cell survival, apoptosis, differentiation, cell cycle, inflammation, immunity, and metabolism (22–26). NR4A1 can bind to DNA in three ways to regulate the expression of target genes: (1) it can form the response element NBRE (sequence: AAAGGTCA); (2) it

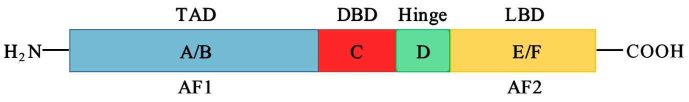


FIGURE 1
The structure of NR4A receptors. NR4A structure has a N-terminal domain containing AF-1, and C-terminal domain with AF-2, they flank a DNA-binding domain (DBD) and a hinge region.

binds to the NurRE element (AAAT(G/A)(C/T)CA, which are related to the NBRE) in the form of homodimer or heterodimer formed with other members of the family; (3) NR4A1 and retinoid X receptors (RXRs) form heterodimers and then binds to the DR5 response element to produce transcriptional activation (sequence: AGGTCA-NNNAA-AGGTCA) (27) (Figure 2). Because of the specific structure, NR4A1 can directly affect the target genes promoter to exhibit transcriptional activity. For example, in prostate cancer, prostaglandin E2 (PGE2) activates NR4A-RXR heterodimer to enhance micrometastasis; this effect can be reversed by cyclooxygenase 2 (COX2) inhibitor in cancer suppression (28). NR4A1 also complexes with Sp1 and p300 on the region of survivin promoter to increase pancreatic cancer cell proliferation and decrease apoptosis (29). In inflammatory diseases, NR4A1 could regulate SerpinA3 through the NBRE in its promoter region (30). NR4A1 transcriptionally inhibits the expression of Dicer to activate downstream Akt/mTORC1 signaling, thereby inducing colon cancer epithelial-to-mesenchymal transition (EMT) (31) (Figure 3).

In addition, increasing evidence indicates that NR4A1 displays non-genomic functions to affect cell biological processes. NR4A1 exerts non-genomic activities by translocating from the nucleus to mitochondria, or endoplasmic reticulum (ER), which triggers apoptosis or autophagy (32). Studies have shown that the overexpressed NR4A1 can activate the Wnt/β-catenin signaling pathway to enhance colon tumor growth, colony formation, and migration (33). It also moves to the cytoplasm to stimulate the dysregulation of β-catenin and the stabilization of HIF-1α under normoxia (34). In mitochondrial, NR4A1 translocated from the nucleus and bound to Bcl-2, converts Bcl-2 to a pro-apoptotic protein, then induces cytochrome C release and apoptosis (35). Another report showed that NR4A1 could induce MDM2 ubiquitination and degradation by blocking p53 acetylation, this effect can enhance p53-depended apoptosis (36). Furthermore, NR4A1 interacts, and blocks binds and sequesters Liver kinase B1 (LKB1) in the nucleus, then releases and shuttles LKB1 to the cytoplasm, thereby attenuating AMP-activated protein kinase (AMPK) activation to treat metabolic diseases such as streptozotocin-induced diabetes (37). In osteoclast, knockout of NR4A1 can promote the differentiation of RAW264.7 by activating the NF-κB signaling pathway, in order to decrease the expression of IκB-α and induce IKK-β (23). Additionally, NR4A1 translocated from nucleus to mitochondria, then interacted with tumor necrosis factor receptor-associated factor 2 (TRAF2), leading to TRAF2 ubiquitination. NR4A1 also interacted with p62/SQSTM1 to sensitize cells to autophagy (38) (Figure 4).

## Metabolism roles of NR4A1 in tumor

NR4A1 is widely involved in the metabolism of tumors, including glucose metabolism, glutamine metabolism, fatty acid metabolism, and amino acid metabolism (39–42).

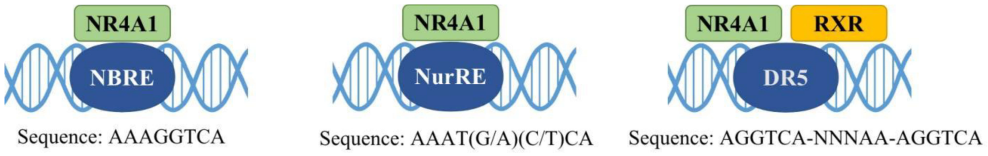


FIGURE 2
Interactions of NR4A1 with different elements. NR4A1 activates target gene expression through binding with NBRE, NurRE, and a DR5 motif (with RXR), respectively.

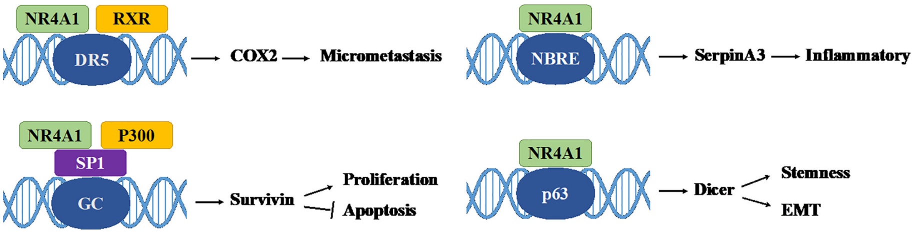


FIGURE 3
Some genomic effects of NR4A1. NR4A1 can bind to different genes promoter to involve in cancer progression. GC: GC-rich promoter regions of Survivin.

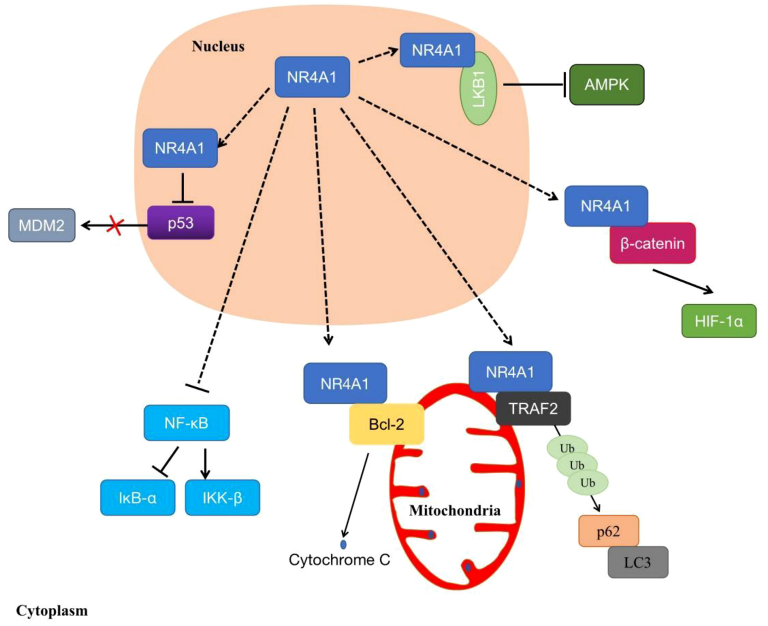


FIGURE 4
The non-genomic regulations of NR4A1. NR4A1 can affect kinds of biological processes through binding to LKB1, β-catenin, Bcl-2, MDM2, TRAF2, or blocking NF-κB.

## NR4A1 and glucose metabolism

Glucose metabolism is the key source to provide metabolic carbon in cells. There are three main ways of glucose metabolism: aerobic oxidation, glycolysis and pentose phosphate pathways. Normally, cells uptake glucose by glucose transporters (GLUTs), then it enters the glycolysis process with the action of hexokinase (HK), fructose phosphate kinase (PFK), and pyruvate kinase (PK) under normal oxygen conditions to produce pyruvate. However, uncontrolled proliferation is a crucial characteristic of cancer. Tumor cells alter their glucose metabolism patter to an efficient aerobic glycolysis rate to sustain vigorous proliferation and other biological activities. In this process, the activities and expression levels of GLUT and glycolytic rate-limiting enzymes such as HK, PFK, and PK were significantly up-regulated to improve glucose uptake, which is called the “Warburg effect” (43). This effect not only provides the energy for tumor cell survival, but also provides biosynthetic raw materials for other metabolic pathways, including the tricarboxylic acid (TCA) cycle, hexosamine pathway, pentose phosphate pathway, glycogen synthesis, and serine biosynthesis pathway (44).

Recently, it has been progressively realized that NR4A1 plays diverse roles in glucose metabolic regulation. GLUTs facilitate the transport of glucose from extracellular to the cellar membrane. Overexpressed NR4A1 has been reported to upregulate GLUT4 production to increase glucose oxidation and glycogen synthesis in muscle L6 cells. NR4A1 also changes the activity of several key glycolytic enzymes; for example, NR4A1 upregulates the expression of HK and PFK in rat muscle cells (45). In HFD-induces obese mice, inhibition of NR4A1 by siRNA could modulate the key rate-limiting enzyme HK2, leading to the disturbed glucose metabolism homeostasis in mice cardiac (46). Furthermore, bis-indole-derived NR4A1 ligands enhanced the accumulation of GLUT4 in the cell membrane and the overall glucose uptake in muscle cells in diabetes (47). NR4A1 is considered a promising therapeutic target for metabolic syndromes.

In addition, the paradoxical roles of NR4A1 in regulating glucose metabolism in cancer were investigated. In hepatocellular carcinoma, low expression of NR4A1 was observed, promoting HCC development. NR4A1 can inhibit glycolysis and elevate gluconeogenesis by interacting with and suppressing the rate-limiting enzyme phosphoenolpyruvate carboxykinase (PEPCK1), leading to ATP depletion and an arrest of cell growth (48). Another research reported that by binding to the promoter of WAP four-disulfide core domain 21 pseudogene (WFDC21P), NR4A1 also inhibited two key glycolysis enzymes, the platelet-type PFK (PFKP) and the M2 isoform of pyruvate kinase (PKM2), to suppress the HCC cell proliferation and tumor metastasis (49). Furthermore, in acute promyelocytic leukemia (APL) cells, silencing NR4A1 can activate glycolytic transporter GLUT1 and decrease the expression of TIGAR (TP53-induced glycolysis and apoptosis regulator) to induce APL development (50). Cytosporone B (CsnB) is an NR4A1 agonist; it induced tumor cell apoptosis and inhibited tumor growth in C57 mice via translocating NR4A1 to mitochondrial to cause cytochrome C release. CsnB also induced gluconeogenesis-related genes, resulting in elevated of blood glucose levels in tumors (51).

In contrast, NR4A1 is overexpressed in many other human malignant tumors, for example, pancreatic cancer, colorectal cancer, and breast cancer. Several studies have revealed that hypoxic exposure results in increased HIF-1α protein stabilization, which has been implicated in promoting the glycolysis of tumor cells. This response can be regulated by NR4A1 through repressing MDM2 expression, suggesting the enhancement of glycolysis induced by HIF-1α was partially attributed to NR4A1 upregulation (52, 53). In colorectal cancer cells, Dong et al. (54, 55) reported the relationship between enhanced glycolysis and the aberrant activation of β-catenin, while our previous study confirmed that β-catenin and NR4A1 could form a mutually feedback control circuit to promote CRC invasion, demonstrating that NR4A1 may be involved in the glycolysis in colorectal cancer. These findings underscore the regulation of NR4A1 on glucose oxidation and glycogen synthesis, indicating that the impact of NR4A1 on glucose metabolism is complex and cell-dependent.

## NR4A1 and glutamine metabolism

Glutamine metabolism is dysregulated in a variety of solid tumor cells, and it is indispensable for cancer cell proliferation. Depletion of glutamine can promote EMT and metastasis, overcome tumor immune evasion (56–58). Therefore, glutamine has become a very attractive target for tumor anti-metabolic therapy. To be better utilized by cells, glutamine is transported into cells through specific transporters and converted into glutamate under the action of glutaminase to enter the TCA cycle and provide energy for the growth and development of tumor cells. Glutamine enters cells via the solute carrier family (SLC) transporters, including SLC1A5 and SLC7A8, which are overexpressed in many cancers. These Na+-independent neutral amino acid transporters can activate mTOR signaling and are controlled by Myc (59). In 2009, Gao et al. (60) reported that Myc inhibits glutamine metabolism by suppressing miR-23a/b expression to generate energy for proliferating cancer cells. In ovarian cancer, miR-145 decreased glutamine metabolism through targeting c-Myc via activating glutaminase 1 (GLS1) transcription expression (61). In pancreatic cancer, the nuclear translocation of β-catenin can increase c-Myc expression, resulting in a rise in glutamine uptake and glutamate release (62). On the other hand, a study indicated that NR4A1 acts as a β-catenin mediator to allow β-catenin to escape degradation in HCC (63). Meanwhile, our previous study reported a positive NR4A1-β-catenin feed-forward loop in cancers (31, 55). Another study reported that NR4A1 inhibition decreased the levels of β-catenin and c-Myc (64); thus, it is reasonable to speculate that NR4A1 may participate in glutamine metabolism through β-catenin/Myc signaling pathway. In terms of mTOR, glutamine upregulated the activity of glutaminase (GLS) and glutamate dehydrogenase (GDH) by inducing mTOR upregulation. This effect can be reversed by mTOR inhibitor rapamycin, leading to a decrease of glutamine-induced cell proliferation in ovarian cancer (65). NR4A1 can regulate mTOR signaling, and knockdown of NR4A1 inhibits mTOR through reactive oxygen species-dependent activation of AMPK (66, 67), so NR4A1 may be involved in glutamine metabolism via mTOR regulation.

Furthermore, reduced oxygen supply increases GLS1 mRNA and protein expression due to transcriptional activation of HIF-1, accelerates glutamine metabolism, and is conducive to the growth, invasion, migration, as well as metastasis in colorectal cancer (68). NR4A1 is an important regulator of HIF-1. The relationship between NR4A1 and HIF-1 has been shown in many tumors (53, 69). In renal cell carcinoma, NR4A1 stabilized and transactivated HIF-1α. Moreover, NR4A1 is highly expressed in acute myeloid leukemia; when truncated protein-encoding for part of the N-terminal domain of NR4A1, the NR4A1 transcript variant still maintains the stability and activity of HIF-1α (70). On the other hand, HIF-1α activated NR4A1 by binding to the putative HIF responsive element in the NR4A1 promoter, then upregulating the expression of NR4A1 (71). Under chronic hypoxia conditions, NR4A1 has low expression in non-small cell lung cancer (NSCLC) cells by the mediation of HIF-1α, involved in hypoxia-induced apoptosis resistance (72). Therefore, NR4A1 and HIF-1α can form an interaction circulus, influencing each other. Since HIF-1 is a vital regulator in glutamine metabolism, NR4A1 is likely to become a potential target of tumor glutamine metabolism.

## NR4A1 and fatty acid metabolism

Lipids are classified as fatty acids, cholesterol, phospholipids, or triacylglycerides, major components of cell membranes. Lipids are widely contributed to energy sources, signaling molecules, and second messengers. As an important component of various lipids, fatty acid synthase (FASN) plays an irreplaceable role in cell proliferation and survival. Various raw materials for fatty acid synthesis synthesize fatty acids from scratch under the catalysis of enzymes such as ATP citrate lyase (ACLY), acetyl CoA carboxylase (ACC), and fatty acid synthase (FAS). Accumulating evidence has shown that dysregulation in lipid metabolism is one of the most abnormal metabolic changes in tumor cells, while the enhancement of de novo fatty acid synthesis is the main manifestation of lipometabolic reprogramming in tumor cells. The FASN process leads to the increased expression of a variety of key enzymes, mainly ACLY, ACC, and FAS (73), which affect multiple aspects of carcinogenesis, such as cell proliferation, differentiation and cell cycle (74).

Cancers drive fatty acid mainly from exogenously microenvironment or endogenously through de novo synthesis by FASN. The most well-characterized transporters include CD36, solute carrier protein family 27 (SCL27), and fatty acid-binding proteins (FABPs). Among them, CD36 is reported to be highly expressed in various malignancies, including breast cancer, ovarian cancer, and gastric cancer (75–77). SCL27 family has six members (SCL27 A1-A6) associated with tumor fatty acid uptake (78, 79). FABPs consists of 12 family members, and they can transport lipid to cellular mitochondria, nuclei, and so on (80). And they are frequently found to be highly expressed in bladder cancer, prostate cancer, and renal cell carcinoma (81, 82).

In the early decade, numerous studies explored the complex roles of NR4A1 in regulating fatty acid metabolism in normal tissues, including liver, skeletal muscle, and adipose. For example, Wang's team reported (83) that NR4A1 could specifically bind to LKB1 in the nucleus and prevent the translocation of LKB1 to the cytosol. This interaction between NR4A1 and LKB1 can be broken by antagonist TPMA, promoting of AMPKα phosphorylation and activating downstream fatty acid enzymes like ACC and CPT1A to inhibit fatty acids synthesis in primary hepatocytes. Sterol regulatory element-binding protein 1 c (SREBP1c) is a well-established transcription factor to regulate FASN (84), to regulate hepatic lipid metabolism, NR4A1 decreased SREBP1c expression by reducing its target genes stearoyl-coA desaturase-1 (SCD1), mitochondrial glycerol-3-phosphate acyltransferase (GPAT), and FASN (85). In skeletal muscle cells, attenuation of NR4A1 expression decreased lipolysis by inhibiting beta-AR and its downstream CD36, adiponectin receptor 2, and caveolin-3 expressions (86). Like in liver cells, Jung's team reported (86, 87) that the interaction of NR4A1 and AMPKα in inhibiting adipogenesis *in vitro* and *in vivo*.

Recently, the lipid metabolic roles of NR4A1 attracted more and more attention in tumor progression, especially fatty acid metabolism. Fatty acid metabolism includes fatty acid synthesis and fatty acid oxidation (FAO). NR4A1 is thought to participate in fatty acid uptake and oxidation to affect cancer cell fate. Fatty acid oxidation provides the ATP and NADPH to overcome metabolic stress. To assess the role of NR4A1 in cancers, a recent study reported that NR4A1 is required in melanoma cells to protecting FAO. The overexpressed NR4A1 can bind to and activate the rate-limiting enzyme trifunctional protein β (TPβ) to maintain ATP and NADPH levels and prevent ROS increase and melanoma cell death. NR4A1 regulated the linkage FAO-NADPH-ROS during metabolic stress to target melanoma (88). Holla et al. (89) reported the pro-oncogenic effect of NR4A1 in regulating the fatty acid oxidation pathway in colon cancer. A high level of PGE2 induced-NR4A2 was reported to bind to NR4A1-binding response elements (NBRE), which can recruit and induce the expression of four genes related to fatty acid metabolism: acyl-CoA oxidase (ACOX), carnitine palmitoyltransferase 1B (CPT1M), fatty acid-binding protein-2 (FABP2) and FABP4. A novel study reported that peroxisome proliferator-activated receptor-γ (PPARγ) acts as an antagonist of NR4A1 and can ubiquitination and degradation of NR4A1 through ubiquitin enzyme tripartite motif 13 (TRIM13); this process interferes with the interaction of NR4A1 and SWI/SNF complex, and recruit to the promoter of fatty acid transporters CD36 and FABP4 to inhibit their transcription, which blocked fatty acid uptake to suppress cancer cell proliferation (41) (Figure 5).

Apart from fatty acid synthesis, NR4A1 also participates in cholesterol metabolite, Dendrogenin A (DDA) is identified as a cholesterol metabolite in mammal cells. There has a complemental effect between cancer cells and DDA; by binding to the liver X receptor (LXR), DDA can activate NR4A1 expression to exhibit an anti-tumor effect on breast cancer and melanoma (90). In acute myeloid leukemia, DDA also partly activates LXR to increase NR4A1, further inhibiting the expression level of cholesterol biosynthesizing enzyme 3β-hydroxysterol-Δ8,7-isomerase (D8D7I), leading to cancer autophagy induction (91). In HepG2 cells, downregulation of NR4A1 induced an increase in total cholesterol (TCHO) levels, low-density lipoprotein receptor (LDLR), and HMGCoA reductase (HMGCR) levels are also increased following the inhibition of NR4A1, suggesting NR4A1 is capable of reducing hepatic cholesterol based on lipid overloading. This evidence is proved that the effect of NR4A1 in regulating lipid metabolism in cancer growth and proliferation.

## NR4A1 and amino acid metabolism

Amino acid is essential for mammalian cells as the substrate for new protein synthesis. However, to drive the continuous proliferation of cancer cells, an abundant supply of amino acids is observed (92). A novel reports demonstrated that amino acid deficiency (AAD) could activate myocyte enhancer factor 2D (MEF2D) and induce the expression of NR4A1, which mediated reticulophagy to maintain intracellular amino acid levels (93). Although amino acid deficiency induces NR4A1, there are rare study focusing on the connection between NR4A1 and amino acid metabolism. Almost 20 years ago, Li and colleagues found (94) that changing the DNA-binding site of NR4A1 at Ser354 with negatively charged amino acids, such as Asp or Glu, can significantly decrease the NR4A1 transactivation activities. A most recent study from Xu et al. (21) first indicated the role of NR4A1 in regulating amino acids. They observed total amino acids compositions and found valine (Val), leucine (Leu), and isoleucine (Ile) were all decreased as well as many other amino acids, including aspartic acid (ASP), glutamic acid (Glu), alanine

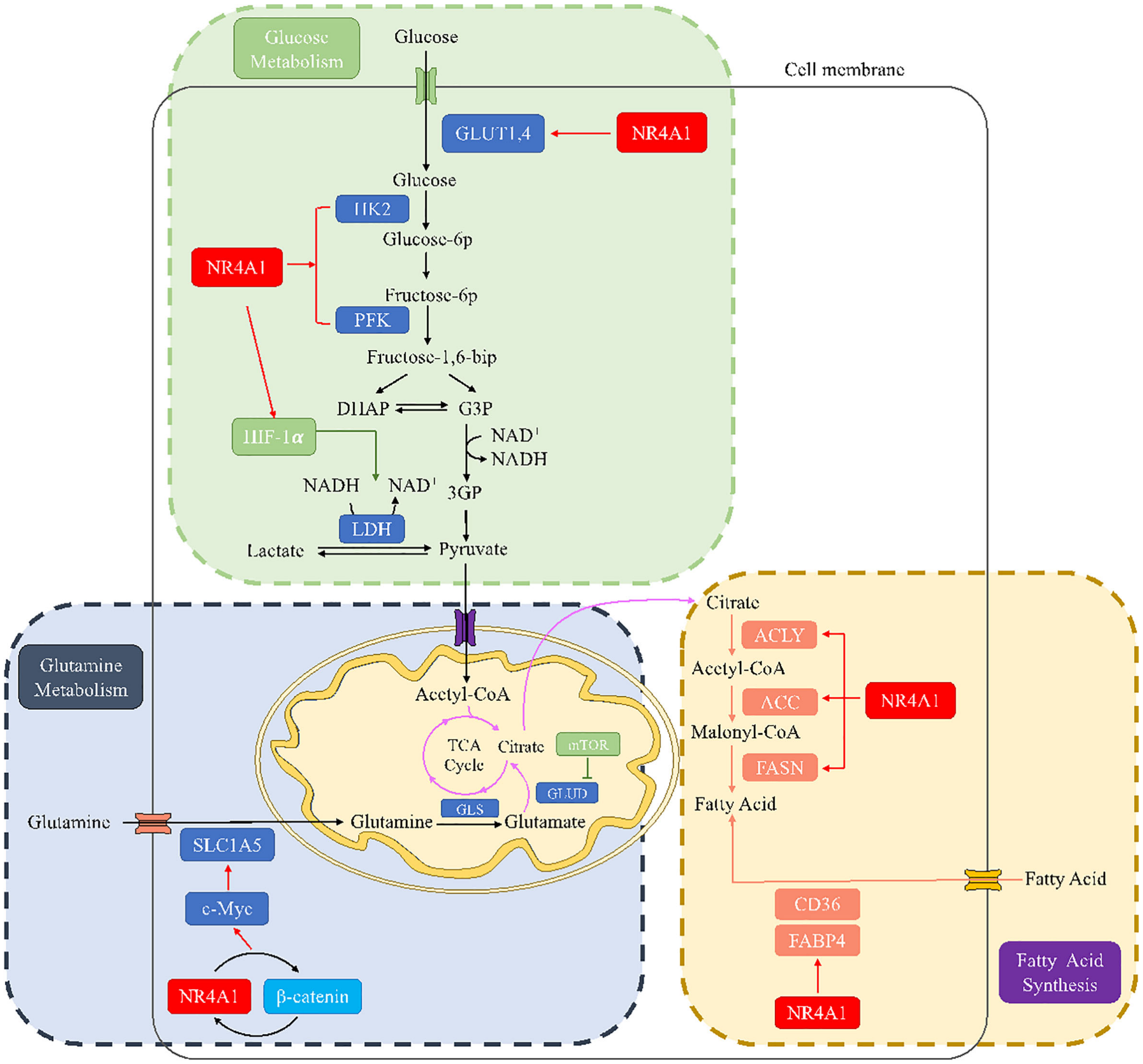


FIGURE 5
The metabolic regulation of NR4A1 in glucose, glutamine, and fatty acid metabolism. In glucose metabolism, NR4A1 can regulate glucose transporter GLUT1 and 4, key enzymes HK and PFK, as well as HIF-1α. NR4A1 mediates glutamine metabolism by regulating SLC1A5 *via* c-Myc. NR4A1 can regulate the fatty acid synthesis of major enzymes ACLY, ACC, and FASN, and it also participates in fatty acid uptake through CD36 and FABP4.

(Ala), tyrosine (Tyr), histidine (His), methionine (Met), proline (Pro), and so on in NR4A1−/− zebrafish larva (Figure 6).

## NR4A1 and tumor microenvironment

Additionally, NR4A1 also participated in cancer immunity by regulating metabolic pathways. In acute myeloid leukaemia (AML), researchers found COX2 inhibition dramatically decreased NR4A1 transcription and the WNT signaling pathway. In AML-mesenchymal stromal cells (MSCs)-CD34+ cells co-cultured system, this a novel COX2/NR4A1/CTNNB1 axis increased leukaemia-reactive T-effector cells and rescued cellular metabolism and anti-leukaemia immunity (95). In the melanoma tumor microenvironment, T-cell receptor (TCR) signaling can trigger its downstream NR4A1 expression, so using NR4A1-GFP indicated that blocking β-AR signaling increased metabolic reprogramming of CD8+ T-cell activation via TCR signaling. This impairment of β-AR on TCR signaling occurs through GLUT-1 downregulation and subsequent increase of glycolysis (96) (Figure 7).

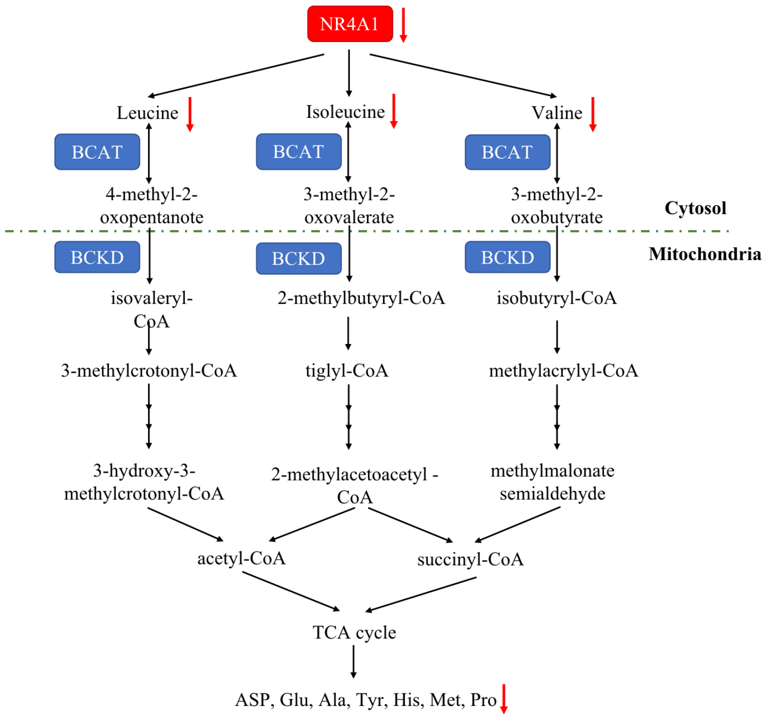


FIGURE 6
The amino acid metabolic pathway of NR4A1. In zebrafish larva, NR4A1 deficiency can decrease key branched-chain amino acids Val, Leu, and Ile, leading to the inhibition of ASP, Glu, Ala, Tyr, His, Met and Pro.

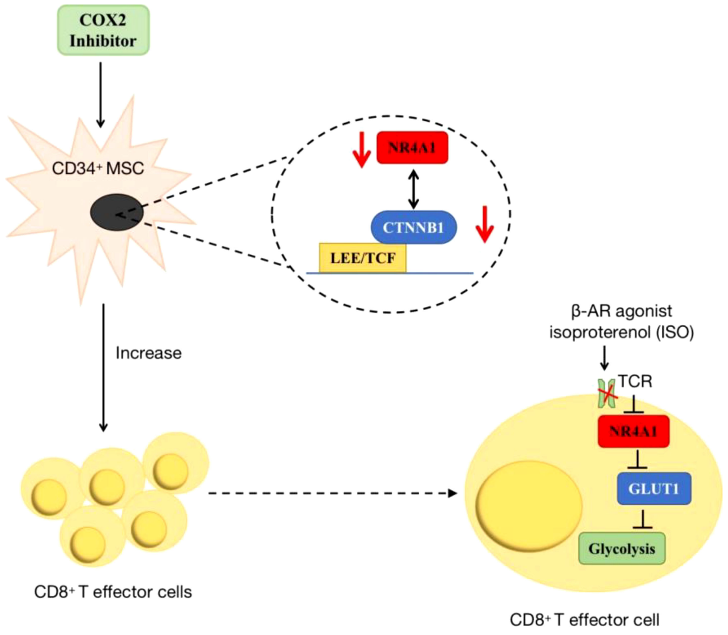


FIGURE 7
The roles of NR4A1 in regulating T cell. In AML, the inhibition of COX2/NR4A1/CTNNB1 signaling pathway can increase the produce of CD8+ T cell to rescue anti-cancer immunity. In melanoma, β-AR agonist inhibits NR4A1 and its downstream glycolysis to mediate CD8+T cell activity.

## Potential for targeting NR4A1

The expression and functions of NR4A1 in cancer metabolism are emerging as a promising area in treating and preventing human cancer malignant evolvement. Developments on mechanisms of NR4A1 silencing or strategies for its activity are leading to the explore of novel therapeutic agents. For example, CsnB is the first described NR4A1 agonist; multiple studies have indicated the CsnB can bind to LBD of NR4A1 and modulate NR4A1 nuclear export to mitochondria, causing Cyto-C release and apoptosis (51). Additionally, CsnB also acts as a candidate to downregulate CD36/FABP4 expression, leading to the inhibition of fatty acid uptake and consequent breast cancer cell proliferation in NR4A1-dependent manner (41). A class of Bisindole-derived (CDIMs) NR4A1 antagonists, such as 1,1-bis (3'-indolyl)-1-(p-hydroxyphenyl) methane (DIM-C-pPhOH), can decrease the expression of NR4A1 in breast, lung, and liver cancer cells to inhibit tumor growth, EMT and stemness (97, 98). Additionally, some natural compounds also act as NR4A1 ligands to exhibit an anti-tumor effect. Kaempferol and Quercetin are flavonoid compounds; they bind to NR4A1 and inhibit NR4A1-dependent transactivation by decreasing PAX3-FOXO1-G9a and mTOR signaling to suppress RMS cell growth (64). 1,3,7-trihydroxy-2,4-diprenylxanthone (CCE9) is a xanthone compound that induces the expression of NR4A1 and the interaction of NR4A1 and Bcl-2, leading to increased apoptosis through p38α/MAPK signaling pathway (99). Celastrol has a potent anti-inflammation effect by binding to NR4A1 and inducing NR4A1 to transport to mitochondria, resulting in sensitivity to autophagy (38). (Table 1 and Figure 8).

## Discussion

Aberrant metabolism reprogramming is a core feature of the tumor; increased cancer metabolism, such as fatty acid synthesis, glycolysis, plays vital roles in tumor proliferation, metastasis, and multidrug resistance. Hence, developing novel therapeutic methods and drug targets are required. Accumulating evidence proves that NR4A1 implicated metabolic processes in regulating various diseases, including obesity, atherosclerosis, liver disorders, and diabetes.

NR4A1 is an orphan nuclear receptor that exhibits pro-oncogene or anti-cancer effects in different cancers. For example, in colorectal cancer, overexpressed NR4A1 promoted cancer cell growth, epithelial-mesenchymal transition (EMT), and cancer stem-like cells (CSCs) properties. However, in HCC, NR4A1 is low expressed, upregulating NR4A1 by CsnB or other compounds, such as 4-(quinoline-4-amino), can inhibit tumor cell growth *in vitro* and *in vivo* (105). In terms of metabolism, it has been found that targeting NR4A1 can regulate glycolytic key enzymes GLUT4, HK2, and PFK in the liver and muscle cells to target metabolic syndromes. Although the metabolic roles of NR4A1 have been reported, the diverse effects in cancer metabolic reprogramming have not been delineated. This review links NR4A1 to metabolic processes in cancers. By altering glucose metabolism, NR4A1 depleted ATP and induced cell cycle arrest in HCC. NR4A1 also inhibited glycolysis enzymes PFKP and PKM2 to block HCC metastasis. The NR4A1 mediator CsnB induced tumor cell apoptosis; this suppressive function of CsnB is associated with the translocation of NR4A1 from the nucleus to mitochondria to release the cytochrome C-depended Bcl-2 apoptotic pathway. Paradoxically, in colorectal, pancreatic, and breast cancer, NR4A1 shows the opposite effect in modulating glycolysis. Aberrant activated β-catenin signaling in colon cancer enhanced glycolysis; meanwhile, an NR4A1-β-catenin feed-forward loop happening in colon cancer cells proves from the side that NR4A1 may be involved in promoting glycolysis. Nonetheless, the two side effects of NR4A1 have been observed on glucose metabolism, underscoring the complex and cell depend on its metabolic regulation, demonstrating NR4A1 acts as a potential therapeutic target in malignant tumors.

Fatty acid metabolism has its particularity and university. Cells run FASN and FAO to supply necessary nutrients. NR4A1

TABLE 1 NR4A1 ligands.

| Type | Name | Target | Applications | Ref |
|---|---|---|---|---|
| Inducer | Cytosporone B | Cyto-C | Breast cancer<br>Colon cancer<br>Lung cancer<br>Bladder cancer | (41, 72, 100, 101) |
| Inhibitor | DIM-C-pPhOH | | Breast cancer Rhabdomyosarcoma Pancreatic cancer | (102–104) |
| | Kaempferol | mTOR | Rhabdomyosarcoma | (64) |
| | Quercetin | mTOR | Rhabdomyosarcoma | (64) |
| | CCE9 | p38α<br>MAPK<br>Bcl-2 | Cervical cancer<br>Liver cancer | (99) |
| | Celastrol | p62 | Cervical cancer | (38) |

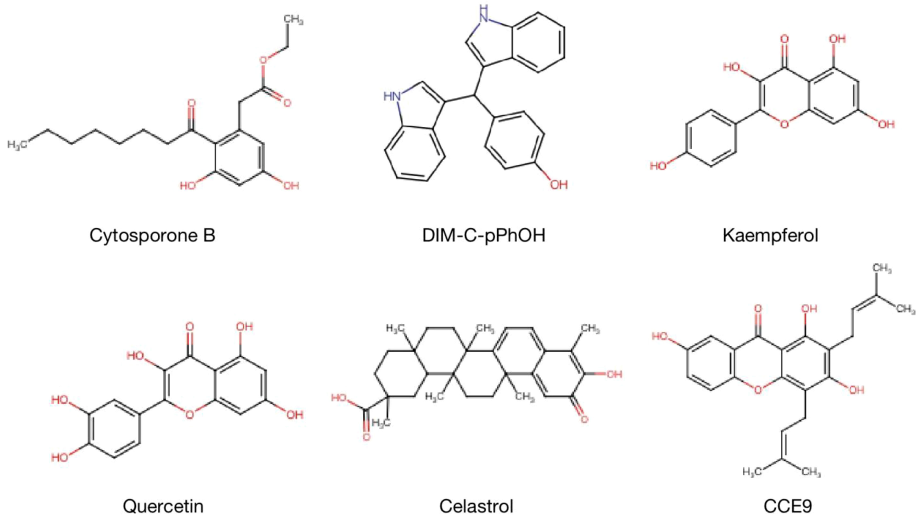


FIGURE 8
Structure of NR4A1 agonists or antagonists. Structure of NR4A1 inducer Cytosporone B, and inhibitors DIM-C-pPhOH, Kaempferol, Quercetin, Celastrol, CCE9.

draws increasing attention to this procedure. On the one hand, in non-cancer tissues, such as liver, muscle, and adipose, NR4A1 can alter the expression levels of fatty acid key enzymes ACC, SCD1, CPT1A, as well as transporters CD36, adiponectin receptor 2 (ADIPOR2), and Caveolin 3 (CAV3) by regulating LKB1-AMPK classic signaling pathway and its downstream SREBP1c and FAS. On the other hand, as a core hallmark of cancer, altered fatty acid synthesis is specifically important. Tumor cells drive this process to provide energy and biological materials for uncontrolled proliferation. Hence, overexpression of NR4A1 binds to TPβ to maintain the FAO-NADPH-ROS loop, leading to the suppression of cancer growth. NR4A1 has been found to bind to the NBRE or coactivator SWI/SNF complex response elements by NR4A2 or PPARγ, resulting in the change of fatty acid-related genes ACOX, CPT1M, FABP2, and FABP4 in melanoma and breast cancer. In addition, NR4A1 exhibits cholesterol regulating function by interacting with LXR in acute myeloid leukemia. All the above evidence suggests the unique and irreplaceable features in lipid metabolism; However, whether NR4A1 displays similar metabolic effects in regulating cancer metastasis, cancer stem cell phenotypes need to be deeply explored in the future.

Furthermore, MYC, mTOR, and HIF-1 are the main mediators regulating SLC transporters and glutamine synthase relative enzymes in cancers. NR4A1 is involved in glutamine metabolism through interaction with β-catenin, which further influences the expression level of Myc. In order to increase the activity of key enzymes GLS and GDH, NR4A1 induced mTOR upregulation by activating the ROS-depended AMPK signaling pathway. Similar to β-catenin, NR4A1 can stabilize and transactivate HIF-1α; meanwhile, HIF-1α bins to the promoter NR4A1 and promotes its transcription; under hypoxia, HIF-1α and NR4A1 form an interaction circuit to affect each other.

Additionally, amino acid regulation is also a function of NR4A1. It has been reported that amino acid deficiency can induce NR4A1 expression. NR4A1−/− zebrafish larva decreased total amino acids and the level of ASP, Glu, Ala, Tyr, His, Met, Pro. However, there are rare studies on tumor cells, so it could be a promising area worth exploring further. Besides the amino acid metabolism, NR4A1 also participates in tumor immunity. The COX2/NR4A1/CTNNB1 axis has been reported to increase CD34+ T effector cells, while TCR- NR4A1-β-AR system can increase metabolic reprogramming of CD8+ T-cell activation through downregulating GLUT-1 expression. Therefore, for one thing, further studies could focus on the diverse functions of NR4A1 on cancer immunity by glycolysis reprogramming. For another, maybe it can be extended the roles of NR4A1 to other aspects of tumor metabolism.

Cancer metabolism provides innovative opportunities for next-generation anticancer therapies that could be further improved using novel NR4A1 agonists or antagonists that simultaneously regulate NR4A1 and its downstream signaling pathways.

## Conclusion and perspectives

NR4A1 is a well-studied transcription factor, and recent researches focus on identifying its genomic and non-genomic effects in cancers, including melanoma (106), breast cancer (107), and colorectal cancer (31). As described above, NR4A1 exhibits important functions in cancer cell metabolic reprogramming. By regulating glucose and fatty acid-related enzymes, such as GLUT4, PEPCK1, ACC, ACLY, NR4A1 exhibits divers metabolic effects by regulating the downstream signaling pathways. NR4A1 acts as a novel application to enable tumor growth, evasion of apoptosis, migration, and invasion. However, whether NR4A1 displays the lipid metabolic functions in cancer metastasis is not fully identified. Although increasingly ligands are found to bind to, and active or inactive NR4A1, leading to cancer cell growth, apoptosis, autophagy, EMT, the study on metabolism is rare. Thus, it is an urgent need to understand the metabolic functions of NR4A1, especially how this receptor mediates fatty acid synthesis, amino acid metabolism and glutamine in tumors. The underlying mechanisms are worthy of exploring. The continued investigation of agents that can modulate NR4A1 is needed. The selective NR4A1 agonists or antagonists against cancer cell metabolism might be potential for cancer treatment.

## Author contributions

SD: Conceptualization, Methodology, Investigation, Writing-original draft. BC: Writing-review & editing. JH and XL: Writing & Revising, Supervision. All authors contributed to the article and approved the submitted version.

## Funding

This work was funded by the Natural Science Foundation of Jiangsu Province (BK20220464, BK20190938), the National Natural Science Foundation of China (No. 82004288), Project of National Clinical Research Base of Traditional Chinese Medicine in Jiangsu Province (No. JD2019SZXYB04) and Jiangsu province TCM leading talent training project (No. SLJ0211) and Natural Science Foundation of Jiangsu Province (BK20190938).

## 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

## References

1. Wolpaw AJ, Dang CV. Exploiting metabolic vulnerabilities of cancer with precision and accuracy. *Trends Cell Biol* (2018) 28(3):201–12. doi: 10.1016/j.tcb.2017.11.006

2. Park JH, Pyun WY, Park HW. Cancer metabolism: Phenotype, signaling and therapeutic targets. *Cells* (2020) 9(10):2308. doi: 10.3390/cells9102308

3. Yu Q, Dai W, Ji J, Wu L, Feng J, Li J, et al. Sodium butyrate inhibits aerobic glycolysis of hepatocellular carcinoma cells *Via* the c-Myc/Hexokinase 2 pathway. *J Cell Mol Med* (2022) 26(10):3031–45. doi: 10.1111/jcmm.17322

4. Carmo F, Silva C, Martel F. Inhibition of glutamine cellular uptake contributes to the cytotoxic effect of xanthohumol in triple-negative breast cancer cells. *Nutr Cancer* (2022) 20:1–18. doi: 10.1080/01635581.2022.2076889

5. Li Z, Zhang H. Reprogramming of glucose, fatty acid and amino acid metabolism for cancer progression. *Cell Mol Life Sci* (2016) 73(2):377–92. doi: 10.1007/s00018-015-2070-4

6. Fan H, Huang W, Guo Y, Ma X, Yang J. Alpha-linolenic acid suppresses proliferation and invasion in osteosarcoma cells *Via* inhibiting fatty acid synthase. *Molecules* (2022) 27(9):2741. doi: 10.3390/molecules27092741

7. Li H, Song J, He Y, Liu Y, Liu Z, Sun W, et al. Crispr/Cas9 screens reveal that hexokinase 2 enhances cancer stemness and tumorigenicity by activating the Acsl4-fatty acid beta-oxidation pathway. *Adv Sci (Weinh)* (2022) 9(21):e2105126. doi: 10.1002/advs.202105126

8. Buel GR, Dang H, Asara JM, Blenis J, Mutvei AP. Prolonged deprivation of arginine or leucine induces Pi3k/Akt-dependent reactivation of Mtorc1. *J Biol Chem* (2022) 298(6):102030. doi: 10.1016/j.jbc.2022.102030

9. Martinez-Gonzalez J, Badimon L. The Nr4a subfamily of nuclear receptors: New early genes regulated by growth factors in vascular cells. *Cardiovasc Res* (2005) 65(3):609–18. doi: 10.1016/j.cardiores.2004.10.002

10. To SK, Zeng JZ, Wong AS. Nur77: A potential therapeutic target in cancer. *Expert Opin Ther Targets* (2012) 16(6):573–85. doi: 10.1517/14728222.2012.680958

11. Kumar R, Thompson EB. The structure of the nuclear hormone receptors. *Steroids* (1999) 64(5):310–9. doi: 10.1016/s0039-128x(99)00014-8

12. Niu G, Lu L, Gan J, Zhang D, Liu J, Huang G. Dual roles of orphan nuclear receptor Tr3/Nur77/Ngfi-b in mediating cell survival and apoptosis. *Int Rev Cell Mol Biol* (2014) 313:219–58. doi: 10.1016/B978-0-12-800177-6.00007-4

13. de Vera IM, Giri PK, Munoz-Tello P, Brust R, Fuhrmann J, Matta-Camacho E, et al. Identification of a binding site for unsaturated fatty acids in the orphan nuclear receptor Nurr1. *ACS Chem Biol* (2016) 11(7):1795–9. doi: 10.1021/acschembio.6b00037

14. Yousefi H, Fong J, Alahari SK. Nr4a family genes: A review of comprehensive prognostic and gene expression profile analysis in breast cancer. *Front Oncol* (2022) 12:777824. doi: 10.3389/fonc.2022.777824

15. Chen J, Lopez-Moyado IF, Seo H, Lio CJ, Hempleman LJ, Sekiya T, et al. Nr4a transcription factors limit car T cell function in solid tumours. *Nature* (2019) 567(7749):530–4. doi: 10.1038/s41586-019-0985-x

16. Hazel TG, Nathans D, Lau LF. A gene inducible by serum growth factors encodes a member of the steroid and thyroid hormone receptor superfamily. *Proc Natl Acad Sci USA* (1988) 85(22):8444–8. doi: 10.1073/pnas.85.22.8444

17. Chang C, Kokontis J, Liao SS, Chang Y. Isolation and characterization of human Tr3 receptor: A member of steroid receptor superfamily. *J Steroid Biochem* (1989) 34(1-6):391–5. doi: 10.1016/0022-4731(89)90114-3

18. Xie P, Yan LJ, Zhou HL, Cao HH, Zheng YR, Lu ZB, et al. Emodin protects against lipopolysaccharide-induced acute lung injury *Via* the Jnk/Nur77/C-jun signaling pathway. *Front Pharmacol* (2022) 13:717271. doi: 10.3389/fphar.2022.717271

19. di Martino O, Niu H, Hadwiger G, Ferris MA, Welch JS. Cytokine exposure mediates transcriptional activation of the orphan nuclear receptor Nur77 in hematopoietic cells. *J Biol Chem* (2021) 297(5):101240. doi: 10.1016/j.jbc.2021.101240

20. Niu B, Liu J, Lv B, Lin J, Li X, Wu C, et al. Interplay between transforming growth factor-beta and Nur77 in dual regulations of inhibitor of differentiation 1 for colonic tumorigenesis. *Nat Commun* (2021) 12(1):2809. doi: 10.1038/s41467-021-23048-5

21. Xu Y, Tian J, Kang Q, Yuan H, Liu C, Li Z, et al. Knockout of Nur77 leads to amino acid, lipid, and glucose metabolism disorders in zebrafish. *Front Endocrinol (Lausanne)* (2022) 13:864631. doi: 10.3389/fendo.2022.864631

22. Wu L, Chen L. Characteristics of Nur77 and its ligands as potential anticancer compounds (Review). *Mol Med Rep* (2018) 18(6):4793–801. doi: 10.3892/mmr.2018.9515

23. Tian H, Chen F, Wang Y, Liu Y, Ma G, Zhao Y, et al. Nur77 prevents osteoporosis by inhibiting the nf-kappab signalling pathway and osteoclast differentiation. *J Cell Mol Med* (2022) 26(8):2163–76. doi: 10.1111/jcmm.17238

24. Yuan R, Zhang W, Nie P, Lan K, Yang X, Yin A, et al. Nur77 deficiency exacerbates macrophage Nlrp3 inflammasome-mediated inflammation and accelerates atherosclerosis. *Oxid Med Cell Longev* (2022) 2022:2017815. doi: 10.1155/2022/2017815

25. Hiwa R, Brooks JF, Mueller JL, Nielsen HV, Zikherman J. Nr4a nuclear receptors in T and b lymphocytes: Gatekeepers of immune tolerance. *Immunol Rev* (2022) 307(1):116–33. doi: 10.1111/imr.13072

26. Dodat F, Mader S, Levesque D. Minireview: What is known about sumoylation among Nr4a family members? *J Mol Biol* (2021) 433(21):167212. doi: 10.1016/j.jmb.2021.167212

27. Maira M, Martens C, Philips A, Drouin J. Heterodimerization between members of the nur subfamily of orphan nuclear receptors as a novel mechanism for gene activation. *Mol Cell Biol* (1999) 19(11):7549–57. doi: 10.1128/MCB.19.11.7549

28. Zheng Y, Comaills V, Burr R, Boulay G, Miyamoto DT, Wittner BS, et al. Cox-2 mediates tumor-stromal prolactin signaling to initiate tumorigenesis. *Proc Natl Acad Sci USA* (2019) 116(12):5223–32. doi: 10.1073/pnas.1819303116

29. Lee SO, Abdelrahim M, Yoon K, Chintharlapalli S, Papineni S, Kim K, et al. Inactivation of the orphan nuclear receptor Tr3/Nur77 inhibits pancreatic cancer cell and tumor growth. *Cancer Res* (2010) 70(17):6824–36. doi: 10.1158/0008-5472.CAN-10-1992

30. Zhao Y, Liu Y, Zheng D. Alpha 1-Antichymotrypsin/Serpina3 is a novel target of orphan nuclear receptor Nur77. *FEBS J* (2008) 275(5):1025–38. doi: 10.1111/j.1742-4658.2008.06269.x

31. Shi Z, To SKY, Zhang S, Deng S, Artemenko M, Zhang M, et al. Hypoxia-induced Nur77 activates Pi3k/Akt signaling *Via* suppression of Dicer/Let-7i-5p to induce epithelial-to-Mesenchymal transition. *Theranostics* (2021) 11(7):3376–91. doi: 10.7150/thno.52190

32. Pawlak A, Strzadala L, Kalas W. Non-genomic effects of the Nr4a1/Nur77/Tr3/Ngfib orphan nuclear receptor. *Steroids* (2015) 95:1–6. doi: 10.1016/j.steroids.2014.12.020

33. Wu H, Lin Y, Li W, Sun Z, Gao W, Zhang H, et al. Regulation of Nur77 expression by beta-catenin and its mitogenic effect in colon cancer cells. *FASEB J* (2011) 25(1):192–205. doi: 10.1096/fj.10-166462

34. Sun Z, Cao X, Jiang MM, Qiu Y, Zhou H, Chen L, et al. Inhibition of beta-catenin signaling by nongenomic action of orphan nuclear receptor Nur77. *Oncogene* (2012) 31(21):2653–67. doi: 10.1038/onc.2011.448

35. Moll UM, Marchenko N, Zhang XK. P53 and Nur77/Tr3 - transcription factors that directly target mitochondria for cell death induction. *Oncogene* (2006) 25(34):4725–43. doi: 10.1038/sj.onc.1209601

36. Zhao BX, Chen HZ, Lei NZ, Li GD, Zhao WX, Zhan YY, et al. P53 mediates the negative regulation of Mdm2 by orphan receptor Tr3. *EMBO J* (2006) 25(24):5703–15. doi: 10.1038/sj.emboj.7601435

37. Zhan YY, Chen Y, Zhang Q, Zhuang JJ, Tian M, Chen HZ, et al. The orphan nuclear receptor Nur77 regulates Lkb1 localization and activates ampk. *Nat Chem Biol* (2012) 8(11):897–904. doi: 10.1038/nchembio.1069

38. Hu M, Luo Q, Alitongbieke G, Chong S, Xu C, Xie L, et al. Celastrol-induced Nur77 interaction with Traf2 alleviates inflammation by promoting mitochondrial ubiquitination and autophagy. *Mol Cell* (2017) 66(1):141–53.e6. doi: 10.1016/j.molcel.2017.03.008

39. Ye Z, Zhuo Q, Hu Q, Xu X, Mengqi L, Zhang Z, et al. Fbw7-Nra41-Scd1 axis synchronously regulates apoptosis and ferroptosis in pancreatic cancer cells. *Redox Biol* (2021) 38:101807. doi: 10.1016/j.redox.2020.101807

40. Koenis DS, Medzikovic L, van Loenen PB, van Weeghel M, Huveneers S, Vos M, et al. Nuclear receptor Nur77 limits the macrophage inflammatory response through transcriptional reprogramming of mitochondrial metabolism. *Cell Rep* (2018) 24(8):2127–40.e7. doi: 10.1016/j.celrep.2018.07.065

41. Yang PB, Hou PP, Liu FY, Hong WB, Chen HZ, Sun XY, et al. Blocking ppargamma interaction facilitates Nur77 interdiction of fatty acid uptake and suppresses breast cancer progression. *Proc Natl Acad Sci USA* (2020) 117(44):27412–22. doi: 10.1073/pnas.2002997117

42. Ye T, Peng J, Liu X, Hou S, Niu G, Li Y, et al. Orphan nuclear receptor Tr3/Nur77 differentially regulates the expression of integrins in angiogenesis. *Microvasc Res* (2019) 122:22–33. doi: 10.1016/j.mvr.2018.10.011

43. Archetti M. Evolutionary dynamics of the warburg effect: Glycolysis as a collective action problem among cancer cells. *J Theor Biol* (2014) 341:1–8. doi: 10.1016/j.jtbi.2013.09.017

44. Schwartz L, Supuran CT, Alfarouk KO. The warburg effect and the hallmarks of cancer. *Anticancer Agents Med Chem* (2017) 17(2):164–70. doi: 10.2174/1871520616666161031143301

45. Kanzleiter T, Preston E, Wilks D, Ho B, Benrick A, Reznick J, et al. Overexpression of the orphan receptor Nur77 alters glucose metabolism in rat muscle cells and rat muscle *in vivo*. *Diabetologia* (2010) 53(6):1174–83. doi: 10.1007/s00125-010-1703-2

46. Men L, Hui W, Guan X, Song T, Wang X, Zhang S, et al. Cardiac transcriptome analysis reveals Nr4a1 mediated glucose metabolism dysregulation in response to high-fat diet. *Genes (Basel)* (2020) 11(7):720. doi: 10.3390/genes11070720

47. Mohankumar K, Lee J, Wu CS, Sun Y, Safe S. Bis-Indole-Derived Nr4a1 ligands and metformin exhibit Nr4a1-dependent glucose metabolism and uptake in C2c12 cells. *Endocrinology* (2018) 159(5):1950–63. doi: 10.1210/en.2017-03049

48. Bian XL, Chen HZ, Yang PB, Li YP, Zhang FN, Zhang JY, et al. Nur77 suppresses hepatocellular carcinoma *Via* switching glucose metabolism toward gluconeogenesis through attenuating phosphoenolpyruvate carboxykinase sumoylation. *Nat Commun* (2017) 8:14420. doi: 10.1038/ncomms14420

49. Guan YF, Huang QL, Ai YL, Chen QT, Zhao WX, Wang XM, et al. Nur77-activated lncrna Wfdc21p attenuates hepatocarcinogenesis *Via* modulating glycolysis. *Oncogene* (2020) 39(11):2408–23. doi: 10.1038/s41388-020-1158-y

50. Corrocher FA, Bueno de Paiva L, Duarte ASS, Ferro KP, Silveira LDR, de Lima TI, et al. Reduced expression of Nr4a1 activates glycolytic pathway in acute promyelocytic leukemia cells. *Leuk Lymphoma* (2018) 59(6):1501–4. doi: 10.1080/10428194.2017.1387900

51. Zhan Y, Du X, Chen H, Liu J, Zhao B, Huang D, et al. Cytosporone b is an agonist for nuclear orphan receptor Nur77. *Nat Chem Biol* (2008) 4(9):548–56. doi: 10.1038/nchembio.106

52. Kierans SJ, Taylor CT. Regulation of glycolysis by the hypoxia-inducible factor (Hif): Implications for cellular physiology. *J Physiol* (2021) 599(1):23–37. doi: 10.1113/JP280572

53. Yoo YG, Yeo MG, Kim DK, Park H, Lee MO. Novel function of orphan nuclear receptor Nur77 in stabilizing hypoxia-inducible factor-1alpha. *J Biol Chem* (2004) 279(51):53365–73. doi: 10.1074/jbc.M408554200

54. Dong S, Liang S, Cheng Z, Zhang X, Luo L, Li L, et al. Ros/Pi3k/Akt and Wnt/Beta-catenin signalings activate hif-1alpha-Induced metabolic reprogramming to impart 5-fluorouracil resistance in colorectal cancer. *J Exp Clin Cancer Res* (2022) 41(1):15. doi: 10.1186/s13046-021-02229-6

55. To SK, Zeng WJ, Zeng JZ, Wong AS. Hypoxia triggers a Nur77-Beta-Catenin feed-forward loop to promote the invasive growth of colon cancer cells. *Br J Cancer* (2014) 110(4):935–45. doi: 10.1038/bjc.2013.816

56. Leone RD, Zhao L, Englert JM, Sun IM, Oh MH, Sun IH, et al. Glutamine blockade induces divergent metabolic programs to overcome tumor immune evasion. *Science* (2019) 366(6468):1013–21. doi: 10.1126/science.aav2588

57. Recouvreux MV, Moldenhauer MR, Galenkamp KMO, Jung M, James B, Zhang Y, et al. Glutamine depletion regulates slug to promote emt and metastasis in pancreatic cancer. *J Exp Med* (2020) 217(9):e20200388. doi: 10.1084/jem.20200388

58. Cetinbas N, Daugaard M, Mullen AR, Hajee S, Rotblat B, Lopez A, et al. Loss of the tumor suppressor Hace1 leads to ros-dependent glutamine addiction. *Oncogene* (2015) 34(30):4005–10. doi: 10.1038/onc.2014.316

59. Feng M, Xiong G, Cao Z, Yang G, Zheng S, Qiu J, et al. Lat2 regulates glutamine-dependent mtor activation to promote glycolysis and chemoresistance in pancreatic cancer. *J Exp Clin Cancer Res* (2018) 37(1):274. doi: 10.1186/s13046-018-0947-4

60. Gao P, Tchernyshyov I, Chang TC, Lee YS, Kita K, Ochi T, et al. C-myc suppression of mir-23a/B enhances mitochondrial glutaminase expression and glutamine metabolism. *Nature* (2009) 458(7239):762–5. doi: 10.1038/nature07823

61. Li J, Li X, Wu L, Pei M, Li H, Jiang Y. Mir-145 inhibits glutamine metabolism through c-Myc/Gls1 pathways in ovarian cancer cells. *Cell Biol Int* (2019) 43(8):921–30. doi: 10.1002/cbin.11182

62. Ganguly K, Bhatia R, Rauth S, Kisling A, Atri P, Thompson C, et al. Mucin 5ac serves as the nexus for beta-Catenin/C-Myc interplay to promote glutamine dependency during pancreatic cancer chemoresistance. *Gastroenterology* (2022) 162(1):253–68.e13. doi: 10.1053/j.gastro.2021.09.017

63. Lu WJ, Chua MS, Wei W, So SK. Ndrg1 promotes growth of hepatocellular carcinoma cells by directly interacting with gsk-3beta and Nur77 to prevent beta-catenin degradation. *Oncotarget* (2015) 6(30):29847–59. doi: 10.18632/oncotarget.4913

64. Shrestha R, Mohankumar K, Martin G, Hailemariam A, Lee SO, Jin UH, et al. Flavonoids kaempferol and quercetin are nuclear receptor 4a1 (Nr4a1, Nur77) ligands and inhibit rhabdomyosarcoma cell and tumor growth. *J Exp Clin Cancer Res* (2021) 40(1):392. doi: 10.1186/s13046-021-02199-9

65. Yuan L, Sheng X, Willson AK, Roque DR, Stine JE, Guo H, et al. Glutamine promotes ovarian cancer cell proliferation through the Mtor/S6 pathway. *Endocr Relat Cancer* (2015) 22(4):577–91. doi: 10.1530/ERC-15-0192

66. Mohankumar K, Li X, Sridharan S, Karki K, Safe S. Nuclear receptor 4a1 (Nr4a1) antagonists induce ros-dependent inhibition of mtor signaling in endometrial cancer. *Gynecol Oncol* (2019) 154(1):218–27. doi: 10.1016/j.ygyno.2019.04.678

67. Safe S, Karki K. The paradoxical roles of orphan nuclear receptor 4a (Nr4a) in cancer. *Mol Cancer Res* (2021) 19(2):180–91. doi: 10.1158/1541-7786

68. Xiang L, Mou J, Shao B, Wei Y, Liang H, Takano N, et al. Glutaminase 1 expression in colorectal cancer cells is induced by hypoxia and required for tumor growth, invasion, and metastatic colonization. *Cell Death Dis* (2019) 10(2):40. doi: 10.1038/s41419-018-1291-5

69. Kim BY, Kim H, Cho EJ, Youn HD. Nur77 upregulates hif-alpha by inhibiting pvhl-mediated degradation. *Exp Mol Med* (2008) 40(1):71–83. doi: 10.3858/emm.2008.40.1.71

70. Koenis DS, Medzikovic L, Vos M, Beldman TJ, van Loenen PB, van Tiel CM, et al. Nur77 variants solely comprising the amino-terminal domain activate hypoxia-inducible factor-1alpha and affect bone marrow homeostasis in mice and humans. *J Biol Chem* (2018) 293(39):15070–83. doi: 10.1074/jbc.RA118.002720

71. Choi JW, Park SC, Kang GH, Liu JO, Youn HD. Nur77 activated by hypoxia-inducible factor-1alpha overproduces proopiomelanocortin in Von hippel-Lindau-Mutated renal cell carcinoma. *Cancer Res* (2004) 64(1):35–9. doi: 10.1158/0008-5472.can-03-0145

72. Wohlkoenig C, Leithner K, Olschewski A, Olschewski H, Hrzenjak A. Tr3 is involved in hypoxia-induced apoptosis resistance in lung cancer cells downstream of hif-1alpha. *Lung Cancer* (2017) 111:15–22. doi: 10.1016/j.lungcan.2017.06.013

73. Fhu CW, Ali A. Fatty acid synthase: An emerging target in cancer. *Molecules* (2020) 25(17):3935. doi: 10.3390/molecules25173935

74. Molendijk J, Robinson H, Djuric Z, Hill MM. Lipid mechanisms in hallmarks of cancer. *Mol Omics* (2020) 16(1):6–18. doi: 10.1039/c9mo00128j

75. Liang Y, Han H, Liu L, Duan Y, Yang X, Ma C, et al. Cd36 plays a critical role in proliferation, migration and tamoxifen-inhibited growth of er-positive breast cancer cells. *Oncogenesis* (2018) 7(12):98. doi: 10.1038/s41389-018-0107-x

76. Ladanyi A, Mukherjee A, Kenny HA, Johnson A, Mitra AK, Sundaresan S, et al. Adipocyte-induced Cd36 expression drives ovarian cancer progression and metastasis. *Oncogene* (2018) 37(17):2285–301. doi: 10.1038/s41388-017-0093-z

77. Pan J, Fan Z, Wang Z, Dai Q, Xiang Z, Yuan F, et al. Cd36 mediates palmitate acid-induced metastasis of gastric cancer *Via* Akt/Gsk-3beta/Beta-Catenin pathway. *J Exp Clin Cancer Res* (2019) 38(1):52. doi: 10.1186/s13046-019-1049-7

78. Yen MC, Chou SK, Kan JY, Kuo PL, Hou MF, Hsu YL. New insight on solute carrier family 27 member 6 (Slc27a6) in tumoral and non-tumoral breast cells. *Int J Med Sci* (2019) 16(3):366–75. doi: 10.7150/ijms.29946

79. Yen MC, Chou SK, Kan JY, Kuo PL, Hou MF, Hsu YL. Solute carrier family 27 member 4 (Slc27a4) enhances cell growth, migration, and invasion in breast cancer cells. *Int J Mol Sci* (2018) 19(11):3434. doi: 10.3390/ijms19113434

80. Furuhashi M, Hotamisligil GS. Fatty acid-binding proteins: Role in metabolic diseases and potential as drug targets. *Nat Rev Drug Discovery* (2008) 7(6):489–503. doi: 10.1038/nrd2589

81. Kawaguchi K, Senga S, Kubota C, Kawamura Y, Ke Y, Fujii H. High expression of fatty acid-binding protein 5 promotes cell growth and metastatic potential of colorectal cancer cells. *FEBS Open Bio* (2016) 6(3):190–9. doi: 10.1002/2211-5463.12031

82. Uehara H, Takahashi T, Oha M, Ogawa H, Izumi K. Exogenous fatty acid binding protein 4 promotes human prostate cancer cell progression. *Int J Cancer* (2014) 135(11):2558–68. doi: 10.1002/ijc.28903

83. Wang X, Li G, Guo C, Zhang J, Kong J, He J, et al. Ethyl 2-[2,3,4-Trimethoxy-6-(1-Octanoyl)Phenyl] acetate (Tmpa) ameliorates lipid accumulation by disturbing the combination of Lkb1 with Nur77 and activating the ampk pathway in Hepg2 cells and mice primary hepatocytes. *Diabetes Metab Syndr Obes* (2021) 14:4165–77. doi: 10.2147/DMSO.S321246

84. Eberle D, Hegarty B, Bossard P, Ferre P, Foufelle F. Srebp transcription factors: Master regulators of lipid homeostasis. *Biochimie* (2004) 86(11):839–48. doi: 10.1016/j.biochi.2004.09.018

85. Pols TW, Ottenhoff R, Vos M, Levels JH, Quax PH, Meijers JC, et al. Nur77 modulates hepatic lipid metabolism through suppression of Srebp1c activity. *Biochem Biophys Res Commun* (2008) 366(4):910–6. doi: 10.1016/j.bbrc.2007.12.039

86. Maxwell MA, Cleasby ME, Harding A, Stark A, Cooney GJ, Muscat GE. Nur77 regulates lipolysis in skeletal muscle cells. evidence for cross-talk between the beta-adrenergic and an orphan nuclear hormone receptor pathway. *J Biol Chem* (2005) 280(13):12573–84. doi: 10.1074/jbc.M409580200

87. Jung YS, Lee HS, Cho HR, Kim KJ, Kim JH, Safe S, et al. Dual targeting of Nur77 and ampkalpha by isoalantolactone inhibits adipogenesis *in vitro* and decreases body fat mass *in vivo*. *Int J Obes (Lond)* (2019) 43(5):952–62. doi: 10.1038/s41366-018-0276-x

88. Li XX, Wang ZJ, Zheng Y, Guan YF, Yang PB, Chen X, et al. Nuclear receptor Nur77 facilitates melanoma cell survival under metabolic stress by protecting fatty acid oxidation. *Mol Cell* (2018) 69(3):480–92.e7. doi: 10.1016/j.molcel.2018.01.001

89. Holla VR, Wu H, Shi Q, Menter DG, DuBois RN. Nuclear orphan receptor Nr4a2 modulates fatty acid oxidation pathways in colorectal cancer. *J Biol Chem* (2011) 286(34):30003–9. doi: 10.1074/jbc.M110.184697

90. Poirot M, Silvente-Poirot S. The tumor-suppressor cholesterol metabolite, dendrogenin a, is a new class of lxr modulator activating lethal autophagy in cancers. *Biochem Pharmacol* (2018) 153:75–81. doi: 10.1016/j.bcp.2018.01.046

91. Segala G, David M, de Medina P, Poirot MC, Serhan N, Vergez F, et al. Dendrogenin a drives lxr to trigger lethal autophagy in cancers. *Nat Commun* (2017) 8(1):1903. doi: 10.1038/s41467-017-01948-9

92. Vettore L, Westbrook RL, Tennant DA. New aspects of amino acid metabolism in cancer. *Br J Cancer* (2020) 122(2):150–6. doi: 10.1038/s41416-019-0620-5

93. Shiozaki Y, Miyazaki-Anzai S, Keenan AL, Miyazaki M. Mef2d-Nr4a1-Fam134b2-Mediated reticulophagy contributes to amino acid homeostasis. *Autophagy* (2022) 18(5):1049–61. doi: 10.1080/15548627.2021.1968228

94. Li Y, Lau LF. Adrenocorticotropic hormone regulates the activities of the orphan nuclear receptor Nur77 through modulation of phosphorylation. *Endocrinology* (1997) 138(10):4138–46. doi: 10.1210/endo.138.10.5464

95. Wu L, Amarachintha S, Xu J, Oley FJr., Du W. Mesenchymal Cox2-pg secretome engages Nr4a-wnt signalling axis in haematopoietic progenitors to suppress anti-leukaemia immunity. *Br J Haematol* (2018) 183(3):445–56. doi: 10.1111/bjh.15548

96. Qiao G, Chen M, Mohammadpour H, MacDonald CR, Bucsek MJ, Hylander BL, et al. Chronic adrenergic stress contributes to metabolic dysfunction and an exhausted phenotype in T cells in the tumor microenvironment. *Cancer Immunol Res* (2021) 9(6):651–64. doi: 10.1158/2326-6066.CIR-20-0445

97. Mohankumar K, Shrestha R, Safe S. Nuclear receptor 4a1 (Nr4a1) antagonists target paraspeckle component 1 (Pspc1) in cancer cells. *Mol Carcinog* (2022) 61(1):73–84. doi: 10.1002/mc.23362

98. Karki K, Mohankumar K, Schoeller A, Martin G, Shrestha R, Safe S. Nr4a1 ligands as potent inhibitors of breast cancer cell and tumor growth. *Cancers (Basel)* (2021) 13(11):2682. doi: 10.3390/cancers13112682

99. Liu J, Wang GH, Duan YH, Dai Y, Bao Y, Hu M, et al. Modulation of the Nur77-Bcl-2 apoptotic pathway by P38alpha mapk. *Oncotarget* (2017) 8 (41):69731–45. doi: 10.18632/oncotarget.19227

100. Ismaiel M, Murphy B, Aldhafiri S, Giffney HE, Thornton K, Mukhopadhya A, et al. Attenuates pro-inflammatory mediators in human colorectal cancer tissue *ex vivo*. *Biochem Biophys Res Commun* (2021) 554:179–85. doi: 10.1016/j.bbrc.2021.03.110

101. Wu J, Liu J, Jia R, Song H. Nur77 inhibits androgen-induced bladder cancer growth. *Cancer Invest* (2013) 31(10):654–60. doi: 10.3109/07357907.2013.853077

102. Hedrick E, Li X, Cheng Y, Lacey A, Mohankumar K, Zarei M, et al. Potent inhibition of breast cancer by bis-Indole-Derived nuclear receptor 4a1 (Nr4a1) antagonists. *Breast Cancer Res Treat* (2019) 177(1):29–40. doi: 10.1007/s10549-019-05279-9

103. Hedrick E, Mohankumar K, Lacey A, Safe S. Inhibition of Nr4a1 promotes ros accumulation and Il24-dependent growth arrest in rhabdomyosarcoma. *Mol Cancer Res* (2019) 17(11):2221–32. doi: 10.1158/1541-7786.MCR-19-0408

104. Hedrick E, Lee SO, Safe S. The nuclear orphan receptor Nr4a1 regulates Beta1-integrin expression in pancreatic and colon cancer cells and can be targeted by Nr4a1 antagonists. *Mol Carcinog* (2017) 56(9):2066–75. doi: 10.1002/mc.22662

105. Li B, Huang J, Liu J, He F, Wen F, Yang C, et al. Discovery of a Nur77-mediated cytoplasmic vacuolation and paraptosis inducer (4-pqbh) for the treatment of hepatocellular carcinoma. *Bioorg Chem* (2022) 121:105651. doi: 10.1016/j.bioorg.2022.105651

106. Chen C, Li Y, Hou S, Bourbon PM, Qin L, Zhao K, et al. Orphan nuclear receptor Tr3/Nur77 biologics inhibit tumor growth by targeting angiogenesis and tumor cells. *Microvasc Res* (2020) 128:103934. doi: 10.1016/j.mvr.2019.103934

107. Zhang X, Li T, Liu S, Xu Y, Meng M, Li X, et al. Beta-glucan from lentinus edodes inhibits breast cancer progression via the Nur77/Hif-1alpha axis. *Biosci Rep* (2020) 40(12):BSsR20201006. doi: 10.1042/BSR20201006

frontiers | Frontiers in Oncology

TYPE Systematic Review
PUBLISHED 20 September 2022
DOI 10.3389/fonc.2022.954751

OPEN ACCESS

EDITED BY
Oliver Diaz,
University of Barcelona, Spain

REVIEWED BY
Yasmein Elbeblawy,
Minia University, Egypt
Matti Sievert,
University Hospital Erlangen, Germany

*CORRESPONDENCE
Jiangfeng Wu
wjfhospital@163.com
Zhengping Wang
zpwang_2016@163.com
Guilong Jin
allon_dy@163.com

†These authors share first authorship

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 27 May 2022
ACCEPTED 30 August 2022
PUBLISHED 20 September 2022

CITATION
Wu J, Zhou Z, Wang X, Jin Y, Wang Z and Jin G (2022) Diagnostic performance of elastosonography in the differential diagnosis of benign and malignant salivary gland tumors: A meta-analysis.
*Front. Oncol.* 12:954751.
doi: 10.3389/fonc.2022.954751

# Diagnostic performance of elastosonography in the differential diagnosis of benign and malignant salivary gland tumors: A meta-analysis

Jiangfeng Wu[1*†], Zhijuan Zhou[2†], Xiaoyun Wang[3], Yun Jin[4], Zhengping Wang[1*] and Guilong Jin[1*]

[1]Department of Ultrasound, Dongyang Hospital of Wenzhou Medical University, Dongyang, China, [2]Department of Ultrasound, Tianxiang East Hospital, Yiwu, China, [3]Department of Nephrology, Dongyang Hospital of Wenzhou Medical University, Dongyang, China, [4]Department of Ultrasound, Dongyang People's Hospital, Dongyang, China

**Purpose:** The clinical practice of elastosonography for the detection of salivary gland tumors is still a controversial issue. The objective of this meta-analysis was to evaluate the effect of elastosonography for the diagnosis of salivary gland tumors and to compare the diagnostic value of elastosonography and conventional ultrasound in the diagnosis of salivary gland tumors.

**Methods:** A comprehensive literature search through PubMed, EMBASE, and Cochrane Library was carried out from inception to November 2021. Two researchers independently extracted the data from the enrolled papers using a standard data extraction form. The pooled sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), diagnostic odds ratio (DOR), and area under the curve (AUC) were calculated to evaluate the diagnostic performance of elastosonography. The Quality Assessment of Diagnostic Accuracy Studies—2 (QUADAS-2) tool was utilized to evaluate the quality of each included study. Meta-DiSc version 1.4, Review Manager 5.3, and StataSE 15 were used.

**Results:** Sixteen studies with a total of 1105 patients with 1146 lesions were included in this meta-analysis. The pooled sensitivity, specificity, PLR, NLR, and DOR of elastosonography for the differentiation between benign and malignant salivary gland tumors were 0.73 (95%CI, 0.66–0.78), 0.64 (95%CI, 0.61–0.67), 2.83 (95%CI, 1.97–4.07), 0.45 (95%CI, 0.32–0.62), and 9.86 (95%CI, 4.49–21.62), respectively, with an AUC of 0.82. Four studies provided data regarding the conventional ultrasound for the differentiation between benign and malignant salivary gland tumors. The pooled sensitivity, specificity, and DOR were 0.62 (95%CI, 0.50–0.73), 0.93 (95%CI, 0.90–0.96), and 25.07 (95%CI, 4.28–146.65), respectively. The meta-regression and subgroup analyses found that assessment methods were associated with significant heterogeneity, and quantitative or semiquantitative elastosonography performed better than the qualitative one.

**Conclusions:** Elastosonography showed a limited value for diagnosing malignant salivary gland tumors; it could be considered as a supplementary diagnostic technology to conventional ultrasound, and quantitative or semiquantitative elastosonography was superior to the qualitative one.

KEYWORDS

**elastosonography, conventional ultrasound, diagnosis, salivary gland tumor, meta-analysis**

## Introduction

Salivary gland tumors are rare, which account for 2%–6% of all head and neck tumors, with an annual incidence ranging from one to five cases per 100,000 population. The most common benign tumors of the salivary glands include pleomorphic adenoma (PA) and Warthin tumor (WT), and 85% of the tumors arise in the parotid gland (1, 2). The treatment strategy of salivary gland tumors depends primarily on its pathology, and preoperative diagnosis of the tumor entity directly affects the selection of surgical procedure; therefore, to determine whether a tumor is benign or malignant is crucial (3–5).

Currently, it is not an easy task to accurately identify benign from malignant salivary tumors because of a broad variety of potential differential diagnoses and the lack of specific imaging characteristics (6–8). Conventional ultrasound (US) is the first-line imaging technique for the diagnosis of the salivary gland tumors as it is a widely available, noninvasive, nonradioactive, and cost-effective method (6). However, the diagnostic accuracy of conventional US depends on the sonographer's diagnostic skill and experience, and there are overlaps of sonographic appearances among different pathological tumors (6, 7). Consequently, the accuracy of conventional US for salivary tumors is less than satisfactory (6–8). Magnetic resonance imaging (MRI) and computed tomography (CT) are also the primary imaging modalities for evaluating salivary gland tumors (9–11). While they can find tumors with high sensitivity, these are less accurate for predicting histology due to an appreciable overlap of imaging findings between different pathological types of salivary gland tumors (8, 11). Thus, acquiring the histopathology of tumors by US-guided fine-needle aspiration cytology (FNAC) or core-needle biopsy (CNB) continues to be necessary before the surgical procedure (12). However, these techniques are invasive and could possibly lead to some complications such as pain and hemorrhage. Thus, an alternative imaging technique providing additional information for identifying salivary gland tumors would be greatly valuable.

Elasticity is an important feature revealing tissue stiffness, which is defined as the rate of change of spatial displacement due to the tensile stress on the tissue under applied pressure (13). Elastosonography is a simple approach that determines tissue stiffness as qualitative, semiquantitative, or quantitative, which has been demonstrated to be useful for the evaluation of thyroid nodules, breast tumors, and cirrhosis (14–17).

The clinical practice of elastosonography for the detection of salivary gland tumors is still a controversial issue, as the diagnostic performance is variable in different studies, with the sensitivity ranging from 38% to 100% and specificity from 26% to 97% (18–21). Thus, we thought it is necessary and timely to summarize currently available data to provide valuable information for clinical practice. The objective of this meta-analysis was to evaluate the effect of elastosonography for the diagnosis of salivary gland tumors and to compare the diagnostic value of elastosonography and conventional US in the diagnosis of salivary gland tumors.

## Materials and methods

This meta-analysis was performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) Statement (22).

---

**Abbreviations:** PA, pleomorphic adenoma; WT, Warthin tumor; US, ultrasound; MRI, magnetic resonance imaging; CT, computed tomography; FNAC, fine-needle aspiration cytology; CNB, core-needle biopsy; MeSH, medical subject heading; PLR, positive likelihood ratio; NLR, negative likelihood ratio; DOR, diagnostic odds ratio; CIs, confidence intervals; SROC, summary receiver operator curve; AUC, area under the curve; $I^2$, inconsistency index; SE, strain elasticity; SWE, shear wave velocity; ARFI, acoustic radiation force impulse.

## Literature search

A comprehensive literature search through PubMed, EMBASE, and Cochrane Library was carried out from inception to November 2021 to identify English-language studies on elastosonography for diagnosing salivary gland tumors. The search strategy was in accordance with the combination of the medical subject heading (MeSH) terms, key words, and word variants for "elastosonography", "elastography", "ultrasound elastography", "ultrasonic elastography", "parotid gland tumor", and "salivary gland tumor". Reference lists of the included papers were also manually screened to detect additional relevant studies. Details of the strategy of searching are provided in Supplementary Table 1.

## Inclusion and exclusion criteria

Two researchers independently scanned the titles and abstracts of the relevant papers. The inclusion and exclusion criteria were defined to increase reproducibility and validity before identifying the studies. All the disagreements were resolved by consensus. All potentially relevant articles satisfying the following criteria were included: (1) diagnostic studies were included; (2) studies assessing the diagnostic performance of elastosonography in differentiating benign from malignant salivary gland tumors were included; and (3) reference standards such as postoperative pathology and/or biopsy results were adopted. The exclusion criteria for the studies were as follows: (1) case reports, reviews, consensus statements, editorial comments, letters, conference reports, and unpublished articles were excluded; (2) studies without sufficient data to construct a 2 × 2 contingency table were excluded; and (3) studies that were not published in English were excluded.

## Data extraction and processing

Two researchers independently extracted the data from the enrolled papers using a standard data extraction form. All the disagreements were resolved by consensus. For included studies, the following items were extracted: author, year of publication, country, study type, sample method, blinding method, sex, number of lesions, age, mean size of tumors, site of lesions, technology, index of elastography, threshold value, reference standard, ultrasound equipment and probe, sensitivity, and specificity.

## Quality assessment

The Quality Assessment of Diagnostic Accuracy Studies—2 (QUADAS-2) tool recommended by the Cochrane collaboration was utilized to evaluate the quality of each included study (23). The QUADAS-2 tool comprises two main categories, namely the risk of bias of four domains and the clinical applicability of three domains. The four domains include patient selection, index test, reference standard, and flow and timing. Every domain is assessed for risk of bias, and the first three domains are assessed for clinical applicability. The quality assessment was performed using the RevMan 5.3 software (Nordic Cochrane Centre, Copenhagen, Denmark).

## Statistical analysis

From the enrolled papers, a bivariate effect model was utilized in this study to calculate the pooled sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), and diagnostic odds ratio (DOR) with corresponding 95% confidence intervals (CIs), which revealed the diagnostic performance of elastosonography in differentiating benign from malignant salivary gland tumors. The presence of a threshold effect was determined by analyzing the Spearman correlation coefficient between sensitivity and the false-positive rate, through a $p < 0.05$ indicating threshold effect. In addition, the summary receiver operator curve (SROC) was developed, and this allowed us to compute the area under the curve (AUC). The AUC values of 0.5–0.7, 0.7–0.9, and >0.9 indicate low, moderate, and perfect diagnostic performance, respectively (24). The Higgins $I^2$ statistic and Q test were utilized to evaluate the heterogeneity of the study with $I^2 >$ 50% showing significant heterogeneity (25). A random-effects model is adopted when the significant heterogeneity is found across studies; otherwise, a fixed-effects model is adopted. The Deeks' funnel plot was generated to evaluate publication bias (26) through a $p < 0.05$ indicating potential publication bias.

Meta-regression and subgroup analyses using several covariates were conducted to investigate the potential factors of heterogeneity: study design (prospective vs. others), year of publication (2010–2013 vs. 2014–2020), diagnostic measurement (quantitative or semiquantitative vs. qualitative), and blinding method (yes vs. unclear). All the above statistical analyses were carried out by Meta-DiSc version 1.4 and StataSE 15 (Stata Corporation, College Station, TX).

# Results

## Literature search

On the basis of the predefined MeSH terms, key words, and word variants, our database search initially identified 210 papers for consideration. PubMed found 95 studies, EMBASE identified 88, and the Cochrane Library discovered 27. After excluding the duplications, the remaining 136 potentially eligible original papers were further reviewed. Furthermore, according to the inclusion criteria in the study selection process, 107 studies were discarded after screening the titles and abstracts. Twenty-nine papers were assessed by reviewing the full text, of which 13 were further excluded. Finally, 16 studies were included in this meta-analysis. Figure 1 shows the detailed flowchart of the literature search.

## Characteristics of included studies

The 16 included studies were published from 2010 to 2020 and written in English (18–21, 27–38). A total of 1105 patients with 1146 lesions were included in these studies. Strain elasticity (SE) was used in 11 studies; shear wave velocity (SWE) was used in one study; acoustic radiation force impulse (ARFI) was used in three studies; and SE and ARFI were used in one study. Ten studies included parotid tumors only, while other studies included parotid, submandibular, or sublingual tumors. Quantitative or semiquantitative methods were utilized in four studies, while qualitative assessment methods were used in 12 studies. More detailed data extracted from the enrolled studies are available in Tables 1 and 2. The histopathological results of the included studies are revealed in Supplementary Table 2.

## Quality assessment

Quality assessment of each study based on the QUADAS-2 tool is shown graphically in Figure 2. Concerning the patient selection domain, five studies were thought to be "unknown" (18–20, 31, 37) because the sample method of patient selection was not definitely mentioned. Concerning the index test domain, four studies (19–21, 35) were thought to be "unknown" because the blinded status of the reference standard was not definitely mentioned; one study was considered as "high" because the sonographer was aware of the histological results of the respective tumors (18). With respect to the reference standard domain, 14 studies (19–21, 27–30, 32–38) were regarded as "unknown" because the blinded status of the elastosonography results was not definitely depicted. Regarding the flow and timing domain, 14 studies were regarded as "unknown" because the authors did not definitely mention the precise duration between the reference standard and the elastosonography examination (18–21, 27–33, 36–38).

With regard to applicability, one study was regarded as "high" for the patient selection domain because 12 benign tumors were all pleomorphic adenomas (19). For the index test and reference standard domains, all studies were thought to have low concerns.

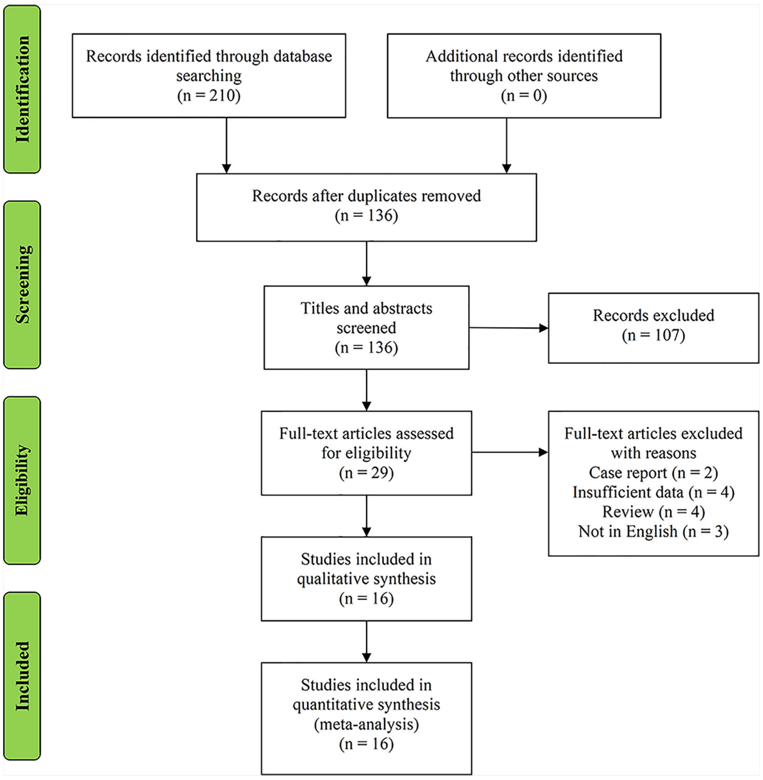


FIGURE 1
Flowchart of study selection.

TABLE 1 Primary data extracted from the included studies for meta-analysis.

| Author | Year | Country | Study type | Sample method | Blinding method | Number of lesions | Male/ female | Age, year (mean or range) | Mean size of tumors, mm | Site of lesions |
|---|---|---|---|---|---|---|---|---|---|---|
| Bhatia et al. (27) | 2010 | China | NR | Consecutive | Single blind | 65 | 48/13 | 60.5 | 26 | Parotid (57) and submandibular (8) |
| Dumitriu et al. (28) | 2011 | Romania | NR | Consecutive | Single blind | 74 | 37/29 | 50.8 | 29.54 | Parotid (63) and submandibular (11) |
| Klintworth et al. (18) | 2012 | Germany | R | NR | Single blind | 57 | 27/30 | 53.3 | NR | Parotid |
| Yerli et al. (29) | 2012 | Turkey | P | Consecutive | Single blind | 36 | NR | NR | 19.5 | Parotid (30) and submandibular (6) |
| Celebi et al. (30) | 2012 | Turkey | P | Consecutive | Single blind | 81 | 36/39 | Men: 44.75; women: 49.44 | NR | Parotid |
| Badea et al. (19) | 2013 | Romania | P | NR | NR | 20 | 15/5 | 40-72 | NR | Parotid |
| Wierzbicka et al. (21) | 2013 | Poland | P | Consecutive | NR | 43 | 16/27 | 54 | NR | Parotid |
| Yu et al. (31) | 2016 | China | NR | NR | Double blind | 51 | NR | 45 | NR | Parotid |
| Zhou et al. (32) | 2016 | China | R | Consecutive | Single blind | 40 | 26/14 | 44 | 24.9 | Parotid (29) and submandibular (11) |
| Cortcu et al. (33) | 2017 | Turkey | P | Consecutive | Single blind | 39 | 22/17 | 52 | PA: 27.8; WT: 33; other benign: 28; malignant: 25.6 | parotid |
| Mansour et al. (20) | 2017 | Germany | P | NR | NR | 202 | NR | 58.6 | NR | Parotid |
| Cantisani et al. (34) | 2017 | Italy | P | Consecutive | Single blind | 63 | 36/29 | 56 | NR | Parotid |
| Altinbas et al. (35) | 2017 | Turkey | P | Consecutive | NR | 54 | 26/20 | 60.01 | 23.68 | Parotid |
| Liu et al. (36) | 2018 | China | P | Consecutive | Single blind | 76 | 40/36 | 47.24 | Benign: 24.53; malignant: 25.05 | Parotid |
| Karaman et al. (37) | 2019 | Turkey | P | NR | Single blind | 60 | 30/30 | 48.8 | 24.36 | Parotid (42) and submandibular (18) |
| Matsuda et al. (38) | 2020 | Japan | R | Consecutive | Single blind | 185 | 103/65 | Benign: 62.8; malignant: 62.7 | Benign: 27.6; malignant: 31.5 | Parotid (169), submandibular (15), and sublingual (1) |

NR, not reported; P, prospective; R, retrospective; PA, pleomorphic adenocarcinoma; WT, Warthin tumor.

## Meta-analysis

The Spearman correlation coefficient was 0.24 (p = 0.37), indicating that no threshold effect existed. The sensitivities of the 16 enrolled studies ranged from 38.0% to 100.0%, and the specificities ranged from 47.0% to 97.0%. Overall, the pooled sensitivity and specificity of elastosonography for the differentiation between benign and malignant salivary gland tumors were 0.73 (95% CI, 0.66–0.78) and 0.64 (95% CI, 0.61–0.67) (Figure 3). The summary estimates of the diagnostic sensitivity and specificity of elastosonography for differentiating benign from malignant salivary gland tumors were analyzed by the random effects method based on significant statistical heterogeneity ($I^2$ = 55.7% for sensitivity, p = 0.00; $I^2$ = 94.1% for specificity, p = 0.00). The pooled PLR, NLR, and DOR of elastosonography for the differentiation between benign and malignant salivary gland tumors were 2.83 (95%CI, 1.97–4.07), 0.45 (95%CI, 0.32–0.62), and 9.86 (95%CI, 4.49–21.62) (Figure 4), respectively. As illustrated in Figure 5, the AUC under the SROC curve for the value of elastosonography in the diagnosis of malignant salivary gland tumors was 0.82.

## Meta-regression and subgroup analyses

As a result of the significant heterogeneity among the 16 included studies, a meta-regression analysis was performed to

TABLE 2 Characteristics of the included studies.

| Author | Technology | Index of elastography | Threshold value | Reference standard | US equipment and probe | Sen (%) | Spe (%) |
|---|---|---|---|---|---|---|---|
| Bhatia et al. (27) | SE | 4-point | ≥ 3 | Surgery or biopsy | Philips IU22 and Siemens Acuson Premium Edition; a 5- to 12-MHz linear probe and a 13.5-MHz linear probe | 83 | 47 |
| Dumitriu et al. (28) | SE | 4-point | ≥ 3 | Surgery | EUB 8500, Hitachi; a 6- to 13-MHz linear probe | 72 | 57 |
| Klintworth et al. (18) | SE | Garland sign or not | Garland sign | Surgery | Acuson S2000; a 9-MHz linear probe | 38 | 96 |
| Yerli et al. (29) | SE | 4-point | ≥ 3 | Surgery or biopsy | EUB-7000 ultrasound system; a 5- to 13-MHz linear probe | 75 | 64 |
| Celebi et al. (30) | SE | 4-point | ≥ 3 | Surgery or biopsy | Siemens Acuson S2000 US; a 13-MHz probe | 59 | 61 |
| Badea et al. (19) | SE or ARFI | NR | NR | Surgery | GE 7, GE 8, GE 9, iU22 Phillips, and Siemens S 2000; a 7- to 11-MHz linear probe | 100 | 50 |
| Wierzbicka et al. (21) | SE | 5-point | ≥ 4 | Surgery | AIXPLORER equipment; a Linear SL-15-4 transducer | 40 | 97 |
| Yu et al. (31) | SWE | SWV | 2.76 | Surgery | ACUSON S2000; a 7- to 12-MHz linear probe | 69 | 97 |
| Zhou et al. (32) | VTI (ARFI) | 6-point | ≥ 4 | Surgery or biopsy | Siemens Acuson S2000; a 9L4 linear probe | 63 | 81 |
| Cortcu et al. (33) | SE | Strain ratio | 2.1 | Surgery or biopsy | Aplio XG SSA-790A; a 12-MHz linear probe | 83 | 97 |
| Mansour et al. (20) | SE | 3-point | ≥ 2 | Surgery | Acuson S2000; a 9- to 14-MHz linear probe | 69 | 26 |
| Cantisani et al. (34) | SE | Elasticity contrast index | >3.5 | Surgery or biopsy | ACCUVIX A30, RS 80 A; a 10- to 18-MHz linear probe | 94 | 89 |
| Altinbas et al. (35) | SE | 0-6 | 3 | Biopsy | Logiq S7 Expert machine; a 9L-D linear probe | 70 | 66 |
| Liu et al. (36) | VTQ (ARFI) | SWV | 2.445 m/s | Surgery or biopsy | Siemens Acuson S2000; 14L5 linear probe and curvilinear probe | 80 | 92 |
| Karaman et al. (37) | SE | 4-point | ≥ 3 | Histopathology | Acuson Antares; a 6- to 13-MHz linear probe | 100 | 66 |
| Matsuda et al. (38) | VTI (ARFI) | 4-point | ≥ 3 | Surgery or biopsy | Siemens Acuson S2000; a 4- to 9-MHz or 14-MHz linear probe | 77 | 64 |

SE, strain elasticity; ARFI, acoustic radiation force impulse; VTI, virtual touch imaging; VTQ, virtual touch quantification; NR, not reported; SWV, shear wave velocity; Sen, sensitivity; Spe, specificity.

explore potential sources of heterogeneity. The covariates included the blinding method (yes vs. unclear), year of publication (2010–2013 vs. 2014–2020), study design (prospective vs. others), and assessment methods (quantitative or semiquantitative vs. qualitative). Among the various potential covariates, the assessment methods were associated with the significant heterogeneity (Table 3).

Four studies (20, 21, 34, 36) provided data regarding the conventional ultrasound for the differentiation between benign and malignant salivary gland tumors. The pooled sensitivity and specificity of conventional ultrasound for malignant salivary gland tumors were 0.62 (95% CI, 0.50–0.73) and 0.93 (95% CI, 0.90–0.96) (Figure 6). The pooled DOR of conventional ultrasound was 25.07 (95%CI, 4.28–146.65) (Figure 7). As illustrated in Figure 8, the AUC under the SROC curve for the value of conventional ultrasound in the diagnosis of malignant salivary gland tumors was 0.74.

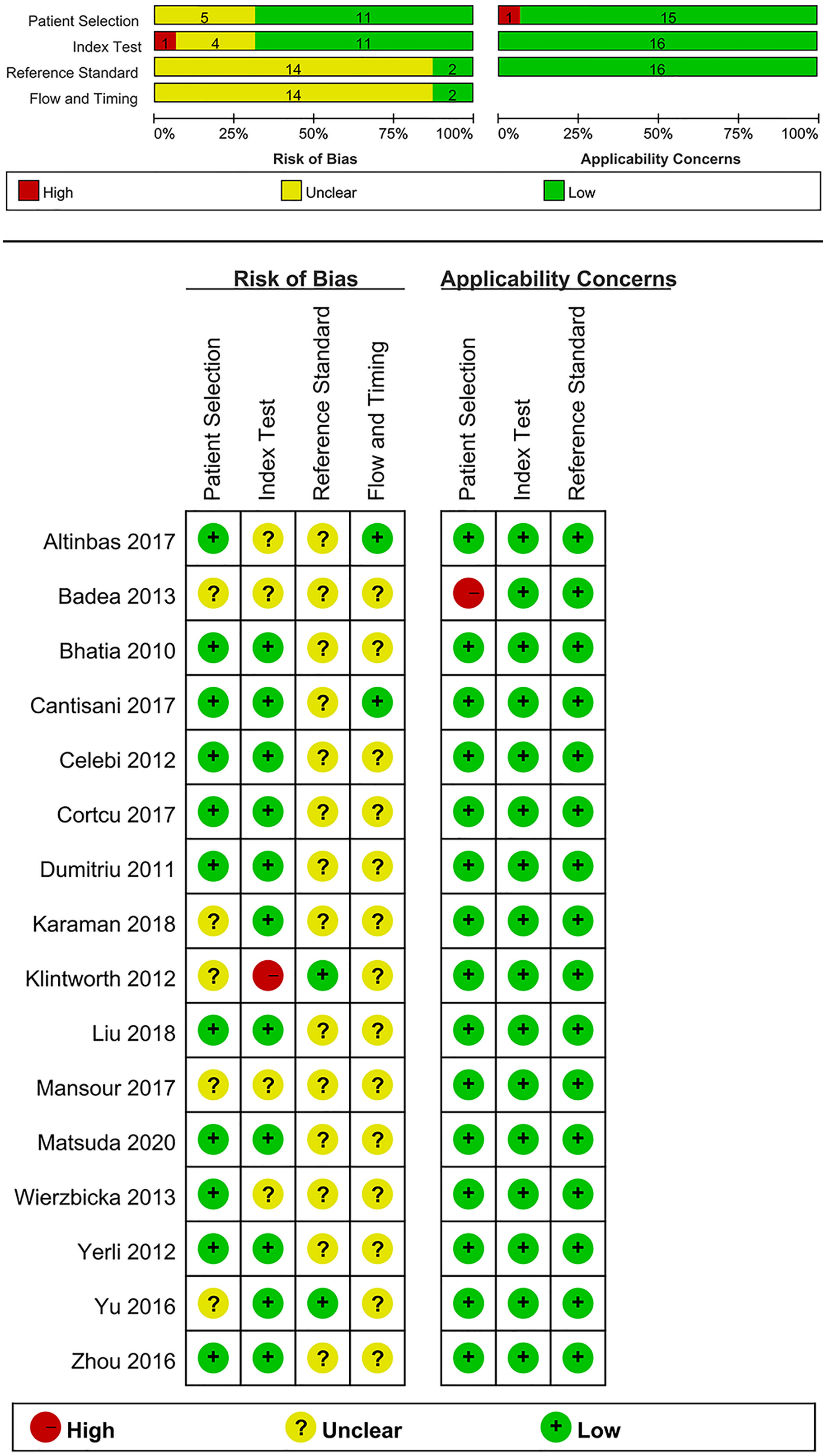
## Sensitivity analysis

A sensitivity analysis was carried out, and the results of the sensitivity analysis found that the meta-analysis results are robust (Figure 8).

## Fagan plot analysis and likelihood matrix

The Fagan diagram was developed for the assessment of clinical application as revealed in Figure 10, indicating that when the pretest probability was 20%, the posttest probability was 46% if the results were positive and 8% if the results were negative for malignant salivary gland tumors (Figure 9).

The likelihood matrix demonstrated that the summary PLR and NLR for the elastosonography diagnosis of malignant salivary gland tumors with 95% confidence intervals were



| | Risk of Bias | | | | Applicability Concerns | | |
|---|---|---|---|---|---|---|---|
| | Patient Selection | Index Test | Reference Standard | Flow and Timing | Patient Selection | Index Test | Reference Standard |
| Altinbas 2017 | + | ? | ? | + | + | + | + |
| Badea 2013 | ? | ? | ? | ? | − | + | + |
| Bhatia 2010 | + | + | ? | ? | + | + | + |
| Cantisani 2017 | + | + | ? | + | + | + | + |
| Celebi 2012 | + | + | ? | ? | + | + | + |
| Cortcu 2017 | + | + | ? | ? | + | + | + |
| Dumitriu 2011 | + | + | ? | ? | + | + | + |
| Karaman 2018 | ? | + | ? | ? | + | + | + |
| Klintworth 2012 | ? | − | + | ? | + | + | + |
| Liu 2018 | + | + | ? | ? | + | + | + |
| Mansour 2017 | ? | ? | ? | ? | + | + | + |
| Matsuda 2020 | + | + | ? | ? | + | + | + |
| Wierzbicka 2013 | + | ? | ? | ? | + | + | + |
| Yerli 2012 | + | + | ? | ? | + | + | + |
| Yu 2016 | ? | + | + | ? | + | + | + |
| Zhou 2016 | + | + | ? | ? | + | + | + |

− High ? Unclear + Low

FIGURE 2
Summary of risk of bias and applicability concerns.

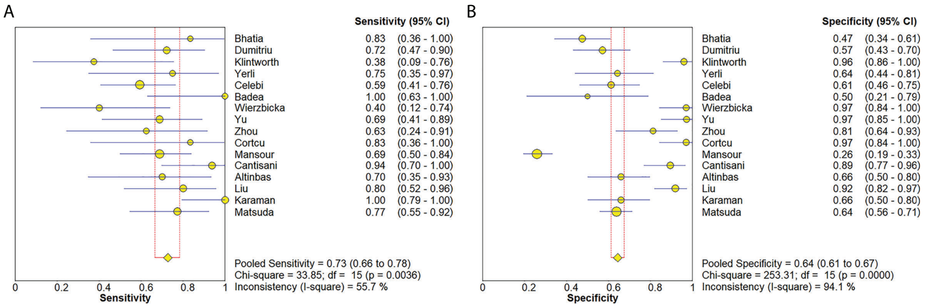


FIGURE 3
Forest plots for sensitivity **(A)** and specificity **(B)** of elastosonography for diagnosis of malignant salivary gland tumors.

concentrated on the right lower quadrant, indicating that elastosonography was not effective for malignant salivary gland tumor confirmation and exclusion (Figure 10). Therefore, elastosonography is a limited value in the diagnosis of malignant salivary gland tumors.

## Publication bias

The Deeks' funnel plot revealed symmetry in scattered points, suggesting that there was no significant publication bias (p = 0.05) (Figure 11).

# Discussion

Our current study found that elastosonography had a pooled moderate sensitivity of 0.73 (95% CI, 0.66–0.78) and a relatively low specificity of 0.64 (95%CI, 0.61–0.67) for the differentiation between benign and malignant salivary gland tumors. The pooled PLR and NLR were 2.83 (95%CI, 1.97–4.07) and 0.45 (95%CI, 0.32–0.62), demonstrating that elastosonography was not effective for malignant salivary gland tumor confirmation and exclusion. The diagnostic odds ratio was 9.86 (95%CI, 4.49–21.62), with an AUC of 0.82. The results indicated that elastosonography revealed a limited value for diagnosing malignant salivary gland tumors.

Four papers supplied the data with regard to conventional ultrasound for the differentiation between benign and malignant salivary gland tumors. The pooled sensitivity, specificity, DOR, and AUC of conventional ultrasound diagnosing malignant salivary gland tumors were 0.62 (95% CI, 0.50–0.73), 0.93 (95% CI, 0.90–0.96), 25.07 (95%CI, 4.28–146.65), and 0.57, respectively. Compared with elastosonography in the diagnosis of salivary gland tumors indirectly, conventional ultrasound had higher specificity (0.93 vs. 0.64), but lower sensitivity (0.62 vs. 0.73), which meant that

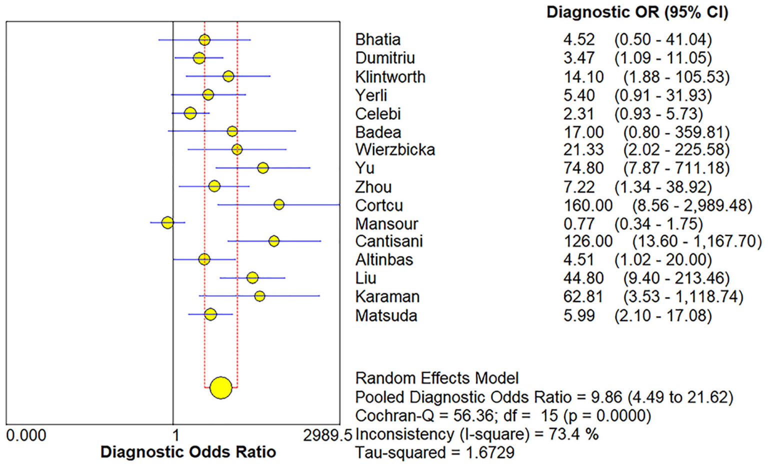


FIGURE 4
Forest plot for diagnostic odds ratio of elastosonography for diagnosis of malignant salivary gland tumors.

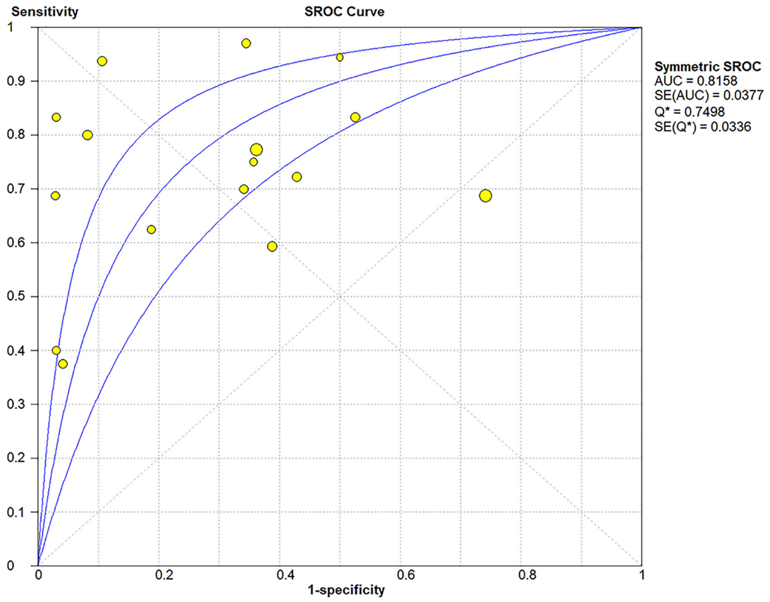


FIGURE 5
Summary receiver operating characteristic (SROC) curve of elastosonography for diagnosis of malignant salivary gland tumors.

conventional ultrasound was more effective in the diagnosis of benign salivary gland tumors than of malignant tumors; in contrast, compared with conventional ultrasound, elastosonography had slightly high sensitivity so that it was more effective in the diagnosis of malignant tumors. Consequently, taking the place of utilizing elastosonography or conventional ultrasound alone, the combined use of the two techniques might result in better diagnostic performance. Hence, we believed that elastosonography could be considered as a supplementary diagnostic technique to conventional ultrasound for the assessment of salivary gland tumors.

A prior meta-analysis by Zhang et al. in 2018 (39) included 10 eligible studies on elastosonography for differential diagnosis between benign and malignant parotid lesions, with a total of 725 parotid lesions, and demonstrated that sonoelastography had a limited value for diagnosing

TABLE 3 Meta-regression and subgroup analyses.

| Covariate | Number of studies | Pooled sensitivity (95% CI) | Pooled specificity (95% CI) | Pooled DOR (95% CI) | AUC | p-Value |
|---|---|---|---|---|---|---|
| **Study design** | | | | | | 0.57 |
| Prospective | 10 | 0.75 (0.67–0.81) | 0.61 (0.57–0.65) | 12.14 (3.60–40.92) | 0.85 | |
| Others | 6 | 0.69 (0.58–0.79) | 0.69 (0.64–0.73) | 7.20 (3.50–14.81) | 0.78 | |
| **Year of publication** | | | | | | 0.23 |
| 2010-2013 | 7 | 0.64 (0.54–0.74) | 0.67 (0.62–0.73) | 4.16 (2.32–7.40) | 0.72 | |
| 2014-2020 | 9 | 0.78 (0.70–0.85) | 0.63 (0.59–0.67) | 16.25 (4.28–61.68) | 0.86 | |
| **Assessment method** | | | | | | 0.0034* |
| Quantitative or semiquantitative | 4 | 0.81 (0.68–0.91) | 0.93 (0.88–0.96) | 73.49 (25.99–207.76) | 0.96 | |
| Qualitative | 12 | 0.70 (0.63–0.77) | 0.58 (0.54–0.61) | 4.72 (2.43–9.17) | 0.74 | |
| **Blinding method** | | | | | | 0.24 |
| Yes | 12 | 0.74 (0.67–0.81) | 0.73 (0.57–0.65) | 12.43 (5.54–27.90) | 0.84 | |
| Unclear | 4 | 0.68 (0.55–0.80) | 0.43 (0.37–0.49) | 4.35 (0.79–23.94) | 0.73 | |

*, Statistical significance ($p < 0.05$); CI, confidence interval; DOR, diagnostic odds ratio; AUC, area under the curve.

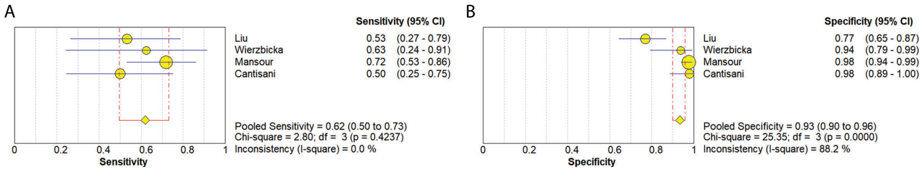


FIGURE 6
Forest plots for sensitivity **(A)** and specificity **(B)** of conventional ultrasound for diagnosis of malignant salivary gland tumors.

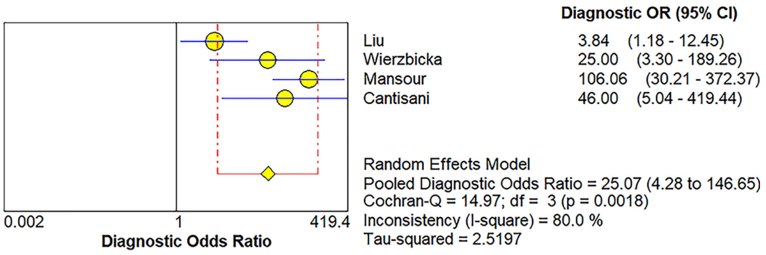


FIGURE 7
Forest plot for diagnostic odds ratio of conventional ultrasound for diagnosis of malignant salivary gland tumors.

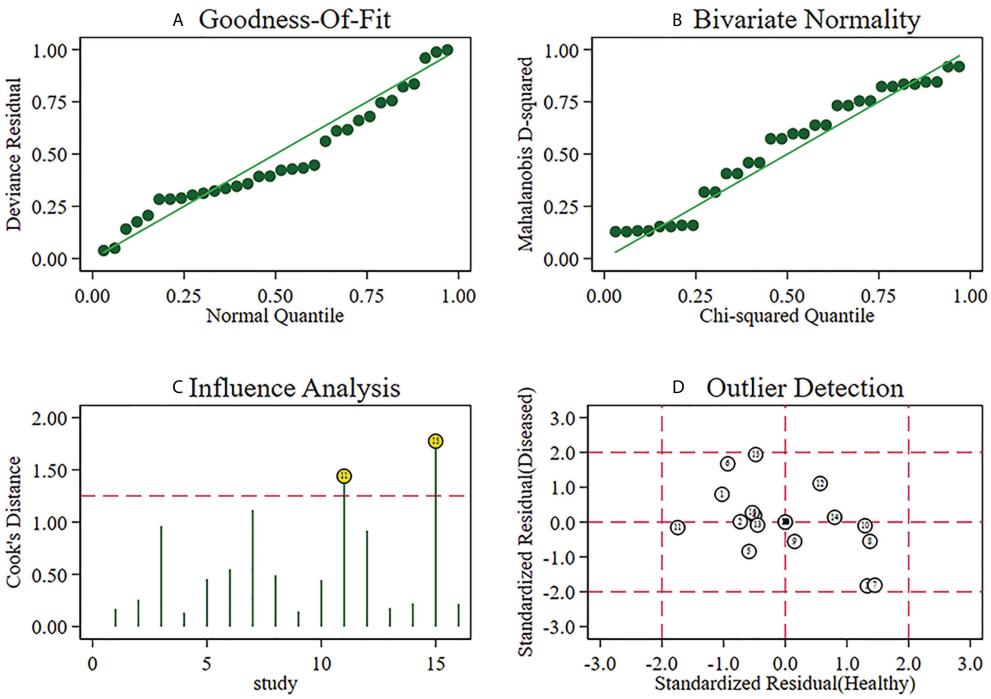


FIGURE 8
Sensitivity analysis of studies.

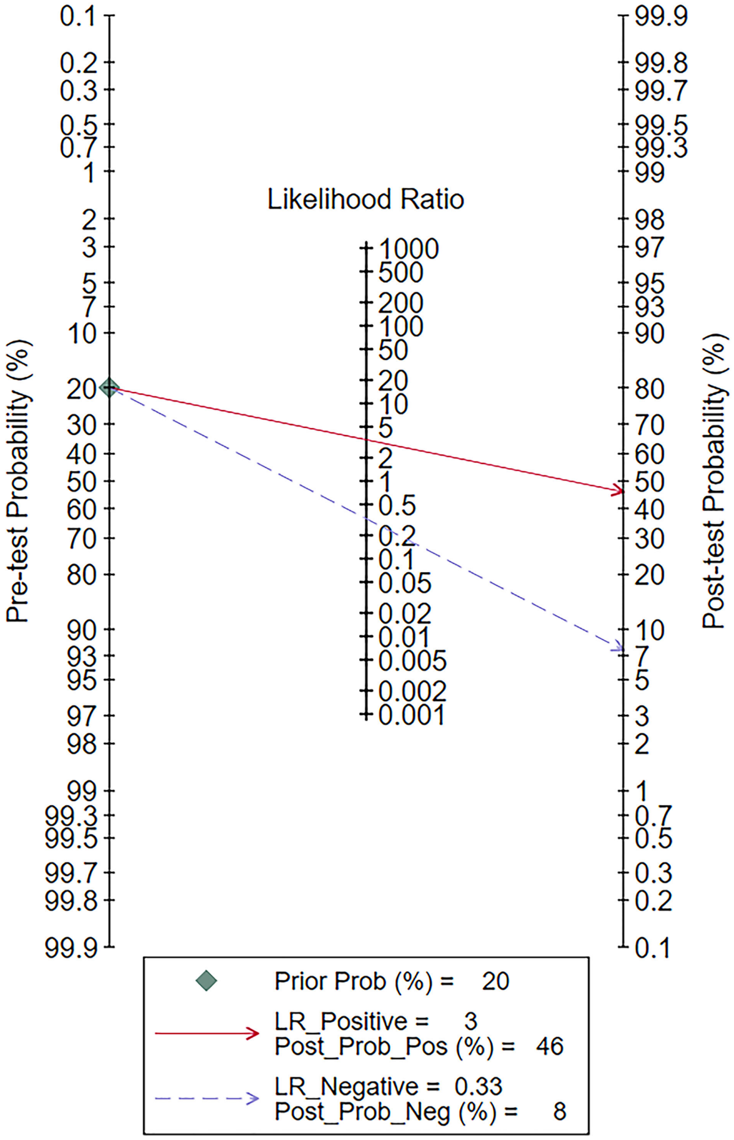


FIGURE 9
Fagan plot of elastosonography by patient analysis for the diagnosis of malignant salivary gland tumors.

malignant parotid lesions with a pooled sensitivity and specificity of 0.67 and 0.64, respectively. Compared with their study, our meta-analysis found that elastosonography had comparable sensitivity (0.73 vs. 0.67) and equal specificity (0.64 vs. 0.64), which confirmed the value of elastosonography in the diagnosis of salivary gland tumors. Furthermore, our study included patients not only with parotid lesions but also with submandibular or sublingual lesions, while Zhang et al.'s study only included patients with parotid lesions. In addition, our meta-analysis enrolled more eligible studies (16 vs. 10 papers). Therefore, we believed that the conclusion of our study might be more generalized.

Another prior meta-analysis by Li et al. (40) included nine eligible articles with 581 tumors using real-time elastography to differentiate benign and malignant salivary gland tumors, and demonstrated moderate diagnostic performance that the pooled sensitivity, specificity, and AUC were 0.76, 0.73, and 0.81, respectively. All the eligible studies adopted strain elastography technology, the traditional form of elastography, which depends on the sonographer's experience and external manual pressure and is a non-quantitative technology, to assess the stiffness of tumors. In contrast, the included studies in our meta-analysis used not only strain elastography but also shear wave elastography (31, 36), which allows an objective and quantitative assessment of the tumor

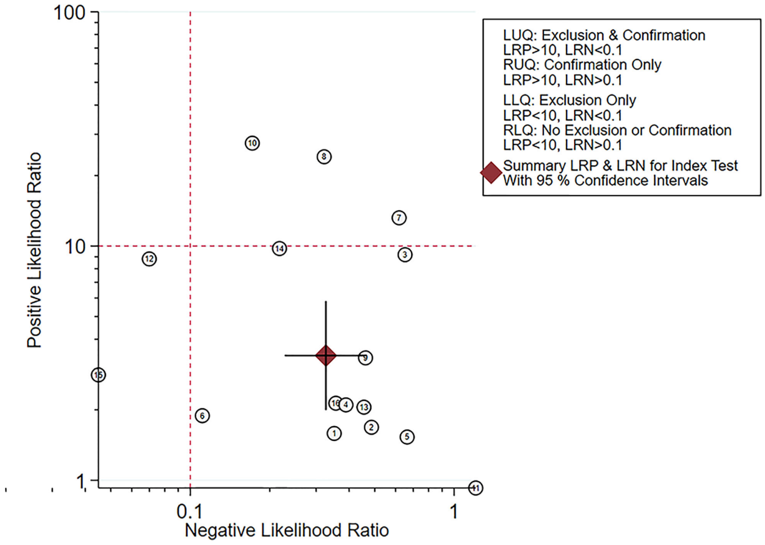


FIGURE 10
Likelihood matrix indicated that summary positive likelihood ratio and negative likelihood ratio for elastosonography in the diagnosis of malignant salivary gland tumors with 95% confidence intervals were concentrated on the right lower quadrant.

stiffness (41). The pooled sensitivity and specificity of our meta-analysis were lower compared with Li et al.'s study, whereas our meta-analysis enrolled more recently published articles that not only enhanced the statistical power of this study but also further supported clinical application of elastosonography for diagnosing malignant salivary tumors.

A study by Dumitriu et al. (28) indicated that the depth of the tumor might be a hindrance for elastosonography, which was particularly true for tumors located in the deep parotid lobe. Yerli and colleagues (29) revealed that the assessment of tumors located in the deep parotid lobe was a limitation of conventional ultrasound and was also a limitation of elastosonography. For tumors located in

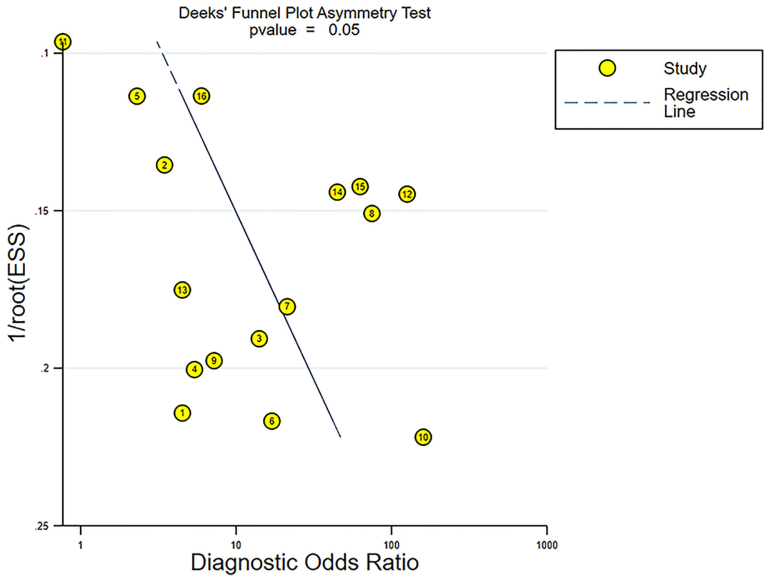


FIGURE 11
Funnel plot for evaluating potential publication bias.

the superficial parotid lobe but relatively deep, the mandibula can hinder the performance of optimal compression in the transverse plane. Furthermore, the mandibula can also affect the performance of optimal longitudinal compression of the submandibular gland. Matsuda et al. (38) found that the sensitivity for differentiating malignant tumors in the superficial parotid lobe was 100%, while the sensitivity was 20% for tumors in the deep lobe, which was attributed to the inability of attenuated acoustic pulses to reach the deep parotid lobe. Therefore, for certain anatomic structures, such as the mandible, the depth of the tumor location and tumors located in different salivary glands might have an effect on the results of elastosonography. However, we were not able to implement meaningful subgroups based on the factors mentioned above, as the data were not recorded in most of the studies.

Although malignant tumors are generally stiffer than benign lesions, a substantial overlap of elastic properties between malignant and benign salivary gland tumors was found in published papers (27, 28, 34). Pleomorphic adenoma, the most common benign salivary gland tumor, is a histologically diverse group of tumors (42), which results in the extremely wide range of elastographic values. In addition, some types of tumors, like Warthin tumors, have variable proportions of solid and cystic components, which would result in a considerable variance in stiffness. Moreover, some benign lesions, inflammatory diseases, as well as abscesses are considered as malignant tumors due to their appearance on elastosonography. It is still significantly difficult to discriminate between benign and malignant salivary gland tumors, and the diagnostic performance of elastosonography is unsatisfying (20, 29, 30). Therefore, other imaging methods complementing elastosonography, such as conventional ultrasound, magnetic resonance imaging, and computed tomography, are needed.

High heterogeneity among the included studies was a major problem in this meta-analysis. The Spearman correlation coefficient was 0.24 ($p = 0.37$), indicating that no threshold effect existed. Further meta-regression and subgroup analyses revealed that the assessment methods (quantitative or semiquantitative vs. qualitative) might play an important role in the heterogeneity. Quantitative or semiquantitative elastosonography, with higher pooled sensitivity (0.81 vs. 0.70), specificity (0.93 vs. 0.58), DOR (73.49 vs. 4.72), and AUC (0.96 vs. 0.74), performed better than the qualitative one, as shown in Table 3. The probable explanation was that compared with qualitative elastosonography, quantitative or semiquantitative elastosonography adopted an algorithm automatically calculated by an ultrasound equipment and was thus less operator-dependent and more objective. Although meta-regression and subgroup analyses excluded the influence of study design, year of publication, and blinding method, other factors such as ultrasound equipment, threshold values, index of elastography, and demographic characteristics would like to be taken into account. Due to the limited included studies, we were not able to perform meaningful subgroups on the basis of other factors mentioned above.

This meta-analysis has some limitations, which should be taken into account while interpreting the conclusions. First, a strict procedure was performed to review the articles and ultimately 16 eligible studies that fulfilled the inclusion criteria were enrolled. There are still relatively rare published studies exploring the value of elastosonography for diagnosis of salivary gland tumors, as the clinical application of elastosonography in the diagnosis of malignant salivary gland tumors was not reported until 2010 (27). Furthermore, only studies written in English were included in our meta-analysis, and then, language bias was inevitable. Second, the comparison between elastosonography and conventional ultrasound was performed indirectly. To determine which imaging modality is superior, a more rigorous research should be carried out adopting these two ultrasound technologies on the same cohort of patients. Finally, methodological limitations in the majority of the included studies were identified, especially in domains including patient selection, index test, reference standard, and flow and timing. Hence, more rigorous studies in the future are needed to address these methodological limitations.

## Conclusions

The existing evidence indicated that elastosonography showed a limited value for diagnosing malignant salivary gland tumors; it could be considered as a supplementary diagnostic technology to conventional ultrasound, and quantitative or semiquantitative elastosonography performed better than the qualitative one. However, large prospective multicenter studies are still needed to validate the conclusion and to further develop the clinical application of elastosonography in salivary gland tumors.

## 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

JW, ZW, and GJ designed this study. JW, ZZ, and YJ acquired and analyzed the data. JW, ZZ and XW wrote and edited the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

This project received support from Jinhua Science and Technology Bureau Scientific Research Project (2022-3-019).

## 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

## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.954751/full#supplementary-material

## References

1. Spiro RH. Salivary neoplasms: Overview of a 35-year experience with 2,807 patients. *Head Neck Surg* (1986) 8(3):177–84. doi: 10.1002/hed.2890080309

2. Pinkston JA, Cole P. Incidence rates of salivary gland tumors: Results from a population-based study. *Otolaryngol Head Neck Surg* (1999) 120(6):834–40. doi: 10.1016/S0194-5998(99)70323-2

3. Lewis AG, Tong T, Maghami E. Diagnosis and management of malignant salivary gland tumors of the parotid gland. *Otolaryngol Clin North Am* (2016) 49 (2):343–80. doi: 10.1016/j.otc.2015.11.001

4. Guzzo M, Locati LD, Prott FJ, Gatta G, McGurk M, Licitra L. Major and minor salivary gland tumors. *Crit Rev Oncol Hematol* (2010) 74(2):134–48. doi: 10.1016/j.critrevonc.2009.10.004

5. Andry G, Hamoir M, Locati LD, Licitra L, Langendijk JA. Management of salivary gland tumors. *Expert Rev Anticancer Ther* (2012) 12(9):1161–8. doi: 10.1586/era.12.92

6. Onkar PM, Ratnaparkhi C, Mitra K. High-frequency ultrasound in parotid gland disease. *Ultrasound Q* (2013) 29(4):313–21. doi: 10.1097/RUQ.0b013e3182a0abe0

7. Cantisani V, David E, Sidhu PS, Sacconi B, Greco A, Pandolfi F, et al. Parotid gland lesions: Multiparametric ultrasound and MRI features. *Ultraschall Med* (2016) 37(5):454–71. doi: 10.1055/s-0042-109171

8. Lee YY, Wong KT, King AD, Ahuja AT. Imaging of salivary gland tumours. *Eur J Radiol* (2008) 66(3):419–36. doi: 10.1016/j.ejrad.2008.01.027

9. Attyé A, Troprès I, Rouchy RC, Righini C, Espinoza S, Kastler A, et al. Diffusion MRI: Literature review in salivary gland tumors. *Oral Dis* (2017) 23 (5):572–5. doi: 10.1111/odi.12543

10. Coudert H, Mirafzal S, Dissard A, Boyer L, Montoriol PF. Multiparametric magnetic resonance imaging of parotid tumors: A systematic review. *Diagn Interv Imaging* (2021) 102(3):121–30. doi: 10.1016/j.diii.2020.08.002

11. Tan TJ, Tan TY. CT features of parotid gland oncocytomas: a study of 10 cases and literature review. *AJNR Am J Neuroradiol* (2010) 31(8):1413–7. doi: 10.3174/ajnr.A2090

12. Kim HJ, Kim JS. Ultrasound-guided core needle biopsy in salivary glands: A meta-analysis. *Laryngoscope* (2018) 128(1):118–25. doi: 10.1002/lary.26764

13. Sarvazyan A, Hall TJ, Urban MW, Fatemi M, Aglyamov SR, Garra BS. An overview of elastography - An emerging branch of medical imaging. *Curr Med Imaging Rev* (2011) 7(4):255–82. doi: 10.2174/157340511798038684

14. Magri F, Chytiris S, Chiovato L. The role of elastography in thyroid ultrasonography. *Curr Opin Endocrinol Diabetes Obes* (2016) 23(5):416–22. doi: 10.1097/MED.0000000000000274

15. Barr RG. The role of sonoelastography in breast lesions. *Semin Ultrasound CT MR* (2018) 39(1):98–105. doi: 10.1053/j.sult.2017.05.010

16. Xiao G, Zhu S, Xiao X, Yan L, Yang J, Wu G. Comparison of laboratory tests, ultrasound, or magnetic resonance elastography to detect fibrosis in patients with nonalcoholic fatty liver disease: A meta-analysis. *Hepatology* (2017) 66(5):1486–501. doi: 10.1002/hep.29302

17. Sigrist RMS, Liau J, Kaffas AE, Chammas MC, Willmann JK. Ultrasound elastography: Review of techniques and clinical applications. *Theranostics* (2017) 7 (5):1303–29. doi: 10.7150/thno.18650

18. Klintworth N, Mantsopoulos K, Zenk J, Psychogios G, Iro H, Bozzato A. Sonoelastography of parotid gland tumours: Initial experience and identification of characteristic patterns. *Eur Radiol* (2012) 22(5):947–56. doi: 10.1007/s00330-011-2344-7

19. Badea AF, Bran S, Tamas-Szora A, Floareş A, Badea R, Baciut G. Solid parotid tumors: an individual and integrative analysis of various ultrasonographic criteria. *A Prospect Observational Study Med Ultrason* (2013) 15(4):289–98. doi: 10.11152/mu.2013.2066.154.afb2

20. Mansour N, Bas M, Stock KF, Strassen U, Hofauer B, Knopf A. Multimodal ultrasonographic pathway of parotid gland lesions. *Ultraschall Med* (2017) 38 (2):166–73. doi: 10.1055/s-0035-1553267

21. Wierzbicka M, Kałużny J, Szczepanek-Parulska E, Stangierski A, Gurgul E, Kopeć T, et al. Is sonoelastography a helpful method for evaluation of parotid tumors? *Eur Arch Otorhinolaryngol* (2013) 270(7):2101–7. doi: 10.1007/s00405-012-2255-5

22. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PMPRISMA-DTA Group, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: The PRISMA-DTA statement. *JAMA* (2018) 319(4):388–96. doi: 10.1001/jama.2017.19163

23. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* (2011) 155(8):529–36. doi: 10.7326/0003-4819-155-8-201110180-00009

24. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* (1982) 143(1):29–36. doi: 10.1148/radiology.143.1.7063747

25. Ryan RCochrane Consumers and Communication Group. *Heterogeneity and subgroup analyses in cochrane consumers and communication group reviews: Planning the analysis at protocol stage* . Available at: http://cccrg.cochrane.org (Accessed June 15, 2018).

26. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* (2005) 58(9):882–93. doi: 10.1016/j.jclinepi.2005.01.016

27. Bhatia KS, Rasalkar DD, Lee YP, Wong KT, King AD, Yuen HY, et al. Evaluation of real-time qualitative sonoelastography of focal lesions in the parotid and submandibular glands: Applications and limitations. *Eur Radiol* (2010) 20 (8):1958–64. doi: 10.1007/s00330-010-1756-0

28. Dumitriu D, Dudea S, Botar-Jid C, Baciut M, Baciut G. Real-time sonoelastography of major salivary gland tumors. *AJR Am J Roentgenol* (2011) 197(5):W924–30. doi: 10.2214/AJR.11.6529

29. Yerli H, Eski E, Korucuk E, Kaskati T, Agildere AM. Sonoelastographic qualitative analysis for management of salivary gland masses. *J Ultrasound Med* (2012) 31(7):1083–9. doi: 10.7863/jum.2012.31.7.1083

30. Celebi I, Mahmutoglu AS. Early results of real-time qualitative sonoelastography in the evaluation of parotid gland masses: A study with histopathological correlation. *Acta Radiol* (2013) 54(1):35–41. doi: 10.1258/ar.2012.120405

31. Yu J, Du Y, Lu Y, Zhang W, Zhang D, Wang G, et al. Application of DTI and ARFI imaging in differential diagnosis of parotid tumours. *Dentomaxillofac Radiol* (2016) 45(6):20160100. doi: 10.1259/dmfr.20160100

32. Zhou H, Zhou XL, Xu HX, He YP, Bo XW, Li XL, et al. Initial experience with ultrasound elastography for diagnosis of major salivary gland lesions. *J Ultrasound Med* (2016) 35(12):2597–606. doi: 10.7863/ultra.15.11093

33. Cortcu S, Elmali M, Tanrivermis Sayit A, Terzi Y. The role of real-time sonoelastography in the differentiation of benign from malignant parotid gland tumors. *Ultrasound Q* (2018) 34(2):52–7. doi: 10.1097/RUQ.0000000000000323

34. Cantisani V, David E, De Virgilio A, Sidhu PS, Grazhdani H, Greco A, et al. Prospective evaluation of quasistatic ultrasound elastography (USE) compared with baseline US for parotid gland lesions: Preliminary results of elasticity contrast index (ECI) evaluation. *Med Ultrason* (2017) 19(1):32–8. doi: 10.11152/mu-923

35. Altinbas NK, Gundogdu Anamurluoglu E, Oz II, Yuce C, Yagci C, Ustuner E, et al. Real-time sonoelastography of parotid gland tumors. *J Ultrasound Med* (2017) 36(1):77–87. doi: 10.7863/ultra.16.02038

36. Liu G, Wu S, Liang X, Cui X, Zuo D. Shear wave elastography improves specificity of ultrasound for parotid nodules. *Ultrasound Q* (2018) 34(2):62–6. doi: 10.1097/RUQ.0000000000000354

37. Karaman CZ, Başak S, Polat YD, Ünsal A, Taşkın F, Kaya E, et al. The role of real-time elastography in the differential diagnosis of salivary gland tumors. *J Ultrasound Med* (2019) 38(7):1677–83. doi: 10.1002/jum.14851

38. Matsuda E, Fukuhara T, Donishi R, Taira K, Koyama S, Morisaki T, et al. Clinical utility of qualitative elastography using acoustic radiation force impulse for differentiating benign from malignant salivary gland tumors. *Ultrasound Med Biol* (2021) 47(2):279–87. doi: 10.1016/j.ultrasmedbio.2020.10.007

39. Zhang YF, Li H, Wang XM, Cai YF. Sonoelastography for differential diagnosis between malignant and benign parotid lesions: A meta-analysis. *Eur Radiol* (2019) 29(2):725–35. doi: 10.1007/s00330-018-5609-6

40. Li C, Zhang C, Li N, Li J. Compression real-time elastography for evaluation of salivary gland lesions: A meta-analysis. *J Ultrasound Med* (2016) 35(5):999–1007. doi: 10.7863/ultra.15.08043

41. Ferraioli G, Barr RG, Farrokh A, Radzina M, Cui XW, Dong Y, et al. How to perform shear wave elastography. *Part I Med Ultrason* (2022) 24(1):95–106. doi: 10.11152/mu-3217

42. Hernandez-Prera JC, Skálová A, Franchi A, Rinaldo A, Vander Poorten V, Zbären P, et al. Pleomorphic adenoma: The great mimicker of malignancy. *Histopathology* (2021) 79(3):279–90. doi: 10.1111/his.14322

frontiers | Frontiers in Oncology

TYPE Systematic Review
PUBLISHED 28 September 2022
DOI 10.3389/fonc.2022.944859

OPEN ACCESS

EDITED BY
Po-Hsiang Tsui,
Chang Gung University, Taiwan

REVIEWED BY
Lian-Ping He,
Taizhou University, China
Ahmad Nazlim Bin Yusoff,
Universiti Kebangsaan Malaysia,
Malaysia
Xiao-Yan Zhang,
Chinese Academy of Medical Sciences
and Peking Union Medical College
Hospital, China

*CORRESPONDENCE
Jun Li
1287424798@qq.com
Xin-Wu Cui
cuixinwu@live.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to
Cancer Imaging and
Image-directed Interventions,
a section of the journal
Frontiers in Oncology

RECEIVED 15 May 2022
ACCEPTED 19 August 2022
PUBLISHED 28 September 2022

CITATION
Zhu P-S, Zhang Y-R, Ren J-Y, Li Q-L, Chen M, Sang T, Li W-X, Li J and Cui X-W (2022) Ultrasound-based deep learning using the VGGNet model for the differentiation of benign and malignant thyroid nodules: A meta-analysis. *Front. Oncol.* 12:944859. doi: 10.3389/fonc.2022.944859

# Ultrasound-based deep learning using the VGGNet model for the differentiation of benign and malignant thyroid nodules: A meta-analysis

Pei-Shan Zhu[1†], Yu-Rui Zhang[1†], Jia-Yu Ren[2], Qiao-Li Li[1], Ming Chen[1], Tian Sang[1], Wen-Xiao Li[1], Jun Li[1,3*] and Xin-Wu Cui[2*]

[1]Department of Ultrasound, the First Affiliated Hospital of Medical College, Shihezi University, Shihezi, China, [2]Department of Medical Ultrasound, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [3]NHC Key Laboratory of Prevention and Treatment of Central Asia High Incidence Diseases, First Affiliated Hospital, School of Medicine, Shihezi University, Shihezi, China

**Objective:** The aim of this study was to evaluate the accuracy of deep learning using the convolutional neural network VGGNet model in distinguishing benign and malignant thyroid nodules based on ultrasound images.

**Methods:** Relevant studies were selected from PubMed, Embase, Cochrane Library, China National Knowledge Infrastructure (CNKI), and Wanfang databases, which used the deep learning-related convolutional neural network VGGNet model to classify benign and malignant thyroid nodules based on ultrasound images. Cytology and pathology were used as gold standards. Furthermore, reported eligibility and risk bias were assessed using the QUADAS-2 tool, and the diagnostic accuracy of deep learning VGGNet was analyzed with pooled sensitivity, pooled specificity, diagnostic odds ratio, and the area under the curve.

**Results:** A total of 11 studies were included in this meta-analysis. The overall estimates of sensitivity and specificity were 0.87 [95% CI (0.83, 0.91)] and 0.85 [95% CI (0.79, 0.90)], respectively. The diagnostic odds ratio was 38.79 [95% CI (22.49, 66.91)]. The area under the curve was 0.93 [95% CI (0.90, 0.95)]. No obvious publication bias was found.

**Conclusion:** Deep learning using the convolutional neural network VGGNet model based on ultrasound images performed good diagnostic efficacy in distinguishing benign and malignant thyroid nodules.

**Systematic Review Registration:** https://www.crd.york.ac.nk/prospero, identifier CRD42022336701.

KEYWORDS

meta-analysis, ultrasound, thyroid nodules, deep learning, VGGNet

## Introduction

Thyroid nodules are the most common diseases of the endocrine system, with an ultrasound population detection rate of about 65%, of which approximately 10% is thyroid cancer (1). Thyroid cancer, despite the low incidence, is one of the fastest growing of all cancer types, having increased approximately 2.4 times in the last 30 years (2). It has become a public health concern in most parts of the world. Therefore, early detection and early accurate diagnosis of benign and malignant thyroid nodules are crucial to develop treatment plans and predict prognosis for patients with thyroid nodules, yet this is a great challenge for radiologists and physicians.

Ultrasound is currently the first-line examination of choice for the clinical diagnosis of thyroid nodules, and it is not only the main method for cancer risk stratification of thyroid nodules, but also usually used for the guidance of fine-needle aspiration biopsy. However, the differential diagnosis of thyroid nodules by 2D ultrasound has certain limitations. The quality of ultrasound images is susceptible to many factors, such as the cooperation of patients, the performance of ultrasound machines, and the operating techniques of radiologists (3). In addition, ultrasound diagnostic results are affected by the experience level of radiologists, and the recognition of ultrasound image characteristics of nodules differs among radiologists with different working experience, which is subjective (4). Therefore, there is an urgent need to explore a diagnostic tool that is noninvasive, accurate, and objective in the differential diagnosis of the benign and malignant thyroid nodules preoperatively.

In 2013, deep learning of artificial intelligence (AI) was ranked as one of the top 10 breakthrough technologies by *MIT Technology Review*, ranking no. 1. From then on, deep learning entered an era of rapid development and played a pivotal role in the medical field, especially in medical image recognition. Some studies used the deep learning convolutional neural network to extract ultrasound features to identify and diagnose benign and malignant thyroid nodules, and some of the studies with diagnostic performance could be comparable to or better than the advanced physicians, which could reduce unnecessary punctures and overtreatment, and help grassroots and inexperienced physicians improve diagnostic efficiency and accuracy (5–7). In addition, Lee et al. (8) explored the use of deep learning convolutional neural networks in predicting the presence of lymph node metastasis in thyroid cancer on ultrasound, and their results indicated good predictive diagnostic accuracy (accuracy of 83.0%). Accordingly, ultrasound-based AI provides a new direction and method for radiologists to accurately and non-invasively identify and diagnose benign and malignant thyroid nodules and predict lymph node metastasis in the neck before surgery.

Previous published AI studies on thyroid disease can be broadly classified into two categories: traditional machine learning (ML) and deep learning (DL). Traditional ML uses manual extraction of image features, but ultrasound images are highly variable and feature extraction is dependent on physician experience; therefore, the accuracy of diagnosing benign and malignant thyroid nodules varies between empirical practitioners. Deep learning is a development of machine learning using automated extraction of image features, which is independent of physician experience (9). Among them, convolutional neural network (CNN) is a well-known deep learning structure in the field of medical image analysis and is a fully trainable deep learning algorithm consisting of an input layer, a hidden layer, and an output layer (10, 11). The hidden layer usually contains a convolutional layer, a pooling layer, and a fully connected layer. Compared with traditional machine learning methods, CNN performs better in target detection and image classification, and can better extract semantic features (12). Nowadays, CNN is considered one of the most advanced methods among many representative algorithms of deep learning, and VGGNet is a widely used model in CNN algorithms (10, 11). This model is the first network structure to reach "deep" in a real sense, as it takes a different research direction from previous CNN models, namely, deepening the network, and proves that the deep network with small filters is superior to the shallow network with large filters (13). Therefore, the deep learning VGGNet model alone was selected as the research subject to avoid selection bias and ensure the stability and reliability of the results.

At present, a number of studies have demonstrated that using the deep learning VGGNet model can differentiate benign and malignant thyroid nodules on ultrasound to assist physicians in making diagnostic results, but the sensitivity of different studies varies. The sensitivity was 93% in the study results of Zhu et al. (5), but only 77% in the study results of Zhou et al. (14). The sensitivity of ultrasound-based deep learning VGGNet in the diagnosis of thyroid nodules was quite different, and no meta-analysis of ultrasound-based deep learning VGGNet models for the determination of the nature of thyroid nodules has been found. Therefore, this meta-analysis aims to evaluate the efficacy of the ultrasound-based deep learning VGGNet model in distinguishing and diagnosing the nature of thyroid nodules to help radiologists make more accurate diagnoses.

## Materials and methods

### Search strategy

This meta-analysis was a study summarizing previously published literature on the differential diagnosis of thyroid nodules with an ultrasound-based deep learning convolutional neural network VGGNet model, thus requiring no ethical confirmation or patient consent. The literature was independently searched in PubMed, Embase, Cochrane Libraries, China National Knowledge Infrastructure (CNKI), and Wanfang databases up to September 2021, updated as of June 2022. The main following keywords were searched: "Deep learning" or "DL" or "Neural network" and "ultrasonography" OR "ultrasound" OR "ultrasonic" or "diagnostic imaging" and "thyroid" or "thyroid gland" or "thyroid nodules". Moreover, references of retrieved topic-related systematic reviews were also manually searched, and other relevant studies were read and identified to make the search more comprehensive.

## Study selection

Inclusion criteria were as follows (1): studies that used the deep learning VGGNet model for the differential diagnosis of benign and malignant thyroid nodules (2); at least one ultrasound imaging modality (3); literature that can provide true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN) (4); test set data or validation set data would be chosen; if both were present at the same time, we chose to use the test set; if there were both external and internal test sets, we also conducted a meta-analysis on the external test sets; if there were more than one external test set results in a paper at the same time, we would remove the highest and lowest diagnostic performance results and select the intermediate results; and (5) the gold standard was fine-needle aspiration (FNA), pathology, or both.

Excluded studies include (1) studies that did not match the gold standard (2); convolutional neural network models unrelated to the deep learning VGGNet model (3); studies that did not provide the necessary 2×2 contingency data (4); literature with only abstracts, reviews, conference report, papers not published in journals, full text that were not accessible online, and so on; and (5) duplicate studies.

## Quality assessment and data extraction

The Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool is a recognized tool for quality assessment of diagnostic accuracy tests, because of its specific problem definition and clinical actionability that is widely used in diagnostic meta-analyses (15, 16). The QUADAS tool was revised in 2011 and was called QUADAS-2, consisting of four main parts: case selection, index test, reference standard, and flow and timing, and all components are evaluated in terms of risk of bias (17). The 11 studies included were independently evaluated by two reviewers using the QUADAS-2 tool, and resolved by discussion between internal members if a disagreement was encountered during the assessment. QUADAS-2 results were output using RevMan 5.3, the dedicated software for the Cochrane Collaboration Network.

In this study, two authors independently read the titles and abstracts to screen eligible papers, and then read the full text to determine the included papers. The information obtained from each study was extracted independently, including first author, year of publication, country, gold standard, training set size, test set size, fourfold table data (TP, FP, FN, and TN), sensitivity, specificity, VGGNet type, and testing objects. If fourfold table data were not available in the literature, they were excluded.

## Statistical analysis

The entire data from the included studies were selected using Excel 2019, and sensitivity, specificity, and diagnostic odds ratio (DOR) [95% confidence interval (CI)] were summarized using STATA software version 16.0. The area under the receiver operating characteristic (ROC) curve and 95% CI were also calculated, and the value of the diagnostic test was assessed by the area under the curve (AUC) value, where AUC < 0.70 means low diagnostic accuracy, 0.70 < AUC < 0.90 indicates moderate diagnostic accuracy, and AUC ≥ 0.90 indicates high diagnostic accuracy. Statistical inconsistency between studies was assessed using the $I^2$ and Cochrane $Q$ tests; if $I^2 < 50\%$, it will choose a fixed-effects model to assess sensitivity and specificity, and if $I^2 > 50\%$, a random-effects model would be used. Meta-regression analysis was used and reasons were given when statistical heterogeneity was large. $p < 0.05$ was considered statistically significant.

# Results

## Literature searches

Through a comprehensive search, 2,495 records were obtained for our study as of September 2021, updated as of June 2022, including 544 papers from PubMed, 1,837 papers from Embase, 40 articles from Cochrane Libraries, 31 papers from CNKI, 37 papers from Wanfang database, and 6 papers from other sources. After preliminarily eliminating duplicate literatures, two researchers independently read the titles and abstracts of the remaining literatures, excluding literature reviews, cases, news, and other research types. The full text of the literature obtained will be further read through and eventually include 11 studies eligible for the meta-analysis. The detailed selection procedure is shown in Figure 1.

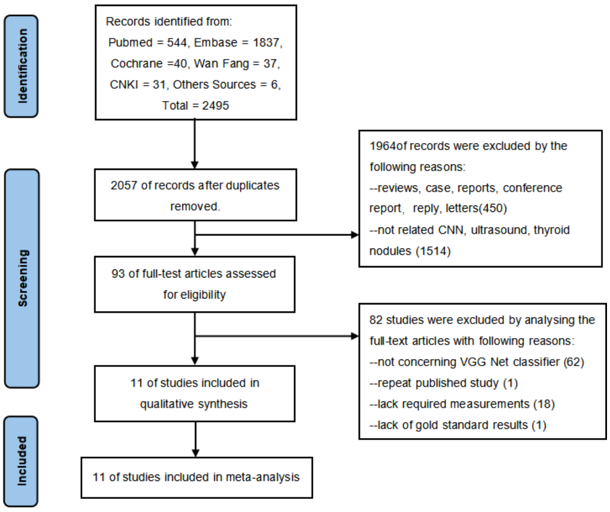


FIGURE 1
Study flowchart. Eleven studies were included in this meta-analysis.

## Study characteristics

We registered this meta-analysis on the PROSPERO website; the registration number is CRD42022336701. Following the PRISMA-Diagnostic studies selection process, we eventually included 11 papers; all studies are included in Table 1. The following are some basic characteristics of the included literature. All studies were published within the last 5 years. Eight papers used the deep learning VGG-16 model (14, 18–24). Four papers clearly indicated retrospective study (5, 6, 14, 19). Two papers did not give an explicit number of training sets (19, 20). Three papers compared the deep learning CNN algorithm with radiologists, and the results were comparable to or better than those of the advanced radiologists (5, 7, 23).

TABLE 1 Characteristics of the included studies.

| Author | Year | Country | Gold standard | Training database | | Test database | | Se (%) | Sp (%) | TP | FP | FN | TN | VGG | Testing objects |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | B | M | B | M | | | | | | | | |
| Kwon S.W et al. (18) | 2020 | Korea | FNA / pathology | 199 | 260 | 62 | 83 | 0.92 | 0.70 | 76 | 19 | 7 | 43 | 16 | Interior |
| Liu Z et al. (19) | 2021 | China | FNA | – | – | 67 | 96 | 0.79 | 0.87 | 76 | 9 | 20 | 58 | 16 | Interior |
| Wu K et al. (20) | 2020 | China | pathology | – | – | 520 | 636 | 0.86 | 0.78 | 547 | 114 | 89 | 406 | 16 | Interior |
| Qin P.L et al. (21) | 2019 | China | pathology | 424 | 484 | 115 | 133 | 0.93 | 0.98 | 123 | 2 | 10 | 113 | 16 | Interior |
| Zhu J.L et al. (7) | 2021 | China | pathology | 6760 | 9641 | 73 | 227 | 0.93 | 0.85 | 212 | 11 | 16 | 62 | 19 | Interior |
| | | | | 6760 | 9641 | 502 | 530 | 0.95 | 0.90 | 503 | 50 | 27 | 452 | 19 | Exterior |
| Zhou H et al. (14) | 2020 | China | FNA / pathology | 719 | 448 | 359 | 224 | 0.84 | 0.88 | 172 | 72 | 52 | 287 | 16 | Interior |
| | | | | 719 | 448 | 802 | 161 | 0.9 | 0.9 | 155 | 80 | 6 | 722 | 16 | Exterior |
| Liang et al. (22) | 2021 | China | pathology | 545 | 530 | 136 | 133 | 0.86 | 0.98 | 114 | 1 | 19 | 133 | 16 | Interior |
| Zhu Y.C et al. (5) | 2020 | China | pathology | 421 | 298 | 57 | 45 | 0.84 | 0.88 | 38 | 7 | 7 | 50 | 19 | Interior |
| Zhu Y.C et al. (23) | 2021 | China | pathology | 300 | 300 | 100 | 100 | 0.85 | 0.79 | 85 | 21 | 15 | 79 | 16 | Exterior |
| Chan W.K et al. (6) | 2021 | China | pathology | 4044 | 3316 | 264 | 204 | 0.81 | 0.8 | 100 | 14 | 24 | 56 | 19 | Interior |
| Kim Y.J et al. (24) | 2022 | Korea | FNA | 9772 | 2555 | 310 | 122 | 0.92 | 0.73 | 122 | 84 | 10 | 226 | 16 | Interior |
| | | | | | | | | 0.87 | 0.68 | 106 | 99 | 16 | 211 | 19 | Interior |
| | | | | 9772 | 2555 | 34 | 25 | 0.79 | 0.77 | 20 | 8 | 5 | 26 | 16 | Exterior |
| | | | | | | | | 0.75 | 0.81 | 19 | 6 | 6 | 28 | 19 | Exterior |

Se, sensitivity; Sp, specificity; M, Malignant; B, Benign; TP, true positives; FP, false positives; FN, false negatives; TN, true negatives; FNA, fine needle aspiration.

Qin et al. (21) extracted both ultrasound image features and ultrasound elastic image features. Zhu et al. (5) only included thyroid nodules in female patients.

## Methodology quality assessment

The results of evaluating the papers' quality assessed by the QUADAS-2 are shown in Figure 2. Most of the studies themselves were of high quality, but a few studies had potential risk of bias in flow and timing. In general, the included studies were considered as eligible.

## Accuracy of the ultrasound-based deep learning VGGNet model in the differential diagnosis of benign and malignant thyroid nodules

The comprehensive Pooled Sensitivity (PSEN) and Pooled Specificity (PESP) of the ultrasound-based deep learning VGGNet model for the differential diagnosis of benign and malignant thyroid nodules were 0.87 [95% CI (0.83, 0.91)] and 0.85 [95% CI (0.79, 0.90)], respectively (Figure 3). Higgins $I^2$ statistics showed significant heterogeneity in terms of sensitivity ($p < 0.05$, $I^2 = 91.09\%$) and specificity ($p < 0.05$, $I^2 = 92.12\%$); therefore, we selected the random-effects model to analyze the sensitivity and specificity; the DOR was 38.79 [95% CI (22.49, 66.91)] (Figure 4). The AUC was 0.93 [95% CI (0.90, 0.95)] (Figure 5). The result of Spearman correlation coefficient by Meta-DiSc version 1.4 ($r = -0.18$, $p = 0.50$) indicated that there was no significant threshold effect ($p > 0.05$), which also showed that other factors may lead to the generation of heterogeneity.

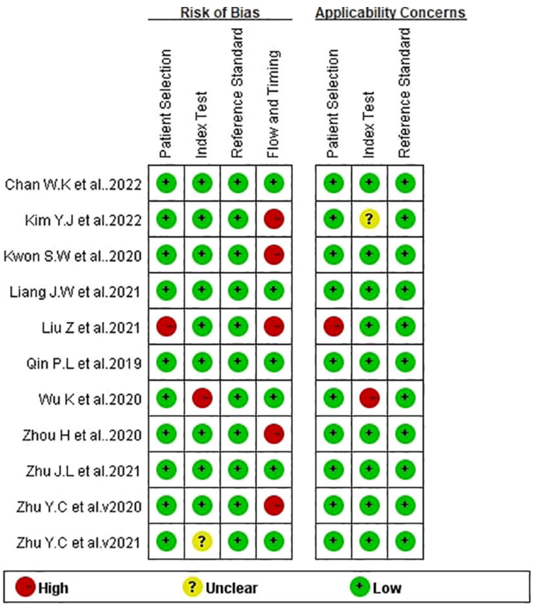


FIGURE 2
Bias risk of the included studies (QUADAS 2 criteria). The authors' assessment of each domain for each included study.

## Publication bias

Deek's funnel plot drawn by STATA16.0 showed no significant asymmetry, with a $p$-value of 0.84 ($p > 0.05$) (Figure 6), which indicated that there was no possibility of significant publication bias.

## Heterogeneity detection

Given the heterogeneity of the studies included in the pooled statistics, this research used regression analysis to analyze several clinically relevant survey variables. The result showed that year of study publication (≤2020 or >2020), number or scale of the region of interest (ROI) (single or multiple), and type of deep learning VGGNet (VGG-16 or VGG-19) were all associated with heterogeneity and were statistically significant for sensitivity ($p < 0.05$). Results of the meta-regression analysis are shown in Table 2. Among these covariates, the pooled sensitivity of studies published in 2020 and before was 0.89 [95% (0.84,0.95)] and 0.86 [95% (0.81, 0.91)] in papers published after 2020; the pooled specificity of papers published in 2020 and before was 0.86 [95% (0.79,0.94)] and 0.85 [95% (0.78,0.92)] in studies published after 2020, both being statistically significant ($p < 0.05$). The pooled sensitivity was 0.87 [95% (0.82, 0.91)] for a single ROI and 0.89 [95% (0.82, 0.96)] for multiple ROIs, the pooled specificity was 0.84 [95% (0.78, 0.90)] for a single ROI and 0.89 [95% (0.80, 0.97)] for multiple ROIs, and the pooled sensitivity difference was statistically significant ($p < 0.05$); the pooled specificity showed no significant differences ($p > 0.05$). The pooled sensitivity of VGG-16 was 0.88 [95% (0.83,0.93)] and VGG-19 was 0.87 [95% (0.80,0.93)], and the pooled specificity of VGG-16 was 0.86 [95% (0.80,0.93)] and VGG-19 was 0.84 [95%(0.75,0.93)], both of which were statistically significant ($p < 0.05$).

## Sensitivity analysis

To explore whether the studies affected the stability of PSEN and PSPE, this study used a method of eliminating the literature one by one, and the results of sensitivity and specificity analysis are shown in Table 3. The results demonstrated that with every single paper excluded, neither PSEN and PSPE nor Higgins $I^2$ had significant changes.

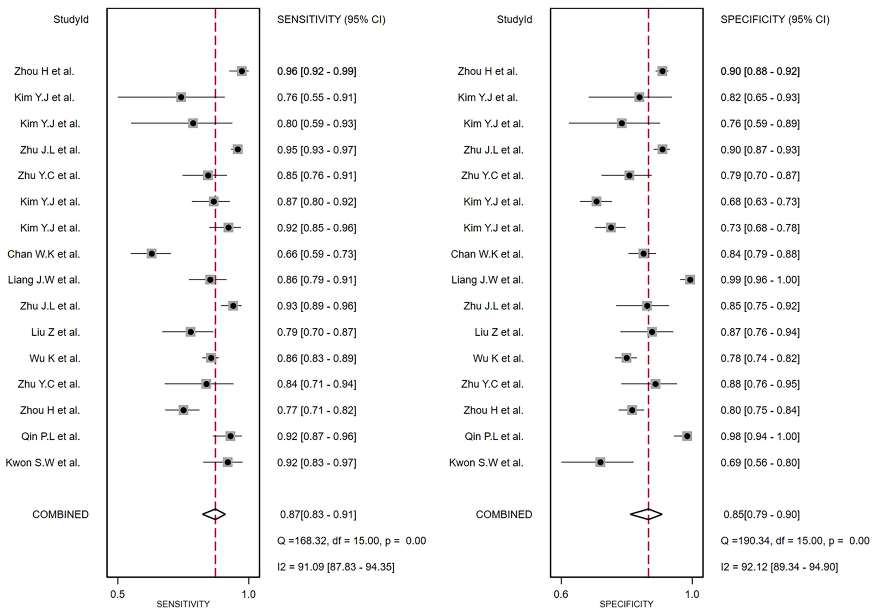


FIGURE 3
The forest plot of sensitivity and specificity for diagnosing thyroid nodules. Horizontal lines illustrate 95% confidence intervals of the individual studies.

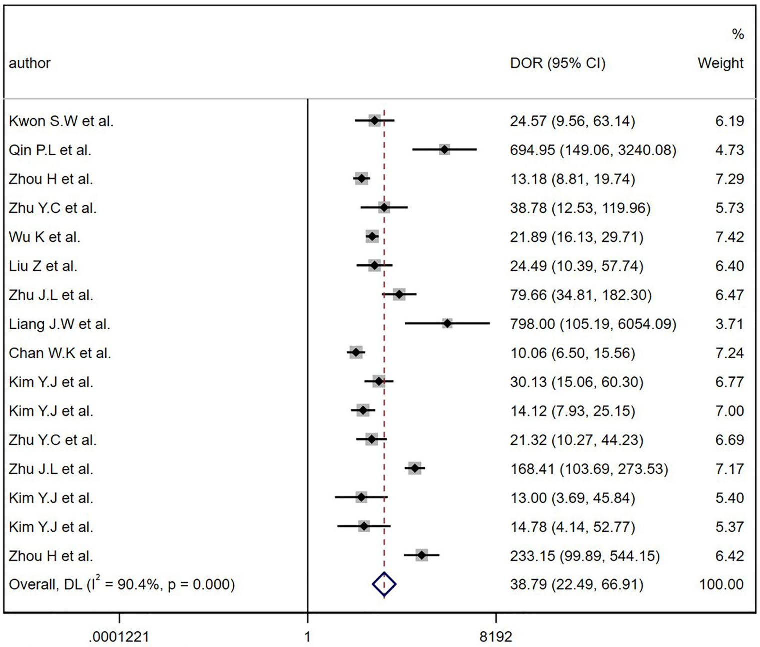


FIGURE 4
The diagnostic odds ratios (DOR) for diagnostic thyroid nodules. Horizontal lines illustrate 95% confidence intervals of the individual studies.

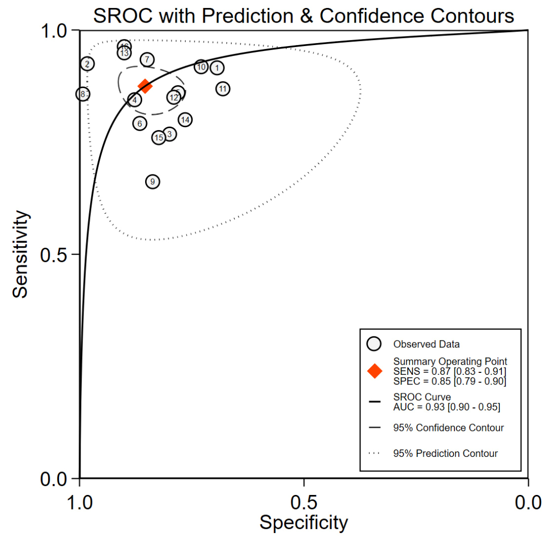


FIGURE 5
The receiver operating characteristic curve (ROC). SENS, sensitivity; SPEC, specificity; SROC, summary receiver operating characteristic curve; AUC, area under the SROC curve.

## Fagan plot analysis

The analysis of the Fagan plots showed that the ultrasound-based deep learning VGGNet model could provide some help for radiologists on the differential diagnosis of the nature of thyroid nodules (Figure 7). When the prior probability was 50%, the posterior probability of the deep learning VGGNet model correctly discriminating malignant nodules as "positive" was 86%, and the posterior probability dropped to 13% when it was "negative". When the prior probabilities were 25% and 75%, the post-test probabilities for positive were 67% and 86%, and the post-test probabilities for negative were 5% and 31%.

## Discussion

This meta-analysis evaluated the efficacy of the ultrasound-based deep learning VGGNet model in the differential diagnosis of benign and malignant thyroid nodules. The results showed that the deep learning VGGNet model achieved satisfactory results in discriminating benign and malignant thyroid nodules on ultrasound images; the pooled sensitivity and specificity were 0.87 [95% CI (0. 83, 0.91)] and 0.85 [95% CI (0.79, 0.90)], respectively, the DOR was 38.79 [95% CI (22.49, 66.91)], and the AUC was 0.93 [95% CI (0.90, 0.95)]. These results indicated that ultrasound-based deep learning VGGNet has high diagnostic accuracy for distinguishing the nature of thyroid nodules.

Traditional machine learning usually involves feature extraction and classification of ROI. Although the popularity of machine learning has gradually increased in recent years, ROI can only be manually selected and analyzed with machine learning using single-area information such as image texture, geometric shape, and statistical distribution (9). Ding et al. (25, 26) extracted statistical and textural features from thyroid elastograms, and then trained support vector machine (SVM)

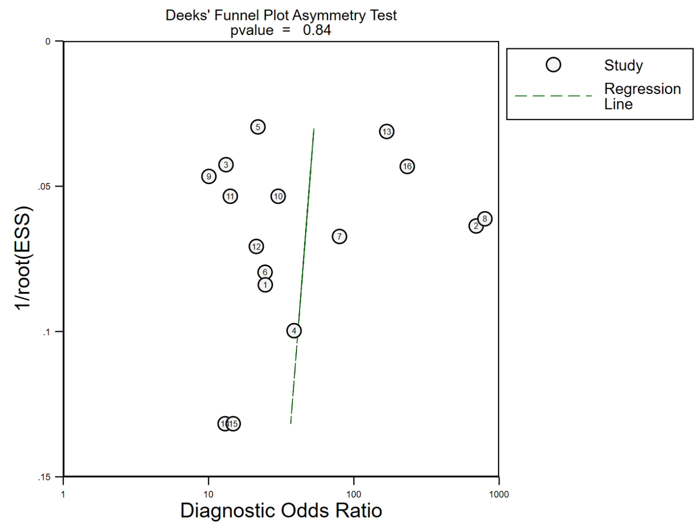


FIGURE 6
The publication bias of the included studies. No significant publication bias was found in the present meta-analysis. Each circle represented eligible research. ESS, effective sample size.

TABLE 2 Meta-regression of ultrasound-based deep learning for differentiating and diagnosing benign and malignant of thyroid nodules.

| Category | N | Se (95% CI) | $p$ | Sp (95%CI) | $p$ |
|---|---|---|---|---|---|
| Year | | | | | |
| ≤2020 | 6 | 0.89 (0.84, 0.95) | <0.05 | 0.86 (0.79, 0.94) | <0.05 |
| >2020 | 10 | 0.86 (0.81, 0.91) | | 0.85 (0.78, 0.92) | |
| ROI | | | | | |
| Single | 12 | 0.87 (0.82, 0.91) | <0.05 | 0.84 (0.78, 0.90) | 0.18 |
| Multiple | 4 | 0.89 (0.82, 0.96) | | 0.89 (0.80, 0.97) | |
| VGG | | | | | |
| VGG-16 | 10 | 0.88 (0.83, 0.93) | <0.05 | 0.86 (0.80, 0.93) | <0.05 |
| VGG-19 | 6 | 0.87 (0.80, 0.93) | | 0.84 (0.75, 0.93) | |

N, number of included studies; Se, sensitivity; Sp, specificity; CI, confidence interval; ROI, region of interest.

to detect malignancy of thyroid nodules with a maximum classification accuracy of 95.2%. However, the classification accuracy was affected by a hard threshold.

Compared with ML, deep learning can automatically extract the multi-level features of the ROI, and learn features from the nodule itself and the difference between the textures of different tissues, which greatly improves the image classification and detection performance (27). Buda et al. (28) used CNN for feature extraction and nodule classification of thyroid nodules, and also compared the diagnosis results with those of nine radiologists; the average sensitivity and average specificity of deep learning for diagnosis were higher than those of the nine radiologists, indicating that deep learning has a good clinical diagnostic value. Vasile et al. (29) used the fusion method of CNN-VGG for thyroid disease feature extraction and image classification, with an overall accuracy of 97.35%, showing that the integrative method is an excellent and stable classifier.

Previously, some meta-analyses were published about cardiovascular disease (30), gastrointestinal disease (31), and colorectal polyposis disease (32), and their combined AUCs were equal to or greater than 0.9, showing the excellent performance of CNN in disease diagnosis. Obviously, meta-analyses of thyroid nodules in ultrasound-based artificial intelligence have been conducted. Zhao et al. (33) included only five studies in meta-analysis. Xu et al. (34) mainly evaluated the overall computer-aided systems (CAD) efficacy of VGGNet, S-Detect, AlexNet, Inception, and so on in meta-analysis. In addition, the number of studies that included various single models was small, and none of them yielded the diagnostic efficacy of single-class models. Through further retrieval and reading of papers, no meta-analysis using ultrasound-based deep learning VGGNet model to differentially diagnose benign and malignant thyroid nodules has been found so far. Therefore, the authors conducted such a study.

TABLE 3 The sensitivity analysis using the method of eliminating papers one by one.

| Delete papers | Se (95% CI) | $I^2$ (95% CI) | $p$ | Sp (95% CI) | $I^2$ (95% CI) | $p$ | AUC (95% CI) |
|---|---|---|---|---|---|---|---|
| Zhou H et al. (14) | 0.87 (0.82, 0.90) | 90.54 (88.89, 94.18) | 0.00 | 0.85 (0.78, 0.90) | 91.16 (87.82, 94.49) | 0.00 | 0.92 (0.90, 0.94) |
| Kim Y.J et al. (24) | 0.88 (0.84, 0.91) | 91.71 (88.64, 94.78) | 0.00 | 0.86 (0.79, 0.90) | 92.74 (90.15, 95.33) | 0.00 | 0.93 (0.91, 0.95) |
| Kin Y.J et al. (24) | 0.88 (0.83, 0.91) | 91.80 (88.78, 94.83) | 0.00 | 0.86 (0.79, 0.91) | 92.75 (90.17, 95.33) | 0.00 | 0.93 (0.91, 0.95) |
| Zhu J.L et al. (7) | 0.87 (0.82, 0.90) | 88.89 (84.40, 93.37) | 0.00 | 0.85 (0.78, 0.90) | 91.68 (88.60, 94.76) | 0.00 | 0.92 (0.90, 0.94) |
| Zhu Y.C et al. (23) | 0.88 (0.83, 0.91) | 91.75 (88.70, 94.80) | 0.00 | 0.86 (0.79, 0.91) | 92.70 (90.09, 95.30) | 0.00 | 0.93 (0.91, 0.95) |
| Kim Y.J et al. (24) | 0.87 (0.83, 0.91) | 91.74 (88.68, 94.79) | 0.00 | 0.86 (0.80, 0.91) | 90.04 (86.15, 93.93) | 0.00 | 0.93 (0.91, 0.95) |
| Kim Y.J et al. (24) | 0.87 (0.82, 0.91) | 91.48 (88.29, 94.66) | 0.00 | 0.86 (0.80, 0.91) | 91.81 (88.79, 94.83) | 0.00 | 0.93 (0.90, 0.95) |
| Chan W.K et al. (6) | 0.89 (0.85, 0.91) | 84.91 (78.27, 91.55) | 0.00 | 0.86 (0.79, 0.90) | 92.25 (89.93, 95.20) | 0.00 | 0.93 (0.91, 0.95) |
| Liang J.W et al. (22) | 0.88 (0.83, 0.91) | 91.58 (88.45, 94.71) | 0.00 | 0.83 (0.78, 0.87) | 91.04 (87.65, 94.44) | 0.00 | 0.92 (0.89, 0.94) |
| Zhu J.L et al. (7) | 0.87 (0.82, 0.90) | 90.89 (87.42, 94.36) | 0.00 | 0.85 (0.79, 0.95) | 92.50 (89.80, 95.19) | 0.00 | 0.93 (0.90, 0.95) |
| Liu Z et al. (19) | 0.88 (0.84, 0.91) | 91.62 (88.51, 94.73) | 0.00 | 0.85 (0.79, 0.90) | 92.67 (90.05, 95.28) | 0.00 | 0.93 (0.91, 0.95) |
| Wu K et al. (20) | 0.88 (0.83, 0.91) | 91.49 (88.31, 94.66) | 0.00 | 0.86 (0.79, 0.91) | 92.40 (89.66, 95.14) | 0.00 | 0.93 (0.91, 0.95) |
| Zhu Y.C et al. (5) | 0.88 (0.83, 0.91) | 91.66 (88.57, 94.76) | 0.00 | 0.85 (0.79, 0.90) | 92.56 (89.89, 95.23) | 0.00 | 0.93 (0.90, 0.95) |
| Zhou H et al. (14) | 0.88 (0.84, 0.91) | 90.89 (87.42, 94.36) | 0.00 | 0.86 (0.79, 0.91) | 92.52 (89.84, 95.21) | 0.00 | 0.93 (0.91, 0.95) |
| Qin P.L et al. (21) | 0.89 (0.84, 0.92) | 91.12 (87.76, 94.47) | 0.00 | 0.87 (0.81, 0.92) | 91.35 (88.10, 94.59) | 0.00 | 0.94 (0.92, 0.96) |
| Kwon S.W et al. (18) | 0.87 (0.83, 0.91) | 91.58 (88.45, 94.71) | 0.00 | 0.84 (0.80, 0.91) | 92.47 (89.75, 95.18) | 0.00 | 0.93 (0.91, 0.95) |

Se, sensitivity; Sp, specificity; CI, confidence interval; AUC, area under the curve.

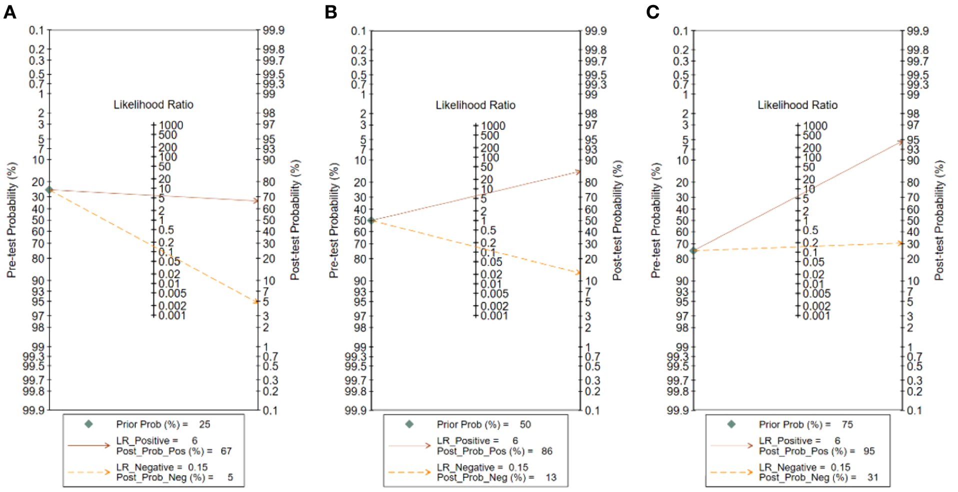


FIGURE 7
Fagan plot analysis for VGGNet model in detecting thyroid nodules: **(A)** Pre-test probability at 25%. **(B)** Pre-test probability at 50%. **(C)** Pre-test probability at 75%. The Fagan plot is composed of the left vertical axis representing the pre-test probability, the middle vertical axis representing the likelihood ratio, and the right vertical axis representing the post-test probability.

All the included studies reported good quality, indicating that most of the included studies were of high quality and did not show significant publication bias. However, a few numbers of included studies did not inform about the continuity and randomization of case selection and the incompleteness of the implementation of the gold standard, resulting in a small number of studies with slightly poorer quality reports, which may lead to implementation bias and measurement bias, resulting in high heterogeneity. Therefore, this study chose meta-regression to explain this high degree of heterogeneity. From the results, we can see that the year of study publication, number or scale of ROI, and type of deep learning VGGNet model may be important reasons for this heterogeneity. The reasons for heterogeneity are analyzed separately in detail below.

Firstly, there were 6 sets of data from five papers published in 2020 and before (5, 14, 18, 20, 21) and 10 sets of data from six papers published after 2020 (6, 7, 19, 22–24); sensitivity and specificity were statistically significant ($p$ < 0.05). The papers published after 2020 had a lower sensitivity than those published in 2020 and before (0.86 *vs.* 0.89). The reason may be that some papers published after 2020 included malignant images of thyroid nodules of different pathological types (6, 19, 23), which reduced the sensitivity of papers published after 2020. In addition, the total number of benign nodules included in papers after 2020 was less than that in 2020 and before, which reduced the specificity.

Secondly, it is easy for the ROI depicted on a single scale to ignore the rich details of ultrasound images of thyroid nodules (35). Therefore, different numbers or scales of ROIs were an important factor affecting study heterogeneity. Among the included studies, the number or scale of different studies in dividing the ROI was varied, 12 sets of data from eight papers delineated one ROI (5–7, 19–21, 23, 24), and 4 sets of data from three papers delineated two or more ROIs at different scales (14, 18, 22); sensitivity was statistically significant ($p$ < 0.05). Among them, Zhou et al. (14) delineated three target regions of thyroid nodules based on average size, which were located roughly inside, around, and outside the thyroid nodule, and all three ROIs contained the nodule, which showed an AUC comparison of classification accuracy between one ROI and three ROIs (0.82 *vs.* 0.87) indicating that the classification accuracy using three ROIs was more accurate. Therefore, it is reasonable to believe that the number or scale of ROIs had an impact on the identification results of thyroid nodules.

Finally, the all included studies used the deep learning VGGNet model. The 10 sets of data from eight papers used the deep learning VGG-16 models (14, 18–24), and 6 sets of data from four papers used the deep learning VGG-19 models (5–7, 24); the paper of Kim et al. (24) had both VGG-16 and VGG-19. Our results suggested that the diagnostic sensitivity and specificity of the VGG-16 model were higher than that of the VGG-19 (0.90 *vs.* 0.79, 0.87 *vs.* 0.83); sensitivity and specificity were $p < 0.05$. *A* study had similar results, Kim et al. (24) used the VGG model to classify benign and malignant thyroid nodules on ultrasound images and compared the diagnostic accuracy of the VGG-16 model with the VGG-19 model. VGG-16 showed higher diagnostic accuracy than VGG-19 in both internal and external test sets.

Moreover, the performance of the DL model is closely connected with the number of training data, and the DL model performs better when the data of the training sample are sufficiently large (36). Based on an analysis of 11 included studies, 2 sets of data from two papers did not give an explicit number of training sets, and 14 sets of data from nine papers did give the number of training sets, but the amount of pre-training varied across studies and the amount of learning varied; thus, it is difficult to know the overfitting results of the model. In addition, some researchers have explored the use of autonomously VGGNet fine-tuned models. Currently, there is no mature deep learning CNN model that can directly differentially diagnose the nature of thyroid nodules on ultrasound, which may inevitably lead to the generation of high heterogeneity.

In addition, the Fagan plot explored the clinical utility of ultrasound-based deep learning VGGNet models. The results showed that the ultrasound-based deep learning VGGNet model had the potential to differentiate benign and malignant thyroid nodules. When a patient was considered to have a 50% chance of developing thyroid cancer after initial clinical assessment, the likelihood of developing thyroid cancer increases from 50% to 86% if the deep learning VGGNet model results appear positive. Therefore, this high probability was highly accurate. In contrast, if the deep learning VGGNet was negative, then patients had a 13% chance of thyroid cancer, which could help our radiologists to exclude thyroid cancer. In real-world clinical practice, a biopsy of masses with a predicted 25% probability of malignancy will be performed regardless of the outcome of deep learning VGGNet. Therefore, the Fagan plot showed that the deep learning VGGNet model can aid in radiologist diagnosis.

This diagnostic meta-analysis has several limitations. Firstly, studies from Europe and America were excluded because they did not meet the inclusion criteria of using the deep learning VGGNet model to differentiate benign from malignant thyroid nodules, which might cause geographic bias. Secondly, this study only included papers published in English and Chinese, which might cause an unavoidable language bias. Thirdly, this meta-analysis only included 11 papers, and the small sample size of the test set in some studies may affect the accuracy of the results of the meta-analysis. To further assess the differential diagnostic efficacy of deep learning VGGNet models, large-scale, prospective, multicenter studies in different regions are necessary.

## Conclusion

This meta-analysis suggests that the ultrasound-based deep learning VGGNet model is a suitable and effective method for radiologists to differentiate and diagnose benign and malignant thyroid nodules. However, given the limitations of the sample size and the varying quality of the studies themselves, additional prospective or multicenter studies are expected to follow for further evaluation to make up for the deficiency.

## 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

P-SZ, JL, and X-WC contributed to the conception and design of the study. P-SZ and Y-RZ searched and reviewed studies, extracted and analyzed the data, and wrote the first draft of the manuscript. Q-LL, J-YR and X-WC reviewed and edited the manuscript. MC, TS, W-XL and JL directed the project and contributed to discussion as well as reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

(No. 2020-PT330-003):Supported by Open Research Fund of NHC Key Laboratory of Prevention and Treatment of Central Asia High Incidence Diseases. Supported by the Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences. (No. 2019DB012): The Corps Science and Technology Key Project.

## 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

## References

1. Durante C, Grani G, Lamartina L, Filetti S, Mandel SJ, Cooper DS. The diagnosis and management of thyroid nodules: A review. *Jama* (2018) 319(9):914–24. doi: 10.1001/jama.2018.0898

2. Ma J, Wu F, Zhu J, Xu D, Kong D. A pre-trained convolutional neural network based method for thyroid nodule diagnosis. *Ultrasonics* (2017) 73:221–30. doi: 10.1016/j.ultras.2016.09.011

3. Liang XW, Cai YY, Yu JS, Liao JY, Chen ZY, Wang NN. Update on thyroid ultrasound: A narrative review from diagnostic criteria to artificial intelligence techniques. *Chin Med J* (2019) 132(16):1974–82. doi: 10.1097/CM9.0000000000000346

4. Li T, Jiang Z, Lu M, Zou S, Wu M, Wei T, et al. Computer-aided diagnosis system of thyroid nodules ultrasonography: Diagnostic performance difference between computer-aided diagnosis and 111 radiologists. *Med (Baltimore)* (2020) 99 (23):e20634. doi: 10.1097/md.0000000000020634

5. Zhu YC, AlZoubi A, Jassim S, Jiang Q, Zhang Y, Wang YB, et al. A generic deep learning framework to classify thyroid and breast lesions in ultrasound images. *Ultrasonics* (2021) 110:8. doi: 10.1016/j.ultras.2020.106300

6. Chan WK, Sun JH, Liou MJ, Li YR, Chou WY, Liu FH, et al. Using Deep Convolutional Neural Networks for Enhanced Ultrasonographic Image Diagnosis of Differentiated Thyroid Cancer. *Biomedicines* (2021) 9(12):1771. doi: 10.3390/biomedicines9121771

7. Zhu JL, Zhang S, Yu RG, Liu ZQ, Gao HY, Yue B, et al. An efficient deep convolutional neural network model for visual localization and automatic diagnosis of thyroid nodules on ultrasound images. *Quant Imaging Med Surg* (2021) 11(4):1368–80. doi: 10.21037/qims-20-538

8. Lee JH, Baek JH, Kim JH, Shim WH, Chung SR, Choi YJ, et al. Deep learning-based computer-aided diagnosis system for localization and diagnosis of metastatic lymph nodes on ultrasound: A pilot study. *Thyroid* (2018) 28(10):1332–8. doi: 10.1089/thy.2018.0082

9. Song WF, Li S, Liu J, Qin H, Zhang B, Zhang SY, et al. Multitask cascade convolution neural networks for automatic thyroid nodule detection and recognition. *IEEE J Biomed Health Inf* (2019) 23(3):1215–24. doi: 10.1109/jbhi.2018.2852718

10. Ko SY, Lee JH, Yoon JH, Na H, Hong E, Han K, et al. Deep convolutional neural network for the diagnosis of thyroid nodules on ultrasound. *Head Neck* (2019) 41(4):885–91. doi: 10.1002/hed.25415

11. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Med Image Anal* (2017) 42:60–88. doi: 10.1016/j.media.2017.07.005

12. Liu T, Guo Q, Lian C, Ren X, Shen D. Automated detection and classification of thyroid nodules in ultrasound images using clinical-knowledge-guided convolutional neural networks. *Med Image Anal* (2019) 58:101555–. doi: 10.1016/j.media.2019.101555

13. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* (2014). doi: 10.48550/arXiv.1409.1556

14. Zhou H, Wang K, Tian J. Online transfer learning for differential diagnosis of benign and malignant thyroid nodules with ultrasound images. *IEEE Trans Biomed Eng* (2020) 67(10):2773–80. doi: 10.1109/tbme.2020.2971065

15. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of quadas: A tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* (2003) 3:25. doi: 10.1186/1471-2288-3-25

16. Wade R, Corbett M, Eastwood A. Quality assessment of comparative diagnostic accuracy studies: Our experience using a modified version of the quadas-2 tool. *Res synthesis Methods* (2013) 4(3):280–6. doi: 10.1002/jrsm.1080

17. Whiting PF, Weswood ME, Rutjes AWS, Reitsma JB, Bossuyt PNM, Kleijnen J. Evaluation of quadas, a tool for the quality assessment of diagnostic accuracy studies. *BMC Med Res Method* (2006) 6:9. doi: 10.1186/1471-2288-6-9

18. Kwon SW, Choi IJ, Kang JY, Jang WI, Lee GH, Lee MC. Ultrasonographic thyroid nodule classification using a deep convolutional neural network with surgical pathology. *J Digital Imaging* (2020) 33(5):1202–8. doi: 10.1007/s10278-020-00362-w

19. Liu Z, Zhong SB, Liu Q, Xie CX, Dai YZ, Peng C, et al. Thyroid nodule recognition using a joint convolutional neural network with information fusion of ultrasound images and radiofrequency data. *Eur Radiol* (2021) 31(7):5001–11. doi: 10.1007/s00330-020-07585-z

20. Kuan W, Pinle Q, Rui C, Jianchao Z. Benign and malignant diagnosis of thyroid nodules based on different ultrasound imaging. *J Comput Appl* (2020) 40 (01):77–82. doi: 10.11772/j.issn.1001-9081.2019061113

21. Qin PL, Wu K, Hu YS, Zeng JC, Chai XF. Diagnosis of benign and malignant thyroid nodules using combined conventional ultrasound and ultrasound elasticity imaging. *IEEE J Biomed Health Inf* (2020) 24(4):1028–36. doi: 10.1109/jbhi.2019.2950994

22. Jiawei L, Taorong Q, Aiyun Z, Pan X, Xuemei X, Hao F. Ensemble of multiscale fine-tuning convolutional neural networks for recognition of benign and malignant thyroid nodules. *J Computer-Aided Design Comput Graphics* (2021) 33 (1):11. doi: 10.3724/SP.J.1089.2021.18254

23. Zhu YC, Jin PF, Bao J, Jiang Q, Wang X. Thyroid ultrasound image classification using a convolutional neural network. *Ann Transl Med* (2021) 9 (20):1526. doi: 10.21037/atm-21-4328

24. Kim YJ, Choi Y, Hur SJ, Park KS, Kim HJ, Seo M, et al. Deep convolutional neural network for classification of thyroid nodules on ultrasound: Comparison of the diagnostic performance with that of radiologists. *Eur J Radiol* (2022) 152:110335. doi: 10.1016/j.ejrad.2022.110335

25. Ma J, Wu F, Zhu J, Xu D, Kong D. A pre-trained convolutional neural network based method for thyroid nodule diagnosis. *Ultrasonics* (2017) 73:221–30. doi: 10.1016/j.ultras.2016.09.011

26. Ding J, Cheng H, Ning C, Huang J, Zhang Y. Quantitative measurement for thyroid cancer characterization based on elastography. *J ultrasound Med Off J Am Institute Ultrasound Med* (2011) 30(9):1259–66. doi: 10.7863/jum.2011.30.9.1259

27. Ma J, Wu F, Jiang T, Zhu J, Kong D. Cascade convolutional neural networks for automatic detection of thyroid nodules in ultrasound images. *Med Phys* (2017) 44(5):1678–91. doi: 10.1002/mp.12134

28. Buda M, Wildman-Tobriner B, Castor K, Hoang JK, Mazurowski MA. Deep learning-based segmentation of nodules in thyroid ultrasound: Improving performance by utilizing markers present in the images. *Ultrasound Med Biol* (2020) 46(2):415–21. doi: 10.1016/j.ultrasmedbio.2019.10.003

29. Vasile CM, Udristoiu AL, Ghenea AE, Popescu M, Gheonea C, Niculescu CE, et al. Intelligent diagnosis of thyroid ultrasound imaging using an ensemble of deep learning methods. *Med Lith* (2021) 57(4):14. doi: 10.3390/medicina57040395

30. Krittanawong C, Virk HUH, Bangalore S, Wang Z, Johnson KW, Pinotti R, et al. Machine learning prediction in cardiovascular diseases: A meta-analysis. *Sci Rep* (2020) 10(1):16057. doi: 10.1038/s41598-020-72685-1

31. Mohan BP, Khan SR, Mba LLK, Ponnada S, Chandan S, Ali T, et al. High pooled performance of convolutional neural networks in computer-aided diagnosis of gi ulcers and/or hemorrhage on wireless capsule endoscopy images: A systematic review and meta-analysis - sciencedirect. *Gastrointest Endoscopy* (2020) 93(2):356–64. doi: 10.1016/j.gie.2020.07.038

32. Xu Y, Ding W, Wang Y, Tan Y, Xu X. Comparison of diagnostic performance between convolutional neural networks and human endoscopists for diagnosis of colorectal polyp: A systematic review and meta-analysis. *PloS One* (2021) 16(2):e0246892. doi: 10.1371/journal.pone.0246892

33. Zhao WJ, Fu LR, Huang ZM, Zhu JQ, Ma BY. Effectiveness evaluation of computer-aided diagnosis system for the diagnosis of thyroid nodules on ultrasound: A systematic review and meta-analysis. *Medicine* (2019) 98(32): e16379. doi: 10.1097/MD.0000000000016379

34. Xu L, Gao J, Wang Q, Yin J, Yu P, Bai B, et al. Computer-aided diagnosis systems in diagnosing malignant thyroid nodules on ultrasonography: A systematic review and meta-analysis. *Eur Thyroid J* (2020) 9(4):186–93. doi: 10.1159/000504390

35. Daoud MI, Bdair TM, Mahasen AN, Rami A. A fusion-based approach for breast ultrasound image classification using multiple-roi texture and morphological analyses. *Comput Math Methods Med* (2016) 2016:6740956. doi: 10.1155/2016/6740956.

36. Yang H-W, Hsu H-C, Yang C-K, Tsai M-J, Kuo Y-F. Differentiating between morphologically similar species in genus cinnamomum (lauraceae) using deep convolutional neural networks. *Comput Electron Agric* (2019) 162:739–48. doi: 10.1016/j.compag.2019.05.003

frontiers | Frontiers in Oncology

TYPE Review
PUBLISHED 12 October 2022
DOI 10.3389/fonc.2022.900450

OPEN ACCESS

EDITED BY
Ryogo Minamimoto,
National Center For Global Health and Medicine, Japan

REVIEWED BY
Joseph Petrulli,
Invicro, United States
Fei Kang,
Fourth Military Medical University, China

*CORRESPONDENCE
Eveline Van De Stadt
e.vandestadt@amsterdamumc.nl

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 19 July 2022
ACCEPTED 28 September 2022
PUBLISHED 12 October 2022

CITATION
Van De Stadt E, Yaqub M, Jahangir AA, Hendrikse H and Bahce I (2022) Radiolabeled EGFR TKI as predictive imaging biomarkers in NSCLC patients – an overview. *Front. Oncol.* 12:900450.
doi: 10.3389/fonc.2022.900450

# Radiolabeled EGFR TKI as predictive imaging biomarkers in NSCLC patients – an overview

Eveline Van De Stadt[1,3]*, Maqsood Yaqub[2,3], A. A. Jahangir[1], Harry Hendrikse[2,3] and Idris Bahce[1,3]

[1]Department of Pulmonology, Amsterdam University Medical Centers (UMC), VU University Medical Center, Amsterdam, Netherlands, [2]Department of Radiology and Nuclear Medicine, Amsterdam University Medical Centers (UMC), VU University Medical Center, Amsterdam, Netherlands, [3]Cancer Center Amsterdam, Amsterdam University Medical Centers (UMC), Amsterdam, Netherlands

Non-small cell lung cancer (NSCLC) has one of the highest cancer-related mortality rates worldwide. In a subgroup of NSCLC, tumor growth is driven by epidermal growth factor receptors (EGFR) that harbor an activating mutation. These patients are best treated with EGFR tyrosine kinase inhibitors (EGFR TKI). Identifying the EGFR mutational status on a tumor biopsy or a liquid biopsy using tumor DNA sequencing techniques is the current approach to predict tumor response on EGFR TKI therapy. However, due to difficulty in reaching tumor sites, and varying inter- and intralesional tumor heterogeneity, biopsies are not always possible or representative of all tumor lesions, highlighting the need for alternative biomarkers that predict tumor response. Positron emission tomography (PET) studies using EGFR TKI-based tracers have shown that EGFR mutational status could be identified, and that tracer uptake could potentially be used as a biomarker for tumor response. However, despite their likely predictive and monitoring value, the EGFR TKI-PET biomarkers are not yet qualified to be used in the routine clinical practice. In this review, we will discuss the currently investigated EGFR-directed PET biomarkers, elaborate on the typical biomarker development process, and describe how the advances, challenges, and opportunities of EGFR PET biomarkers relate to this process on their way to qualification for routine clinical practice.

KEYWORDS
NSCLC, EGFR TKI, PET/CT, radiolabeled EGFR TKI, molecular imaging

## 1 Introduction

Lung cancer is one of the most prevalent cancer types worldwide (1). Lung cancer accounts for approximately 22% of all cancer-related mortality, emphasizing that lung cancer is not only a highly prevalent cancer type, but also one of the deadliest (1). For decades, the standard of care treatment for advanced stage non-small cell lung cancer

(NSCLC) was only chemotherapy (2–5). The introduction of tyrosine kinase inhibitors (TKI) directed against the epidermal growth factor receptor (EGFR), an oncogenic driver pathway promoting cell growth and division, led to a shift in the treatment paradigm of EGFR mutation positive NSCLC, and, ultimately to an acceleration of the development of targeted therapies against other oncogenic driver mutation targets (2–5). Wild type EGFR activation is ligand-dependent, i.e., the EGFR kinase function only activates if an EGF ligand is bound at the extracellular binding site of the receptor (6). However, with activating mutations in the EGFR kinase domain, activation occurs in the absence of a ligand, leading to tumor cell proliferation and growth (6). EGFR TKIs bind with high affinity at the kinase domain of the mutated EGFR and block its function (6, 7). As a result, patients harboring activating EGFR mutations achieve higher tumor responses on EGFR TKI than on conventional chemotherapy (2–4, 8).

The iPASS trial was the first trial that clearly showed the superior clinical efficacy of EGFR TKI as compared to conventional chemotherapy. In this study, Mok et al. demonstrated that the first-generation EGFR TKI gefitinib achieved a higher progression-free survival (PFS) in the intention-to-treat population (HR 0.74; 95%CI 0.65 to 0.85; P<0.001) (3). Many other first-line phase 3 clinical studies using the first-generation EGFR TKI gefitinib or erlotinib, showed comparable results (2, 4, 9, 10). In contrast to the first-generation TKIs, the second-generation TKIs afatinib and dacomitinib were characterized by an irreversible binding of the TKI to the EGFR kinase domain and by multi-kinase targeting (5, 10–15). These second-generation TKIs had possibly a superior efficacy as compared to first-generation TKI at the cost of slightly higher toxicities (10, 16). The third-generation TKI osimertinib was primarily designed to target the secondary resistance mutation T790M (17–21). In the AURA3 trial, patients with T790M secondary mutations, occurring as resistance mutations on an initial treatment with gefitinib or erlotinib, were randomized between osimertinib versus conventional chemotherapy (17). Osimertinib showed superior PFS (10.1m vs. 4.4m; HR 0.30; 95%CI 0.23 to 0.41; P<0.001). The objective response rate was also significantly better with osimertinib (71%; 95% CI, 65 to 76) than with chemotherapy (31%; 95% CI, 24 to 40) (OR 5.39; 95%CI 3.47 to 8.48; P<0.001) (17). Surprisingly, osimertinib also performed above expectations as a first-line therapy. In the FLAURA study, treatment-naïve EGFR mutation positive patients were randomized to osimertinib versus a first-generation EGFR TKI (22). Osimertinib showed superior PFS (18.9m vs. 10.2m; HR 0.46; 95%CI 0.37 to 0.57; P<0.001). In a recent update of the study results, osimertinib also showed OS superiority as compared to the first-generation TKI (38.6m vs. 31.8m; HR 0.80; 95%CI 0.64 to 1.00; P=0.046) (23). These developments illustrate that over the course of approximately a decade, significant advances have been made in the treatment of EGFR mutation positive NSCLC, and that the identification of these patients is of paramount importance.

Diagnosis through next-generation sequencing of tumor DNA, obtained through a histological biopsy, is the gold standard for identifying tumor EGFR mutations (24). Unfortunately, taking biopsies is invasive, at risk for complications and not always possible due to difficult to reach tumor sites. Also, biopsies may not always be representative for all the tumor lesions due to varying intra- and interlesional heterogeneity, this may especially be of importance when resistance occurs and mapping the residual sensitivity for TKI treatment is needed (24). To overcome these limitations new biomarkers have been investigated. Liquid biopsies are ever more used in situations when representative tumor biopsies cannot be obtained. Even though the current sensitivity of liquid biopsies is approximately 70% with specificities above 90%, not all patients can be diagnosed using liquid biopsies alone (25, 26). Also, liquid biopsies do not address the limitation of tumor heterogeneity. Alternatively, in recent years, imaging studies using radiolabeled EGFR TKI have shown that PET could potentially be of value for identifying EGFR mutation positive patients and predicting tumor sensitivity to EGFR TKI (27–31).

In this review, we will discuss the current EGFR-directed PET tracers that have been investigated in EGFR mutated NSCLC. The special focus will lie with radiolabeled EGFR TKI: inertly labeled EGFR TKI used as a PET tracer in NSCLC patients. In addition, we will discuss the framework of the PET biomarker development process, highlighting the different contexts of use to better elucidate the stage in which these EGFR TKI PET biomarkers are at. We will describe the challenges, but also the recent advances and opportunities that could help EGFR PET on its path to generating qualified predictive biomarkers for clinical use.

# 2 Current EGFR PET biomarkers

## 2.1 PET biomarker background

PET is a molecular imaging technique, widely in use in the staging and treatment monitoring schedules in cancer patients. A radioactively labeled compound used as a tracer, which is expected to accumulate at the site of a specific target in the tumor lesion, is injected into the body and its distribution is then imaged. When using a validated tracer, its accumulation in the tumor and other sites is expected to be sensitive and quantifiable. The tracer accumulation or the so-called tracer uptake can be measured using different metrics, which can serve as biomarkers.

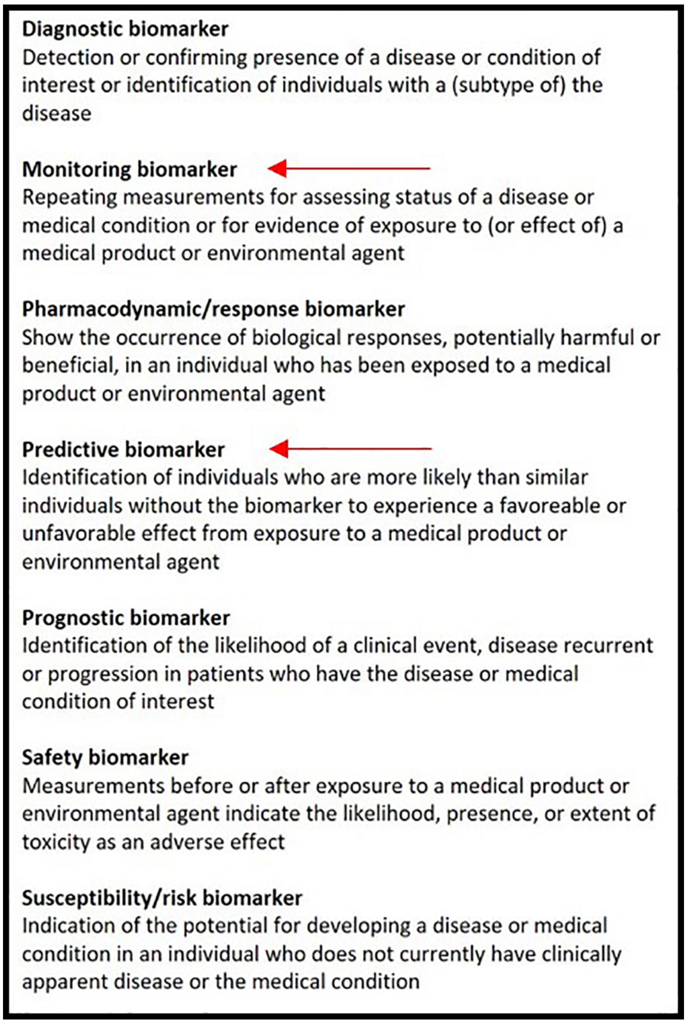
In general, a biomarker is a measurable indicator of a biological process and in case of PET imaging, this can be a measure derived from the tracer uptake in tumors or in healthy tissues, e.g., the Standardized Uptake Value (SUV) or the Distribution Volume ($V_T$).

Also, depending on their aims, biomarkers will have different 'contexts of use'. The evidence that is necessary to support qualification towards clinical practice is dependent on the specific context of use. The FDA Qualification Framework recommends categorizing biomarkers using the BEST biomarker categories according to their aims, as described in Figure 1 (32).

Considering EGFR, PET should provide a *predictive biomarker*, which is most relevant for the clinical practice. The presence of common EGFR mutations (i.e., exon19 deletions, exon21 L858R) are highly predictive for response to TKI therapy; however, in case of uncommon mutations, less is known regarding their clinical relevance and tumor TKI responses may vary greatly between different uncommon mutations. A predictive PET biomarker would therefore be most interesting.

EGFR directed PET biomarkers will *de facto* never be able to diagnose an activating EGFR mutation, as this requires tumor DNA sequencing on tumor tissue or liquid biopsies. Therefore, a PET imaging biomarker could never be a *diagnostic biomarker* that replaces DNA sequencing. On the other hand, PET imaging biomarkers could very well become qualified as predictive biomarkers to predict tumor sensitivity to EGFR TKI as mentioned before.

A *monitoring biomarker* is also of interest, as all tumors eventually develop resistance to EGFR TKI, in which case it could be of clinical importance to know whether lesions or parts of lesions remain TKI sensitive to decide whether TKI should be continued beyond progression.



**Diagnostic biomarker**
Detection or confirming presence of a disease or condition of interest or identification of individuals with a (subtype of) the disease

**Monitoring biomarker** ←
Repeating measurements for assessing status of a disease or medical condition or for evidence of exposure to (or effect of) a medical product or environmental agent

**Pharmacodynamic/response biomarker**
Show the occurrence of biological responses, potentially harmful or beneficial, in an individual who has been exposed to a medical product or environmental agent

**Predictive biomarker** ←
Identification of individuals who are more likely than similar individuals without the biomarker to experience a favoreable or unfavorable effect from exposure to a medical product or environmental agent

**Prognostic biomarker**
Identification of the likelihood of a clinical event, disease recurrent or progression in patients who have the disease or medical condition of interest

**Safety biomarker**
Measurements before or after exposure to a medical product or environmental agent indicate the likelihood, presence, or extent of toxicity as an adverse effect

**Susceptibility/risk biomarker**
Indication of the potential for developing a disease or medical condition in an individual who does not currently have clinically apparent disease or the medical condition

FIGURE 1
The biomarker classification according to the BEST biomarker categories. The red arrows indicate in which category EGFR-directed PET tracers could be included.

The current PET biomarkers can be categorized into 2 categories, i.e., those based on non-EGFR-directed tracers and those that are derived from EGFR TKI-based tracers.

## 2.2 Non-EGFR PET biomarkers in EGFR mutated NSCLC

The most widely-used tracer is $^{18}$F-fluorodeoxyglucose ($^{18}$F-FDG), a radioactive analogue to glucose, that can quantify metabolic activity. In the past decade, many clinical studies attempted to establish the role of $^{18}$F-FDG in evaluating the EGFR mutational status (33). A meta-analysis by Du et al. looked at studies that compared the lesional maximum of standardized uptake value ($SUV_{max}$) of $^{18}$F-FDG uptake between wild-type and mutant EGFR and evaluated its value for predicting the EGFR status in NSCLC patients (33). In 15 studies (3574 patients), the pooled sensitivity and specificity was found to be low. The authors concluded that $^{18}$F-FDG based $SUV_{max}$ should be used with caution when predicting EGFR mutations in NSCLC (33). However, new studies are exploring the potential outcome of radiomics and artificial intelligence (AI) algorithms as biomarkers to assess the predictive capacity of $^{18}$F-FDG PET. For example, Yin et al. demonstrated in a training data set of 198 NSCLC patients with a testing data set of 103 patients that their algorithm could predict EGFR mutations automatically with a ROC-AUC of 0.84 (95% CI, 0.75–0.90) (34). These developments may indicate an increasing role for radiomics and AI as new $^{18}$F-FDG based biomarkers in the future, albeit, these algorithms need optimization and validation using larger cohorts.

In recent years, 3-deoxy-3-$^{18}$F-fluorothymidine ($^{18}$F-FLT) PET scans have generated interest in oncology. As opposed to $^{18}$F-FDG, $^{18}$F-FLT PET reflects cell proliferation (10, 35). This tracer is trapped intracellularly in the S-phase of the cell cycle (35). Elevated $^{18}$F-FLT uptake of lesions could therefore be indicative of tumor cell proliferation and treatment-resistance. This supports the notion that $^{18}$F-FLT could serve to generate treatment monitoring biomarkers. Indeed, studies using $^{18}$F-FLT in EGFR mutation positive NSCLC have shown that a decrease of $^{18}$F-FLT uptake in tumor lesions is associated with response to EGFR TKI treatment (10, 36, 37). As $^{18}$F-FLT is nonspecific to EGFR mutations, the validation of $^{18}$F-FLT-based monitoring biomarkers could be of interest for many cancer types as well.

Other non-EGFR PET tracers that have been investigated in EGFR mutation positive NSCLC patients, are $^{11}$C-choline and *O*-(2-[$^{18}$F]fluoroethyl)-L-tyrosine($^{18}$F-FET). $^{11}$C-choline, a tracer mainly used in diagnostics of prostate cancer, is a component of phospholipids in the cell membrane (38). Phosphorylation of choline is upregulated in cancers through choline-kinase (38). Although $^{11}$C-choline PET is used in the routine practice in other cancer types, results in NSCLC are discouraging (39–41). $^{18}$F-FET has been used in diagnostics of brain tumors, including brain metastases of NSCLC, however, no studies were published on $^{18}$F-FET in extracranial NSCLC tumors (42, 43).

## 2.3 EGFR PET biomarkers

### 2.3.1 Characteristics of EGFR PET tracers

For radiolabeling target-specific drugs such as EGFR TKI, the characteristics of the radionuclide that is used for labeling needs to be aligned with the pharmacokinetic properties of the parent compound. For example, using radionuclides with long-lived isotopes such as zirconium-89 ($t_{1/2}$ 78 hours) are best suited to label large molecules with slow pharmacokinetics like monoclonal antibodies, e.g., $^{89}$Zr-cetuximab, however, inappropriate for labeling EGFR TKI. Since EGFR TKI are small molecules with relatively fast pharmacokinetics, i.e., fast target binding and rapid clearance from the circulation, using short-lived isotopes such as carbon-11 (t½ 20 min) or fluorine-18 (t½ 110 min) is more appropriate.

Also, instead of adding the radionuclide on the parent compound, substituting an existing carbon or fluorine atom of the TKI molecule will maintain the original pharmacokinetic (PK) behavior of the TKI resulting in a tracer that is equally specific as the original TKI. The choice whether carbon-11 or fluorine-18 is used for this inert substitution is based on the chemical structure of the parent compound (27, 31, 44).

Although tracers based on EGFR TKI that are in clinical use, when labeled inertly, provide the best PK behavior metrics to investigate tumor sensitivity to the respective TKI, the development of such tracers is inherently delayed, as clinical safety and efficacy data of the parent TKI need to be established. Moreover, the fast development of subsequent generations of TKI could disrupt the development of early generation TKI tracers and make them redundant. To illustrate this, a timeline indicating the approval of the 3 generations of EGFR TKI used in the clinical and their tracer counterparts is shown in Figure 2.

Clinical PET studies are not only being performed using EGFR PET tracers based on EGFR TKI, but also on tracers without treatment analogue. Many of these tracers without direct treatment analogue have been specifically developed for the purpose of imaging. These tracers, e.g., $^{18}$F-MPG, $^{11}$C-PD153035 and $^{18}$F-IRS, show significant differences amongst themselves in kinetic characteristics, mainly in the binding affinity to the kinase domain (45–47).

### 2.3.2 Present EGFR TKI PET biomarkers

An overview of published clinical studies using EGFR PET tracers is given in Table 1. For $^{11}$C-erlotinib and $^{18}$F-afatinib, studies have shown that EGFR mutation positive patients can be identified and that tumor response to

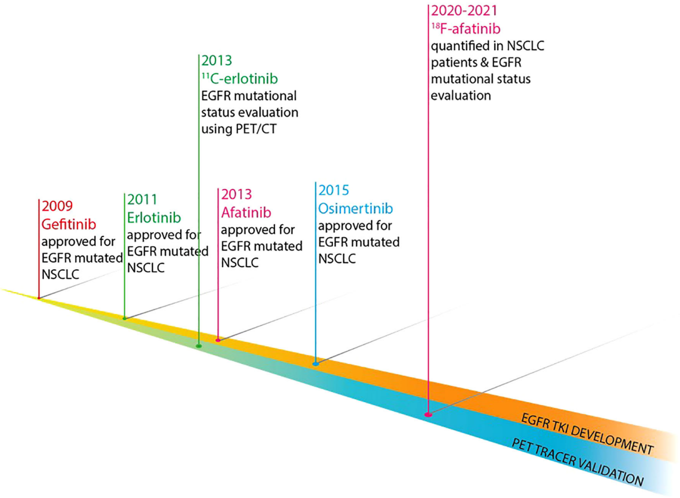


FIGURE 2
Development timeline of EGFR TKI and their respective EGFR-directed PET tracers. .

treatment using the corresponding EGFR TKI (27, 31) could be predicted using PET biomarkers. This was seen in patients with common and uncommon EGFR mutations. For $^{11}$C-osimertinib, the clinical studies investigating its predictive value are still ongoing.

For EGFR PET tracers without treatment analogue, e.g., $^{18}$F-MPG, $^{11}$C-PD153035 and $^{18}$F-IRS, studies have shown that tumor tracer uptake could be quantified and that this was predictive for the presence of an EGFR mutation and for TKI therapy response (45–47). Both $^{18}$F-IRS and $^{11}$C-PD153035

TABLE 1 Overview of clinical EGFR TKI PET studies.

| Year | Tracer | N | Uptake parameter | Kinetic modeling? | Used as biomarker for EGFR status? | EGFR mutation in studies | Study |
|---|---|---|---|---|---|---|---|
| 2008 | $^{11}$C-PD153035 | 11 | SUV | No | No | Exon 19 & exon 21 mutations | Yu et al. (48) |
| 2009 | $^{11}$C-PD153035 | 14 | $SUV_{max}$ | No | Yes | Exon 19 & exon 21 mutations | Yu et al. (49) |
| 2010 | $^{11}$C-PD153035 | 19 | $SUV_{max}$ | No | No | Unknown | Liu et al. (50) |
| 2011 | $^{11}$C-PD153035 | 21 | $SUV_{max}$ | No | No | Unknown | Meng et al. (45) |
| 2011 | $^{11}$C-Erlotinib | 13 | Radioactivity per mL tissue | No | No | Unknown | Memon et al. (51) |
| 2013 | $^{11}$C-Erlotinib | 10 | $V_T$ | Yes | Yes | Exon 19 del | Bahce et al. (27) |
| 2017 | $^{18}$F-IRS | 3 | $SUV_{max}$ | No | Yes | Exon 19 del | Song et al. (47) |
| 2018 | $^{18}$F-MPG | 75 | $SUV_{max}$ | No | Yes | Unknown | Sun et al. (46) |
| 2018 | $^{18}$F-ODS2004436 | 20 | $SUV_{ratio}$ | No | Yes | Unknown | Cochet et al. (52) |
| 2021 | $^{18}$F-Afatinib | 12 | $TBR\_WB_{60-90}$ | Yes | Yes | Exon 19 deletion, exon 19 L747P insertion, exon 18 G719A point mutation, exon 18 G709T deletion | van de Stadt et al. (30) |
| 2021 | $^{11}$C-erlotinib | 10 | $V_T$ & $SUV_{mean}$ | Yes | Yes | Exon 19 deletion, L858R point mutation, G719S + S768I mutation, L861Q mutation | Petrulli et al. (53) |

The EGFR mutational status as described in the study is shown.
SUV, standardized uptake value; $V_T$, volume of distribution; $TBR\_WB_{60-90}$, tumor-to-whole-blood ratio in the time interval 60-90 minutes post-injection.

showed a close relation between tracer uptake ($SUV_{max}$) and EGFR expression, and for all three tracers a correlation between uptake ($SUV_{max}$) and treatment response was observed (45–47).

The overview in Table 1, comprising approximately 200 NSCLC patients, summarizes several study characteristics. When new tracers are introduced, the pharmacokinetic behavior of this tracer needs to be established by performing kinetic modeling. Kinetic modeling allows to better understand the obtained PET images and to quantify the tracer uptake using optimal dynamic parameters of uptake such as 'Distribution Volume' ($V_T$). For some tracers, this has been performed, as indicated in Table 1. In the absence of dynamic uptake parameters, usually simplified static uptake parameters such as 'Standardized Uptake Values' (SUVs) are used. For some tracers such as $^{11}$C-erlotinib and $^{18}$F-afatinib, the pharmacokinetic modeling has been published and, in these tracers, uptake parameters other than SUV have been suggested (29, 30, 53). In Table 2, tracer targets are listed for each tracer.

While this overview highlights the efforts done to investigate and discover the potential of the existing EGFR PET tracers and their biomarkers, it also highlights that data is scarce. From a clinical point of view, the question rises on what would be needed for EGFR PET biomarkers to be able to qualify in the routine clinical practice. To better understand the framework in which such a qualification occurs, we will below elaborate on the typical biomarker development process and how the current state of these tracers and their respective biomarkers relate to this process.

# 3 Challenges and opportunities in the development of EGFR PET biomarkers

## 3.1 Development process of PET biomarkers

To be able to qualify for use in the clinical practice, there are 3 main phases of development that a PET imaging biomarker must transition. See Figure 3, which is based on the consensus paper of the CRUK and the EORTC (59). In transitioning from one phase into another, biomarkers need to bridge several gaps. The first gap for a biomarker is to be able to enter the validation phase as a potential biomarker, fit to be tested for performance. In the validation phase, a biomarker needs to proof it is reliable and 'fit for purpose'. For the development of PET biomarkers, the 3 main validation tracks (analytical, clinical and cost-effectiveness validation) are typically developed in parallel and in an iterative manner. In the qualification phase, sufficient evidence will be needed to support the qualification of a biomarker for a specific context of use in drug development or routine clinical care. support qualification of a biomarker.

## 3.2 EGFR PET biomarker validation challenges

### 3.2.1 Analytical validation

The analytical validation track evaluates the measures related to biomarker precision, e.g., repeatability, reproducibility and technical bias, and the measures related to biomarker availability in the targeted patient group. The analytical validation, generally, does not consider the clinical utility of the biomarker, however, poor analytical features will hamper the clinical validation and qualification (59).

Ideally, new EGFR PET tracers for biomarking EGFR that are used in humans will undergo full kinetic modeling. This is an elaborate dynamic PET scanning procedure with arterial blood sampling and measurement of blood radioactivity and blood tracer metabolites. A dynamic PET scan is a continuous scan of 1 section of the body, where both the tumor and a large blood pool or vessel is included in the field of view (FoV), as depicted in Figure 4. Since conventional PET scanners have a limited (e.g., 18 cm) FoV, only a small part of the body where the tumor is located will be scanned continuously. The pharmacokinetic behavior over time of the tumor tracer concentration will be measured to produce a time-activity-concentration curve (TAC). Additionally, the radioactivity

TABLE 2 Key tracer targets for each tracer are shown.

| Tracer | Key tracer targets | |
|---|---|---|
| $^{11}$C-PD153035 | EGFR (wild-type and mutations), HER2 | (54) |
| $^{11}$C-erlotinib | 1st-generation TKI: common EGFR mutations (exon 19del, exon 21 L858R), partly wild-type EGFR, not T790M | (7) |
| $^{18}$F-IRS | Comparable to 1st-generation TKI: common EGFR mutations, no T790M | (48) |
| $^{18}$F-MPG | Common EGFR mutations, not wild type EGFR, not T790M | (47) |
| $^{18}$F-ODS2004436 | Limited data is publicly available, targets wild type and exon 21 L858R, not T790M | (55, 56) |
| $^{18}$F-afatinib | 2nd-generation TKI: common EGFR mutations (exon 19del, exon 21 L858R) + other ERBB family kinases, partly T790M | (57–59) |
| $^{11}$C-osimertinib | 3rd-generation TKI: specifically developed for EGFR T790M mutation, common EGFR mutations, also uncommon non-exon20 insertions, not wild type EGFR | (20, 23) |

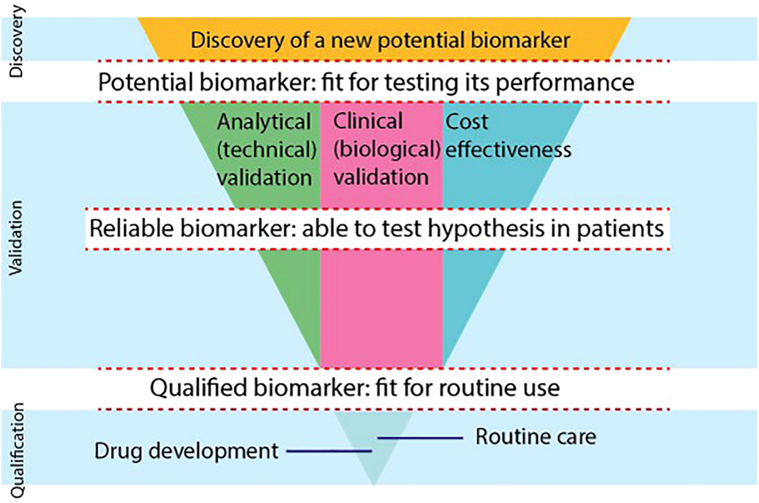


FIGURE 3
The biomarker development process is shown using a modified scheme, based on the consensus statement on biomarker development of the CRUK and EORTC (59). There are 3 phases of development (discovery, validation, qualification) that biomarkers go through. Biomarkers need to overcome gaps to become potential biomarkers, reliable biomarkers and qualified biomarkers. In the validation phase, 3 separate tracks will be evaluated in parallel and iteratively, i.e., the analytical, clinical and cost-effectiveness track. To be able to use a biomarker in drug development or in routine clinical care, biomarkers need to provide qualification evidence. (*) The FDA Evidentiary Framework provides recommendations that guide the evidence needed to support qualification, bridging the final gap to routine care and drug development in the qualification phase.

concentrations of the arterial blood pool over time will be measured to calculate the so-called blood 'input functions' using both blood samples from an arterial cannula, and PET image-derived blood pool data. Also, metabolites will be measured repeatedly *via* arterial blood samples to calculate the true parent tracer concentrations over time. Using the TACs, the blood input function and the metabolites data, the pharmacokinetic model that best describes the pharmacokinetic behavior of the tracer in the tumor will be established. This pharmacokinetic model yields various physiologic parameters, which can be used to select the optimal tracer uptake parameter to quantify the tracer uptake. These dynamic uptake parameters are considered the most precise biomarkers for tracer uptake. Only a few EGFR PET tracers such as $^{11}$C-erlotinib and $^{18}$F-afatinib have undergone full kinetic modeling.

To evaluate intra-patient repeatability is another step in the analytical validation of a biomarker to assure that biomarkers produce similar results when repeatedly measured in the same circumstances. This has been shown for tumor $^{11}$C-erlotinib $V_T$, however, this crucial step is lacking in many other tracers.

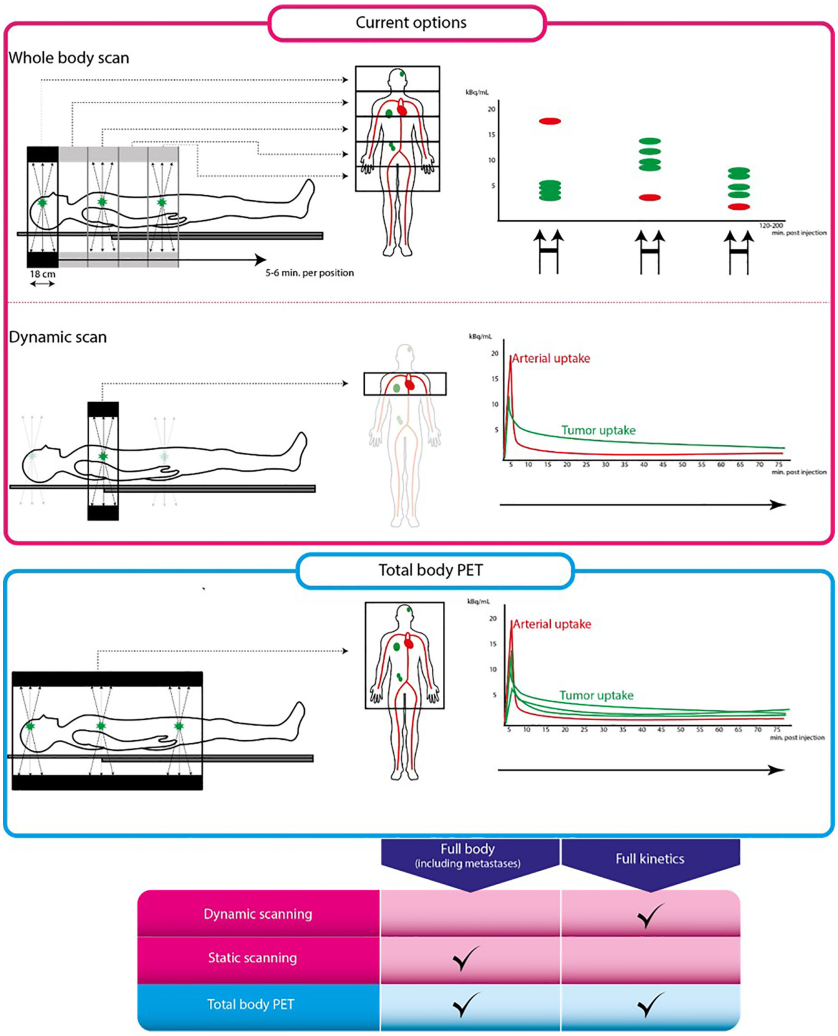
Availability of short-lived EGFR PET tracers is limited due to the short half-life of their radionuclides. For examples, the half-life of carbon-11 is approximately 20 minutes, meaning that the scan must be performed in the same center where the tracer is produced and cannot be exported to other centers. The half-life of fluorine-18 is approximately 5 times longer (t ½ ~110 min), which allows shipping to external not-too-distant centers. Another factor limiting the availability is the scarcity of expertise to apply the complex algorithms used to interpret uptake. In the same vein, dedicated software with intuitive user-friendly interfaces are lacking.

### 3.2.2 Clinical validation

The clinical validation is a process in which the relationship of a biomarker to a clinical feature is evaluated. Biomarkers are typically linked to biological mechanisms of action at the tumor microenvironment. Ultimately, depending on the context of use, the clinical validation should lead to the identification of biomarkers that benefit clinical outcomes or improve the prevention, screening, staging, diagnosis, therapies, or care of patients (59).

Insights obtained in clinical validation studies will feedback into the analytical validation process in order to further optimize the technical aspect of the biomarker. This positive feedback loop highlights the interdependency between these two tracks. Another time-consuming factor in this (clinical) track is the fact that large, prospective clinical PET studies will only be initiated after analytical validation studies have established the precision and accuracy of the tracer as an EGFR biomarker.

The prompt introduction of new EGFR TKI therapy options and the rapid changes in the standard of care for these patients pose a risk on the EGFR PET tracer development, as most TKI-



| | Full body (including metastases) | Full kinetics |
|---|---|---|
| Dynamic scanning | | ✓ |
| Static scanning | ✓ | |
| Total body PET | ✓ | ✓ |

FIGURE 4
Conventional PET scan versus total body PET scan. From left to right: schematic representation of scan procedure, illustrations on the left are parts of the body that can be scanned using each scanning technique. Illustrations on the right are tracer uptake quantification differences for each technique. The pink box represent conventional PET scanning, the blue box represents total body PET. Table below shows characteristics of each scanning technique. Full kinetics indicates whether quantification using pharmacokinetic modeling is possible using this technique.

based tracers have a few years of delay vis-à-vis their therapeutic parents, which can lead to tracers become obsolete. This is highlighted by the timeline depicted in Figure 2: approval of afatinib dates back to 2013, whereas research regarding $^{18}$F-afatinib was first published in 2020, a 7-year delay. In contrast, osimertinib was approved for clinical use in 2015, only 2 years after afatinib entered the market and 5 years before the first publication of $^{18}$F-afatinib.

### 3.2.3 Cost-effectiveness

In the cost-effectiveness track, the costs associated with the use of biomarkers need to be assessed. To become a qualified biomarker for clinical use, these costs need to compare favorably to the existing alternative biomarkers such as bio specimen-derived biomarkers, e.g., liquid biopsies. Costs may become lower at a later stage after broad-scaled implementation (59).

The added advantage of the EGFR PET is to evaluate tumor EGFR TKI sensitivity when regular biopsies are not informative enough or for obtaining spatial insights in the tumor TKI sensitivity to guide decision-making. This technique is therefore used in addition to regular biopsy-techniques. Consequently, evaluating the cost-effectiveness for these situations is difficult. With further analytical and clinical optimization supported by upcoming PET technology and improved data processing algorithms, EGFR PET biomarkers hold promise to provide value for their costs. However, at the current stage, no EGFR PET tracer could be considered cost-effective, especially when compared to biopsy-techniques already widely-used in clinical practice.

## 3.3 Opportunities

The clinical implementation of EGFR PET biomarkers have been limited by the abovementioned challenges, however, recent developments in emerging new technologies are promising to help the biomarker validation process. Although technological advancements may seem to mainly benefit the technical validation and cost-effectiveness tracks, these optimizations feedback positively to the clinical track as well and therefore improve the full validation process. One of the developments that will advance the validation of EGFR PET biomarkers will be the large-scaled introduction of the so-called ‘total body PET’.

### 3.3.1 Total body PET

The total body PET scanner refers to a new generation of commercially available PET-CT scanners that have a much larger axial FoV as compared to conventional state-of-the-art PET-CT systems with an axial FOV of less than 20cm. These new large-FoV PET-CT systems achieve ultra-high (40-to-200-fold higher) sensitivity and allow to visualize and quantify tracer uptake in all major internal organs in the body simultaneously (60–64). This provides numerous new imaging opportunities for patient care and research, since these total body PET-CT scanners will speed up the validation of EGFR PET biomarkers by optimizing their analytical validation and by supporting the clinical validation.

One of the advantages of the ultra-high sensitivity will be the possibility to use lower amounts of radioactivity per tracer injection, which will enable to lower the radiation burden to the patients (60, 64). This could make EGFR PET imaging biomarkers more suited for therapy monitoring through performing multiple PET-CT scans longitudinally.

For static tracer uptake parameters such as SUV, another advantage of the ultra-high sensitivity will be the shorter scan durations (currently 30-40 min per $^{18}$F-FDG PET scan), which in turn will improve patient comfort. The optimal scan duration per EGFR tracer on the total body PET-CT scanner is not clear yet, but this could be as short as 20 seconds (a breath-hold) for some tracers. Short acquisition times could also significantly decrease possible *partial volume effects* caused by smearing the PET signal by the movement of small lesions, e.g., due to breathing-motions (60, 64). Also, this will reduce co-registration mismatch of the PET and CT data, e.g., because of patients moving on the scanner while scanning, which generates artefacts in the reconstructed PET data due to faulty CT-attenuation correction (60, 64). These improvements will increase the resolution and precision of the PET biomarkers, broadening their applicability.

For dynamic tracer uptake parameters, combining the large-axial FOV and the ultra-high specificity of the PET-CT system could greatly improve biomarker specificity, repeatability, and reproducibility. As compared to static PET studies, using dynamic PET studies allows to better characterize the pharmacokinetic (PK) behavior of short-lived tracers by generating dynamic tracer uptake parameters (i.e., biomarkers) that are more target specific and accurate than simplified static parameters (60–64). Typically, the limited axial FoV of the conventional PET-CT systems limits most dynamic scans to single organ studies. Also, for dynamic kinetic analysis a so-called ‘arterial input function’ is needed to describe the bioavailability of the radiotracer in blood. The total body PET-CT, covering all major organs and arterial blood pools (eliminating the need for an arterial cannula) could not only dynamically scan most tumor lesions and all major organs at once but could also provide a reliable image-based arterial input function, non-invasively and automatically, which could generate easily-accessible dynamic uptake parameters with higher specificity and precision (60, 64). Also, the large-FOV coverage will generate new insights on biodistribution in healthy organs, which may open avenues for discovering new PET biomarkers to predict toxicity or biomarkers to guide drug dosing.

Using the total body PET-CT would allow to address many of the analytical validation steps in a single PET study, while this would require many studies using the conventional PET system. Speeding up the analytical validation would significantly fasten the clinical validation as well. As less patients would be needed in the various validation steps of a biomarker, this would ultimately be more cost-effective, through shortening the delay between the introduction of a new EGFR TKI and its validation testing. As most of the tumor lesions, all the major organs and a significant part of the blood pool will be included in the dynamic scans, more comprehensive and automatable scanning and data processing algorithms will be developed. With such algorithms, uptake parameters will be produced more easily, and may require less effort from the PET physics personnel.

### 3.3.2 Further optimizations

With the advent of new PET technologies and improved data processing algorithms, radiolabeling new EGFR TKI could be of interest for pharmaceutical companies to learn about the biodistribution and PK behavior of their new EGFR TKI therapies at an early stage of development. For example, variations in the brain tissue penetration and uptake of TKI in the brain metastases could be of interest as there is quite some variability in the brain penetration of different TKIs (65). Also, blocking studies could be used to explore the optimal dosing to saturate all targets to support the optimal dosing strategy of a TKI (66, 67). The analytical validation associated with these pharmacological drug development projects could support the clinical validation effort as well.

## 4 Conclusion

The use of EGFR TKI PET tracers can generate predictive biomarkers to identify and monitor patients who will respond to

EGFR TKI therapies. Current EGFR TKI tracer biomarkers are still in a validation phase, where clinical and analytical improvements loop back iteratively. New developments such as the availability of large-FoV total body PET systems and more improved data processing algorithms can optimize the EGFR TKI PET biomarker validation process. Nevertheless, more evidence is needed for their qualification as predictive and monitoring biomarkers in drug development and routine clinical practice.

## Author contributions

MY: supervision, edit and rewriting. HH: supervision, edit and rewriting. IB: supervision, edit and rewriting. All authors contributed to the article and approved the submitted version.

## 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

## References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2021. *CA Cancer J Clin* (2021) 71(1):7–33. doi: 10.3322/caac.21654

2. Maemondo M, Inoue A, Kobayashi K, Sugawara S, Oizumi S, Isobe H, et al. Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR. *N Engl J Med* (2010) 362(25):2380–8. doi: 10.1056/NEJMoa0909530

3. Mok TS, Wu YL, Thongprasert S, Yang CH, Chu DT, Saijo N, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* (2009) 361 (10):947–57. doi: 10.1056/NEJMoa0810699

4. Rosell R, Carcereny E, Gervais R, Vergnenegre A, Massuti B, Felip E, et al. Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): A multicentre, open-label, randomised phase 3 trial. *Lancet Oncol* (2012) 13(3):239–46. doi: 10.1016/S1470-2045(11)70393-X

5. Yang JC, Wu YL, Schuler M, Sebastian M, Popat S, Yamamoto N, et al. Afatinib versus cisplatin-based chemotherapy for EGFR mutation-positive lung adenocarcinoma (LUX-lung 3 and LUX-lung 6): Analysis of overall survival data from two randomised, phase 3 trials. *Lancet Oncol* (2015) 16(2):141–51. doi: 10.1016/S1470-2045(14)71173-8

6. Eck MJ, Yun CH. Structural and mechanistic underpinnings of the differential drug sensitivity of EGFR mutations in non-small cell lung cancer. *Biochim Biophys Acta* (2010) 1804(3):559–66. doi: 10.1016/j.bbapap.2009.12.010

7. Mitchell RA, Luwor RB, Burgess AW. Epidermal growth factor receptor: Structure-function informing the design of anticancer therapeutics. *Exp Cell Res* (2018) 371(1):1–19. doi: 10.1016/j.yexcr.2018.08.009

8. Paez JG, Janne PA, Lee JC, Tracy S, Greulich H, Gabriel S, et al. EGFR mutations in lung cancer: Correlation with clinical response to gefitinib therapy. *Science* (2004) 304(5676):1497–500. doi: 10.1126/science.1099314

9. Carey KD, Garton AJ, Romero MS, Kahler J, Thomson S, Ross S, et al. Kinetic analysis of epidermal growth factor receptor somatic mutant proteins shows increased sensitivity to the epidermal growth factor receptor tyrosine kinase inhibitor, erlotinib. *Cancer Res* (2006) 66(16):8163–71. doi: 10.1158/0008-5472.CAN-06-0453

10. Scheffler M, Kobe C, Zander T, Nogova L, Kahraman D, Thomas R, et al. Monitoring reversible and irreversible EGFR inhibition with erlotinib and afatinib in a patient with EGFR-mutated non-small cell lung cancer (NSCLC) using sequential [18F]fluorothymidine (FLT-)PET. *Lung Cancer* (2012) 77(3):617–20. doi: 10.1016/j.lungcan.2012.05.110

11. Stopfer P, Marzin K, Narjes H, Gansser D, Shahidi M, Uttereuther-Fischer M, et al. Afatinib pharmacokinetics and metabolism after oral administration to healthy male volunteers. *Cancer Chemother Pharmacol* (2012) 69(4):1051–61. doi: 10.1007/s00280-011-1803-9

12. Wind S, Schnell D, Ebner T, Freiwald M, Stopfer P. Clinical pharmacokinetics and pharmacodynamics of afatinib. *Clin Pharmacokinet* (2017) 56(3):235–50. doi: 10.1007/s40262-016-0440-1

13. Yang JC, Shih JY, Su WC, Hsia TC, Tsai CM, Ou SH, et al. Afatinib for patients with lung adenocarcinoma and epidermal growth factor receptor mutations (LUX-lung 2): a phase 2 trial. *Lancet Oncol* (2012) 13(5):539–48. doi: 10.1016/S1470-2045(12)70086-4

14. Lau SCM, Batra U, Mok TSK, Loong HH. Dacomitinib in the management of advanced non-Small-Cell lung cancer. *Drugs* (2019) 79(8):823–31. doi: 10.1007/s40265-019-01115-y

15. Reckamp KL, Giaccone G, Camidge DR, Gadgeel SM, Khuri FR, Engelman JA, et al. A phase 2 trial of dacomitinib (PF-00299804), an oral, irreversible pan-HER (human epidermal growth factor receptor) inhibitor, in patients with advanced non-small cell lung cancer after failure of prior chemotherapy and erlotinib. *Cancer* (2014) 120(8):1145–54. doi: 10.1002/cncr.28561

16. Soria JC, Felip E, Cobo M, Lu S, Syrigos K, Lee KH, et al. Afatinib versus erlotinib as second-line treatment of patients with advanced squamous cell carcinoma of the lung (LUX-lung 8): An open-label randomised controlled phase 3 trial. *Lancet Oncol* (2015) 16(8):897–907. doi: 10.1016/S1470-2045(15)00006-6

17. Mok TS, Wu YL, Ahn MJ, Garassino MC, Kim HR, Ramalingam SS, et al. Osimertinib or platinum-pemetrexed in EGFR T790M-positive lung cancer. *N Engl J Med* (2017) 376(7):629–40. doi: 10.1056/NEJMoa1612674

18. Mok TS, Wu YL, Papadimitrakopoulou VA. Osimertinib in EGFR T790M-positive lung cancer. *N Engl J Med* (2017) 376(20):1993–4. doi: 10.1056/NEJMc1703339

19. Oxnard GR, Hu Y, Mileham KF, Husain H, Costa DB, Tracy P, et al. Assessment of resistance mechanisms and clinical implications in patients with EGFR T790M-positive lung cancer and acquired resistance to osimertinib. *JAMA Oncol* (2018) 4(11):1527–34. doi: 10.1001/jamaoncol.2018.2969

20. Santos ES, Kaplan B, Kirshner E, Croft EF, Sequist LV, Chau M, et al. Osimertinib for previously treated patients with advanced EGFR T790M mutation-positive NSCLC: Tolerability and diagnostic methods from an expanded access program. *Oncol Ther* (2018) 6(1):45–58. doi: 10.1007/s40487-018-0061-y

21. Papadimitrakopoulou VA, Han JY, Ahn MJ, Ramalingam SS, Delmonte A, Hsia TC, et al. Epidermal growth factor receptor mutation analysis in tissue and plasma from the AURA3 trial: Osimertinib versus platinum-pemetrexed for T790M mutation-positive advanced non-small cell lung cancer. *Cancer* (2020) 126(2):373–80. doi: 10.1002/cncr.32503

22. Soria JC, Ohe Y, Vansteenkiste J, Reungwetwattana T, Chewaskulyong B, Lee KH, et al. Osimertinib in untreated EGFR-mutated advanced non-Small-Cell lung cancer. *N Engl J Med* (2018) 378(2):113–25. doi: 10.1056/NEJMoa1713137

23. Ramalingam SS, Vansteenkiste J, Planchard D, Cho BC, Gray JE, Ohe Y, et al. Overall survival with osimertinib in untreated, EGFR-mutated advanced NSCLC. *N Engl J Med* (2020) 382(1):41–50. doi: 10.1056/NEJMoa1913662

24. Kalemkerian GP, Narula N, Kennedy EB, Biermann WA, Donington J, Leighl NB, et al. Molecular testing guideline for the selection of patients with lung cancer for treatment with targeted tyrosine kinase inhibitors: American society of clinical oncology endorsement of the college of American Pathologists/ International association for the study of lung Cancer/Association for molecular pathology clinical practice guideline update. *J Clin Oncol* (2018) 36(9):911–9. doi: 10.1200/JCO.2017.76.7293

25. Rolfo C, Mack P, Scagliotti GV, Aggarwal C, Arcila ME, Barlesi F, et al. Liquid biopsy for advanced NSCLC: A consensus statement from the international

association for the study of lung cancer. *J Thorac Oncol* (2021) 16(10):1647–62. doi: 10.1016/j.jtho.2021.06.017

26. Rolfo C, Mack PC, Scagliotti GV, Baas P, Barlesi F, Bivona TG, et al. Liquid biopsy for advanced non-small cell lung cancer (NSCLC): A statement paper from the IASLC. *J Thorac Oncol* (2018) 13(9):1248–68 doi: 10.1016/j.jtho.2018.05.030.

27. Bahce I, Smit EF, Lubberink M, van der Veldt AA, Yaqub M, Windhorst AD, et al. Development of [(11)C]erlotinib positron emission tomography for *in vivo* evaluation of EGF receptor mutational status. *Clin Cancer Res* (2013) 19 (1):183–93 doi: 10.1158/1078-0432.CCR-12-0289.

28. Bahce I, Yaqub M, Smit EF, Lammertsma AA, van Dongen GA, Hendrikse NH. Personalizing NSCLC therapy by characterizing tumors using TKI-PET and immuno-PET. *Lung Cancer* (2017) 107:1–13 doi: 10.1016/j.lungcan.2016.05.025.

29. Yaqub M, Bahce I, Voorhoeve C, Schuit RC, Windhorst AD, Hoekstra OS, et al. Quantitative and simplified analysis of 11C-erlotinib studies. *J Nucl Med* (2016) 57(6):861–6. doi: 10.2967/jnumed.115.165225

30. van de Stadt Y. Quantification of [(18)F]afatinib using PET/CT in NSCLC patients: A feasibility study. *EJNMMI Res* (2020) 10(1):97.

31. van de Stadt EA, Yaqub M, Lammertsma AA, Poot AJ, Schuit RC, Remmelzwaal S, et al. Identifying advanced stage NSCLC patients who benefit from afatinib therapy using (18)F-afatinib PET/CT imaging. *Lung Cancer* (2021) 155:156–62. doi: 10.1016/j.lungcan.2021.03.016

32. Group F-NBW. *BEST (Biomarkers, EndpointS, and other tools) resource*. Bethesda: Silver Springs (2016).

33. Du B, Wang S, Cui Y, Liu G, Li X, Li Y. Can (18)F-FDG PET/CT predict EGFR status in patients with non-small cell lung cancer? a systematic review and meta-analysis. *BMJ Open* (2021) 11(6):e044313. doi: 10.1136/bmjopen-2020-044313

34. Yin G, Wang Z, Song Y, Li X, Chen Y, Zhu L, et al. Prediction of EGFR mutation status based on (18)F-FDG PET/CT imaging using deep learning-based model in lung adenocarcinoma. *Front Oncol* (2021) 11:709137. doi: 10.3389/fonc.2021.709137

35. Bollineni VR, Kramer GM, Jansma EP, Liu Y, Oyen WJ. A systematic review on [(18)F]FLT-PET uptake as a measure of treatment response in cancer patients. *Eur J Cancer* (2016) 55:81–97. doi: 10.1016/j.ejca.2015.11.018

36. Ullrich RT, Zander T, Neumaier B, Koker M, Shimamura T, Waerzeggers Y, et al. Early detection of erlotinib treatment response in NSCLC by 3'-deoxy-3'-[F]-fluoro-L-thymidine ([F]FLT) positron emission tomography (PET). *PloS One* (2008) 3(12):e3908. doi: 10.1371/journal.pone.0003908

37. Iqbal R, Kramer GM, Frings V, Smit EF, Hoekstra OS, Boellaard R, et al. Validation of [(18)F]FLT as a perfusion-independent imaging biomarker of tumour response in EGFR-mutated NSCLC patients undergoing treatment with an EGFR tyrosine kinase inhibitor. *EJNMMI Res* (2018) 8(1):22. doi: 10.1186/s13550-018-0376-6

38. de Jong IJ, Pruim J, Elsinga PH, Vaalburg W, Mensink HJ. 11C-choline positron emission tomography for the evaluation after treatment of localized prostate cancer. *Eur Urol* (2003) 44(1):32–8. doi: 10.1016/s0302-2838(03)00207-0

39. Khan N, Oriuchi N, Zhang H, Higuchi T, Tian M, Inoue T, et al. A comparative study of 11C-choline PET and [18F]fluorodeoxyglucose PET in the evaluation of lung cancer. *Nucl Med Commun* (2003) 24(4):359–66. doi: 10.1097/00006231-200304000-00004

40. Tian M, Zhang H, Oriuchi N, Higuchi T, Endo K. Comparison of 11C-choline PET and FDG PET for the differential diagnosis of malignant tumors. *Eur J Nucl Med Mol Imaging* (2004) 31(8):1064–72. doi: 10.1007/s00259-004-1496-y

41. Hara T, Inagaki K, Kosaka N, Morita T. Sensitive detection of mediastinal lymph node metastasis of lung cancer with 11C-choline PET. *J Nucl Med* (2000) 41 (9):1507–13.

42. Abdulla DSY, Scheffler M, Brandes V, Ruge M, Kunze S, Merkelbach-Bruse S, et al. Monitoring treatment response to erlotinib in EGFR-mutated non-small-cell lung cancer brain metastases using serial O-(2-[(18)F]fluoroethyl)-L-tyrosine PET. *Clin Lung Cancer* (2019) 20(2):e148–e51 doi: 10.1016/j.cllc.2018.10.011.

43. Langen KJ, Stoffels G, Filss C, Heinzel A, Stegmayr C, Lohmann P, et al. Imaging of amino acid transport in brain tumours: Positron emission tomography with O-(2-[(18)F]fluoroethyl)-L-tyrosine (FET). *Methods* (2017) 130:124–34 doi: 10.1016/j.ymeth.2017.05.019.

44. Varrone A, Varnas K, Jucaite A, Cselenyi Z, Johnstrom P, Schou M, et al. A PET study in healthy subjects of brain exposure of (11)C-labelled osimertinib - a drug intended for treatment of brain metastases in non-small cell lung cancer. *J Cereb Blood Flow Metab* (2020) 40(4):799–807. doi: 10.1177/0271678X19843776

45. Meng X, Loo BWJr., Ma L, Murphy JD, Sun X, Yu J. Molecular imaging with 11C-PD153035 PET/CT predicts survival in non-small cell lung cancer treated with EGFR-TKI: A pilot study. *J Nucl Med* (2011) 52(10):1573–9. doi: 10.2967/jnumed.111.092874

46. Sun X, Xiao Z, Chen G, Han Z, Liu Y, Zhang C, et al. A PET imaging approach for determining EGFR mutation status for improved lung cancer patient management. *Sci Transl Med* (2018) 10(431):443–55. doi: 10.1126/scitranslmed.aan8840

47. Song Y, Xiao Z, Wang K, Wang X, Zhang C, Fang F, et al. Development and evaluation of (18)F-IRS for molecular imaging mutant EGF receptors in NSCLC. *Sci Rep* (2017) 7(1):3121. doi: 10.1038/s41598-017-01443-7

48. Yu JM, Liu N, Yang G, Guo H, Ma L, Zhao S, et al. 11C-PD153035 PET/CT for molecular imaging of EGFR in patients with non-small cell lung cancer (NSCLC). *J Clin Oncol* (2008) 26(15_suppl):3503. doi: 10.1200/jco.2008.26.15_suppl.3503

49. Yu J, Liu N, Hu M, Song X, Xie L, Meng X, et al. Further evaluation of 11C-PD153035 as a molecular imaging probe for the assessment of the epidermal growth factor receptor status in non-small cell lung cancer patients. *J Clin Oncol* (2009) 27(15_suppl):3590. doi: 10.1200/jco.2009.27.15_suppl.3590

50. Liu N, Li M, Li X, Meng X, Yang G, Zhao S, et al. PET-based biodistribution and radiation dosimetry of epidermal growth factor receptor-selective tracer 11C-PD153035 in humans. *J Nucl Med* (2009) 50(2):303–8. doi: 10.2967/jnumed.108.056556

51. Memon AA, Weber B, Winterdahl M, Jakobsen S, Meldgaard P, Madsen HH, et al. PET imaging of patients with non-small cell lung cancer employing an EGF receptor targeting drug as tracer. *Br J Cancer* (2011) 105(12):1850–5. doi: 10.1038/bjc.2011.493

52. Cochet A, Isambert N, Foucher P, Bertaut A, Berthet C, Coudert BP, et al. Phase 0/1 of positron emission tomography (PET) imaging agent [18F]-ODS2004436 as a marker of EGFR mutation in patients with non-small cell lung cancer (NSCLC). *J Clin Oncol* (2018) 36(15_suppl):e24184. doi: 10.1200/JCO.2018.36.15_suppl.e24184

53. Petrulli JR, Zheng M, Huang Y, Nabulsi NB, Goldberg SB, Contessa JN, et al. Evaluation of quantitative modeling methods in whole-body, dynamic [(11)C]-erlotinib PET. *Am J Nucl Med Mol Imaging* (2021) 11(2):143–53.

54. Philippe Genne CB, Raguin O, Chalon S, Tizon X, Serriere S. Preclinical proof of concept for the first nanocyclix TKI-PET radiotracer targeting activated EGFR positive lung tumors. *Cancer Res* (2017) 77. doi: 10.1158/1538-7445.AM2017-1875A

55. Bos M, Mendelsohn J, Kim YM, Albanell J, Fry DW, Baselga J. PD153035, a tyrosine kinase inhibitor, prevents epidermal growth factor receptor activation and inhibits growth of cancer cells in a receptor number-dependent manner. *Clin Cancer Res* (1997) 3(11):2099–106.

56. Slobbe P, Windhorst AD, Stigter-van Walsum M, Schuit RC, Smit EF, Niessen HG, et al. Development of [18F]afatinib as new TKI-PET tracer for EGFR positive tumors. *Nucl Med Biol* (2014) 41(9):749–57. doi: 10.1016/j.nucmedbio.2014.06.005

57. Solca F, Dahl G, Zoephel A, Bader G, Sanderson M, Klein C, et al. Target binding properties and cellular activity of afatinib (BIBW 2992), an irreversible ErbB family blocker. *J Pharmacol Exp Ther* (2012) 343(2):342–50. doi: 10.1124/jpet.112.197756

58. Wang DD, Lee VH, Zhu G, Zou B, Ma L, Yan H. Selectivity profile of afatinib for EGFR-mutated non-small-cell lung cancer. *Mol Biosyst* (2016) 12 (5):1552–63. doi: 10.1039/C6MB00038J

59. O'Connor JP, Aboagye EO, Adams JE, Aerts HJ, Barrington SF, Beer AJ, et al. Imaging biomarker roadmap for cancer studies. *Nat Rev Clin Oncol* (2017) 14 (3):169–86. doi: 10.1038/nrclinonc.2016.162

60. Badawi RD, Shi H, Hu P, Chen S, Xu T, Price PM, et al. First human imaging studies with the EXPLORER total-body PET scanner. *J Nucl Med* (2019) 60(3):299–303. doi: 10.2967/jnumed.119.226498

61. Vandenberghe S, Moskal P, Karp JS. State of the art in total body PET. *EJNMMI Phys* (2020) 7(1):35. doi: 10.1186/s40658-020-00290-2

62. Badawi RD, Karp JS, Nardo L, Pantel AR. Total body PET: Exploring new horizons. preface. *PET Clin* (2021) 16(1):xvii–xviii doi: 10.1016/j.cpet.2020.09.005.

63. Cherry SR, Badawi RD, Karp JS, Moses WW, Price P, Jones T. Total-body imaging: Transforming the role of positron emission tomography. *Sci Transl Med* (2017) 9(381) doi: 10.1126/scitranslmed.aaf6169.

64. Cherry SR, Jones T, Karp JS, Qi J, Moses WW, Badawi RD. Total-body PET: Maximizing sensitivity to create new opportunities for clinical research and patient care. *J Nucl Med* (2018) 59(1):3–12 doi: 10.2967/jnumed.116.184028.

65. Colclough N, Chen K, Johnstrom P, Strittmatter N, Yan Y, Wrigley GL, et al. Preclinical comparison of the blood-brain barrier permeability of osimertinib with other EGFR TKIs. *Clin Cancer Res* (2021) 27(1):189–201. doi: 10.1158/1078-0432.CCR-19-1871

66. Bauer M, Karch R, Wulkersdorfer B, Philippe C, Nics L, Klebermass EM, et al. A proof-of-Concept study to inhibit ABCG2- and ABCB1-mediated efflux transport at the human blood-brain barrier. *J Nucl Med* (2019) 60(4):486–91. doi: 10.2967/jnumed.118.216432

67. Verheijen RB, Yaqub M, Sawicki E, van Tellingen O, Lammertsma AA, Nuijen B, et al. Molecular imaging of ABCB1 and ABCG2 inhibition at the human blood-brain barrier using elacridar and (11)C-erlotinib PET. *J Nucl Med* (2018) 59 (6):973–9. doi: 10.2967/jnumed.117.195800

frontiers | Frontiers in Oncology

TYPE Review
PUBLISHED 20 October 2022
DOI 10.3389/fonc.2022.930917

OPEN ACCESS

EDITED BY
Min Wu,
Sichuan University, China

REVIEWED BY
Binghui Zhao,
Tongji University, China
Eugene Demidenko,
Dartmouth College, United States

*CORRESPONDENCE
Haitao Liu
haitao.liu@shgh.cn
Wei Shen
wei.shen@sjtu.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 28 April 2022
ACCEPTED 30 September 2022
PUBLISHED 20 October 2022

CITATION
Li M, Jiang Z, Shen W and Liu H (2022) Deep learning in bladder cancer imaging: A review.
*Front. Oncol.* 12:930917.
doi: 10.3389/fonc.2022.930917

# Deep learning in bladder cancer imaging: A review

Mingyang Li[1†], Zekun Jiang[2†], Wei Shen[2*] and Haitao Liu[1*]

[1]Department of Urology, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, [2]Ministry of Education (MoE) Key Lab of Artificial Intelligence, Artificial Intelligence (AI) Institute, Shanghai Jiao Tong University, Shanghai, China

Deep learning (DL) is a rapidly developing field in machine learning (ML). The concept of deep learning originates from research on artificial neural networks and is an upgrade of traditional neural networks. It has achieved great success in various domains and has shown potential in solving medical problems, particularly when using medical images. Bladder cancer (BCa) is the tenth most common cancer in the world. Imaging, as a safe, noninvasive, and relatively inexpensive technique, is a powerful tool to aid in the diagnosis and treatment of bladder cancer. In this review, we provide an overview of the latest progress in the application of deep learning to the imaging assessment of bladder cancer. First, we review the current deep learning approaches used for bladder segmentation. We then provide examples of how deep learning helps in the diagnosis, staging, and treatment management of bladder cancer using medical images. Finally, we summarize the current limitations of deep learning and provide suggestions for future improvements.

KEYWORDS

**bladder cancer, deep learning, artificial intelligence, medical imaging, computed tomography, magnetic resonance imaging**

## Introduction

According to the latest statistics from Global Cancer, bladder cancer (BCa) is the tenth most common cancer in the world, with approximately 573,000 new cases and 213,000 deaths in 2020 (1). Early diagnosis and treatment are key to reducing morbidity and mortality associated with BCa (2, 3). In current clinical practice, pathological examination following transurethral resection of bladder tumor (TURBT) and cystoscopy are the gold standard for diagnosing BCa (4). However, these methods are expensive and invasive, making it difficult for many patients to afford them, which may delay diagnosis (5). Therefore, as a noninvasive and inexpensive method, imaging techniques play an increasingly important role in the diagnosis of BCa (6). At present, magnetic resonance imaging (MRI), positron emission tomography (PET), and computed tomography (CT) are the conventional imaging methods for diagnosis before treatment (7). However, due to the complex and variable imaging features of BCa, it is difficult for radiologists to make an accurate BCa diagnosis based only on their experience. Therefore, there is an urgent need for better imaging methods to achieve a noninvasive and accurate diagnosis of BCa.

Deep learning (DL) is a rapidly developing field in machine learning (ML). Compared with classical ML algorithms, manual selection of features is not necessarily required in advance in DL. In contrast, the algorithm can learn the most relevant features for classification or prediction (8). In addition, it easily takes advantage of increases in the amount of available computation and data, with very little engineering by hand. This makes DL particularly useful for solving complex computational problems involving large-scale image classification, speech recognition, and many other domains (9, 10).

Medical images contain a vast amount of data with extremely valuable signals and information, which is far beyond the ability of human beings to analyze. ML is naturally and rapidly used in this field because of its unique ability to integrate, analyze, and make predictions based on large amounts of data (11). As an emerging technology in recent years, DL has the potential to make better use of a large amount of data and provide better results (12, 13). In this review, we describe the research status of DL in the image segmentation, diagnosis, staging, and treatment response prediction of BCa (Figure 1). We are the first comprehensive review to present the current state of research on DL in BCa imaging. We focus on the purpose, DL methods, advantages, and limitations of the current research and discuss possible future directions in the field.

## Methods

We conducted a literature search in the PubMed, Web of Science, and IEEE Xplore databases using the term "Bladder Cancer," combined with the terms "Deep Learning", "Diagnostic Imaging", and "Medical Imaging". In order to obtain articles that met the requirements of this review, we applied the following eligibility criteria: ① The paper is written in English; ② the paper is not a review article or editorial; ③ the paper is mainly related to BCa; ④ the paper discusses DL; and ⑤ the paper discusses imaging data. Figure 2 illustrates the process of selecting articles based on the PRISMA criteria. To conduct our review, we extracted the names of the papers, authors, year of publication, DL modules, number of patients included, performance evaluation parameters, and many other features.

## Deep learning in bladder cancer segmentation

Medical image segmentation plays an important role in current medical imaging systems (14). In BCa, the accurate segmentation of normal bladder structures and tumor regions is an important step in tumor diagnosis and tumor stage evaluation (15). Figure 3 illustrates the workflow of bladder cancer image segmentation using deep learning. The deep learning model is first trained by the training dataset and the ground truth label. Then the model can automatically analyze the input validation images and output the corresponding segmented images of all regions and compare them with ground truth for verification. However, as a hollow organ, the bladder undergoes various changes in position, shape, and volume. In addition, complex noise and artifacts are prevalent in medical images, which makes segmentation difficult (17–19). To date, many DL studies have focused only on the segmentation of the bladder wall (20–24). This is due to the

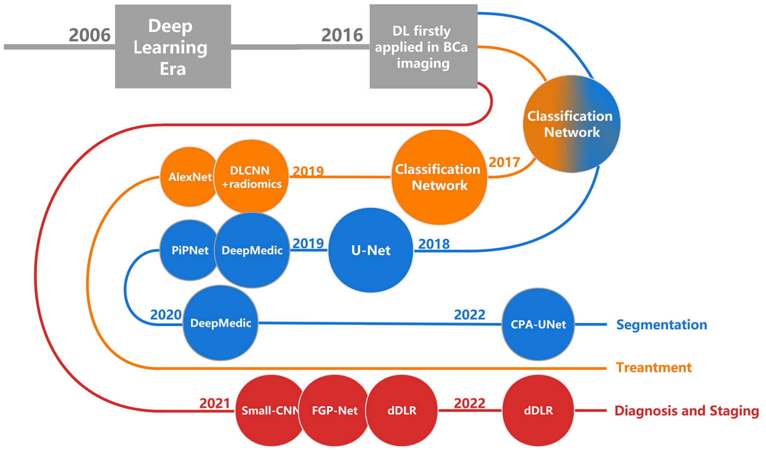


FIGURE 1
The development history of DL in BCa imaging. Each node corresponds to a research, named after the DL architecture that the research primarily used. DL, deep learning; BCa, bladder cancer.

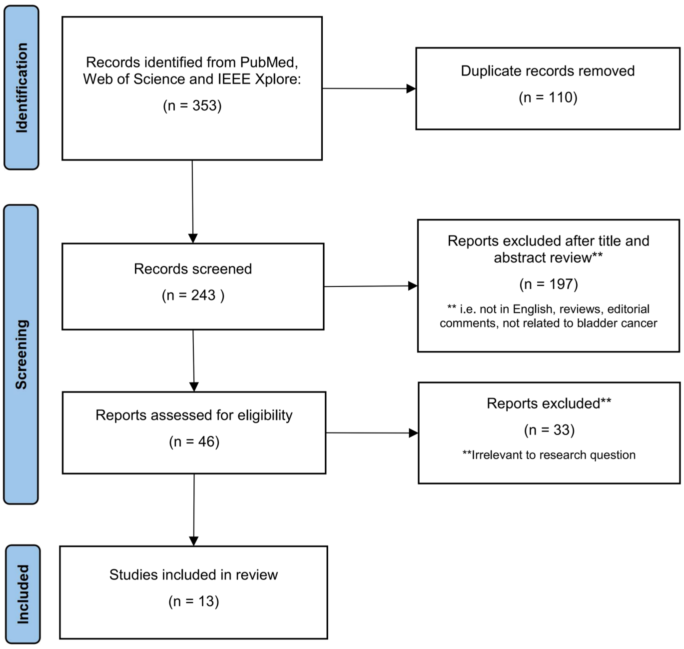


FIGURE 2
Summary of study selection process.

high variability in tumor shape and intensity, making it difficult to distinguish between the bladder wall and a tumor. Therefore, it is more difficult to obtain accurate segmentation results than with normal bladder segmentation. In this review, we focus only on the literature that contains the segmentation of tumor regions (Table 1).

In 2016, Cha et al. (25) developed a network consisting of two convolution layers, two locally connected layers, and one fully connected layer, which is based on the well-known AlexNet (30) backbone. They then used level sets to perform minor refinements to the contour to identify the tumor boundary. However, these methods have many limitations, including a considerably slow process, sensitivity to initialization and image intensity, and independent pixel prediction. The achieved results were not significantly improved when compared with manual segmentation; therefore, they were quickly replaced by fully convolutional architectures.

U-Net (31) is undoubtedly one of the most successful methods in the fully convolutional architectures in image segmentation tasks, serving as the backbone of many new medical image segmentation methods. In 2018, Dolz et al. (26) added dilated convolutions to the U-Net model, where the

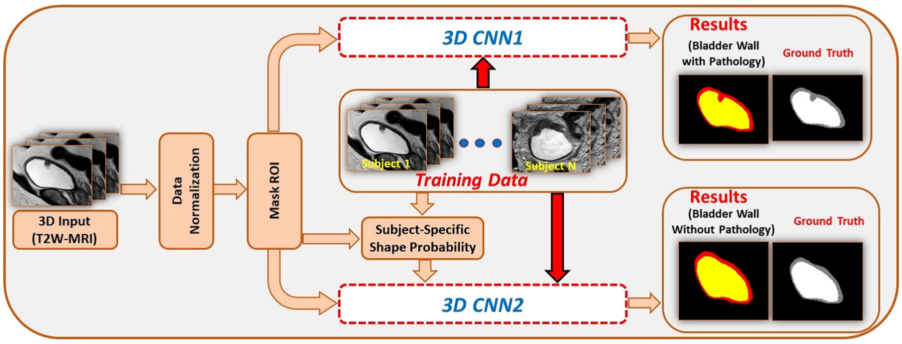


FIGURE 3
An example for bladder cancer image segmentation using deep learning. Image from Ref (16). Copyright © 2020, IEEE.

TABLE 1 Studies using deep learning approach for bladder cancer segmentation.

| Author | Year | Modality | Number of patients (Train/Val/Test) | CNN structure | Target | Performance (validation or testing dataset) |
|---|---|---|---|---|---|---|
| Cha et al. (25) | 2016 | CT | 62, LOOCV | A network contains 2 convolution layers, 2 locally connected layers, and 1 fully connected layer with level sets | Tumor | AVDIST = 4.7mm<br>JACCARD = 36.3% |
| Dolz et al. (26) | 2018 | T2W MRI | 40/5/15, LOOCV | U-Net with progressive dilated convolutional modules, 2D | IW/<br>OW/<br>Tumor | DSC (IW) = 0.9836<br>DSC (OW) = 0.8391<br>DSC (Tumor) = 0.6856<br>ASSD (IW) = 0.3517mm<br>ASSD (OW) = 0.4299mm<br>ASSD (Tumor) = 2.8352mm |
| Liu et al. (27) | 2019 | T2W MRI | 40/-/7, n-fold CV | PiPNet (U-Net with progressive dilated convolutional modules and three prediction masks), 2D | OW/<br>Tumor | DSC (OW) = 0.8874<br>DSC (Tumor) = 0.9543 |
| Hammouda et al. (28) | 2019 | T2W MRI | 20, LOOCV | DeepMedic (a dual pathway CNN with a learnable adaptive shape prior model), 2D | IW/<br>OW/<br>Tumor | DSC (IW) = 0.9895<br>DSC (OW) = 0.9775<br>DSC (Tumor) = 0.9705<br>HD (IW) = 0.17mm<br>HD (OW) = 0.18mm<br>HD (Tumor) = 0.25mm |
| Hammouda et al. (16) | 2020 | T2W MRI | 17, LOOCV | DeepMedic (two CNN network with a learnable adaptive shape prior model and CRF), 3D | IW/<br>OW/<br>Tumor | DSC (IW) = 0.9802<br>DSC (OW) = 0.9742<br>DSC (Tumor) = 0.9566<br>HD (IW) = 0.13mm<br>HD (OW) = 0.19mm<br>HD (Tumor) = 0.35mm |
| Yu et al. (29) | 2022 | T2W MRI | 220/-/25, | CPA-Unet (a Unet for rough segmentation,a path augmentation structure for fine segmentation) | IW/<br>OW/<br>Tumor | DSC (IW) = 0.9819<br>DSC (OW) = 0.8224<br>DSC (Tumor) = 0.8740 |

AVDST, average distance; JACCARD, Jaccard similarity coefficient; DSC, Dice similarity coefficient; ASSD, average symmetric surface distance; HD, Hausdorff distance; IW, bladder inner wall; OW, bladder outer wall; LOOCV, leave-one-out cross-validation.

dilation rate within each module progressively increased. The dilated convolutions can provide a larger receptive field that can leverage more contextual information. The increasing dilation rate allows the use of multi-scale information to better meet the segmentation requirements for both small and large objects. The model was trained and evaluated on T2-weighted (T2W) MR image datasets of 60 BCa patients and compared with the original U-Net, E-Net (32), and ERF-Net (33). The mean Dice similarity coefficient (DSC) values of their model were 0.98, 0.84, and 0.69 for the segmentation of the bladder inner wall, bladder outer wall, and tumor region, respectively, which were the best values of all the models trained. In addition, even though U-Net was improved with progressive dilated convolutional modules to avoid too much computation, the model's inference time for the entire 3D volume is still less than 1 s. In 2019, Liu et al. (27) proposed a CNN architecture called the Pyramid in Pyramid Network (PiPNet), which is based on the U-Net model. The proposed PiPNet consists of a pyramid backbone similar to that of U-Net and adopts atrous spatial pyramid pooling (ASPP) of four parallel atrous convolutions with increasing dilation rates. In addition, the proposed model generates three prediction masks for the segmentation in the feature map of the last three layers to compute an overall loss function to extract multi-scale features. Depthwise separable convolution was used to improve the efficiency and performance of the model. The model was trained and evaluated on T2W MR images of 47 patients with BCa and compared with SegNet (34), U-Net, and Dolz's (26) model. The DSC values were 0.89 and 0.95 for the outer wall and tumor, respectively, which were better than those of other models. Interestingly, in this study, Dolz et al.'s (26) model also achieved better results than the original, with DSCs of 0.86 and 0.92 for the outer wall and tumor, respectively. All models achieved better segmentation accuracy on tumors than on the bladder wall, contrary to the findings of Dolz et al. (26). Therefore, we believe that in the case of less data, different dataset quality and ground truth annotation methods have a greater impact on the performance of the trained model. Yu et al. (29) developed a Cascade Path Augmentation Unet (CPA-Unet) in 2022. They proposed a two-stage segmentation strategy and a hybrid loss function to improve the segmentation results. They first used U-Net for rough segmentation and then used the segmented image with the original image concatenated as a sample with two channels and input into the path augmentation structure (PA-Unet) for fine segmentation. The PA-Unet was based on the Path Aggregation Network (35), and the hybrid loss function incorporated the dice and cross-entropy losses, which can improve the performance (36). The CPA-Unet extracts multi-scale features more accurately, improves small target classification, and achieves better segmentation results than the U-Net, Prog Dilated (26), and PiPNet (27) networks.

These methods based on U-net improve the network performance through a more elaborate network design. However, these methods do not take advantage of the unique characteristics of BCa data and only improve the results by increasing network's robustness. The advantages of these methods include better network characteristics and improved prediction results, which prove their effectiveness. However, as these methods are not specific in nature, which is not fundamentally different from other methods and networks in medical imaging, they do not make good use of data specificity when designing methods.

In addition to U-Net, another well-known CNN architecture for medical image segmentation, DeepMedic (37), has also been used for BCa segmentation. It can make better use of the geometric information of the bladder. Hammouda et al. (28) adopted a dual pathway 2D CNN to segment T2-weighted MRI images. In addition to inputting MRI image data, they also input subject-specific shape information that is adaptively built during segmentation. The adaptive shape prior (ASP) information comes from the results of co-aligning MRI images and ground truth images using an Affine transformation followed by a B-spline based transformation. The use of adaptive shape and contextual information significantly enhanced the segmentation performance, with DSC values of 0.99, 0.98 and 0.97 for the bladder inner wall, outer wall, and tumor, respectively. In 2020, Hammouda et al. (16) further improved their study. They extended their work to 3D bladder segmentation using T2W MRI. The proposed 3D CNN contains two branch networks. The first network aimed to segment the bladder wall with the tumor, and the second network only extracted the bladder. They used a 3D ASP model mixed with the original training data to feed the second network, and the outputs were refined using a fully connected conditional random field (CRF). The CRF can effectively reduce isolated small regions or small holes caused by local minima during training and noise in the input images. The performance of the proposed model significantly outperformed that of U-Net. These methods improved the results because the novelty of these methods changed from a simple network layer design to combining geometric information for segmentation.

When comparing the results of the existing segmentation works, we found that different literature often adopted different evaluation metrics. Most articles used the popular evaluation metric in medical image segmentation, the Dice coefficient (DSC). It can be computed as follows:

$$\mathrm{DSC} = \frac{2|A \cap B|}{|A| + |B|}$$

DSC is a metric to assess the similarity between the predicted area and ground truth area based on the number of pixels of the overlapping region. A similar evaluation metric to it is the Jaccard index, which can be defined as:

$$\mathrm{JACCARD} = \frac{A \cap B}{A \cup B}$$

However, region-based evaluation metrics are not sufficient to evaluate the segmentation of the bladder wall or to evaluate the contour consistency between the predicted area and ground truth area. Therefore, some articles included distance-based evaluation metrics, such as the average distance (AVDIST), the average symmetric surface distance (ASSD) and the Hausdorff distance (HD). AVDIST (25) is the average of the distances between the closest points of contours A and B and can be calculated as follows:

$$\begin{aligned}&\mathrm{AVDIST}^{3D}(\mathrm{A,B})\\&= \frac{1}{2}\left(\frac{\sum_{a\in A}\min\limits_{b\in B} d(a,b)}{N_A}\right) + \left(\frac{\sum_{b\in B}\min\limits_{a\in A} d(b,a)}{N_B}\right)\end{aligned}$$

$N_A$ and $N_B$ denote the number of voxels on A and B, respectively. The function $d$ is the Euclidean distance. The ASSD is also used to calculate the average distance between 2 contours, which can be defined as follows:

$$\mathrm{ASSD}(\mathrm{A, B}) = \frac{1}{|A| + |B|}\left(\sum_{a\in A}\min_{b\in B} d(a,b) + \sum_{b\in B}\min_{a\in A} d(b,a)\right)$$

The HD is also a commonly used distance-based evaluation metric that is sensitive to segmentation boundaries. It can be computed using the following equation:

$$\mathrm{HD}(\mathrm{A, B}) = \max\left\{\max_{a\in A}\left\{\min_{b\in B}\{d(a,b)\}\right\}, \max_{b\in B}\left\{\min_{a\in A}\{d(a,b)\}\right\}\right\}$$

However, the use of diverse evaluation metrics makes it difficult to directly compare the performance of different models. In addition, metrics that are closely related to the clinical application such as model computation time should also be included. We believe that the adoption of consistent and comprehensive evaluation metrics, such as DSC and HD, can help us recognize the effects of different methods and make reasonable improvements.

In summary, these researches use different deep learning networks and algorithms to significantly improve the segmentation accuracy. Before deep learning methods were widely used, early literature used methods including Markov Random Fields, region growing, mathematical morphology, level-set, Chan-Vese model, geodesic active contour (GAC) and continuous max-flow algorithm for bladder segmentation (17–19, 38–47). And most of these researches were not able to segment tumor regions due to the limitations of algorithm and dataset size. In the only article that segmented the tumor region and used JACCARD as an evaluation criterion, they adopted a level-set-based method on a small dataset of ten patients, and the JACCARD of tumor regions extracted by it was 86.3% (45). The best DSC of tumor segmentation among the deep learning methods, on the other hand, reached 97.05% (28). For the segmentation of the bladder wall, the best DSC achieved by the method before deep learning was 87.28% (47). In contrast, the DSC of bladder wall segmentation of deep learning methods generally achieves over 90%. Deep learning methods have different innovations and produce satisfactory results that beyond traditional methods.

## Deep learning in bladder cancer diagnosis and staging

BCa is divided into non-muscle-invasive bladder cancer (NMIBC) and muscle-invasive bladder cancer (MIBC) according to whether the cancer invades the muscle (4). NMIBC accounts for approximately 75% of BCa cases and MIBC accounts for approximately 25%. MIBC is associated with a high degree of malignancy and a poor prognosis. The 5-year survival rate of MIBC patients after radical cystectomy is approximately 45-68%, whereas the survival time of MIBC patients with metastases generally does not exceed 2 years (48). Therefore, early and accurate diagnosis of BCa and assessment of the tumor stage are crucial for guiding clinical treatment and evaluating patient prognosis (49, 50).

In the past, the combination of artificial intelligence and radiomics has replaced traditional methods of manually defining the region of interest (ROI) and extracting image features and has achieved good results in the diagnosis and staging of BCa (51). However, DL can perform the above tasks automatically and achieve better results (Table 2). Yang et al. (52) proposed a small DL-CNN containing four convolutional and max-pooling layers to differentiate NMIBC from MIBC. The small DL-CNN was trained on their own database of 369 patients. In contrast, they developed eight well-known models that were pretrained on the ImageNet dataset. The results show that the possibility of overfitting for the small-CNN is minimized with a sensitivity of 0.722 and a specificity of 1.000. This may be because of the relatively low complexity of the model. Among the eight pretrained DL-CNNs, VGG16, VGG19, etc. (56) showed high performance, with an AUROC of 0.997-0.762. In general, DL-CNNs can achieve a favorable performance. However, in this study, an additional artificial enhancement step was required before the data were fed into the DL-CNN model rather than being fully automatic. This prevents the fully automated processing capability of DL from being fully exploited. Zhang et al. (53) used CT urography images of 441 patients from two medical centers to predict the muscular invasiveness of BCa. To date, this is a rare multicenter study of DL in BCa with a large dataset. The model is based on a novel 3D DL-CNN, a Filter-guided Pyramid Network (FGP-Net) (57). Dense blocks were applied to the network to enhance the transmission of features and alleviate vanishing-gradient problems, and discriminative filter learning (DFL) modules were used to enhance the mid-level representation by learning a bank of convolutional filters that capture class-specific discriminative patches. The network adopted a 2-channel input, and the input data consisted of a vertical superposition of the original and masked tumor regions. They compared the evaluation results of the model with those of two radiologists. Notably, they applied an external cohort evaluation to assess performance more

TABLE 2 Studies using deep learning approach for bladder cancer diagnosis and staging.

| Author | Year | Modality | Number of patients (Train/Val/Test) | CNN structure | Performance (validation or testing dataset) |
|---|---|---|---|---|---|
| Yang et al. (52) | 2021 | CT | 369 patients,1200 images (70%/15%/15%) | A small convolutional network contains four conv_layer +max_pooling_layer stages/eight pretrained models, 2D | Accuracy (small) = 0.861<br>AUROC (small) = 0.998<br>Accuracy (VGG16) = 0.939<br>AUROC (VGG16) = 0.997 |
| Zhang et al. (53) | 2021 | CT | 183/110/73 (internal)/75 (external) | FGP-Net (a novel convolutional network contains Dense Blocks and DFL modules), 3D | AUC (internal) = 0.861<br>Accuracy (internal) = 0.795<br>AUC (external) = 0.791<br>Accuracy (external) = 0.747 |
| Liu et al. (54) | 2022 | T2W MRI | 51/8/16 | ResNet18 with the super-resolution module and the Non-local attention module, 2D | Sensitivity = 94.74 |
| Taguchi et al. (55) | 2021 | T2W MRI | 68 | The denoising Deep Learning Reconstruction (dDLR) | – |

AUC, area under curve; Sensitivity=TP/(TP+ FN).

rigorously (58). Although its final performance is not satisfactory and needs to be improved, the DL model can obtain slightly better, more objective, and more stable results compared with the results of the two radiologists. However, the objective results had another advantages. Radiologists may subjectively improve tumor staging in some ambiguous patients because of concerns about the negative consequences of losing MIBC, which may help in early clinical intervention. Liu et al. (54) adopted the ResNet18 (59) network for the diagnosis and staging of BCa based on MRI. They applied the super-resolution module and non-local attention module to improve the quality of MRI images and enhance the model's ability to perceive features at longer distances.

In addition to diagnosis, DL can be used to improve other parts of the imaging workflow, such as removing image noise and indirectly improving diagnostic capabilities in conjunction with other systems. The vesical imaging reporting and data system (VI-RADS) (60) is a tool for evaluating BCa staging using MRI images. Taguchi et al. (55) used a convolutional neural network to improve the signal-to-noise ratio in high-spatial-resolution images. Although they did not develop the network themselves, this study also showed the potential of DL in assisting in BCa diagnosis.

## Deep learning in bladder cancer treatment assessment

Neoadjuvant chemotherapy has been shown to improve overall survival for patients with BCa (61). However, not all patients benefit from neoadjuvant treatment and instead suffer from severe side effects (62). Therefore, it is important to assess changes in tumor size and treatment response early to help doctors make personalized treatment plans. Nevertheless, there are two major problems with the current clinical treatment assessment. First, although accurate, surgery may not be appropriate for patients undergoing chemotherapy. Second, the current World Health Organization (WHO) criteria (63) and Response Evaluation Criteria in Solid Tumors (RECIST) (64) are inaccurate. Neither set of criteria address three-dimensional (3D) measurements, and the results are heavily influenced by observer experience, especially for tumors with complex and irregular shapes (65). At the same time, because organs and tumors are not rigid bodies, they will have different deformations in the human body, making the design of direct networks for ML very difficult. These problems make ordinary ML methods not particularly adaptable, and therefore drive the progress of DL methods in this field. DL has been recognized as a powerful tool to solve these problems (Table 3).

Cha et al. (25) used the network they developed to segment and measure the gross tumor volume (GTV) from CT images to predict treatment response. As described in the bladder segmentation section, classification-based networks cannot accurately segment tumors because of their limitations, particularly those that shrink after treatment. Their DL-CNN was comparable to radiologists' manual predictions. In 2017, Cha et al. (66) developed a DL-CNN with a structure similar to that in previous studies. However, DL-CNN was used to predict the response to neoadjuvant chemotherapy in this study. They first used their auto-initialized cascaded level set (AI-CALS) (69) system to segment the tumor region. They then paired ROIs extracted from pre- and post-treatment tumor regions of the same patient's scans to form 6700 image pairs. They compared the model with two radiomic feature-based approaches. Owing to their relatively simple DL-CNN structure, the three methods they tested achieved similar results and were also similar to the manual methods. However, it also demonstrates the potential of DL techniques in predicting the treatment response. In 2019, Wu et al. (67) developed seven DL-CNNs based on a previous study (66) and adopted the same image-processing method (66). They modified the filter size, filter stride, and padding type of convolutions and max pooling performed in layers C1 and C2 to develop three different models, and developed two models by freezing different layers. Furthermore, they pretrained the model on the CIFAR10 (70) image set. Only one network variation (DL-CNN-2, C1 convolution filter stride 1→2, C2 max pooling size 3×3→2×2, stride 2→1) exhibited significant performance improvements. The performance of the DL-CNN generally decreased as more layers were frozen, but there was a slight improvement in performance when the C1 layers were frozen. This may be because the subsequent layers are designed to capture more specific features, such as bladder lesions. The pretrained network achieved better performance, but it was better to pretrain with data related to the training images. Overall, they demonstrated that the use of DL-CNN can match or even exceed the level of doctors, and using deeper DL-CNN models and making more effective adjustments to network structures can further improve its performance in the future. Recently, Cha et al. (68) developed a computerized CT-based decision-support system for MIBC treatment response assessment (CDSS-T) based on their previous work (56). They followed the segmentation system and their previously developed DL-CNN combined with a radiomics assessment model. A combined score from the DL-CNN and radiomic model was used to assist physicians in the assessment of the treatment response. With the help of the CDSS-T, 12 physicians improved the assessment accuracy for evaluating the neoadjuvant chemotherapy response in MIBC. This is the first observer study to use a CAD system for this purpose. Interestingly, the accuracy rate of the CDSS-T alone was higher than that of using CDSS-T to assist physicians in assessment. This shows that doctors' experience and trust in using the system still needs to be cultivated, which is also one of the key issues to be overcome in the future clinical application of DL.

TABLE 3 Studies using deep learning approach for bladder cancer treatment.

| Author | Year | Modality | Number of patients (Train/Val/Test) | CNN structure | Performance (validation or testing dataset) |
|---|---|---|---|---|---|
| Cha et al. (25) | 2016 | CT | 62, LOOCV | A network contains 2 convolution layers, 2 locally connected layers, and 1 fully connected layer. | AUC = 0.73 |
| Cha et al. (66) | 2017 | CT | 82 | A network contains 2 convolution layers, 2 locally connected layers, and 1 fully connected layer. Each layer contains 16 kernals. | AUC = 0.73 |
| Wu et al. (67) | 2019 | CT | 73/9/41 | The basic network contains 2 convolution layers, 2 locally connected layers, and 1 fully connected layer. | AUC (basic-random weights) = 0.73<br>AUC (basic-pretrained weights) = 0.79<br>AUC (DL-CNN-1) = 0.72<br>AUC (DL-CNN-2) = 0.86<br>AUC (DL-CNN-3) = 0.69<br>AUC (C1 Frozen) = 0.81<br>AUC (C1,C2 Frozen) = 0.78<br>AUC (C1,C2,L3 Frozen) = 0.71 |
| Cha et al. (68) | 2019 | CT | 123, LOOCV | DL-CNN with a radiomics assessment model | AUC (CDSS-T only) = 0.80<br>AUC (with CDSS-T) = 0.77<br>AUC (No CDSS-T) = 0.74 |

AUC, area under the curve; LOOCV, leave-one-out cross-validation.

## Challenges and future directions

DL is a state-of-the-art technology and popular research area in medical imaging. Its performance is comparable to that of human experts in many studies and applications and it has good development prospects and potential (71). However, research on DL in BCa is still in its infancy, and there are still many shortcomings compared to other fields with mature applications.

### For data

The imaging diagnosis of BCa by clinicians often requires the integration of various imaging data, such as CT and different sequences of MRI images. Although CT is the most commonly used imaging technique for the diagnosis of BCa, MRI has been shown to be more effective, especially in staging, because of the increased soft-tissue contrast resolution. Diffusion-weighted imaging (DWI) and dynamic contrast enhancement (DCE) are far more useful for assessing tumor invasiveness and infiltration into surrounding structures. However, most of the current DL studies on BCa imaging still use CT as the original data. Moreover, all studies using MRI have chosen T2WI sequences, and there is a lack of studies on DWI and DCE sequences. Combining DL with the most appropriate as well as the most advanced techniques in BCa imaging will be a research direction. In addition, based on CT or MRI, most data currently used in BCa studies focus on only one modality of medical imaging. In recent years, many studies have shown that processing multiple modalities simultaneously can significantly improve the performance of DL models (26, 72, 73).

We can also attempt to improve performance by combining imaging-based assessment with other available clinical data, such as genomics and pathology. Multimodal approaches have been shown to outperform unimodal ones (74). In fact, in both natural and medical image processing, multimodal fusion is becoming a mainstream and effective trend. BCa are heterogeneous at the molecular level, and different molecular classifications may be useful to stratify patients for prognosis or response to treatment. The inclusion of multimodal information helps to complement the shortcomings of BCa imaging in these areas. However, due to various reasons, such as the small number of BCa open datasets, there are not many multi-modality processing methods used in the research of DL in BCa. In addition, the limited quantity of medical image data restricts the development of DL. The amount of data significantly affects the performance of DL models. Transfer learning (75) and data augmentation can improve performance to some extent, but they cannot replace the need for a large dataset. To date, the datasets of many studies of DL in BCa have been so small that they do not even have independent validation or test sets, which biases the assessment of the model performance. In addition, the different scanning methods and equipment adopted by different hospitals make the established models difficult to use across institutions, which also limits the clinical application of DL. In this case, it is necessary to use semi-supervised or self-supervised methods to process data. However, the application of these methods for BCa is limited, highlighting the need for future research. In this case, we expect increasing data diversity, multimodal methods, and more comprehensive BCa datasets including multi-center data or a nationwide BCa imaging database to significantly advance the field.

### For algorithm

Most of the DL models used in the current research only stay in the application of existing networks and lack optimization of the imaging characteristics of BCa. The BCa data have many unique structures, including their unique geometry, empty structure, and other characteristics. However, in the current research field on BCa, these characteristics are not well utilized. Compared with other ML methods, DL is a complex black box. To optimize this model in the future, it is important to reflect doctors' ideas and experiences in the diagnosis and treatment of diseases in the DL model and improve its interpretability. Only when the doctor can understand the reason why the DL model makes the assessment can the model better assist the doctor in decision-making. Furthermore, many state-of-the-art results in the field of DL, such as self-supervised learning, pre-training models, transformers, and contrastive learning, have not yet been applied in the field of BCa research, which could be the subject of our future research.

### For application

There are many application scenarios and research directions of DL that people can explore in BCa. For example, there are various pathological types of BCa, including urothelial carcinoma and squamous cell carcinoma. NMIBC and MIBC can also be divided into many molecular subtypes according to the MD Anderson Cancer Center (MDA) (76), Cancer Genome Atlas (TCGA) (77), and other classification criteria. Based on the above criteria, a more complex classification of BCa can be attempted using medical imaging. In addition, DL can be used to predict patient prognosis through medical imaging. Whether DL can predict the outcome of surgical treatment for BCa or be applied to ROI extraction, feature extraction, and feature modelling in radiomics remains unclear. At present, a large amount of research is still focused on image segmentation, and we believe that the development of DL can help doctors in more ways.

## Conclusions

This study reviews the applications of DL in BCa imaging. As a potential technology, DL has extremely broad application prospects in BCa. Limited by the small number of studies in this field, we provide a detailed review of the existing studies, but lack more evidence to demonstrate more possibilities of DL in BCa imaging. However, in the era of increasing emphasis on precision medicine and individualized diagnosis and treatment, how to give full play to the advantages of DL and transform it into a means that can effectively help physicians in clinical diagnosis and treatment will be the direction of our future research. The powerful potential demonstrated by DL is expected to bring about a new revolution in BCa management.

## Author contributions

ML conceived and reviewed the manuscript. ZJ helped with the revising and provided insightful suggestions on the manuscript. All authors have contributed to the manuscript and approved the submitted version.

## Funding

This work was supported by a project of the National Facility for Translational Medicine (Shanghai) (TMSK-2021-118).

## 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

# References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660

2. Chang SS, Bochner BH, Chou R, Dreicer R, Kamat AM, Lerner SP, et al. Treatment of non-metastatic muscle-invasive bladder cancer: Aua/Asco/Astro/Suo guideline. *J Urol* (2017) 198(3):552–9. doi: 10.1016/j.juro.2017.04.086

3. Milowsky MI, Rumble RB, Booth CM, Gilligan T, Eapen LJ, Hauke RJ, et al. Guideline on muscle-invasive and metastatic bladder cancer (European association of urology guideline): American society of clinical oncology clinical practice guideline endorsement. *J Clin Oncol* (2016) 34(16):1945–52. doi: 10.1200/JCO.2015.65.9797

4. Lenis AT, Lec PM, Chamie K, Mshs MD. Bladder cancer: A review. *JAMA* (2020) 324(19):1980–91. doi: 10.1001/jama.2020.17598

5. Botteman MF, Pashos CL, Redaelli A, Laskin B, Hauser R. The health economics of bladder cancer: A comprehensive review of the published literature. *Pharmacoeconomics* (2003) 21(18):1315–30. doi: 10.1007/BF03262330

6. Svatek RS, Hollenbeck BK, Holmang S, Lee R, Kim SP, Stenzl A, et al. The economics of bladder cancer: Costs and considerations of caring for this disease. *Eur Urol* (2014) 66(2):253–62. doi: 10.1016/j.eururo.2014.01.006

7. Wong VK, Ganeshan D, Jensen CT, Devine CE. Imaging and management of bladder cancer. *Cancers (Basel)* (2021) 13(6):1396. doi: 10.3390/cancers13061396

8. Jang HJ, Cho KO. Applications of deep learning for the analysis of medical data. *Arch Pharm Res* (2019) 42(6):492–504. doi: 10.1007/s12272-019-01162-9

9. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* (2015) 521(7553):436–44. doi: 10.1038/nature14539

10. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. *Nat Med* (2019) 25(1):24–9. doi: 10.1038/s41591-018-0316-z

11. Ravi D, Wong C, Deligianni F, Berthelot M, Andreu-Perez J, Lo B, et al. Deep learning for health informatics. *IEEE J BioMed Health Inform* (2017) 21(1):4–21. doi: 10.1109/JBHI.2016.2636665

12. Wainberg M, Merico D, Delong A, Frey BJ. Deep learning in biomedicine. *Nat Biotechnol* (2018) 36(9):829–38. doi: 10.1038/nbt.4233

13. Zou J, Huss M, Abid A, Mohammadi P, Torkamani A, Telenti AA. Primer on deep learning in genomics. *Nat Genet* (2019) 51(1):12–8. doi: 10.1038/s41588-018-0295-5

14. Smistad E, Falch TL, Bozorgi M, Elster AC, Lindseth F. Medical image segmentation on gpus–a comprehensive review. *Med Image Anal* (2015) 20(1):1–18. doi: 10.1016/j.media.2014.10.012

15. Bandyk MG, Gopireddy DR, Lall C, Balaji KC, Dolz J. Mri and ct bladder segmentation from classical to deep learning based approaches: Current limitations and lessons. *Comput Biol Med* (2021) 134:104472. doi: 10.1016/j.compbiomed.2021.104472

16. Hammouda K, Khalifa F, Soliman A, Abdeltawab H, Ghazal M, Abou El-Ghar M, et al. A 3d cnn with a learnable adaptive shape prior for accurate segmentation of bladder wall using Mr images. *2020 IEEE 17th Int Symposium Biomed Imaging (Isbi 2020)* (2020) 935–8. doi: 10.1109/ISBI45749.2020.9098733

17. Duan C, Liang Z, Bao S, Zhu H, Wang S, Zhang G, et al. A coupled level set framework for bladder wall segmentation with application to Mr cystography. *IEEE Trans Med Imaging* (2010) 29(3):903–15. doi: 10.1109/TMI.2009.2039756

18. Qin X, Li X, Liu Y, Lu H, Yan P. Adaptive shape prior constrained level sets for bladder Mr image segmentation. *IEEE J BioMed Health Inform* (2014) 18 (5):1707–16. doi: 10.1109/JBHI.2013.2288935

19. Duan C, Yuan K, Liu F, Xiao P, Lv G, Liang Z. An adaptive window-setting scheme for segmentation of bladder tumor surface *Via* Mr cystography. *IEEE Trans Inf Technol BioMed* (2012) 16(4):720–9. doi: 10.1109/TITB.2012.2200496

20. Cha KH, Hadjiiski L, Samala RK, Chan HP, Caoili EM, Cohan RH. Urinary bladder segmentation in ct urography using deep-learning convolutional neural network and level sets. *Med Phys* (2016) 43(4):1882. doi: 10.1118/1.4944498

21. Gsaxner C, Pfarrkirchner B, Lindner L, Pepe A, Roth PM, Egger J, et al. Pet-train: Automatic ground truth generation from pet acquisitions for urinary bladder segmentation in ct images using deep learning. *BioMed Eng Int Conf* (2018) 1–5. doi: 10.1109/BMEiCON.2018.8609954

22. Xu X, Zhou F, Liu B. Automatic bladder segmentation from ct images using deep cnn and 3d fully connected crf-rnn. *Int J Comput Assist Radiol Surg* (2018) 13 (7):967–75. doi: 10.1007/s11548-018-1733-7

23. Gordon MN, Hadjiiski LM, Cha KH, Samala RK, Chan HP, Cohan RH, et al. Deep-learning convolutional neural network: Inner and outer bladder wall segmentation in ct urography. *Med Phys* (2019) 46(2):634–48. doi: 10.1002/mp.13326

24. Ma X, Hadjiiski LM, Wei J, Chan HP, Cha KH, Cohan RH, et al. U-Net based deep learning bladder segmentation in ct urography. *Med Phys* (2019) 46 (4):1752–65. doi: 10.1002/mp.13438

25. Cha KH, Hadjiiski LM, Samala RK, Chan HP, Cohan RH, Caoili EM, et al. Bladder cancer segmentation in ct for treatment response assessment: Application of deep-learning convolution neural network-a pilot study. *Tomography* (2016) 2 (4):421–9. doi: 10.18383/j.tom.2016.00184

26. Dolz J, Xu X, Rony J, Yuan J, Liu Y, Granger E, et al. Multiregion segmentation of bladder cancer structures in mri with progressive dilated convolutional networks. *Med Phys* (2018) 45(12):5482–93. doi: 10.1002/mp.13240

27. Liu JX, Liu LB, Xu BL, Hou XX, Liu BZ, Chen X, et al. Bladder cancer multi-class segmentation in mri with pyramid-in-Pyramid network. *I S BioMed Imaging* (2019) 28–31. doi: 10.1109/ISBI.2019.8759422

28. Hammouda K, Khalifa F, Soliman A, Ghazal M, Abou El-Ghar M, Haddad A, et al. A deep learning-based approach for accurate segmentation of bladder wall using Mr images. *IEEE Conf Imaging Sy* (2019) 1–6. doi: 10.1109/IST48021.2019.9010233

29. Yu J, Cai L, Chen C, Fu X, Wang L, Yuan B, et al. Cascade path augmentation unet for bladder cancer segmentation in mri. *Med Phys* (2022) 49 (7):4622–31. doi: 10.1002/mp.15646

30. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep Convolutional Neural networks. *Adv Neural Inf Process Syst* (2012) 2:1097–105.

31. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. *Lect Notes Comput Sc* (2015) 9351:234–41. doi: 10.1007/978-3-319-24574-4_28

32. Paszke A, Chaurasia A, Kim S, Culurciello E. ENet: A deep neural network architecture for real-time semantic segmentation. *arXiv preprint arXiv* (2016) 1606:02147. doi: 10.48550/arXiv.1606.02147

33. Romera E, Alvarez JM, Bergasa LM, Arroyo R. Erfnet: Efficient residual factorized convnet for real-time semantic segmentation. *IEEE T Intell Transp* (2018) 19(1):263–72. doi: 10.1109/Tits.2017.2750080

34. Badrinarayanan V, Kendall A, Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans Pattern Anal Mach Intell* (2017) 39(12):2481–95. doi: 10.1109/TPAMI.2016.2644615

35. Liu S, Qi L, Qin HF, Shi JP, Jia JY. Path aggregation network for instance segmentation. *Proc Cvpr IEEE* (2018) 8759–68. doi: 10.1109/Cvpr.2018.00913

36. Ni ZL, Bian GB, Zhou XH, Hou ZG, Xie XL, Wang C, et al. Raunet: Residual attention U-net for semantic segmentation of cataract surgical instruments. *Lect Notes Comput Sc* (2019) 11954:139–49. doi: 10.1007/978-3-030-36711-4_13

37. Kamnitsas K, Ledig C, Newcombe VFJ, Simpson JP, Kane AD, Menon DK, et al. Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation. *Med Image Anal* (2017) 36:61–78. doi: 10.1016/j.media.2016.10.004

38. Li LH, Wang ZG, Li X, Wei XZ, Adler HL, Huang W, et al. A new partial volume segmentation approach to extract bladder wall for computer aided detection in virtual cystoscopy. *Pro BioMed Opt Imag* (2004) 5(23):199–206. doi: 10.1117/12.535913

39. Li LH, Liang ZR, Wang S, Lu HY, Wei XZ, Wagshul M, et al. Segmentation of multispectral bladder Mr images with inhomogeneity correction for virtual cystoscopy. *Proc Spie* (2008) 6916. doi: 10.1117/12.769914

40. Chi JW, Brady SM, Moore NR, Schnabel JA. Segmentation of the bladder wall using coupled level set methods. *I S BioMed Imaging* (2011) 1653–6. doi: 10.1109/ISBI.2011.5872721

41. Garnier C, Ke W, Dillenseger JL. Bladder segmentation in mri images using active region growing model. *IEEE Eng Med Bio* (2011, 5702–5. doi: 10.1109/IEMBS.2011.6091380

42. Ma Z, Jorge RN, Mascarenhas T, Tavares JMRS. Novel approach to segment the inner and outer boundaries of the bladder wall in T2-weighted magnetic resonance images. *Ann BioMed Eng* (2011) 39(8):2287–97. doi: 10.1007/s10439-011-0324-3

43. Chai XF, van Herk M, Betgen A, Hulshof M, Bel A. Automatic bladder segmentation on cbct for multiple plan art of bladder cancer using a patient-specific bladder model. *Phys Med Biol* (2012) 57(12):3945–62. doi: 10.1088/0031-9155/57/12/3945

44. Han H, Li LH, Duan CJ, Zhang H, Zhao Y, Liang ZR. A unified em approach to bladder wall segmentation with coupled level-set constraints. *Med Image Anal* (2013) 17(8):1192–205. doi: 10.1016/j.media.2013.08.002

45. Xiao D, Zhang GP, Liu Y, Yang ZY, Zhang X, Li LH, et al. 3d detection and extraction of bladder tumors *Via* Mr virtual cystoscopy. *Int J Comput Ass Rad* (2016) 11(1):89–97. doi: 10.1007/s11548-015-1234-x

46. Pinto JR. Tavares JMRS. a versatile method for bladder segmentation in computed tomography two-dimensional images under adverse conditions. *P I Mech Eng H* (2017) 231(9):871–80. doi: 10.1177/0954411917714294

47. Xu XP, Zhang X, Liu Y, Tian Q, Zhang GP, Yang ZY, et al. Simultaneous segmentation of multiple regions in 3d bladder mri by efficient convex optimization of coupled surfaces. *Lect Notes Comput Sc* (2017) 10667:528–42. doi: 10.1007/978-3-319-71589-6_46

48. Nishiyama H, Habuchi T, Watanabe J, Teramukai S, Tada H, Ono Y, et al. Clinical outcome of a Large-scale multi-institutional retrospective study for locally advanced bladder cancer: A survey including 1131 patients treated during 1990-2000 in Japan. *Eur Urol* (2004) 45(2):176–81. doi: 10.1016/j.eururo.2003.09.011

49. Babjuk M, Bohle A, Burger M, Capoun O, Cohen D, Comperat EM, et al. Eau guidelines on non-Muscle-Invasive urothelial carcinoma of the bladder: Update 2016. *Eur Urol* (2017) 71(3):447–61. doi: 10.1016/j.eururo.2016.05.041

50. Alfred Witjes J, Lebret T, Comperat EM, Cowan NC, De Santis M, Bruins HM, et al. Updated 2016 eau guidelines on muscle-invasive and metastatic bladder cancer. *Eur Urol* (2017) 71(3):462–75. doi: 10.1016/j.eururo.2016.06.020

51. Ge L, Chen Y, Yan C, Zhao P, Zhang P, A R, et al. Study progress of radiomics with machine learning for precision medicine in bladder cancer management. *Front Oncol* (2019) 9:1296. doi: 10.3389/fonc.2019.01296

52. Yang Y, Zou X, Wang Y, Ma X. Application of deep learning as a noninvasive tool to differentiate muscle-invasive bladder cancer and non-Muscle-Invasive bladder cancer with ct. *Eur J Radiol* (2021) 139:109666. doi: 10.1016/j.ejrad.2021.109666

53. Zhang G, Wu Z, Xu L, Zhang X, Zhang D, Mao L, et al. Deep learning on enhanced ct images can predict the muscular invasiveness of bladder cancer. *Front Oncol* (2021) 11:654685. doi: 10.3389/fonc.2021.654685

54. Liu D, Wang S, Wang J. The effect of ct high-resolution imaging diagnosis based on deep residual network on the pathology of bladder cancer classification and staging. *Comput Methods Programs BioMed* (2022) 215:106635. doi: 10.1016/j.cmpb.2022.106635

55. Taguchi S, Tambo M, Watanabe M, Machida H, Kariyasu T, Fukushima K, et al. Prospective validation of vesical imaging-reporting and data system using a next-generation magnetic resonance imaging scanner-is denoising deep learning reconstruction useful? *J Urol* (2021) 205(3):686–92. doi: 10.1097/JU.0000000000001373

56. Simonyan K, Zisserman A. Very deep convolutional networks for Large-scale image recognition. *arXiv preprint arXiv* (2014) 1409:1556. doi: 10.48550/arXiv.1409.1556

57. Huang C, Lv W, Zhou C, Mao L, Xu Q, Li X, et al. Discrimination between transient and persistent subsolid pulmonary nodules on baseline ct using deep transfer learning. *Eur Radiol* (2020) 30(12):6913–23. doi: 10.1007/s00330-020-07071-6

58. Kleppe A, Skrede OJ, De Raedt S, Liestol K, Kerr DJ, Danielsen HE. Designing deep learning studies in cancer diagnostics. *Nat Rev Cancer* (2021) 21 (3):199–211. doi: 10.1038/s41568-020-00327-9

59. He KM, Zhang XY, Ren SQ, Sun J. Deep residual learning for image recognition. *2016 IEEE Conf Comput Vision Pattern Recognition (Cvpr)* (2016) 770–8. doi: 10.1109/Cvpr.2016.90

60. Panebianco V, Narumi Y, Altun E, Bochner BH, Efstathiou JA, Hafeez S, et al. Multiparametric magnetic resonance imaging for bladder cancer: Development of vi-rads (Vesical imaging-reporting and data system). *Eur Urol* (2018) 74(3):294–306. doi: 10.1016/j.eururo.2018.04.029

61. Yin M, Joshi M, Meijer RP, Glantz M, Holder S, Harvey HA, et al. Neoadjuvant chemotherapy for muscle-invasive bladder cancer: A systematic review and two-step meta-analysis. *Oncologist* (2016) 21(6):708–15. doi: 10.1634/theoncologist.2015-0440

62. Witjes JA, Wullink M, Oosterhof GO, de Mulder P. Toxicity and results of mvac (Methotrexate, vinblastine, adriamycin and cisplatin) chemotherapy in advanced urothelial carcinoma. *Eur Urol* (1997) 31(4):414–9. doi: 10.1159/000474499

63. World Health Organization. *WHO handbook for reporting results of cancer treatment* . Geneva: World Health Organization. Available at: http://www.who.int/iris/handle/10665/37200 (Accessed September 12, 2016).

64. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: Revised recist guideline (Version 1.1). *Eur J Cancer* (2009) 45(2):228–47. doi: 10.1016/j.ejca.2008.10.026

65. Husband JE, Schwartz LH, Spencer J, Ollivier L, King DM, Johnson R, et al. Evaluation of the response to treatment of solid tumours - a consensus statement of the international cancer imaging society. *Br J Cancer* (2004) 90(12):2256–60. doi: 10.1038/sj.bjc.6601843

66. Cha KH, Hadjiiski L, Chan HP, Weizer AZ, Alva A, Cohan RH, et al. Bladder cancer treatment response assessment in ct using radiomics with deep-learning. *Sci Rep* (2017) 7(1):8738. doi: 10.1038/s41598-017-09315-w

67. Wu E, Hadjiiski LM, Samala RK, Chan HP, Cha KH, Richter C, et al. Deep learning approach for assessment of bladder cancer treatment response. *Tomography* (2019) 5(1):201–8. doi: 10.18383/j.tom.2018.00036

68. Cha KH, Hadjiiski LM, Cohan RH, Chan HP, Caoili EM, Davenport MS, et al. Diagnostic accuracy of ct for prediction of bladder cancer treatment response with and without computerized decision support. *Acad Radiol* (2019) 26(9):1137–45. doi: 10.1016/j.acra.2018.10.010

69. Hadjiiski L, Chan HP, Caoili EM, Cohan RH, Wei J, Zhou C. Auto-initialized cascaded level set (Ai-cals) segmentation of bladder lesions on multidetector row ct urography. *Acad Radiol* (2013) 20(2):148–55. doi: 10.1016/j.acra.2012.08.012

70. Krizhevsky A. *Learning multiple layers of features from tiny images [Master's thesis]*. Toronto: University of Toronto (2009). Available at: http://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf.

71. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. *Annu Rev BioMed Eng* (2017) 19:221–48. doi: 10.1146/annurev-bioeng-071516-044442

72. Dolz J, Gopinath K, Yuan J, Lombaert H, Desrosiers C, Ben Ayed I. Hyperdense-net: A hyper-densely connected cnn for multi-modal image segmentation. *IEEE Trans Med Imaging* (2019) 38(5):1116–26. doi: 10.1109/TMI.2018.2878669

73. Wang L, Nie D, Li G, Puybareau E, Dolz J, Zhang Q, et al. Benchmark on automatic 6-Month-Old infant brain segmentation algorithms: The iseg-2017 challenge. *IEEE Trans Med Imaging* (2019) 38(9):2219–30. doi: 10.1109/TMI.2019.2901712

74. Stahlschmidt SR, Ulfenborg B, Synnergren J. Multimodal deep learning for biomedical data fusion: A review. *Brief Bioinform* (2022) 23(2):bbab569. doi: 10.1093/bib/bbab569

75. Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks ? *Adv Neur In* (2014) 27.

76. Choi W, Porten S, Kim S, Willis D, Plimack ER, Hoffman-Censits J, et al. Identification of distinct basal and luminal subtypes of muscle-invasive bladder cancer with different sensitivities to frontline chemotherapy. *Cancer Cell* (2014) 25 (2):152–65. doi: 10.1016/j.ccr.2014.01.009

77. Robertson AG, Kim J, Al-Ahmadie H, Bellmunt J, Guo G, Cherniack AD, et al. Comprehensive molecular characterization of muscle-invasive bladder cancer. *Cell* (2017) 171(3):540–56.e25. doi: 10.1016/j.cell.2017.09.007

frontiers | Frontiers in Oncology

TYPE Review
PUBLISHED 27 October 2022
DOI 10.3389/fonc.2022.1028493

OPEN ACCESS

EDITED BY
Xin-Wu Cui,
Huazhong University of Science
and Technology, China

REVIEWED BY
Guohui Liu,
Huazhong University of Science
and Technology, China
Mohammad Hojjat-Farsangi,
Karolinska Institutet (KI), Sweden
Zhaogang Teng,
Nanjing University of Posts and
Telecommunications, China

*CORRESPONDENCE
Aixi Yu
yuaixi@whu.edu.cn
Xiang Hu
shawnhu2002@whu.edu.cn

†These authors have contributed
equally to this work

SPECIALTY SECTION
This article was submitted to
Cancer Imaging and
Image-directed Interventions,
a section of the journal
Frontiers in Oncology

RECEIVED 26 August 2022
ACCEPTED 17 October 2022
PUBLISHED 27 October 2022

CITATION
Yi X, Wang Z, Hu X and Yu A (2022)
Affinity probes based on small-molecule inhibitors for tumor imaging.
*Front. Oncol.* 12:1028493.
doi: 10.3389/fonc.2022.1028493

# Affinity probes based on small-molecule inhibitors for tumor imaging

Xinzeyu Yi†, Zheng Wang†, Xiang Hu* and Aixi Yu*

Department of Orthopedics Trauma and Microsurgery, Zhongnan Hospital of Wuhan University, Wuhan, China

Methods for molecular imaging of target areas, including optical imaging, radionuclide imaging, magnetic resonance imaging and other imaging technologies, are helpful for the early diagnosis and precise treatment of cancers. In addition to cancer management, small-molecule inhibitors are also used for developing cancer target probes since they act as the tight-binding ligands of overexpressed proteins in cancer cells. This review aims to summarize the structural designs of affinity probes based on small-molecule inhibitors from the aspects of the inhibitor, linker, dye and radionuclide, and discusses the influence of the modification of these structures on affinity and pharmacokinetics. We also present examples of inhibitor affinity probes in clinical applications, and these summaries will provide insights for future research and clinical translations.

KEYWORDS

**inhibitor, affinity probe, near-infrared, radiotracer, tumor imaging**

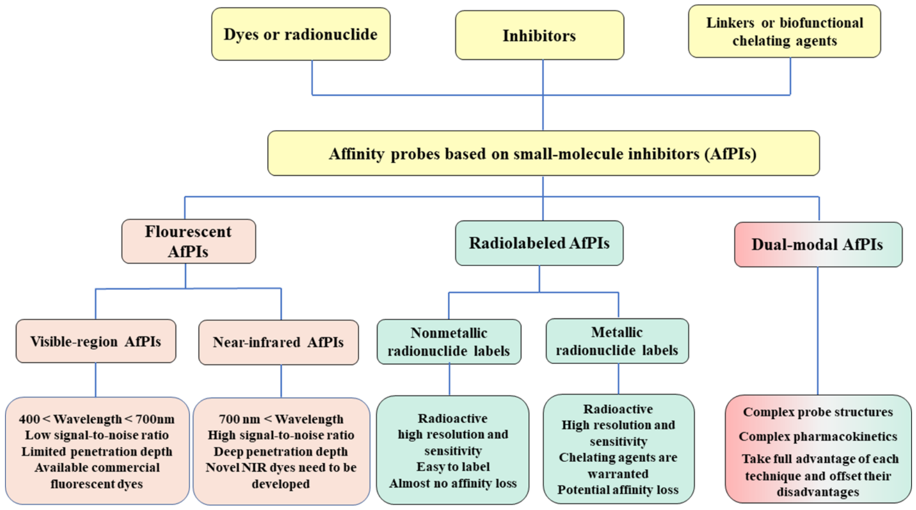
## Introduction

Multiple enzymes and receptor proteins in organisms are involved in life processes such as cell metabolism, proliferation, differentiation, migration, and apoptosis by regulating biochemical reactions or signaling pathways. Small-molecule inhibitors can regulate protein function by reversibly or irreversibly binding with these proteins (1, 2). By specifically binding to highly expressed proteins in cancer cells and producing effects, many small-molecule inhibitors have been used in targeted cancer therapy. Moreover, new targets and subtype-selective inhibitors have also been developed in response to the problems of cancer resistance and potential side effects (3–6). On this basis, affinity probes based on small-molecule inhibitors (AfPIs) for targeted cancer imaging have become research areas of major interest in recent years. Despite their severe metabolic problems, like peptide probes (7, 8), AfPIs not only have the advantages of non-immunogenicity, easy structure modification, fast target recognition, and strong affinity, but also have a broader biodistribution and a higher signal-to-noise ratio than antibody-conjugated probes or peptide probes (9, 10). Hence, they are efficient tools for cancer research and have broad application prospects in early diagnosis, prognosis assessment, surgery navigation and drug delivery monitoring (11, 12).

This review summarizes the tumor-targeting AfPIs emerging in recent years and aims to provide design strategies for developing novel AfPIs. The key challenges and corresponding solutions in the design of such probes are discussed below. Herein, we classify AfPIs into traditional visible-region, near-infrared, radiolabeled and dual-modal probes for comparison. We specifically focus on near-infrared and radiolabeled probes with promising clinical applications, and reveal the characteristics of the two probe types and provide references for future clinical translation. Scheme 1 summarises the classifications of AfPIs and their features.

## Visible-region AfPIs

Fluorescence imaging is an excellent, noninvasive imaging method that allows the visualization of cell status and many biochemical reactions (13, 14). The introduction of inhibitor structures enhances the targeting ability of probes to distinguish cancer from the normal region. This section focuses on the fluorescent AfPIs in the visible region (wavelength below 700 nm), mainly used for targeted imaging of cells or tissues. As shown in Figure 1, the recognition group Polo-like kinase 1 (PLK1) inhibitor **SBE13** (15) was conjugated with linker and coumarin derivatives chosen for fluorophores, forming two kinds of PLK1 affinity probes, **1** and **2**, with emission wavelengths of 480 nm and 660 nm, respectively. Modifying the coumarin structure



SCHEME 1
The classifications of AfPIs and their features. In this review, we classified AfPIs into visible-region, near-infrared, radiolabeled and dual-modal probes, and introduced them from three aspects: the *inhibitors, linkers* and *dyes* or *radionuclides*.

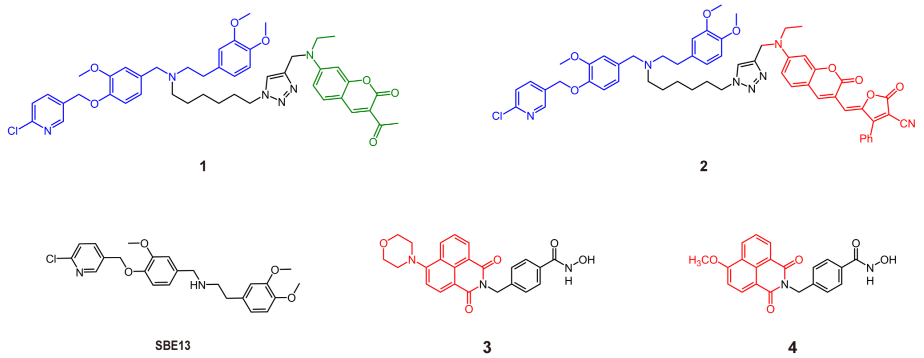


FIGURE 1
Some structures of traditional visible AfPIs and their parent inhibitors (Blue, inhibitor structure; green and red, fluorophores).

in **2** resulted in intramolecular charge transfer (ICT), and a redshift close to the near-infrared region in its emission could be imaged *in vivo* (9). Overexpression of PLK1 in some human tumor cells makes it a target for antitumor drug treatment (16). By binding with PLK1, the probe is concentrated in the PLK1 kinase-rich region to distinguish it from normal regions. Although the imaging effect of **2** was demonstrated *in vivo*, 660 nm is insufficient to meet the needs of *in vivo* detection. In addition, inhibitors with unique structures can also serve as fluorescent moieties; hence, no extra dye conjugation is warranted. For instance, histone deacetylase 6 (HDAC6) inhibitors containing a naphthalimide skeleton, which is intrinsically fluorescent, were synthesized as inhibitor-based affinity probes (**3** and **4**) to detect the expression of HDAC6 in tumor cells (Figure 1) (17, 18). Moreover, there are affinity probes based on the biotin-avidin system that conjugate inhibitors and biotin for proteomic analysis and imaging in cells (19). However, these probes without an OFF-ON function will lead to false positives and phototoxicity because they will be retained in normal tissue regions and release fluorescence. Furthermore, their low signal-to-noise ratio (SNR) blurs the tumor location (20, 21).

Hence, smart probes with an "OFF-ON" design appear more attractive. Because affinity probes bind to proteins directly, the "trigger" can be activated by changing the spatial conformation rather than an enzymatic or chemical reduction stimulus (22). Photoinduced electron transfer (PeT) involves a-PeT and d-PeT processes. In the a-PeT process, the inhibitor provides electrons to the highest occupied molecular orbital (HOMO) of the adjacent fluorophore. In contrast, the fluorophore donates its electrons to the lowest unoccupied molecular orbital (LUMO) of the inhibitor in the d-PeT process. Finally, the electrons in the LUMO of the fluorophore fail to return to the HOMO, resulting in fluorescence quenching (23). When the inhibitor binds to the target, changes in the spatial structure or electronic energy levels will disrupt the process, releasing fluorescence (Figure 2). Based on this principle, Peng et al. (24) used the intramolecular spatial folding effect caused by small-molecule inhibitors and dyes to design the fluorescence probe **5** targeting the Golgi apparatus of cancer cells based on the cyclooxygenase 2 (COX-2) inhibitor indomethacin (IMC). When IMC binds to the amino acid residues Arg120, Tyr355 and Glu522 of the COX-2 molecule, its folded structure is open, and the PeT effect disappears, resulting in the release of fluorescence with a maximum excitation wavelength of 547 nm. Although the two-photon property of the probe has improved its tissue penetration to a certain extent, its emission wavelength still limits its application in biological imaging *in vivo*. Based on 5-bromobenzofuran-2-carboxylic acid, an inhibitor of Pim-1 kinase, Guo designed probe **6** with a PeT effect, whose emission wavelength reached the red light level and achieved live animal imaging of tumor xenograft mice (25). Similar to the COX-2 probe, probe **6** changes from the folded state to the unfolded state by binding with Pim-1 kinase, thereby removing the fluorescence quenching and releasing the fluorescence. Compared with traditional non-OFF-ON probes, this type of probe utilizes the conformational changes of inhibitors and dyes to exhibit a higher SNR, reduce the phototoxicity of nontargeted areas, and significantly reduce the false-positive phenomenon during imaging.

Many commercial fluorescent dyes in the visible region have been developed (26), and less steric hindrance and better pharmacokinetics can be easily obtained by modifying the structure of dyes. However, due to the short wavelength of these probes, it is difficult to obtain good results for *in vivo* imaging, so they are more suitable for qualitative or semiquantitative research at the molecular level and imaging at the level of cells or tissue slices. Designing near-infrared probes with near-infrared dyes is the future trend in the clinical translation of AfPIs.

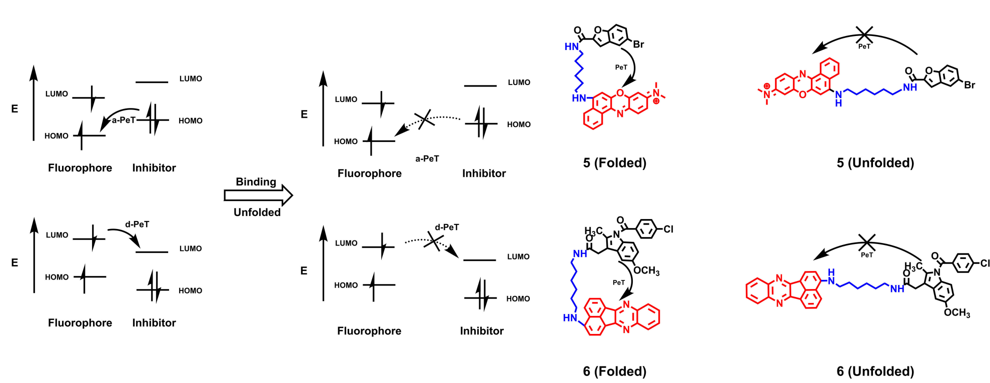


FIGURE 2
The quenching mechanism of PeT effects and AfPIs are designed based on PeT effects. When probes do not bind to the proteins, the fluorescence is quenched by Pet effects. After binding to proteins, the folded structure is open and the PeT effect disappears, resulting in the release of fluorescence (HOMO, highest occupied molecular orbital; LUMO, lowest unoccupied molecular orbital).

## Near-infrared AfPIs

The near-infrared (NIR) band can be roughly divided into near-infrared window I (700-1000 nm) and near-infrared window II (1000-1700 nm) (27) and exhibits a higher penetrating capability than visible fluorescence in tissues. In addition, compared with traditional visible-light imaging, near-infrared imaging is less affected by biological matrix scattering and tissue autofluorescence, which gives it a higher signal-to-noise ratio and better spatial resolution. Therefore, near-infrared imaging is more suitable for *in vivo* imaging, and NIR AfPIs are also ideal for early diagnosis, surgery navigation and photothermal therapy of tumors (28–30). Near-infrared inhibitor probes mainly include three structures: inhibitors, linkers and near-infrared dyes. The influences of these three structures on the affinity and metabolism of the probe and the design strategy of the probe are discussed in the following.

### Inhibitor structure in AfPIs

The presence or absence of the inhibitor structure in the probe and the modification of crucial groups in the inhibitor structure will significantly impact the probe's affinity and selectivity. Taking the monoamine oxidase (MAO) series of probes as an example, MAO is an important enzyme that regulates some biochemical reactions in the body, controlling the metabolism of catecholamines and serotonin. It plays a crucial role in the progression of tumors and Parkinson's disease. MAO contains two isoforms: MAO-A and MAO-B. The original design of the MAO-A targeting probe **7** only contains a fluorophore and propylamine group as the recognition moiety. When propylamine meets MAO, the propylamine group undergoes a continuous oxidation/β-elimination reaction and is removed, releasing free fluorescent groups and producing fluorescence (31). However, this probe shows no subtype selectivity and has insufficient affinity. Based on this probe structure, Wu et al. (32) introduced the structure of the MAO-A selective inhibitor clorgyline to the probe (**8**), which gave the probe higher MAO-A affinity and selectivity. Replacement of the chlorine substituent on the benzene ring, such as the methoxy group (probe **9**), drastically decreased the selectivity of the probe to lower than that of **8** but still higher than that of the previous generation probe **7**, which lacked an inhibitor structure. Similarly, the clorgyline derivative probes **10** and **11** based on the dicyanomethylene-4H-pyran chromophore (DCM) structure developed by Yang et al. (33) had a higher selectivity for MAO-A than MAO-B, with relative fluorescence intensity of approximately 42-fold. However, when the halogen substituent was changed, the affinity of the unsubstituted (H atom) probe **12** decreased slightly, and the selectivity decreased by approximately 20-fold. The resulting product lost selectivity and affinity if it was substituted with methoxy or methyl. Comparing the performance of these probes shows that in addition to the fact that the halogen element chlorine plays a key role in binding, steric hindrance may also have a certain effect. This potential effect is consistent with previous molecular docking results for MAO-A and clorgyline (34). When clorgyline undergoes docking with MAO-A, two chlorine atoms form hydrogen bonds with the Cys323 and Thr326 residues of MAO-A (Figure 3A). These hydrogen bonds help stabilize the binding between the inhibitor and the protein. Wu et al. (35) chose to connect the NIR dye to the other end of the clorgyline to synthesize **13**, protecting two chlorine atoms so that the probe had a more potent antitumor ability than the parent compound. Although the mitochondrial-targeting effect of the NIR dyes here contributes to

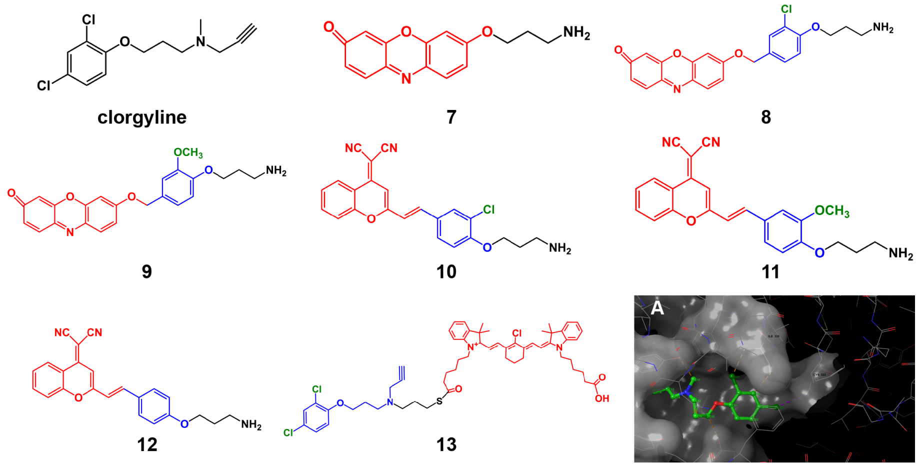


FIGURE 3
Structures of clorgyline and clorgyline-derived AfPIs (Blue, inhibitor structure; red fluorophores). **(A)** Molecular docking shows that clorgyline forms hydrogen bonds between its chlorine atoms and the Cys323 and Thr326 residues of MAO-A (PDB ID: 2BXR).

the antitumor ability, it also illustrates the importance of protecting key groups.

The above studies indicate that the interaction between the targets of some key groups of AfPIs and the steric hindrance of some groups play critical roles in the performance of probe affinity. In designing AfPIs, the groups of the inhibitor that play a vital role in binding to the target must be protected to avoid diminishing the overall affinity of the probe. However, the loss of certain key groups does not necessarily or directly lead to the failure of probe imaging. For example, in the aforementioned OFF-ON probe based on the Pim-1 inhibitor, the carboxyl group on its parent inhibitor structure can form a salt bridge and hydrogen bond with Pim-1 kinase, which is crucial for binding kinase. And when the carboxyl group is destroyed, this will lead to an apparent loss of affinity (36). This result shows that imaging can still be achieved in the case of the loss of some key groups, possibly because the benzene ring still contains a bromine atom to help stabilize the binding, and the OFF-ON imaging mechanism avoids the fluorescence of probes when they are not bound to the kinase. This also illustrates the imaging advantages of OFF-ON probes from another aspect, which can avoid the problem that the tumor cannot be distinguished sufficiently from the surrounding normal tissues due to a loss of affinity.

Containing multiple inhibitor structures or co-targeting through multiple regions can also help probes more easily gather in the target region. Prostate-specific membrane antigen (PSMA), a peripheral glutamate carboxypeptidase, is a biomarker highly expressed by prostate cancer cells. PSMA is located on the cell membrane surface, and its active site faces the outside of the cell; this enzyme has become a common target for AfPIs (37). Its representative inhibitor structure is glutamate-urea-lysine. Based on this structure, the NIR dye can be connected to achieve targeted prostate cancer imaging (38, 39). On this basis, Kwon et al. (40) established two bivalent AfPIs, **15** and **16,** with two GLU units, and these probes exhibited a higher tumor uptake rate than that with only one GLU unit (**14**). Later, 2-nitroimidazole, which has a targeted hypoxia effect, was introduced onto the other end of the structure to synthesize **17** (41) so that the dual-targeting effect of hypoxia and PSMA was achieved without significant loss of the original affinity of PSMA(Figure 4A). There was a partial loss of affinity in compound **18** with the introduction of two 2-nitroimidazole

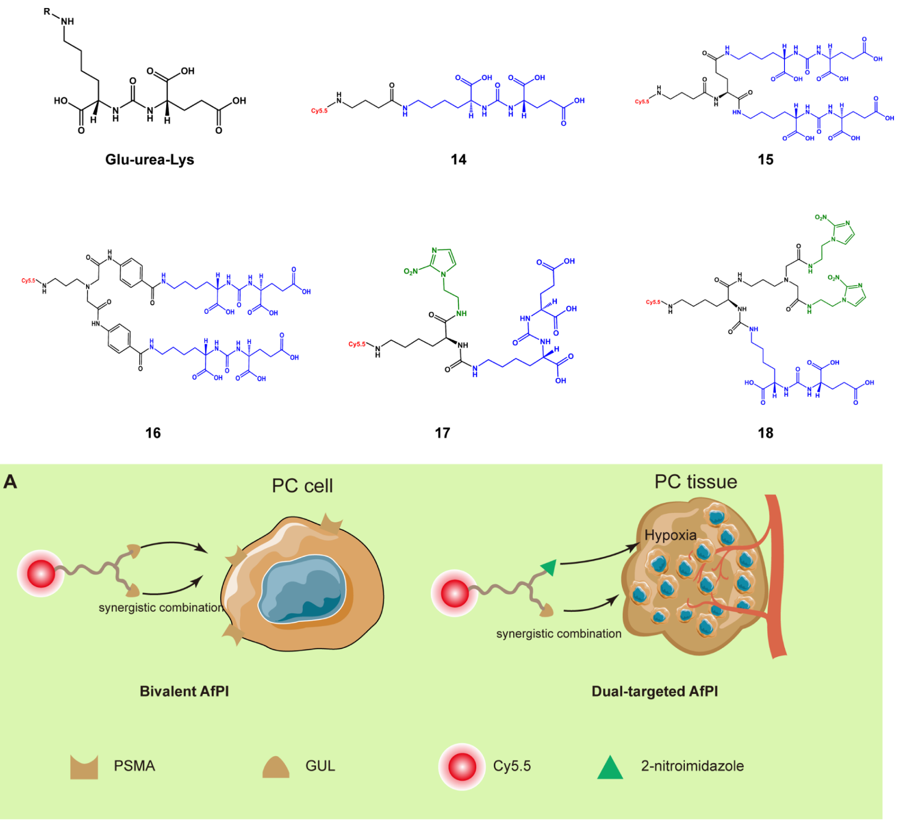


FIGURE 4
Structures of the PSMA inhibitor and its derived AfPIs (Blue, GLU units Green, 2-nitroimidazole group). **(A)** Schematic of bivalent and dual-targeted AfPIs for prostate cancer (PC, prostate cancer).

groups simultaneously, which may be ascribed to the increased steric hindrance. The simultaneous existence of multiple recognition groups further enhances the imaging effect, reducing the false-negative rate and thus identifying tumor regions more clearly. At the same time, attention needs to be paid to the increase in steric hindrance caused by introducing new groups.

When designing a novel AfPI, the factors affecting inhibitor affinity must be considered, and the probe should be designed as a new "inhibitor". For example, when designing the structure of CYP1B1 targeted AfPI, Meng et al. (42) excluded areas bound to the enzyme and made modifications in a relatively safe area *via* molecular docking (Figure 5A). Wang et al. (46) avoided the sulfonamide structure of celecoxib and chose to modify the pyrazole ring position to reduce the loss of affinity. In this approach, determining the inhibitor's crucial structure, attempting to protect these structures in connecting dyes and linkers, and performing molecular dynamics simulation on these structures is conducive to predicting whether the synthesized probe can bind to the target protein.

## Dyes and linkers in AfPIs

Near-infrared dyes can be roughly classified into two types: nonorganic and organic. Nonorganic dyes include single-walled carbon nanotubes, quantum dots, and rare-earth materials (26) (Figure 5B). Similar to antibodies, inhibitors can be introduced into these inorganic dyes through covalent or noncovalent binding to achieve targeted imaging (47, 48), in which

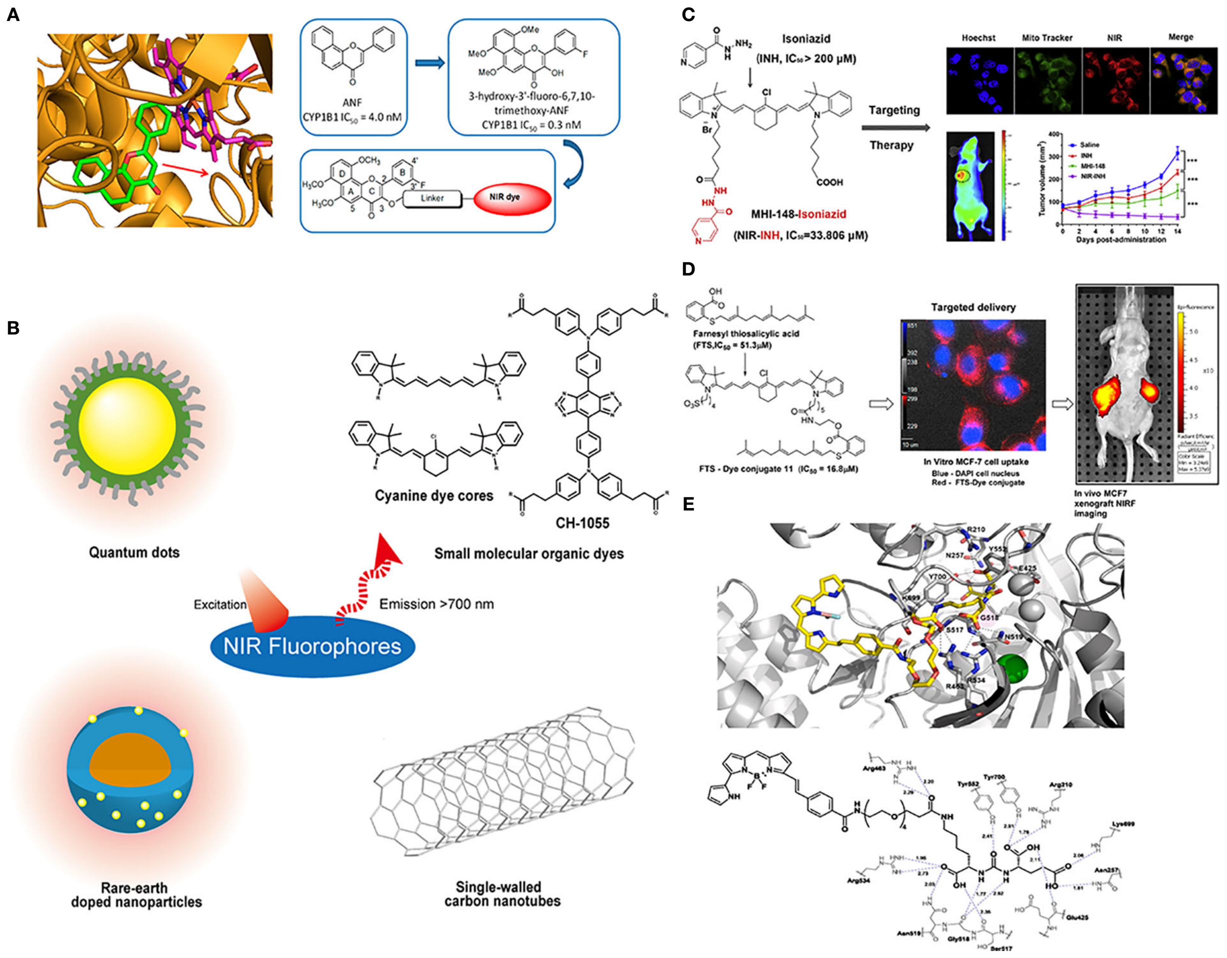


FIGURE 5
Conjugations between inhibitors and NIR fluorophores. **(A)** Conducting molecular docking analysis of CYP1B1 inhibitor and its target before conjugation to avoid the loss of affinity. Reprinted with permission (42). Copyright 2018 American Chemical Society. **(B)** The common types of NIR fluorophores. **(C)** Conjugating with the heptamethine cyanine dye MHI148 can improve the antitumor effect of the MAO-A inhibitor isoniazid. Reprinted with permission (43). Copyright 2018 Elsevier. **(D)** The conjugation of FTS with cancer-targeting heptamethine cyanine dye improved its pharmacological profile. Reprinted with permission (44). Copyright 2017 American Chemical Society. **(E)** Molecular docking results demonstrated a 20 Å tunnel region in PSMA. Reprinted with permission (45). Copyright 2020 Elsevier.

covalent binding is more stable, and these inorganic materials can also be used to deliver targeted drugs to achieve the integration of diagnosis and treatment (49). However, these inorganic materials need to be functionalized in advance (50), and the limitations of water solubility, photothermal stability, immunity uptake and biological clearance in the body must be addressed (51).

Organic dyes have lower molecular weight and higher plasticity, biocompatibility and safety than inorganic dyes. Additionally, some of them have been approved for clinical use or have started in clinical trials, such as indocyanine green (ICG). Some heptamethine cyanine dyes can also preferentially accumulate in the mitochondria of tumor cells through the high expression of organic anion transporter peptides (OATPs) in tumors and the higher transmembrane potential of tumor cells (52, 53), and they can achieve tumor seeking, accumulation and retention *via* the covalent binding of *meso*-chlorine and albumin (54). The conjugation of these dyes and small-molecule inhibitors provides a way to change the pharmacokinetics (55). In addition, the overall properties of organic dyes, such as excitation/emission wavelengths, water-solubility and photostability, can be easily adjusted by chemical modification.

The introduction of dyes and linkers is related to the affinity and pharmacokinetics of the probe, and the differences in some substituents on these dyes will alter the probe metabolism and accumulation of the tumor area. Generally, when choosing dyes, better water solubility and higher emission wavelengths are pursued because these characteristics will be conducive to clinical translation. However, in the process of conjugating dyes, due to steric hindrance or changes in functional groups, the overall affinity of the probe will decrease, which is not conducive to later targeted imaging. Therefore, suitable dyes and synthetic routes should be chosen to avoid loss of affinity. Additionally, better imaging results can be achieved if improvements can be made to synthesize probes that overcome parent inhibitors' deficiencies. Genistein has limited clinical antitumor applications because of its low oral bioavailability and poor pharmacokinetics. Guan et al. (56) conjugated genistein with the near-infrared dye IR-783 and used the advantage that IR-783 could be transported by OATPs and enriched in breast cancer cells to improve its antitumor effect and achieve NIR imaging. Similarly, Lv et al. (43)conjugated the MAO-A inhibitor isoniazid with the heptamethine cyanine dye MHI148 and used its mitochondrial targeting ability to obtain a theranostic probe for prostate cancer (Figure 5C), which showed a more potent antitumor effect than the parent inhibitor isoniazid. Similar designs have been reported in many other studies. When S-trans-trans-farnesyl salicylic acid (FTS), an RAS and mammalian target of rapamycin (mTOR) inhibitor, was connected with the heptamethine cyanine dye, the inhibitory effect on mTOR and antitumor effect of the probe was better than FTS, and the $EC_{50}$ was reduced from 51.3 nm to 16.8 nm (Figure 5D) (44). These results may be ascribed to the fact that the sulfonate group and the tumor-targeting ability of the dye improve its dose distribution. Inorganic dyes can also achieve this effect. Hu (7) combined carbon quantum dots with an ST14 (suppressor of tumorigenicity 14) inhibitor to improve renal clearance and increase the retention of the inhibitor in tumors, which is beneficial for its antitumor effect and imaging. These studies demonstrated that the improved pharmacokinetics ascribed to introducing dyes and linkers could enhance the tumor targeting and antitumor ability of AfPIs.

Although there have been many studies on AfPIs in the first NIR window, AfPIs whose emission wavelength falls in the second NIR window are just emerging (57, 58), and the wavelength of the existing inhibitor probes is generally low, possibly because it is relatively difficult to design novel dyes. To achieve a redshift of the wavelength, extended π conjugation is required (59). After the probe is combined with small molecules, the resulting structure will become more complex, and the binding effect will be more uncontrollable, so existing dyes are conjugated in most studies. Furthermore, when the wavelength of dyes redshifts to the second NIR window, their quantum yields drop sharply (60). Other issues that NIR dyes share, including water solubility and probe biocompatibility, are challenges that still need to be overcome in studying inhibitor NIR-II window probes.

The linker is also critical to the properties of the probes. It can avoid the effect of steric hindrance of the dye on the affinity of the inhibitor and can improve the metabolic kinetics of the probe through modifications, such as with polyethylene glycol (PEG). Taking prostate cancer as an example, Son et al. used the PEG chain as a linker to conjugate 4,4-difluoro-4-bora-3a,4a-diaza-*s*-indacene (BODIPY) and Glu-CO-Lys to construct probes (45). The molecular docking results showed that the PEG linker was located in the tunnel region, with a length of approximately 20 Å (Figure 5E), which is consistent with previous findings (61). This design allows the entire fluorophore molecule to be outside the target molecule and avoids steric hindrance caused by the introduction of the bulky dye. PEG improves the water solubility and biocompatibility of the probe and eliminates adverse effects of the lipophilic dye BODIPY so that its metabolic kinetics *in vivo* are improved, and the overall affinity is also ensured. When Kwon et al. (40) attempted to change the glutamine structure in the linker to a benzene structure to obtain **16** based on the structure of **15**, the probe showed slower clearance and lower affinity than **15** because of the introduction of the benzene ring structure on the linker. The same is true for the principle of designing radionuclide probes and modifications in the linker area can significantly improve the tumor uptake rate and *in vivo* pharmacokinetics (37, 62).

## Clinical applications

The probe tool should be based on actual clinical problems and converted into clinical applications, which is our ultimate goal in designing AfPIs. Zhu et al. (63) used two AfPIs to perform dual-target imaging of BCL2 and MDM2, simultaneously detecting the activities/expression of apoptosis markers. Arlauckas et al. (64) designed and synthesized JAS239, a novel AfPIs targeting choline kinase alpha (ChoKα), and realized the goals of breast cancer imaging, antitumor therapy and monitoring choline metabolism in breast cancer. Osada et al. (65) took heat shock protein 90 (Hsp90) as the target and used the inhibitor SNX-5422 to connect the near-infrared dye with the PEG chain as the linker to image the target area of the subtype estrogen receptor-positive luminal invasive lobular carcinoma. Their study was representative of the use of imaging to detect a histological subtype that is difficult to diagnose early. This application reflects the advantages of inhibitor probe imaging at the molecular level, which can achieve subtype classification and higher sensitivity than traditional imaging examination (66). It is also possible to use heat shock protein inhibitors to target and inhibit the overexpression of heat shock proteins in tumor cells, thereby enhancing the efficiency of NIR photothermal therapy (67). In addition, there are applications such as surgery navigation and postoperative reconfirmation of the tumor area (57, 68). The design of these probes is based on an actual clinical problem rather than simple imaging and diagnosis of tumors, so they have a promising application prospect in the clinic.

This section mainly discusses the three key elements, inhibitors, dyes and linkers, and their novel applications in the design of NIR AfPIs, with MAO and PSMA inhibitors as examples. Each element may have a significant impact on the fundamental properties of the probe. When designing the structure, not only the properties of the three elements but also the interactions between them must be considered to improve the pharmacokinetics and avoid adverse effects such as decreased affinity caused by the increased steric hindrance.

# Radiolabeled AfPIs

According to the imaging principle, radiolabeled AfPIs can be classified into single-photon-emission computed tomography (SPECT) and positron emission tomography (PET) probes. Compared with SPECT, PET has a lower radiation dose and higher resolution and sensitivity, but the high costs limit its application in primary medical institutions (69). SPECT probes can provide longer image acquisition time due to a longer half-life (a few hours to a few days). Unlike PET, which emits two 511-eV photons, SPECT probes can emit photons with different energies, allowing multiple probes to be imaged simultaneously (70, 71). Our focus is on the imaging effect of radioactive probes based on inhibitors, and due to the differences in radionuclides, the design ideas of the probes will differ significantly. Radioactive elements commonly used in labeling inhibitors include nonmetallic **C**, **F**, **Br** and **I**, while metal elements include **Ga**, **Cu**, **Tc** and **Zr**. Depending on their isotopes, **Ga** and **I** can be used for PET or SPECT imaging.

## Nonmetallic radionuclide labels

Nonmetallic radionuclide labels can be introduced with a low influence on the affinity of inhibitors because nonradioactive carbon, nitrogen and oxygen atoms are inherently present in various biomolecules and compounds. As a result, compared with the nonlabeled inhibitors, only minimal changes occur in the chemical properties of the final probes. In the PET imaging [$^{11}$C] NMS-E973 probe constructed by Vermeulen et al. (72), the carbon atoms on the methyl group of the Hsp90 inhibitor NMS-E973 (**19**) (73) were replaced with $^{11}$**C** (**20**) to conduct *in vivo* melanoma imaging. The time of synthesis and purification should be limited to 2-3 half-lives to ensure the effectiveness of the radiolabeled APSMI (74), and the half-life of $^{11}$**C** is short, which limits its clinical application. However, the introduction of $^{11}$**C** generally does not change the pharmacological properties of the parent inhibitor, and it can be used to study the fate of the inhibitor *in vivo*. Brown et al. (75) used the $^{11}$C-labeled focal adhesion kinase (FAK) inhibitor GSK2256098 to study the pharmacokinetics of parent inhibitor *in vivo* and compared the distribution of probes in normal brain and tumor tissues to study the impact of tumors on the blood-brain barrier. Yu et al. (76) labeled the transient receptor potential channel subfamily member 5 (TRPC5) inhibitor HC608 (**21**) to obtain **22** to study its metabolism *in vivo* and the effect of targeting TrpC5. Moreover, the half-lives of $^{13}$**N** and $^{15}$**O**, at 10 min and 2 min, respectively, are too short to be used for labeling inhibitors.

Probes labeled with halogen radionuclides have been widely used to diagnose tumors and metastases in the clinic. $^{18}$**F**-labeled fludeoxyglucose (FDG) as a PET probe has been used particularly often (77), but due to the active glucose metabolism in the brain and inflammation, it still has limitations in tumor imaging (78–80). Such probes based on small-molecule inhibitors can reduce the false-positive rate because AfPIs can specifically bind to the target, and some of them have entered clinical trials (81). For inhibitors with fluorine in the structure, the loss of affinity caused by radiolabeling can be avoided, such as by replacing the fluorine atom (**24**) or carbon atom (**25**) on the benzene ring of the ROS1 inhibitor lorlatinib (**23**) (82). For inhibitors that do not contain fluorine atoms, $^{18}$**F** can be substituted for a hydrogen atom or hydroxyl group through electrophilic or nucleophilic reactions, which will not cause significant steric hindrances because of their similar sizes. Additionally, the **C-F** bond formed is stronger than the **C-H** bond and thus is not easily destroyed in the body (83), which can decrease false positives in imaging. In the [$^{18}$**F**] labeled tropomyosin receptor kinase (Trk) inhibitor (**27**) synthesized by Bernard-Gauthier et al. (84), the hydrogen atom on the benzene ring was replaced on the parent inhibitor (**26**), resulting in a loss of affinity. However, this loss is acceptable because it does not considerably affect the imaging effect of the probe (Figures 6A, B). Another method to add **F** to the noncritical area of the inhibitor, such as the PEG chain (**28**), which can also avoid the damage of steric hindrance to the affinity, can improve the metabolic kinetics of the probe and is conducive to the imaging effect (62, 85). However, it is necessary to verify the affinity of probes by molecular docking and affinity experiments.

Radiobromine and radioiodine are also commonly used labeling inhibitors. $^{76}$**Br** (**29**) (86) and $^{124}$**I** (87) are used for PET imaging and $^{123}$**I**/$^{125}$**I** (**30**) (88, 89) for SPECT Imaging. Although the steric hindrance of **I** and **Br** is greater than that of **F**, inhibitors can still be introduced through the abovementioned strategy, with a slight loss of affinity. In addition, these radionuclides exhibit a longer half-life than $^{18}$**F**, facilitating the final synthesis of the AfPIs. When these halogen radionuclides are introduced, they may have greater affinity than the parent inhibitors (90), possibly ascribed to the electronegativity of the halogen radionuclides and the extra hydrogen bond formed between the radionuclides and the target receptor.

In general, introducing nonmetallic radionuclide labels to inhibitors to realize tumor imaging is relatively simple. Direct replacement of the original nonradioactive atoms or adding radionuclide with a linker, such as PEG, can avoid diminishing the affinity.

## Metallic radionuclide labels

Unlike nonmetallic radionuclide labels, metallic radionuclides cannot be directly introduced into the inhibitor, so the aid of a metal chelating agent is required. To allow the inhibitors to be labeled without considerably changing their physicochemical properties, bifunctional chelating agents are ideal candidates, which can conjugate with metal ions and inhibitors and can easily react with common functional groups (such as carboxyl, amino and alkyne/azide groups) on inhibitors to form stable covalent bonds (91).

Bifunctional chelating agents can be roughly classified into acyclic and macrocyclic, and the latter is more stable in complexation than the former (92). As a part of the linker in the probe, the chelating agent should be chosen after considering the following factors. The first requirement is that it does not affect the affinity of the original inhibitor and ensures that the ligand can bind to the target later. The design is the same as other AfPIs: The chelating agent cannot affect the critical binding group, and the change in steric hindrance needs to be considered. Second, the thermodynamic stability and kinetic inertness of the chelating agent should be ideal to avoid the release of metal ions to cause biological toxicity (93). During the synthetic process, the production of isomers should also be circumvented to avoid

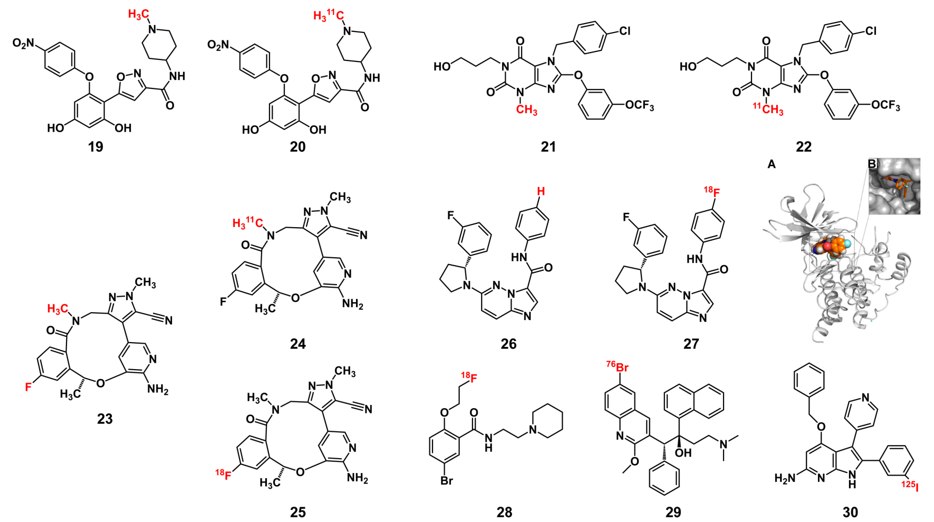


FIGURE 6
Some structures of radiolabeled AfPIs with nonmetallic radionuclide labels(Red: radionuclide labels). **(A, B)** The molecular docking result of [$^{18}$F] labeled TrkA inhibitor with TrkA protein showed that the labeled inhibitor could bind with Trk. Reprinted with permission (84). Copyright 2018 American Chemical Society.

affecting the overall physicochemical properties of the probe. Based on the chelating agent, the metabolism of the probe can be improved by inserting hydrophobic/hydrophilic groups to achieve a careful balance, obtaining the optimal imaging effect. In addition, the insertion of PEG can improve water solubility and promote metabolism, which is governed by the same principle described above. The properties of metal radionuclides, such as size, shape and coordination number, also affect the choice of the chelating agent (94, 95). Therefore, when choosing a chelating agent, the nature of the metal radionuclides should be considered. Commonly used chelating agents for a given radionuclide are often not bad choices.

In summary, the design of radiolabeled AfPIs differs according to the kind of radionuclide. For nonmetal radionuclides, the atoms or groups in the parent inhibitor can be substituted directly, while for metal radionuclides, bifunctional chelating agents are warranted to reduce the loss of affinity. Regardless of the type, the main idea is to complete radionuclide labeling without lowering the affinity of the parent inhibitor while considering the metabolism and biological safety of the final product.

## Dual-modal AfPIs

The advantage of dual-modal probes is that they combine the two imaging technologies to take full advantage of each technique and offset their disadvantages, achieving the goals of high sensitivity and high resolution simultaneously. The most direct examples are PET/CT, SPECT/CT and PET/MRI, which use the anatomical information provided by CT or MRI technology to offset the insufficient spatial resolution of PET/SPECT, and these approaches have also been widely used in the clinic. PET/optical imaging (OI) or SPECT/OI can overcome not only the insufficient tissue penetration of fluorescent probes but also provide higher imaging resolution than PET and SPECT. Based on the connection of the PSMA inhibitor to the Cy3 fluorescent dye, Kommidi et al. (96) introduced $^{18}$**F** through the click reaction at the distal end of the linker to achieve dual-modal imaging. PET imaging is helpful for preoperative planning, while fluorescence imaging can help surgery navigation for tumor resection and reconfirm the edge after surgery. Metal radionuclides can also be labeled on inhibitors with fluorescent dyes using bifunctional chelating agents. Baranski et al. (68) used Glu-urea-Lys-HBED-CC as the core structure to connect $^{68}$**Ga** and various fluorescent dyes, and performed fluorescence-guided tumor resection in mice using a probe connected with IRDye 800CW. Near-infrared dyes provide a greater imaging depth for fluorescence imaging, making fluorescence-guided surgical resection possible, and deeper tumor tissues need to be positioned by PET before surgery. In addition to diagnosis and surgery navigation, PET/OI can be used to observe the administration and metabolism of the inhibitor by labeling the parent structure. Wang et al. (11) designed a PET/OI dual-modal dasatinib probe to compare the effect of convection-enhanced delivery on bypassing the blood-brain barrier and delivering it to glioma by intravenous administration. Fluorescence imaging overcomes the shortcoming that PET cannot monitor drug delivery at the cellular level. These applications are examples of solving clinical needs through the combination of radionuclide imaging and optical imaging. In addition, by adding functional groups, such as amino groups, to $\alpha$-$Fe_2O_3$ nanoparticles, inhibitors and fluorescent dyes can be labeled to achieve MRI/OI dual-modal imaging (97).

As with single-modal AfPIs, the affinity, metabolism and tissue distribution should also be considered for dual-modal AfPIs. Conducting multiple labeling at the same time will inevitably cause more significant potential damage to affinity because it may alter more groups or cause greater steric hindrance, so additional dyes and radionuclide labeling should be as far as possible from the target area when the probe structure is being designed. Metabolism and tissue distribution need to be modified according to the *in vivo* performance of the core structure of probes. For example, Kimura et al. (98) used hydrophobic Cy5.5 dye to enhance tumor retention and reduce the impact of $^{64}$**Cu** labeling on the imaging effect. Alternatively, increasing the number of sulfonate groups could improve the water solubility of the probe and switch the hepatobiliary to renal elimination, and a more concentrated signal at the tumor was obtained (99). Therefore, further structure modification can improve the metabolism and tissue distribution and eliminate the influence of multiple labels on the imaging effect of the probe.

Although much progress has been made in dual-modal imaging in recent years, the bimodal imaging probes including the dual-modal AfPIs still fall short of applicability in the clinic due to the limitations of the development of imaging instruments and software. However, these dual-modal or trimodal probes can provide more anatomical or functional information, and this considerable advantage is worthy of more research and development.

## Conclusion

The design and synthesis of AfPIs involve interdisciplinary research, and numerous issues need to be considered, including the affinity, distribution and pharmacokinetics *in vivo* of probes; the synthetic route; and translation to clinical applicability. It is necessary to perform molecular docking before designing probes to determine the effect of changes in steric hindrance and modification of moieties on their affinity. Moreover, the probe's fate *in vivo* is crucial for imaging, and appropriate dyes and linkers can significantly improve the pharmacokinetic and imaging efficacy of the probe.

Over the last decade, there have been tremendous advances in the research of AfPIs. The AfPIs have been proved to have better specificity, smaller molecular weight, lower immunogenicity and faster targeting than protein and peptide probes. Although there are many reports on radiolabeled AfPIs, and some have entered into clinical trials, there is considerable room for improvement in NIR and dual-modal AfPIs, especially in the second NIR window. This challenge is related to the lack of suitable dyes and the greater difficulty of design and synthesis in the second NIR window probe. Therefore, problems such as developing new-generation inhibitors, NIR dyes and bifunctional chelators, improving quantum yield, and biological safety are still hindering the clinical application of AfPIs and are warranted to be solved. Finally, the probe is only a tool, and the ultimate objective is to solve medical needs. Hence, the final product should be convenient for clinical application in disease diagnosis or treatment.

In conclusion, it remains to be seen whether AfPIs can be applied in the clinic. However, with the development of more economical imaging instruments and new-generation inhibitors with fewer side effects and better selectivity, and the urgent need for more reliable detection methods and more efficient and safer treatment for cancer, AfPIs have broad prospects for cancer diagnosis and treatment monitoring.

## Author contributions

XY: initiating and writing the manuscript. ZW: initiating and writing the manuscript. XH: revising and editing the manuscript. AY: revising and editing the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

This research is funded by the Open Competition Mechanism Science and Technology Project of Hubei Province (2022, NO.028).

## 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

## References

1. Sharpe AH, Pauken KE. The diverse functions of the Pd1 inhibitory pathway. *Nat Rev Immunol* (2018) 18(3):153–67. doi: 10.1038/nri.2017.108

2. Lv PC, Jiang AQ, Zhang WM, Zhu HL. Fak inhibitors in cancer, a patent review. *Expert Opin Ther Pat* (2018) 28(2):139–45. doi: 10.1080/13543776.2018.1414183

3. Gorecki L, Andrs M, Rezacova M, Korabecny J. Discovery of atr kinase inhibitor berzosertib (Vx-970, M6620): Clinical candidate for cancer therapy. *Pharmacol Ther* (2020) 210:107518. doi: 10.1016/j.pharmthera.2020.107518

4. Huang Y, Su R, Sheng Y, Dong L, Dong Z, Xu H, et al. Small-molecule targeting of oncogenic fto demethylase in acute myeloid leukemia. *Cancer Cell* (2019) 35(4):677–91 e10. doi: 10.1016/j.ccell.2019.03.006

5. Inoue-Yamauchi A, Jeng PS, Kim K, Chen HC, Han S, Ganesan YT, et al. Targeting the differential addiction to anti-apoptotic bcl-2 family for cancer therapy. *Nat Commun* (2017) 8:16078. doi: 10.1038/ncomms16078

6. Han H, Jain AD, Truica MI, Izquierdo-Ferrer J, Anker JF, Lysy B, et al. Small-molecule myc inhibitors suppress tumor growth and enhance immunotherapy. *Cancer Cell* (2019) 36(5):483–97 e15. doi: 10.1016/j.ccell.2019.10.001

7. Hu P, Shang L, Chen J, Chen X, Chen C, Hong W, et al. A nanometer-sized protease inhibitor for precise cancer diagnosis and treatment. *J Mater Chem B* (2020) 8(3):504–14. doi: 10.1039/c9tb02081k

8. Werle M, Bernkop-Schnurch A. Strategies to improve plasma half life time of peptide and protein drugs. *Amino Acids* (2006) 30(4):351–67. doi: 10.1007/s00726-005-0289-3

9. Hou JT, Ko KP, Shi H, Ren WX, Verwilst P, Koo S, et al. Plk1-targeted fluorescent tumor imaging with high signal-to-Background ratio. *ACS Sens* (2017) 2(10):1512–6. doi: 10.1021/acssensors.7b00544

10. Loktev A, Lindner T, Mier W, Debus J, Altmann A, Jager D, et al. A tumor-imaging method targeting cancer-associated fibroblasts. *J Nucl Med* (2018) 59 (9):1423–9. doi: 10.2967/jnumed.118.210435

11. Wang M, Kommidi H, Tosi U, Guo H, Zhou Z, Schweitzer ME, et al. A murine model for quantitative, real-time evaluation of convection-enhanced delivery (Rt-ced) using an (18)[F]-positron emitting, fluorescent derivative of dasatinib. *Mol Cancer Ther* (2017) 16(12):2902–12. doi: 10.1158/1535-7163.MCT-17-0423

12. Chou PH, Luo CK, Wali N, Lin WY, Ng SK, Wang CH, et al. A chemical probe inhibitor targeting Stat1 restricts cancer stem cell traits and angiogenesis in colorectal cancer. *J BioMed Sci* (2022) 29(1):20. doi: 10.1186/s12929-022-00803-4

13. Markiewicz R, Litowczenko J, Gapinski J, Wozniak A, Jurga S, Patkowski A. Nanomolar nitric oxide concentrations in living cells measured by means of fluorescence correlation spectroscopy. *Molecules* (2022) 27(3):1010. doi: 10.3390/molecules27031010

14. Danylchuk DI, Jouard PH, Klymchenko AS. Targeted solvatochromic fluorescent probes for imaging lipid order in organelles under oxidative and mechanical stress. *J Am Chem Soc* (2021) 143(2):912–24. doi: 10.1021/jacs.0c10972

15. Keppner S, Proschak E, Schneider G, Spankuch B. Identification and validation of a potent type ii inhibitor of inactive polo-like kinase 1. *ChemMedChem* (2009) 4(11):1806–9. doi: 10.1002/cmdc.200900338

16. Yu Z, Deng P, Chen Y, Liu S, Chen J, Yang Z, et al. Inhibition of the Plk1-coupled cell cycle machinery overcomes resistance to oxaliplatin in colorectal cancer. *Adv Sci (Weinh)* (2021) 8(23):e2100759. doi: 10.1002/advs.202100759

17. Ho YH, Wang KJ, Hung PY, Cheng YS, Liu JR, Fung ST, et al. A highly Hdac6-selective inhibitor acts as a fluorescent probe. *Org Biomol Chem* (2018) 16 (42):7820–32. doi: 10.1039/c8ob00966j

18. Zhang Y, Yan J, Yao TP. Discovery of a fluorescent probe with Hdac6 selective inhibition. *Eur J Med Chem* (2017) 141:596–602. doi: 10.1016/j.ejmech.2017.10.022

19. Chen WL, Li DD, Wang ZH, Xu XL, Zhang XJ, Jiang ZY, et al. Design, synthesis, and initial evaluation of affinity-based small molecular probe for detection of Wdr5. *Bioorg Chem* (2018) 76:380–5. doi: 10.1016/j.bioorg.2017.11.018

20. Zhang Y, Zhang X-F, Chen Q, Cao X-Q, Shen S-L. A novel near-infrared fluorescence off-on probe for imaging hypoxia and nitroreductase in cells and in vivo. *Sens Actuators B: Chem* (2022) 353:131145. doi: 10.1016/j.snb.2021.131145

21. Zhu S, Lu Y, Jin J, Yu J, Lu W. An Hsp90 inhibitor based fluorescent probe for selective tumor targeting. *Dyes Pigm* (2021) 196:109783. doi: 10.1016/j.dyepig.2021.109783

22. Bian J, Li X, Xu L, Wang N, Qian X, You Q, et al. Affinity-based small fluorescent probe for Nad(P)H:Quinone oxidoreductase 1 (Nqo1). design, synthesis and pharmacological evaluation. *Eur J Med Chem* (2017) 127:828–39. doi: 10.1016/j.ejmech.2016.10.062

23. Sun W, Li M, Fan J, Peng X. Activity-based sensing and theranostic probes based on photoinduced electron transfer. *Acc Chem Res* (2019) 52(10):2818–31. doi: 10.1021/acs.accounts.9b00340

24. Zhang H, Fan J, Wang J, Zhang S, Dou B, Peng X. An off-on cox-2-Specific fluorescent probe: Targeting the golgi apparatus of cancer cells. *J Am Chem Soc* (2013) 135(31):11663–9. doi: 10.1021/ja4056905

25. Guo S, Fan J, Wang B, Xiao M, Li Y, Du J, et al. Highly selective red-emitting fluorescent probe for imaging cancer cells in situ by targeting pim-1 kinase. *ACS Appl Mater Interfaces* (2018) 10(2):1499–507. doi: 10.1021/acsami.7b14553

26. Kenry, Duan Y, Liu B. Recent advances of optical imaging in the second near-infrared window. *Adv Mater* (2018) 30(47):e1802394. doi: 10.1002/adma.201802394

27. Zhu S, Tian R, Antaris AL, Chen X, Dai H. Near-Infrared-Ii molecular dyes for cancer imaging and surgery. *Adv Mater* (2019) 31(24):e1900321. doi: 10.1002/adma.201900321

28. Tang C, Du Y, Liang Q, Cheng Z, Tian J. Development of a novel histone deacetylase-targeted near-infrared probe for hepatocellular carcinoma imaging and fluorescence image-guided surgery. *Mol Imaging Biol* (2020) 22(3):476–85. doi: 10.1007/s11307-019-01389-4

29. Mukkamala R, Lindeman SD, Kragness KA, Shahriar I, Srinivasarao M, Low PS. Design and characterization of fibroblast activation protein targeted pan-cancer imaging agent for fluorescence-guided surgery of solid tumors. *J Mater Chem B* (2022) 10(12):2038–46. doi: 10.1039/d1tb02651h

30. Wang W, Zhang X, Ni X, Zhou W, Xie C, Huang W, et al. Semiconducting polymer nanoparticles for nir-ii fluorescence imaging-guided Photothermal/Thermodynamic combination therapy. *Biomater Sci* (2022) 10(3):846–53. doi: 10.1039/d1bm01646f

31. Albers AE, Rawls KA, Chang CJ. Activity-based fluorescent reporters for monoamine oxidases in living cells. *Chem Commun (Camb)* (2007) 44):4647–9. doi: 10.1039/b713190a

32. Wu X, Shi W, Li X, Ma H. A strategy for specific fluorescence imaging of monoamine oxidase a in living cells. *Angew Chem Int Ed Engl* (2017) 56 (48):15319–23. doi: 10.1002/anie.201708428

33. Yang Z, Li W, Chen H, Mo Q, Li J, Zhao S, et al. Inhibitor structure-guided design and synthesis of near-infrared fluorescent probes for monoamine oxidase a (Mao-a) and its application in living cells and in vivo. *Chem Commun (Camb)* (2019) 55(17):2477–80. doi: 10.1039/c8cc10084e

34. Ma J, Yoshimura M, Yamashita E, Nakagawa A, Ito A, Tsukihara T. Structure of rat monoamine oxidase a and its specific recognitions for substrates and inhibitors. *J Mol Biol* (2004) 338(1):103–14. doi: 10.1016/j.jmb.2004.02.032

35. Wu JB, Lin TP, Gallagher JD, Kushal S, Chung LW, Zhau HE, et al. Monoamine oxidase a inhibitor-near-Infrared dye conjugate reduces prostate tumor growth. *J Am Chem Soc* (2015) 137(6):2366–74. doi: 10.1021/ja512613j

36. Xiang Y, Hirth B, Asmussen G, Biemann HP, Bishop KA, Good A, et al. The discovery of novel benzofuran-2-Carboxylic acids as potent pim-1 inhibitors. *Bioorg Med Chem Lett* (2011) 21(10):3050–6. doi: 10.1016/j.bmcl.2011.03.030

37. Krishnan MA, Pandit A, Sharma R, Chelvam V. Imaging of prostate cancer: Optimizing affinity to prostate specific membrane antigen by spacer modifications in a tumor spheroid model. *J Biomol Struct Dyn* (2021) 39:1–22. doi: 10.1080/07391102.2021.1936642

38. Matsuoka D, Watanabe H, Shimizu Y, Kimura H, Yagi Y, Kawai R, et al. Structure-activity relationships of succinimidyl-Cys-C(O)-Glu derivatives with different near-infrared fluorophores as optical imaging probes for prostate-specific membrane antigen. *Bioorg Med Chem* (2018) 26(9):2291–301. doi: 10.1016/j.bmc.2018.03.015

39. Overchuk M, Damen MPF, Harmatys KM, Pomper MG, Chen J, Zheng G. Long-circulating prostate-specific membrane antigen-targeted nir phototheranostic agent. *Photochem Photobiol* (2020) 96(3):718–24. doi: 10.1111/php.13181

40. Kwon YD, Chung HJ, Lee SJ, Lee SH, Jeong BH, Kim HK. Synthesis of novel multivalent fluorescent inhibitors with high affinity to prostate cancer and their biological evaluation. *Bioorg Med Chem Lett* (2018) 28(4):572–6. doi: 10.1016/j.bmcl.2018.01.047

41. Kwon YD, Oh JM, La MT, Chung HJ, Lee SJ, Chun S, et al. Synthesis and evaluation of multifunctional fluorescent inhibitors with synergistic interaction of prostate-specific membrane antigen and hypoxia for prostate cancer. *Bioconjug Chem* (2019) 30(1):90–100. doi: 10.1021/acs.bioconjchem.8b00767

42. Meng Q, Wang Z, Cui J, Cui Q, Dong J, Zhang Q, et al. Design, synthesis, and biological evaluation of cytochrome P450 1b1 targeted molecular imaging probes for colorectal tumor detection. *J Med Chem* (2018) 61(23):10901–9. doi: 10.1021/acs.jmedchem.8b01633

43. Lv Q, Yang X, Wang M, Yang J, Qin Z, Kan Q, et al. Mitochondria-targeted prostate cancer therapy using a near-infrared fluorescence dye-monoamine oxidase a inhibitor conjugate. *J Control Release* (2018) 279:234–42. doi: 10.1016/j.jconrel.2018.04.038

44. Guan Y, Zhang Y, Xiao L, Li J, Wang JP, Chordia MD, et al. Improving therapeutic potential of farnesylthiosalicylic acid: Tumor specific delivery *Via* conjugation with heptamethine cyanine dye. *Mol Pharm* (2017) 14(1):1–13. doi: 10.1021/acs.molpharmaceut.5b00906

45. Son SH, Kwon H, Ahn HH, Nam H, Kim K, Nam S, et al. Design and synthesis of a novel bodipy-labeled psma inhibitor. *Bioorg Med Chem Lett* (2020) 30(3):126894. doi: 10.1016/j.bmcl.2019.126894

46. Wang X, Wang L, Xie L, Xie Z, Li L, Bui D, et al. Design and synthesis of a novel nir celecoxib-based fluorescent probe for cyclooxygenase-2 targeted bioimaging in tumor cells. *Molecules* (2020) 25(18):4037. doi: 10.3390/molecules25184037

47. Kulkarni NS, Parvathaneni V, Shukla SK, Barasa L, Perron JC, Yoganathan S, et al. Tyrosine kinase inhibitor conjugated quantum dots for non-small cell lung cancer (Nsclc) treatment. *Eur J Pharm Sci* (2019) 133:145–59. doi: 10.1016/j.ejps.2019.03.026

48. Asha Krishnan M, Yadav K, Roach P, Chelvam V. A targeted near-infrared nanoprobe for deep-tissue penetration and imaging of prostate cancer. *Biomater Sci* (2021) 9(6):2295–312. doi: 10.1039/d0bm01970d

49. Lin R, Huang J, Wang L, Li Y, Lipowska M, Wu H, et al. Bevacizumab and near infrared probe conjugated iron oxide nanoparticles for vascular endothelial growth factor targeted Mr and optical imaging. *Biomater Sci* (2018) 6(6):1517–25. doi: 10.1039/c8bm00225h

50. Gao MX, Yang L, Zheng Y, Yang XX, Zou HY, Han J, et al. "Click" on alkynylated carbon quantum dots: An efficient surface functionalization for specific biosensing and bioimaging. *Chem - A Eur J* (2017) 23(9):2171–8. doi: 10.1002/chem.201604963

51. Liu J, Zhao X, Xu H, Wang Z, Dai Z. Amino acid-capped water-soluble near-infrared region Cuins2/Zns quantum dots for selective cadmium ion determination and multicolor cell imaging. *Anal Chem* (2019) 91(14):8987–93. doi: 10.1021/acs.analchem.9b01183

52. Yi X, Cao Z, Yuan Y, Li W, Cui X, Chen Z, et al. Design and synthesis of a novel mitochondria-targeted osteosarcoma theranostic agent based on a Pim1 kinase inhibitor. *J Control Release* (2021) 332:434–47. doi: 10.1016/j.jconrel.2021.02.028

53. Zhang C, Liu T, Luo P, Gao L, Liao X, Ma L, et al. Near-infrared oxidative phosphorylation inhibitor integrates acute myeloid leukemia-targeted imaging and therapy. *Sci Adv* (2021) 7(1):eabb6104. doi: 10.1126/sciadv.abb6104

54. Usama SM, Lin CM, Burgess K. On the mechanisms of uptake of tumor-seeking cyanine dyes. *Bioconjug Chem* (2018) 29(11):3886–95. doi: 10.1021/acs.bioconjchem.8b00708

55. Usama SM, Park GK, Nomura S, Baek Y, Choi HS, Burgess K. Role of albumin in accumulation and persistence of tumor-seeking cyanine dyes. *Bioconjug Chem* (2020) 31(2):248–59. doi: 10.1021/acs.bioconjchem.9b00771

56. Guan Y, Zhang Y, Zou J, Huang LP, Chordia MD, Yue W, et al. Synthesis and biological evaluation of genistein-Ir783 conjugate: Cancer cell targeted delivery in mcf-7 for superior anti-cancer therapy. *Molecules* (2019) 24(22):4120. doi: 10.3390/molecules24224120

57. Zhang L, Shi X, Li Y, Duan X, Zhang Z, Fu H, et al. Visualizing tumors in real time: A highly sensitive psma probe for nir-ii imaging and intraoperative tumor resection. *J Med Chem* (2021) 64(11):7735–45. doi: 10.1021/acs.jmedchem.1c00444

58. Wu J, Lee HJ, You L, Luo X, Hasegawa T, Huang KC, et al. Functionalized nir-ii semiconducting polymer nanoparticles for single-cell to whole-organ imaging of psma-positive prostate cancer. *Small* (2020) 16(19):e2001215. doi: 10.1002/smll.202001215

59. Gayton J, Autry SA, Meador W, Parkin SR, Hill GAJr., Hammer NI, et al. Indolizine-cyanine dyes: Near infrared emissive cyanine dyes with increased stokes shifts. *J Org Chem* (2019) 84(2):687–97. doi: 10.1021/acs.joc.8b02521

60. Cosco ED, Caram JR, Bruns OT, Franke D, Day RA, Farr EP, et al. Flavylium polymethine fluorophores for near- and shortwave infrared imaging. *Angew Chem Int Ed Engl* (2017) 56(42):13126–9. doi: 10.1002/anie.201706974

61. Chen Y, Pullambhatla M, Banerjee SR, Byun Y, Stathis M, Rojas C, et al. Synthesis and biological evaluation of low molecular weight fluorescent imaging agents for the prostate-specific membrane antigen. *Bioconjug Chem* (2012) 23 (12):2377–85. doi: 10.1021/bc3003919

62. Zhou Z, Zalutsky MR, Chitneni SK. Fluorine-18 labeling of the Mdm2 inhibitor Rg7388 for pet imaging: Chemistry and preliminary evaluation. *Mol Pharm* (2021) 18(10):3871–81. doi: 10.1021/acs.molpharmaceut.1c00531

63. Zhu D, Guo H, Chang Y, Ni Y, Li L, Zhang ZM, et al. Cell- and tissue-based proteome profiling and dual imaging of apoptosis markers with probes derived from venetoclax and idasanutlin. *Angew Chem Int Ed Engl* (2018) 57(30):9284–9. doi: 10.1002/anie.201802003

64. Arlauckas SP, Kumar M, Popov AV, Poptani H, Delikatny EJ. Near infrared fluorescent imaging of choline kinase alpha expression and inhibition in breast tumors. *Oncotarget* (2017) 8(10):16518–30. doi: 10.18632/oncotarget.14965

65. Osada T, Kaneko K, Gwin WR, Morse MA, Hobeika A, Pogue BW, et al. *In vivo* detection of Hsp90 identifies breast cancers with aggressive behavior. *Clin Cancer Res* (2017) 23(24):7531–42. doi: 10.1158/1078-0432.CCR-17-1453

66. Zhang K, Liu Z, Yao Y, Qiu Y, Li F, Chen D, et al. Structure-based design of a selective class I histone deacetylase (Hdac) near-infrared (Nir) probe for epigenetic regulation detection in triple-negative breast cancer (Tnbc). *J Med Chem* (2021) 64 (7):4020–33. doi: 10.1021/acs.jmedchem.0c02161

67. Gao G, Jiang YW, Sun W, Guo Y, Jia HR, Yu XW, et al. Molecular targeting-mediated mild-temperature photothermal therapy with a smart albumin-based nanodrug. *Small* (2019) 15(33):e1900501. doi: 10.1002/smll.201900501

68. Baranski AC, Schafer M, Bauder-Wust U, Roscher M, Schmidt J, Stenau E, et al. Psma-11-Derived dual-labeled psma inhibitors for preoperative pet imaging and precise fluorescence-guided surgery of prostate cancer. *J Nucl Med* (2018) 59 (4):639–45. doi: 10.2967/jnumed.117.201293

69. Reinfelder J, Kuwert T, Beck M, Sanders JC, Ritt P, Schmidkonz C, et al. First experience with Spect/Ct using a 99mtc-labeled inhibitor for prostate-specific membrane antigen in patients with biochemical recurrence of prostate cancer. *Clin Nucl Med* (2017) 42(1):26–33. doi: 10.1097/RLU.0000000000001433

70. Zhang L, Suksanpaisan L, Jiang H, DeGrado TR, Russell SJ, Zhao M, et al. Dual-isotope spect imaging with nis reporter gene and duramycin to visualize tumor susceptibility to oncolytic virus infection. *Mol Ther Oncolytics* (2019) 15:178–85. doi: 10.1016/j.omto.2019.10.002

71. Tamarappoo B, Otaki Y, Manabe O, Hyun M, Cantu S, Arnson Y, et al. Simultaneous Tc-99m Pyp/Tl-201 dual-isotope spect myocardial imaging in patients with suspected cardiac amyloidosis. *J Nucl Cardiol* (2020) 27(1):28–37. doi: 10.1007/s12350-019-01753-5

72. Vermeulen K, Naus E, Ahamed M, Attili B, Siemons M, Luyten K, et al. Evaluation of [(11)C]Nms-E973 as a pet tracer for in vivo visualisation of Hsp90. *Theranostics* (2019) 9(2):554–72. doi: 10.7150/thno.27213

73. Brasca MG, Mantegani S, Amboldi N, Bindi S, Caronni D, Casale E, et al. Discovery of nms-E973 as novel, selective and potent inhibitor of heat shock protein 90 (Hsp90). *Bioorg Med Chem* (2013) 21(22):7047–63. doi: 10.1016/j.bmc.2013.09.018

74. Pimlott SL, Sutherland A. Molecular tracers for the pet and spect imaging of disease. *Chem Soc Rev* (2011) 40(1):149–62. doi: 10.1039/b922628c

75. Brown NF, Williams M, Arkenau HT, Fleming RA, Tolson J, Yan L, et al. A study of the focal adhesion kinase inhibitor Gsk2256098 in patients with recurrent glioblastoma with evaluation of tumor penetration of [11c]Gsk2256098. *Neuro Oncol* (2018) 20(12):1634–42. doi: 10.1093/neuonc/noy078

76. Yu Y, Liang Q, Liu H, Luo Z, Hu H, Perlmutter JS, et al. Development of a carbon-11 pet radiotracer for imaging Trpc5 in the brain. *Org Biomol Chem* (2019) 17(22):5586–94. doi: 10.1039/c9ob00893d

77. Nestle U, Schimek-Jasch T, Kremp S, Schaefer-Schuler A, Mix M, Küsters A, et al. Imaging-based target volume reduction in chemoradiotherapy for locally advanced non-Small-Cell lung cancer (Pet-plan): A multicentre, open-label, randomised, controlled trial. *Lancet Oncol* (2020) 21(4):581–92. doi: 10.1016/s1470-2045(20)30013-9

78. Descamps L, Olagne L, Merlin C, Cachin F, Soubrier M, Mathieu S. Utility of Pet/Ct in the diagnosis of inflammatory rheumatic diseases: A systematic review and meta-analysis. *Ann Rheum Dis* (2018) 77(11):e81. doi: 10.1136/annrheumdis-2017-212660

79. Dolan RD, McLees NG, Irfan A, McSorley ST, Horgan PG, Colville D, et al. The relationship between tumor glucose metabolism and host systemic inflammatory responses in patients with cancer: A systematic review. *J Nucl Med* (2019) 60(4):467–71. doi: 10.2967/jnumed.118.216697

80. Jamadar SD, Ward PG, Li S, Sforazzini F, Baran J, Chen Z, et al. Simultaneous task-based bold-fmri and [18-f] fdg functional pet for measurement of neuronal metabolism in the human visual cortex. *Neuroimage* (2019) 189:258–66. doi: 10.1016/j.neuroimage.2019.01.003

81. Pan KH, Wang JF, Wang CY, Nikzad AA, Kong FQ, Jian L, et al. Evaluation of 18f-dcfpyl psma Pet/Ct for prostate cancer: A meta-analysis. *Front Oncol* (2020) 10:597422. doi: 10.3389/fonc.2020.597422

82. Collier TL, Normandin MD, Stephenson NA, Livni E, Liang SH, Wooten DW, et al. Synthesis and preliminary pet imaging of (11)C and (18)F isotopologues of the Ros1/Alk inhibitor lorlatinib. *Nat Commun* (2017) 8:15761. doi: 10.1038/ncomms15761

83. O'Hagan D. Understanding organofluorine chemistry. an introduction to the c-f bond. *Chem Soc Rev* (2008) 37(2):308–19. doi: 10.1039/b711844a

84. Bernard-Gauthier V, Mossine AV, Mahringer A, Aliaga A, Bailey JJ, Shao X, et al. Identification of [(18)F]Track, a fluorine-18-Labeled tropomyosin receptor kinase (Trk) inhibitor for pet imaging. *J Med Chem* (2018) 61(4):1737–43. doi: 10.1021/acs.jmedchem.7b01607

85. Yang D, Comeau A, Bowen WD, Mach RH, Ross BD, Hong H, et al. Design and investigation of a [(18)F]-labeled benzamide derivative as a high affinity dual sigma receptor subtype radioligand for prostate tumor imaging. *Mol Pharm* (2017) 14(3):770–80. doi: 10.1021/acs.molpharmaceut.6b01020

86. Ordonez AA, Carroll LS, Abhishek S, Mota F, Ruiz-Bedoya CA, Klunk MH, et al. Radiosynthesis and pet bioimaging of (76)Br-bedaquiline in a murine model of tuberculosis. *ACS Infect Dis* (2019) 5(12):1996–2002. doi: 10.1021/acsinfecdis.9b00207

87. Nie P, Kalidindi T, Nagle VL, Wu X, Li T, Liao GP, et al. Imaging of cancer gamma-secretase activity using an inhibitor-based pet probe. *Clin Cancer Res* (2021) 27(22):6145–55. doi: 10.1158/1078-0432.CCR-21-0940

88. Tang L, Peng C, Tang B, Li Z, Wang X, Li J, et al. Radioiodinated small-molecule tyrosine kinase inhibitor for Her2-selective spect imaging. *J Nucl Med* (2018) 59(9):1386–91. doi: 10.2967/jnumed.117.205088

89. Hirata M, Yao T, Fujimura S, Kanai Y, Yoshimoto M, Sato T, et al. Development of a P38alpha-selective radioactive probe for qualitative diagnosis of cancer using spect. *Ann Nucl Med* (2019) 33(5):333–43. doi: 10.1007/s12149-019-01341-0

90. Maresca KP, Hillier SM, Femia FJ, Keith D, Barone C, Joyal JL, et al. A series of halogenated heterodimeric inhibitors of prostate specific membrane antigen (Psma) as radiolabeled probes for targeting prostate cancer. *J Med Chem* (2009) 52 (2):347–57. doi: 10.1021/jm800994j

91. Chen KT, Nguyen K, Ieritano C, Gao F, Seimbille Y. A flexible synthesis of (68)Ga-labeled carbonic anhydrase ix (Caix)-targeted molecules *Via* Cbt/1,2-aminothiol click reaction. *Molecules* (2018) 24(1):23. doi: 10.3390/molecules24010023

92. Lattuada L, Barge A, Cravotto G, Giovenzana GB, Tei L. The synthesis and application of polyamino polycarboxylic bifunctional chelating agents. *Chem Soc Rev* (2011) 40(5):3019–49. doi: 10.1039/c0cs00199f

93. Guillou A, Lima LMP, Esteban-Gomez D, Le Poul N, Bartholoma MD, Platas-Iglesias C, et al. Methylthiazolyl tacn ligands for copper complexation and their bifunctional chelating agent derivatives for bioconjugation and copper-64 radiolabeling: An example with bombesin. *Inorg Chem* (2019) 58(4):2669–85. doi: 10.1021/acs.inorgchem.8b03280

94. Li L, de Guadalupe Jaraquemada-Pelaez M, Aluicio-Sarduy E, Wang X, Barnhart TE, Cai W, et al. Coordination chemistry of [Y(Pypa)](-) and comparison immuno-pet imaging of [(44)Sc]Sc- and [(86)Y]Y-Pypa-Phenyl-Trc105. *Dalton Trans* (2020) 49(17):5547–62. doi: 10.1039/d0dt00437e

95. Kelly JM, Amor-Coarasa A, Nikolopoulou A, Kim D, Williams CJr., Vallabhajosula S, et al. Assessment of psma targeting ligands bearing novel chelates with application to theranostics: Stability and complexation kinetics of (68)Ga(3+), (111)in(3+), (177)Lu(3+) and (225)Ac(3). *Nucl Med Biol* (2017) 55:38–46. doi: 10.1016/j.nucmedbio.2017.10.001

96. Kommidi H, Guo H, Nurili F, Vedvyas Y, Jin MM, McClure TD, et al. (18)F-positron Emitting/Trimethine cyanine-fluorescent contrast for image-guided prostate cancer management. *J Med Chem* (2018) 61(9):4256–62. doi: 10.1021/acs.jmedchem.8b00240

97. Mekawy MM, Saito A, Sumiyoshi A, Riera JJ, Shimizu H, Kawashima R, et al. Hybrid magneto-fluorescent nano-probe for live apoptotic cells monitoring at brain cerebral ischemia. *Mater Sci Eng C Mater Biol Appl* (2019) 100:485–92. doi: 10.1016/j.msec.2019.03.032

98. Kimura RH, Miao Z, Cheng Z, Gambhir SS, Cochran JR. A dual-labeled knottin peptide for pet and near-infrared fluorescence imaging of integrin expression in living subjects. *Bioconjug Chem* (2010) 21(3):436–44. doi: 10.1021/bc9003102

99. Schwegmann K, Hohn M, Hermann S, Schafers M, Riemann B, Haufe G, et al. Optimizing the biodistribution of radiofluorinated barbiturate tracers for matrix metalloproteinase imaging by introduction of fluorescent dyes as pharmacokinetic modulators. *Bioconjug Chem* (2020) 31(4):1117–32. doi: 10.1021/acs.bioconjchem.9b00817

frontiers | Frontiers in Oncology

TYPE Systematic Review
PUBLISHED 10 January 2023
DOI 10.3389/fonc.2022.973104

EDITED BY
Chen Liu,
Army Medical University, China

REVIEWED BY
Chuyi Huang,
Shanghai Jiao Tong University, China
Archana Hinduja,
Wexner Medical Center, The Ohio State University, United States

*CORRESPONDENCE
Chun-Mei Chen
cmchen2009@sina.com

†These authors have contributed equally to this work

This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

SPECIALTY SECTION
RECEIVED 20 June 2022
ACCEPTED 20 December 2022
PUBLISHED 10 January 2023

CITATION
Jiang Y-W, Xu X-J, Wang R and Chen C-M (2023) Efficacy of non-enhanced computer tomography-based radiomics for predicting hematoma expansion: A meta-analysis. *Front. Oncol.* 12:973104.
doi: 10.3389/fonc.2022.973104

# Efficacy of non-enhanced computer tomography-based radiomics for predicting hematoma expansion: A meta-analysis

Yan-Wei Jiang†, Xiong-Jei Xu†, Rui Wang and Chun-Mei Chen*

Department of Neurosurgery, Fujian Medical University Union Hospital, Fuzhou, Fujian, China

**Background:** This meta-analysis aimed to assess the efficacy of radiomics using non-enhanced computed tomography (NCCT) for predicting hematoma expansion in patients with spontaneous intracerebral hemorrhage.

**Methods:** Throughout the inception of the project to April 11, 2022, a comprehensive search was conducted on PubMed, Embase, and Cochrane Central Register of Controlled Trials. The methodological quality of studies in this analysis was assessed by the radiomics quality scoring system (RQS). A meta-analysis of radiomic studies based on NCCT for predicting hematoma expansion in patients with intracerebral hemorrhage was performed. The efficacy of the radiomics approach and non-contrast CT markers was compared using network meta-analysis (NMA).

**Results:** Ten articles comprising a total of 1525 patients were quantitatively analyzed for hematoma expansion after cerebral hemorrhage using radiomics. Based on the included studies, the mean RQS was 14.4. The AUC value (95% confidence interval) of the radiomics model was 0.80 (0.76-0.83). Five articles comprising 846 patients were included in the NMA. The results synthesized according to Bayesian NMA revealed that the predictive ability of the radiomics model outperformed most of the NCCT biomarkers.

**Conclusions:** The NCCT-based radiomics approach has the potential to predict hematoma expansion. Compared to NCCT biomarkers, we recommend a radiomics approach. Standardization of the radiomics approach is required for further clinical implementation.

**Systematic review registration:** https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=324034, identifier [CRD42022324034].

KEYWORDS

**non-enhanced computer tomography, radiomics, hematoma expansion, meta-analysis, spontaneous intracerebral hemorrhage**

# 1 Introduction

Intracerebral hemorrhage is a life-threatening and costly disorder that accounts for 10–15% of all strokes (1). Hematoma expansion is an independent risk factor for poor neurological outcomes. Predictions of hematoma expansion risks can help to stratify patients. Previous studies have reported that spot signs are a good predictor of hematoma expansion (2, 3). Nevertheless, the application of spot signs is limited because computed tomography angiography (CTA) and contrast-enhanced CT are not routinely performed in the emergency department. Non-enhanced CT (NCCT) is most commonly used for intracerebral hemorrhage imaging. Several studies have reported that radiological markers extracted from NCCT, including the black hole, satellite, and blend signs, are related to hematoma expansion (2). However, the extraction of radiomic markers is time-consuming and heterogeneous. Further, the accuracy of radiomic markers may depend on the experience of the clinician who reads the medical images.

Radiomics is a new method for the quantitative analysis of medical images (4). Radiomics analysis was initially implemented in the mining of medical images related to oncology. Recently, the radiomics approach has been applied in non-oncological fields (5). An increasing number of studies have used an NCCT-based radiomics approach to predict hemorrhage expansion(Figure 1) (4, 6, 7). However, data on the predictive efficacy of radiomics methods remain insufficient for further implementation.

This meta-analysis aimed to determine whether NCCT-based radiomics approaches are effective for predicting hematoma expansion. Radiomics quality scoring (RQS) was used to determine the quality of the studies included in the meta-analysis (8). Network meta-analysis (NMA) was employed to synthesize diagnostic test accuracy data in order to assess the efficacy of different diagnostic tests (9, 10). We compared the efficacy of common NCCT markers and radiomics approaches for predicting intracerebral hemorrhage expansion using NMA.

# 2 Methods

## 2.1 Literature search and study selection

This study was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement (eTable 1) (11). This study was registered with PROSPERO (CRD42022324034). PubMed, Embase, and the Cochrane Central Register of Controlled Trials were searched thoroughly from inception to April 11, 2022 (eTable 2) for articles in English. References to relevant published articles were also searched to obtain the desired articles.

After pooling the search results from the three databases and removing duplicate articles, the abstracts and titles of the articles were screened independently by two researchers. Eligible articles were identified by a comprehensive reading of the full text. We included all eligible radiomics articles that used non-enhanced CT to assess hematoma expansion in patients with intra-cerebral hemorrhage. Articles that met one or more of the following criteria were excluded: (1) conference abstracts, reviews, letters, case reports, and case series studies with sample sizes < 10; (2) in multiple studies using the same population, only the study with

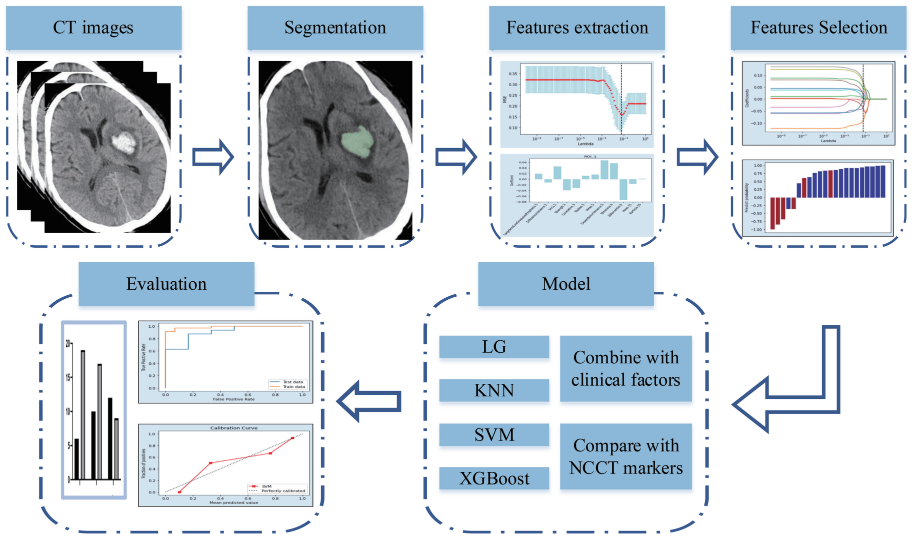


FIGURE 1
Flowchart of (NCCT-based radiomics. NCCT: non-enhanced computed tomography.

the largest dataset was included; (3) non-human studies; (4) secondary intra-cerebral hemorrhage; (5) intraventricular hemorrhage; and (6) studies without comparison. All inconsistencies were resolved by negotiation or by a third investigator.

## 2.2 Data extraction

Eligible articles contained information that was independently extracted by two researchers, including authors, year of publication, sample size, number of cases in the training and validation sets, study population, study design, study country, number of institutions, composition of model construction, mode of visualization, interval image examination, research topic, segmentation software, method of extraction of imaging histology features, validation method, method of screening variables, final study characteristics, sensitivity (Se), specificity (Sp), true positives (TPs), false positives (FPs), true negatives (TNs), false negatives (FNs), diagnostic accuracy rate (DAR), diagnostic odds ratio (DOR), number of hematoma expansions, and non-expansions. All inconsistencies were resolved by negotiation or by a third investigator.

## 2.3 Quality assessment

All eligible studies were assessed for bias using the QUADAS-2 tool for diagnostic meta-analyses (12). Four key domains were assessed, including flow and timing, reference standards, index tests, and patient selection. Three main domains were assessed using the Applicability Concerns Test. Risk of bias was categorized as low, high, or unclear. When all domains were rated as yes, the risk was considered low. A potential risk of bias existed when any of the domains was rated no. The unclear classification only applied when there were insufficient data to report. The 16 components of RQS were used to assessed the quality of radiomics studies (8). Reviewers scored each component and summed up the scores. The procedures for scoring each component have been described previously.

## 2.4 Outcome measures

We performed a synthetic analysis of TP, FP, TN, and FN indicators of eligible articles using a diagnostic meta-analysis. Comparative analyses were performed for Se, Sp, positive predictive value (PPV), negative predictive value (NPV), DAR, and DOR. Articles that did not provide the four indicators TP, FP, TN, and FN were calculated using the number of cases of hematoma expansion and non-expansion, combined with Se and Sp, using Review Manager 5.4.1. (eTable 3).

Se refers to the proportion of positive cases detected within the group diagnosed with disease by the gold standard; a higher Se indicates a lower chance of a missed diagnosis. Sp refers to the proportion of negative tests within the group diagnosed as disease-free by the gold standard; a higher Sp indicates a lower chance of misdiagnosis. PPV reflects the proportion of individuals with a positive screening test result who are actually sick. NPV reflects the proportion of individuals with a negative screening test result who do not actually have the disease. DAR is defined as the proportion of all cases detected as TPs and TNs by clinical diagnostics within all cases. A higher DOR value indicates that the diagnostic test is more effective at distinguishing between patients and non-patients.

# 3 Data Synthesis

## 3.1 Diagnostic meta-analysis to evaluate diagnostic test accuracy

Diagnostic test accuracy indicators, such as Se, Sp, PPV, NPV, DAR, and DOR, were synthesized using a meta-analysis based on a random-effects model. Forest plots were used to represent the effect values (odds ratio, OR) and 95% confidence intervals (CIs). Evaluation of the screening biomarkers (radiological features or radiomics) was based on summary receiver operating characteristic (sROC) curves and areas under the curve (AUCs), whereby a larger AUC indicates better model performance. The Cochrane Q test and $I^2$ were used to measure the heterogeneity of the outcomes. The robustness of the results was evaluated, and sources of heterogeneity were explored by omitting each included article one by one in the pooled analysis. Publication bias was evaluated using funnel plots. A p-value $< 0.05$ for the Q test or $I^2 > 50\%$ indicated the possibility of significant heterogeneity.

## 3.2 NMA

Studies that included a comparison of radiomics and radiological markers were used for the NMA. We used NMA to evaluate the diagnostic value of all radiological features and radiomics evaluating hematoma expansion in cerebral hemorrhage in all eligible studies to estimate the OR and 95% CI for predicting hematoma expansion for Se, Sp, PPV, and NPV in eligible articles, and to summarize the rank order for all screening biomarkers.

The implementation of the NMA was based on a Bayesian model using Markov chain Monte Carlo simulation methods (MCMC), where the calculated prior distribution and likelihood values were substituted into MCMC, and the parameters were adjusted to three chains and 5000 burn-ins using a random-effects model with 50,000 iterations and an interval of 5. An optimal fit state of the convergent posterior distribution was obtained, minimizing the variation of the MCMC error and deviation information criterion to stabilize the ending (13, 14). Trace plots and density distribution plots were used to assess aggregation. We constructed network plots for each outcome measure separately. The plot points represented different screening biomarkers, point sizes indicated the total sample size for each feature, and line thickness represented the number of studies that were conducted between the two points connected. The OR values and 95% CIs between different predictors were represented using forest plots. To better compare diagnostic efficacy, surface under the cumulative ranking (SUCRA) was used to calculate the rankings of the predictors (15). SUCRA values ranged between 0 and 1, with larger values representing higher rankings and diagnostic efficacies.

In hypothesis testing for the NMA, we determined homogeneity and transferability by assessing the baseline characteristics of the included articles, methodological and statistical approaches, and agreement of the same predictor across articles. For consistency tests, we explored local heterogeneity using nodal splits and the deviance information criterion for consistency and inconsistency models (DIC) (16). A smaller difference between DIC values of the two models indicated good agreement.

As most of the articles constructed the model using the training set, we only used data from the validation set for synthetic analysis. Model performance was verified using the validation set. Articles that were not categorized into the training and validation sets were analyzed as validation sets.

All data synthesis was conducted using R version 4.1.2 ("meta" package for diagnostic meta-analysis to evaluate diagnostic test accuracy; "gemtc" package for NMA) and Review Manager 5.4.1.

# 4 Results

A comprehensive search of 2114 articles was conducted, and 2037 articles were screened after excluding duplicates. Of the 77 full-text articles searched, 67 were excluded according to the inclusion and exclusion criteria (Figure 2). Ten articles comprising a total of 4929 patients were finally included in the meta-analysis, five of which were used for the NMA. (Table 1). The causes of cerebral hemorrhage were spontaneous intracerebral hemorrhage and hypertensive intraparenchymal hematoma. A total of 1391 (28.22%) patients had hematoma expansion. The radiological features or radiomics analysis performed included radiomics model, radiological model, radiomics-radiological model, black hole sign, blend sign, heterogeneity, hypodensity, irregular shape, island sign, midline shift, satellite sign, and swirl sign.

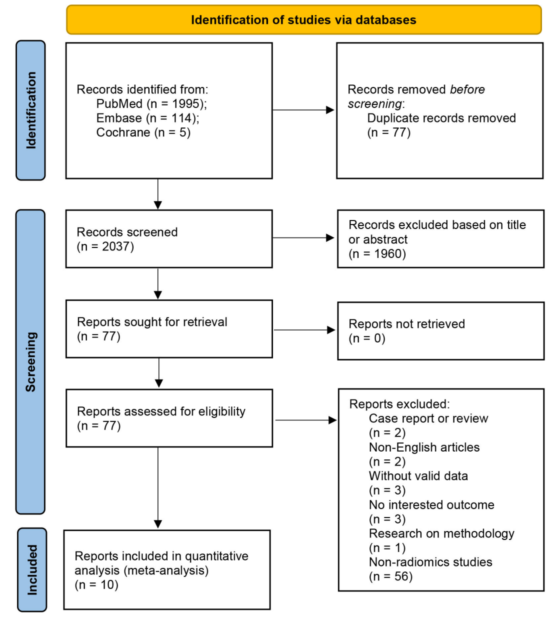


FIGURE 2
Flowchart for study selection.

Based on the QUADAS-2 tool for assessing bias and applicability (eFigure 1), the overall quality of the included articles was satisfactory. For some studies, we could not determine whether the included patients were consecutive or randomized (n=8) and whether a blinded approach and diagnostic thresholds were used for analysis (n=3); these articles were therefore categorized as unclear.

Based on RQSs (Table 2), the included articles were generally of low quality. In ten studies, the mean score was 40% (range, 30.1–69.4%), and one study scored above 50%. The protocols for image acquisition reported in most studies were well-documented. Most studies (70%) used manual segmentation (which is usually performed by an expert drawing ROIs), and three (30%) used semi-automatic segmentation (which combines manual segmentation with some algorithms). Two (20%) of the studies integrated clinical data into radiomic models and suggested that this could improve prediction accuracy.

## 4.1 Diagnostic test meta-analysis

### 4.1.1 Radiomics model

Ten studies comprising a total of 1525 patients were quantitatively analyzed for hematoma expansion after cerebral hemorrhage using the radiomics method (4, 6, 7, 17–23). The pooled Se, Sp, PPV, NPV, and DAR were 0.771 (0.710-0.832), 0.743 (0.684-0.801), 0.612 (0.448-0.737), 0.863 (0.815-0.912), and 0.748 (0.707-0.788), respectively (Figure 3). The synthetic DOR was 9.85 (6.01-16.12) (eFigure 2).

TABLE 1 Baseline characteristics of included studies.

| Author | Year | Sample size | HE | non-HE | Study design | Study location | Imaging modality | Research question |
|---|---|---|---|---|---|---|---|---|
| Chen et al | 2021 | 1153 | 236 | 917 | Retrospectively | China | NCCT | To compare the predictive performance for HE among clinical model, radiomics model, and hybrid model. |
| Duan et al | 2022 | 108 | 54 | 54 | Retrospectively | China | NCCT | To predict HE by using different machine learning methods and to determine the best radiomics model. |
| Li et al | 2019 | 167 | 42 | 125 | Retrospectively | China | NCCT | To quantify the heterogeneity of hematomas in order to find more quantitative, sensitive, and accurate indicators for predicting HE. |
| Ma et al | 2019 | 254 | 58 | 196 | Retrospectively | China | NCCT | To explore the feasibility of predicting hematoma expansion at acute phase *via* a radiomics approach. |
| Pszczolkowski et al | 2021 | 1732 | 474 | 1258 | Retrospectively | England | NCCT | To investigate the use of NCCT radiomics-based features and generalized linear models for prediction of both HE and poor functional outcome |
| Song et al | 2021 | 261 | 110 | 151 | Retrospectively | China | NCCT | To determine whether NCCT) models based on multivariable, radiomics features, and machine learning (ML) algorithms could further improve the discrimination of early hematoma expansion (HE) in patients with spontaneous intracerebral hemorrhage. |
| Xia et al | 2022 | 376 | 108 | 268 | Retrospectively | China | NCCT | To identify supratentorial spontaneous intracerebral hemorrhage patients with HE on admission |
| Xie et al | 2020 | 251 | 108 | 143 | Retrospectively | China | NCCT | To predict hematoma expansion and to compare the predictive performance with conventional radiological feature-based model |
| Zhan et al | 2021 | 313 | 44 | 269 | Retrospectively | China | NCCT | To predict HE and the short- term outcomes in patients with small hematomas. |
| Zhu et al | 2021 | 314 | 157 | 157 | Retrospectively | China | NCCT | To evaluate HE prediction in the perihematomal region using radiomics technology and compare its predictive performance with the intra-hematomal radiomics signature. |

HE: hematoma expansion; NCCT: non-enhanced computed tomography.

TABLE 2 Radiomics quality scores.

| Author | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ | ⑨ | ⑩ | ⑪ | ⑫ | ⑬ | ⑭ | ⑮ | ⑯ | Total score | Mean score (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Song | 1 | 1 | 0 | 0 | 3 | 1 | 0 | 0 | 1 | 1 | 0 | 2 | 2 | 2 | 0 | 0 | 14 | 38.9 |
| Li | 1 | 1 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 0 | 0 | 11 | 30.6 |
| Ma | 1 | 1 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 0 | 0 | 11 | 30.6 |
| Xie | 1 | 1 | 0 | 0 | 3 | 1 | 1 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 0 | 0 | 13 | 36.1 |
| Chen | 1 | 1 | 0 | 0 | 3 | 1 | 1 | 0 | 1 | 1 | 0 | 2 | 2 | 2 | 0 | 0 | 15 | 41.7 |
| Pszczolkowski | 0 | 0 | 0 | 0 | 3 | 1 | 1 | 1 | 1 | 1 | 7 | 4 | 2 | 2 | 0 | 2 | 25 | 69.4 |
| Zhan | 1 | 0 | 0 | 0 | 3 | 1 | 1 | 0 | 1 | 0 | 0 | 2 | 2 | 2 | 0 | 0 | 13 | 36.1 |
| Zhu | 0 | 1 | 0 | 0 | 3 | 1 | 1 | 0 | 1 | 1 | 0 | 2 | 2 | 2 | 0 | 0 | 14 | 38.9 |
| Duan | 1 | 1 | 0 | 0 | 3 | 1 | 0 | 0 | 1 | 0 | 0 | 2 | 2 | 2 | 0 | 0 | 13 | 36.1 |
| Xia | 1 | 1 | 0 | 0 | 3 | 1 | 0 | 0 | 1 | 1 | 0 | 3 | 2 | 2 | 0 | 0 | 15 | 41.7 |

①:Image protocol quality; ②: Multiple segmentations; ③: Phantom study on all scanners; ④: Imaging at multiple time points; ⑤: Feature reduction or adjustment for multiple testing; ⑥: Multivariable analysis with non-radiomics features; ⑦: Detect and discuss biological correlates; ⑧: Cutoff analyses; ⑨: Discrimination statistics; ⑩: Calibration statistics; ⑪: Prospective study registered in a trial database; ⑫: Validation; ⑬: Comparison to gold standard; ⑭: Potential clinical utility; ⑮: Cost effectiveness analysis; ⑯: Open science and data.

## 4.2 sROCs and AUCs

The sROC curves demonstrated similar model performance for the radiomics and radiomics-radiological models, and better performance than that of the other screening biomarkers (eFigure 3). The AUC value (95% CI) of the radiomics model was 0.80 (0.76-0.83) (eFigure 4). However, other biomarkers were not available due to the limited number of studies that obtained AUC values and 95% CIs.

## 4.3 Sensitivity analysis and publication bias

For most diagnostic indictors, the Cochrane's Q ($p<0.05$) and $I^2$ ($I^2>50\%$) tests revealed significant heterogeneity. However, no significant changes were observed in the Se, Sp, PPV, NPV, and DAR values after article-by-article exclusion, suggesting the robustness of the outcomes and relatively low potential heterogeneity (eFigures 5). Funnel plots for different diagnostic indicators of the radiomics model suggested publication bias (eFigure 6).

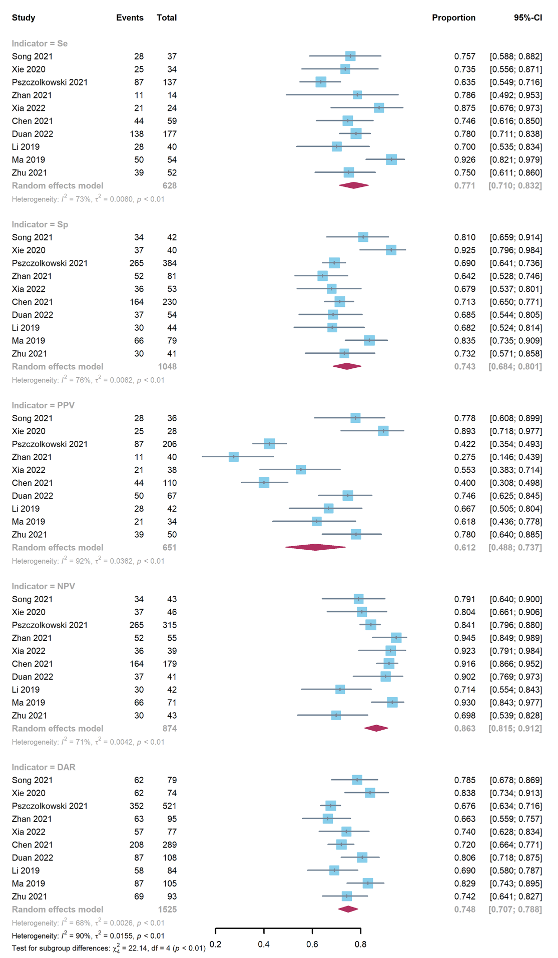


FIGURE 3

Forest plot of radiomics model. CIs: confidence intervals; DAR: diagnostic accuracy rate; NPV: negative predictive value; PPV: positive predictive value; Se: sensitivity; Sp: specificity.

## 4.4 NMA

Figure 4A presents a network plot of the indicators involved in the composition. In the NMA, 846 patients from five articles were included (4, 6, 7, 21, 22). The results synthesized according to Bayesian NMA revealed that the predictive ability of the radiomics model outperformed most of the NCCT biomarkers (Figure 5). According to SUCRA (eTable 4), both radiomics and radiomics-radiological models were ranked in the top two for Se, Sp, PPV, NPV, and DAR. SUCRA curves are presented in Figures 4B–F.

The results of the node-splitting method revealed good consistency ($p > 0.05$) in most of the direct or indirect comparisons (eFigure 7). The difference between the DIC values of both the consistent and inconsistent models did not exceed 5 and exhibited good consistency (eTable 5).

# 5 Discussion

This meta-analysis examined the utility of NCCT-based radiomics methods to predict hematoma expansion. Our analysis indicated that the radiomics approach demonstrated potential for the prediction of hematoma expansion. Despite these promising results, the relatively low RQSs of the included studies revealed that the radiomics approach was suboptimal for clinical application. Additionally, our analysis revealed that the aggregated Se, Sp, and AUC of the radiomics model outperformed those of the radiological biomarkers.

The results of our meta-analysis demonstrated that NCCT-based radiomics is a feasible approach for stratifying the risk of spontaneous intracerebral hemorrhage (21, 24–27). Hematoma expansion is associated with clinical outcomes of spontaneous intracerebral hemorrhage. Though there is currently no definitive therapeutic strategy for prevention of hematoma expansion, we believe that the HE is an appealing target for medical intervention, as it may ultimately help some patients with intracerebral hemorrhages. The CTA spot sign is useful for stratifying patient risk and providing appropriate treatment (3, 28). However, in most medical centers in China, immediate CTA is not routinely performed, thus limiting the implementation of spot signs. NCCT, which is cheaper and more convenient, is the most commonly used method for diagnosing intracerebral hemorrhage. Previous studies have reported that NCCT biomarkers, including the blend sign, black hole sign, and satellite sign, can be used to predict the risk of hematoma expansion. According to Li et al. (29), the blend sign

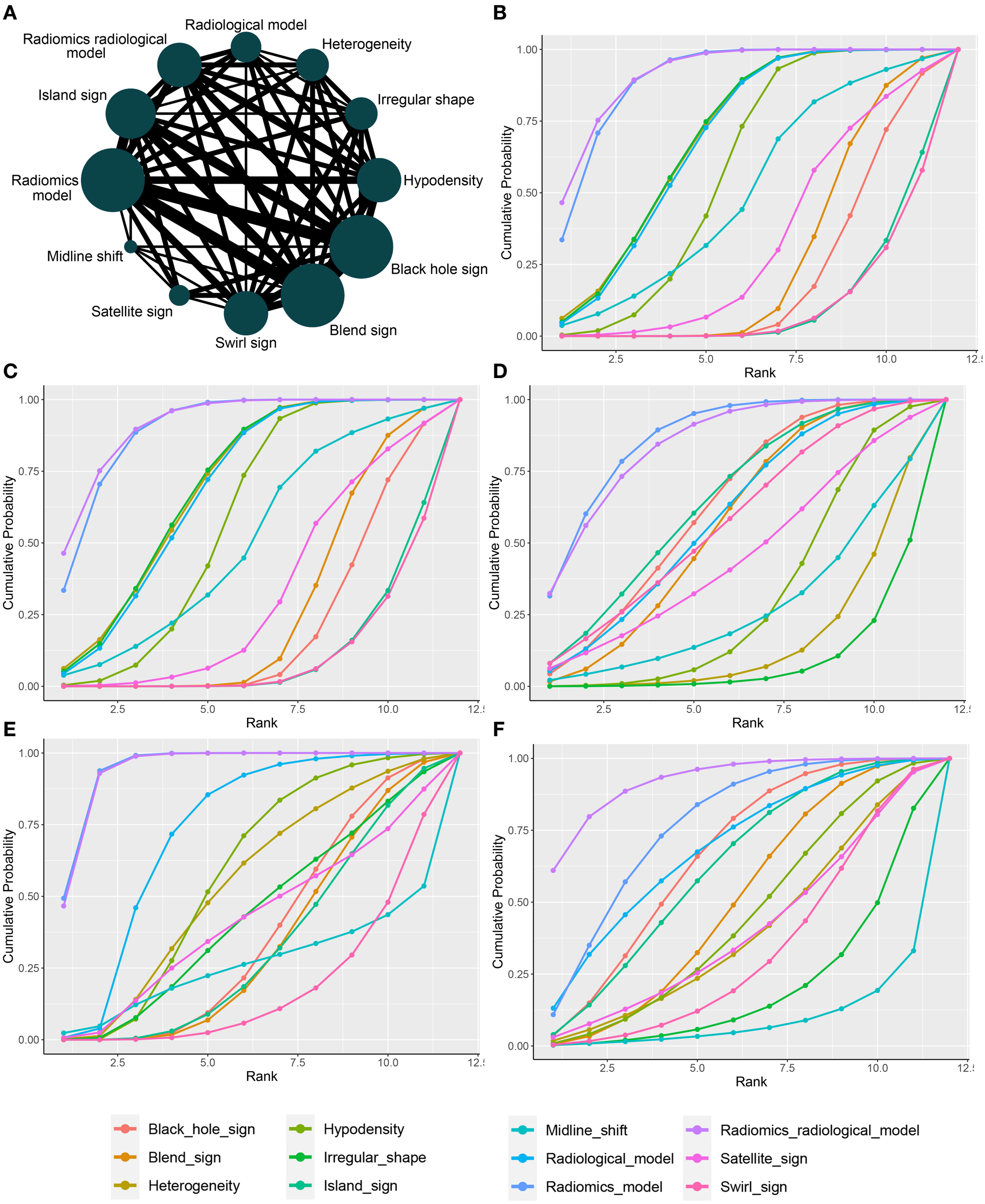


FIGURE 4
Network plot and SUCRA of biomarkers for ranking the accuracy of diagnosis. **(A)** The network plot of Bayesian network meta-analysis; **(B)** sensitivity; **(C)** specificity; **(D)** positive predictive value; **(E)** negative predictive value; **(F)** diagnosis accuracy rate. SUCRA: Surface under the cumulative ranking curve.

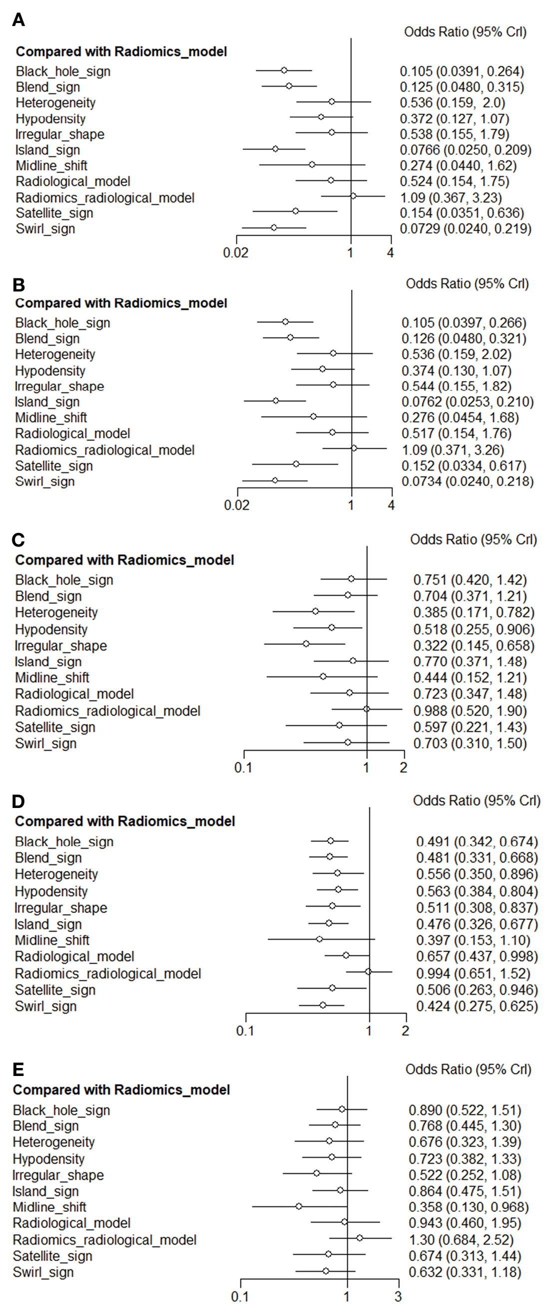


FIGURE 5
Network forest plot for biomarkers compared with radiomics model. CrI: credible interval.

(Figure 6A), which is defined as an area that has relatively low attenuation adjacent to an area with high attenuation, showed 39.3% sensitivity and 95.5% specificity for predicting hematoma growth. The black hole sign (Figure 6B), which represents a low-density area within a hematoma with high density, is reported to have sensitivity of 31.9% and specificity of 94.1% (30). One of the attributes of the satellite sign (Figure 6C) is shape irregularity. A comparative study by Shakya (31) showed that the areas under the curve for the black hole sign and the satellite sign were 63.4% and 67%, respectively. The relatively low efficacy of the predictive ability of a single NCCT biomarker restricts their clinical utility. In contrast, the radiomics studies included in our meta-analysis exhibited superior performance.

Compared to radiological markers, the radiomics method, which is based on mathematical calculation, is a more stable method to predict the risk of hematoma expansion (4, 32). The definition of radiological markers can be inconsistent, which has hindered clinical implementation of NCCT biomarkers. In this regard, a degree of overlap exists between NCCT markers (33). Moreover, the identification of radiological signs depends on the reader's experience. Our meta-analysis demonstrated that the efficacy of NCCT markers was suboptimal for implementation in clinical practice. Notably, radiomics features may quantitatively reflect the corresponding NCCT biomarkers. Although a limited number of studies was included, our results demonstrated that the radiomics method outperformed radiological biomarkers for predicting hematoma expansion.

Despite its potential, the radiomics method is relatively novel, and non-standardized imaging protocols remain commonplace. RQS was designed to measure the quality of radiomics research (8, 34). The RQSs, which includes 16 items, can be used to assess the quality of radiomics studies. Although evolving rapidly, research applying radiomics must comply with certain basic principles. For instance, data obtained from other institutions is considered to be more independent and therefore more reliable when compared to data obtained internally. External validation of models is crucial for ensuring their generalizability. Indeed, the lack of external validation is the main factor for a low RQS (35). Standardization of high-quality image-extracted data may be helpful for clinical decision support systems (36, 37).

This study had several limitations. Meta-analysis had the limitation of heterogeneity among studies included. Based on the methods used for image reconstruction, feature extraction, and algorithms used, there were considerable differences between the included studies. Second, there was a limited number of eligible studies in the meta-analysis, possible because the relative improvement in performance of the radiomics method was overestimated, and the radiological markers were understated. Third, radiomics studies are generally of low quality, most lack external validation, and promising results from radiomics should be interpreted with caution. Higher-level evidence from clinical trials is necessary for clinical implementation of radiomics approaches.

In conclusion, our meta-analysis highlights the potential of NCCT-based radiomics approaches to predict hematoma expansion. In this regard, we recommend a radiomics

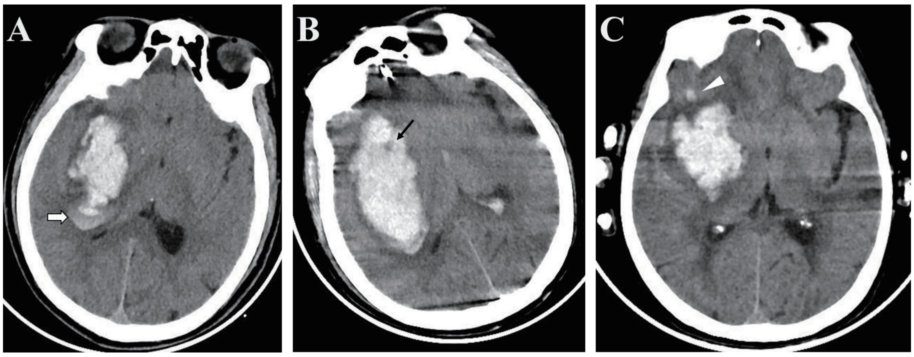


FIGURE 6
Examples of biomarkers on non-enhanced computed tomography **(A)** The blend sign is defined as an area with relatively low attenuation adjacent to an area with high attenuation (white arrow); **(B)** The black hole sign represents a low-density area within a hematoma with high density (black arrow); **(C)** The satellite sign represents a small lesion completely separate from the main hemorrhage visible in at least one slice (white triangle).

approach over NCCT markers. Nevertheless, standardization of radiomics approaches is necessary for further clinical application, and further multicenter prospective studies with stricter designs are warranted to verify 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.

## Author contributions

Y-WJ and X-JX contributed to the study conception and design. Material preparation, data collection and analysis were performed by Y-WJ, X-JX, and C-MC. The first draft of the manuscript was written by Y-WJ and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

## Funding

This work was supported by Joint Funds for the innovation of science and Technology, Fujian province (Grant number: 2020Y9055).

## Acknowledgments

We thank the colleagues in our department for their help in our 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

## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.973104/full#supplementary-material

# References

1. Romero JM, Hito R, Dejam A, Ballesteros LS, Cobos CJ, Liévano JO, et al. Negative spot sign in primary intracerebral hemorrhage: Potential impact in reducing imaging. *Emerg Radiol* (2017) 24:1–6. doi: 10.1007/s10140-016-1428-8

2. Boulouis G, Morotti A, Charidimou A, Dowlatshahi D, Goldstein JN. Noncontrast computed tomography markers of intracerebral hemorrhage expansion. *Stroke* (2017) 48:1120–5. doi: 10.1161/STROKEAHA.116.015062

3. Phan TG, Krishnadas N, Lai VWY, Batt M, Slater L-A, Chandra RV, et al. Meta-analysis of accuracy of the spot sign for predicting hematoma growth and clinical outcomes. *Stroke* (2019) 50:2030–6. doi: 10.1161/STROKEAHA.118.024347

4. Pszczolkowski S, Manzano-Patrón JP, Law ZK, Krishnan K, Ali A, Bath PM, et al. Quantitative CT radiomics-based models for prediction of haematoma expansion and poor functional outcome in primary intracerebral haemorrhage. *Eur Radiol* (2021) 31:7945–59. doi: 10.1007/s00330-021-07826-9

5. Chen Q, Zhang L, Mo X, You J, Chen L, Fang J, et al. Current status and quality of radiomic studies for predicting immunotherapy response and outcome in patients with non-small cell lung cancer: A systematic review and meta-analysis. *Eur J Nucl Med Mol Imaging* (2021) 49:345–60. doi: 10.1007/s00259-021-05509-7

6. Xie H, Ma S, Wang X, Zhang X. Noncontrast computer tomography-based radiomics model for predicting intracerebral hemorrhage expansion: preliminary findings and comparison with conventional radiological model. *Eur Radiol* (2020) 30:87–98. doi: 10.1007/s00330-019-06378-3

7. Zhan C, Chen Q, Zhang M, Xiang Y, Chen J, Zhu D, et al. Radiomics for intracerebral hemorrhage: are all small hematomas benign? *Br J Radiol* (2021) 94:20201047. doi: 10.1259/bjr.20201047

8. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong ECE, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* (2017) 14:749–62. doi: 10.1038/nrclinonc.2017.141

9. Jiang X, Hui F, Qin X, Wu Y, Liu H, Gao J, et al. Diagnosis accuracy and prognostic significance of the dickkopf-1 protein in gastrointestinal carcinomas: Systematic review and network meta-analysis. *J Cancer* (2020) 11:7091–100. doi: 10.7150/jca.49970

10. Laukhtina E, Shim SR, Mori K, D'Andrea D, Soria F, Rajwa P, et al. Diagnostic accuracy of novel urinary biomarker tests in non-muscle-invasive bladder cancer: a systematic review and network meta-analysis. *Eur Urol Oncol* (2021) 4:927–42. doi: 10.1016/j.euo.2021.10.003

11. 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: 10.1186/s13643-021-01626-4

12. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* (2011) 155:529–36. doi: 10.7326/0003-4819-155-8-201110180-00009

13. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* (2004) 23:3105–24. doi: 10.1002/sim.1875

14. Shim SR, Kim SJ, Lee J, Rucker G. Network meta-analysis: Application and practice using r software. *Epidemiol Health* (2019) 41:e2019013. doi: 10.4178/epih.e2019013

15. Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: An overview and tutorial. *J Clin Epidemiol* (2011) 64:163–71. doi: 10.1016/j.jclinepi.2010.03.016

16. van Valkenhoef G, Dias S, Ades AE, Welton NJ. Automated generation of node-splitting models for assessment of inconsistency in network meta-analysis. *Res Synth Methods* (2016) 7:80–93. doi: 10.1002/jrsm.1167

17. Chen Q, Zhu D, Liu J, Zhang M, Xu H, Xiang Y, et al. Clinical-radiomics nomogram for risk estimation of early hematoma expansion after acute intracerebral hemorrhage. *Acad Radiol* (2021) 28:307–17. doi: 10.1016/j.acra.2020.02.021

18. Duan C, Liu F, Gao S, Zhao J, Niu L, Li N, et al. Comparison of radiomic models based on different machine learning methods for predicting intracerebral hemorrhage expansion. *Clin Neuroradiol* (2022) 32:215–23. doi: 10.1007/s00062-021-01040-2

19. Li H, Xie Y, Wang X, Chen F, Sun J, Jiang X. Radiomics features on non-contrast computed tomography predict early enlargement of spontaneous intracerebral hemorrhage. *Clin Neurol Neurosurg* (2019) 185:105491. doi: 10.1016/j.clineuro.2019.105491

20. Ma C, Zhang Y, Niyazi T, Wei J, Guocai G, Liu J, et al. Radiomics for predicting hematoma expansion in patients with hypertensive intraparenchymal hematomas. *Eur J Radiol* (2019) 115:10–5. doi: 10.1016/j.ejrad.2019.04.001

21. Song Z, Guo D, Tang Z, Liu H, Li X, Luo S, et al. Noncontrast computed tomography-based radiomics analysis in discriminating early hematoma expansion after spontaneous intracerebral hemorrhage. *Korean J Radiol* (2021) 22:415–24. doi: 10.3348/kjr.2020.0254

22. Xia X, Ren Q, Cui J, Dong H, Huang Z, Jiang Q, et al. Radiomics for predicting revised hematoma expansion with the inclusion of intraventricular hemorrhage growth in patients with supratentorial spontaneous intraparenchymal hematomas. *Ann Transl Med* (2022) 10:8. doi: 10.21037/atm-21-6158

23. Zhu D, Zhang M, Li Q, Liu J, Zhuang Y, Chen Q, et al. Can perihaematomal radiomics features predict haematoma expansion? *Clin Radiol* (2021) 76:629.e1–e9. doi: 10.1016/j.crad.2021.03.003

24. Morotti A, Arba F, Boulouis G, Charidimou A. Noncontrast CT markers of intracerebral hemorrhage expansion and poor outcome: A meta-analysis. *Neurology* (2020) 95:632–43. doi: 10.1212/WNL.0000000000010660

25. Morotti A. CT markers of intracerebral hemorrhage expansion: Different sides of the same coin? *Neurocrit Care* (2019) 31:451–2. doi: 10.1007/s12028-019-00768-7

26. Wei Y, Zhu G, Gao Y, Chang J, Zhang H, Liu N, et al. Island sign predicts hematoma expansion and poor outcome after intracerebral hemorrhage: A systematic review and meta-analysis. *Front Neurol* (2020) 11:429. doi: 10.3389/fneur.2020.00429

27. Yogendrakumar V, Moores M, Sikora L, Shamy M, Ramsay T, Fergusson D, et al. Evaluating hematoma expansion scores in acute spontaneous intracerebral hemorrhage: a systematic scoping review. *Stroke* (2020) 51:1305–8. doi: 10.1161/STROKEAHA.119.028574

28. Lei C, Geng J, Chen C, Chang X. Accuracy of the blend sign on computed tomography as a predictor of hematoma growth after spontaneous intracerebral hemorrhage: a systematic review. *J Stroke Cerebrovasc Dis* (2018) 27:1705–10. doi: 10.1016/j.jstrokecerebrovasdis.2018.01.032

29. Li Q, Zhang G, Huang YJ, Dong M-X, Lv F-J, Wei X, et al. Blend sign on computed tomography: novel and reliable predictor for early hematoma growth in patients with intracerebral hemorrhage. *Stroke* (2015) 46:2119–23. doi: 10.1161/STROKEAHA.115.009185

30. Li Q, Zhang G, Xiong X, Wang X-C, Yang W-S, Li K-W, et al. Black hole sign: novel imaging marker that predicts hematoma growth in patients with intracerebral hemorrhage. *Stroke* (2016) 47:1777–81. doi: 10.1161/STROKEAHA.116.013186

31. Shakya MR, Fu F, Zhang M, Shan Y, Yu F, Sun S, et al. Comparison of black hole sign, satellite sign, and iodine sign to predict hematoma expansion in patients with spontaneous intracerebral hemorrhage. *BioMed Res Int* (2021) 2021:3919710. doi: 10.1155/2021/3919710

32. Shen Q, Shan Y, Hu Z, Chen W, Yang B, Han J, et al. Quantitative parameters of CT texture analysis as potential markersfor early prediction of spontaneous intracranial hemorrhage enlargement. *Eur Radiol* (2018) 28:4389–96. doi: 10.1007/s00330-018-5364-8

33. Morotti A, Boulouis G, Dowlatshahi D, Li Q, Barras CD, Delcourt C, et al. Standards for detecting, interpreting, and reporting noncontrast computed tomographic markers of intracerebral hemorrhage expansion. *Ann Neurol* (2019) 86:480–92. doi: 10.1002/ana.25563

34. Sollini M, Antunovic L, Chiti A, Kirienko M. Towards clinical application of image mining: a systematic review on artificial intelligence and radiomics. *Eur J Nucl Med Mol Imaging* (2019) 46:2656–72. doi: 10.1007/s00259-019-04372-x

35. Ursprung S, Beer L, Bruining A, Woitek R, Stewart GD, Gallagher FA, et al. Radiomics of computed tomography and magnetic resonance imaging in renal cell carcinoma-a systematic review and meta-analysis. *Eur Radiol* (2020) 30:3558–66. doi: 10.1007/s00330-020-06666-3

36. Arba F, Rinaldi C, Boulouis G, Fainardi E, Charidimou A, Morotti A. Noncontrast computed tomography markers of cerebral hemorrhage expansion: diagnostic accuracy meta-analysis. *Int J Stroke* (2021) 17 (8), 17474930211061639. doi: 10.1177/17474930211061639

37. Li Z, You M, Long C, Bi R, Xu H, He Q, et al. Hematoma expansion in intracerebral hemorrhage: an update on prediction and treatment. *Front Neurol* (2020) 11:702. doi: 10.3389/fneur.2020.00702

frontiers | Frontiers in Oncology

TYPE Systematic Review
PUBLISHED 31 January 2023
DOI 10.3389/fonc.2022.960944

# The diagnostic performance of radiomics-based MRI in predicting microvascular invasion in hepatocellular carcinoma: A meta-analysis

Gao Liang[1], Wei Yu[1], Shuqin Liu[1], Mingxing Zhang[1], Mingguo Xie[1]*, Min Liu[2] and Wenbin Liu[1]

[1]Department of Radiology, Hospital of Chengdu University of Traditional Chinese Medicine, Chengdu, Sichuan, China, [2]Toxicology Department, West China-Frontier PharmaTech Co., Ltd. (WCFP), Chengdu, Sichuan, China

EDITED BY
Jakub Nalepa,
Silesian University of Technology, Poland

REVIEWED BY
Llewellyn Padayachy,
University of Pretoria, South Africa
Renguo Guan,
Sun Yat-sen University Cancer Center (SYSUCC), China

*CORRESPONDENCE
Mingguo Xie
✉ xmg6806@163.com

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 03 June 2022
ACCEPTED 23 December 2022
PUBLISHED 31 January 2023

CITATION
Liang G, Yu W, Liu S, Zhang M, Xie M, Liu M and Liu W (2023) The diagnostic performance of radiomics-based MRI in predicting microvascular invasion in hepatocellular carcinoma: A meta-analysis. *Front. Oncol.* 12:960944.
doi: 10.3389/fonc.2022.960944

**Objective:** The aim of this study was to assess the diagnostic performance of radiomics-based MRI in predicting microvascular invasion (MVI) in hepatocellular carcinoma (HCC).

**Method:** The databases of PubMed, Cochrane library, Embase, Web of Science, Ovid MEDLINE, Springer, and Science Direct were searched for original studies from their inception to 20 August 2022. The quality of each study included was assessed according to the Quality Assessment of Diagnostic Accuracy Studies 2 and the radiomics quality score. The pooled sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), and diagnostic odds ratio (DOR) were calculated. The summary receiver operating characteristic (SROC) curve was plotted and the area under the curve (AUC) was calculated to evaluate the diagnostic accuracy. Sensitivity analysis and subgroup analysis were performed to explore the source of the heterogeneity. Deeks' test was used to assess publication bias.

**Results:** A total of 15 studies involving 981 patients were included. The pooled sensitivity, specificity, PLR, NLR, DOR, and AUC were 0.79 (95%*CI*: 0.72–0.85), 0.81 (95%*CI*: 0.73–0.87), 4.1 (95%*CI*:2.9–5.9), 0.26 (95%*CI*: 0.19–0.35), 16 (95%*CI*: 9–28), and 0.87 (95%*CI*: 0.84–0.89), respectively. The results showed great heterogeneity among the included studies. Sensitivity analysis indicated that the results of this study were statistically reliable. The results of subgroup analysis showed that hepatocyte-specific contrast media (HSCM) had equivalent sensitivity and equivalent specificity compared to the other set. The least absolute shrinkage and selection operator method had high sensitivity and specificity than other methods, respectively. The investigated area of the region of interest had high specificity compared to the volume of interest. The imaging-to-surgery interval of 15 days had higher sensitivity and slightly low specificity than the others. Deeks' test indicates that there was no publication bias ($P$=0.71).

**Conclusion:** Radiomics-based MRI has high accuracy in predicting MVI in HCC, and it can be considered as a non-invasive method for assessing MVI in HCC.

KEYWORDS

**hepatocellular carcinoma, microvascular invasion, MRI, radiomics, meta-analysis**

# Introduction

Hepatocellular carcinoma (HCC) is the most common primary liver malignant tumor, which is also the third leading cause of cancer death (1, 2). Hepatectomy and liver transplantation are still the main treatments for HCC (3, 4). Despite curative therapies, the prognosis of HCC patients remains poor, with 5-year recurrence rates reaching 50%–70% after hepatectomy and <35% after liver transplantation (5–7). It was proven that 15.0%–57.1% of patients presented microvascular invasion (MVI) after hepatectomy, which is a well-established risk factor for postoperative recurrence (8–10). In addition, the 5-year survival rate of patients with MVI significantly declined (11). For the MVI-positive patients, a wide resection margin is recommended. Therefore, an accurate prediction of MVI before operation is of great importance for clinical treatment decision and prognosis evaluation.

MVI is defined as the cancer cell nest in small vessels lined with endothelium, which is visible only under microscopy (12). Conventional imaging methods are of limited value and pose a challenge for non-invasive diagnosis in assessing MVI in HCC. In recent years, radiomics has been widely applied in the tumor diagnosis, the evaluation of response to treatment, and prognosis prediction. As a new and non-invasive technology, radiomics can high-throughput-extract features from large quantities of images to improve diagnostic or prognostic accuracy, which is also effective to preoperatively predict MVI (13). As imaging markers, the extracted radiomics feature can reflect the microscopic pathological changes of the tumor (Supplementary Figure S1), which is promising in the diagnosis of carcinomas (14).

MRI can also provide better soft-tissue resolution, multiparameters, and more stable features for assessing tumor heterogeneity. Previous similar studies have included CT-, MRI-, and US-combined radiomics original studies (13–15). Although they made a subgroup analysis of different imaging modalities, the number of MRI-based radiomics studies included was small. There is no unified conclusion regarding the accuracy of radiomics-based MRI for predicting MVI in HCCs. The current meta-analysis aimed to comprehensively and systematically assess the accuracy of radiomics-based MRI in evaluating the MVI of HCCs.

# Materials and methods

Patients, public-involvement patients, and the public were not involved in this study.

## Searching strategies

The literature search was independently performed by two radiologists. The databases were searched from their inception to 20 August 2022 including PubMed, Cochrane Library, Embase, Web of Science, Ovid MEDLINE, Springer, and ScienceDirect. The search terms were "hepatocellular carcinoma," "liver malignant tumor," "liver cancer," "liver cell carcinoma," "texture analysis," "radiomics," "advanced analysis," etc. The titles and abstracts were searched for their relevance. Disagreements were discussed and resolved to reach a consensus. In addition, the search strategy is presented in detail in Supplementary File 1.

## Study selection

Studies were selected according to the following criteria: (1) original research studies. (2) HCC patients with MVI were confirmed by biopsy or histopathology. (3) Data were available and could be extracted for calculating the true-positive (TP), false-positive (FP), true-negative (TN), and false-negative (FN) values. (4) MRI-based radiomics was applied to predict MVI in HCC. (5) English literature: the excluding criteria were case reports, reviews, abstracts, meta-analyses, insufficient calculable data, or animal studies.

## Data extraction

The relevant information extracted from the original study was as follows: the first author, the year of publication, country and language, sample size, research type, gold standard, the age of patients, TP, FP, FN, TN, MRI field strengths, and radiomics software. When there is a disagreement in the process of document screening and data extraction, the third radiologist will discuss and resolve it.

## Quality assessment of included studies

The quality of each study was assessed on the basis of the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) guideline and the radiomics quality score (RQS) (16, 17), which is recommended by the Cochrane collaboration web. The QUADAS-2 tool consists of four parts: (1) patient selection; (2) index test; (3) reference standard; and (4) flow and timing. The RQS checklist is described in Supplemental Table S1.

## Statistical analysis

Meta-analysis was performed by Stata version 15.1, and Review Manager software, version 5.3. We adopted a bivariate random effects model to calculate the pooled estimates in advance. The Cochran-Q method and inconsistency index ($I^2$) were used to investigate heterogeneity among the studies. If $I^2 > 50\%$, $P < 0.05$, the observed heterogeneity was significant. If $I^2 < 50\%$, $P > 0.05$, the observed heterogeneity was not significant. If there were obvious heterogeneity, the Spearman's correlation coefficient was used to assess the threshold effect between the sensitivity logit and the specificity logit. If there were no threshold effect, sensitivity analysis and subgroup analysis were performed to further investigate the cause of the heterogeneity.

Pooled sensitivity (Sen), specificity (Spec), PLR, NLR, and DOR were calculated to assess the diagnostic performance of radiomics-based MRI. The summary receiver operating characteristic (SROC) curve was plotted, and the area under the curve (AUC) was calculated. Deeks' test was used to evaluate publication bias, and $P > 0.05$, which indicates that there was no significant bias.

## Clinical utility

A Fagan plot was used to evaluate the clinical utility, which demonstrated the posttest probability (P post) of MVI when pretest probabilities were calculated.

# Results

## Research and selection of studies

A total of 661 relevant studies were initially identified from multidatabases, and 229 duplicated articles were excluded. Additionally, 385 records were removed after reading their titles and abstracts and being deemed irrelevant. Subsequently, after reading the full texts, 28 articles were found to be reviews or not related to the MVI of HCC, and 4 articles were unavailable for data extraction. Ultimately, 15 articles were included (18–32). The literature search process is shown in Figure 1.

## Study characteristics

The characteristics of the included studies are shown in Tables 1, 2. All 15 studies were retrospective cohort studies. The total number of patients was 981. From the included studies, the number of MVIs and no MVIs were reported and the pathological histology was used as reference standards. Six studies used hepatocyte-specific contrast media (HSCM). The LASSO method and other methods were used as the method for selection in 11 studies and 4 studies, respectively.

## Quality assessment and publication bias

The quality of the included studies was evaluated according to the QUADAS-2 checklist, and the results are shown in detail in Table 3. It was observed that the 'index test' in the 'risk of bias' and 'applicability concerns' revealed uncertain shortcomings, which may suggest bias regarding inclusion. Overall, the quality of all included studies was satisfactory. Deeks' funnel plot asymmetry test was used to assess the potential publication bias. The results indicated that there was no significant bias ($P = 0.71$), which are shown in Figure 2. The 15 studies reached a mean ± standard deviation RQS of 14.80 ± 1.57, median 16, and range 12–17. The average percentage RQS was 20.6% with a maximum of 47.2%. The RQS individual scores and inter-rater agreement are presented in Supplemental Tables S2, S3.The RQS was reached with good inter-rater agreement (ICC 0.977, 95% CI 0.934–0.992).

## Meta-analysis

The results of the meta-analysis are presented in Figures 3, 4. Pooled sensitivity and specificity were 0.79 (95% *CI* 0.72–0.85) and 0.81 (95% *CI* 0.73–0.87), respectively. The values of PLR, NLR, and DOR were 4.1 (95% *CI* 2.9–5.9), 0.26 (95% *CI* 0.19–0.35), and 16 (95% *CI* 9–28), respectively. The AUC of SROC was 0.87 (95% *CI* 0.84–

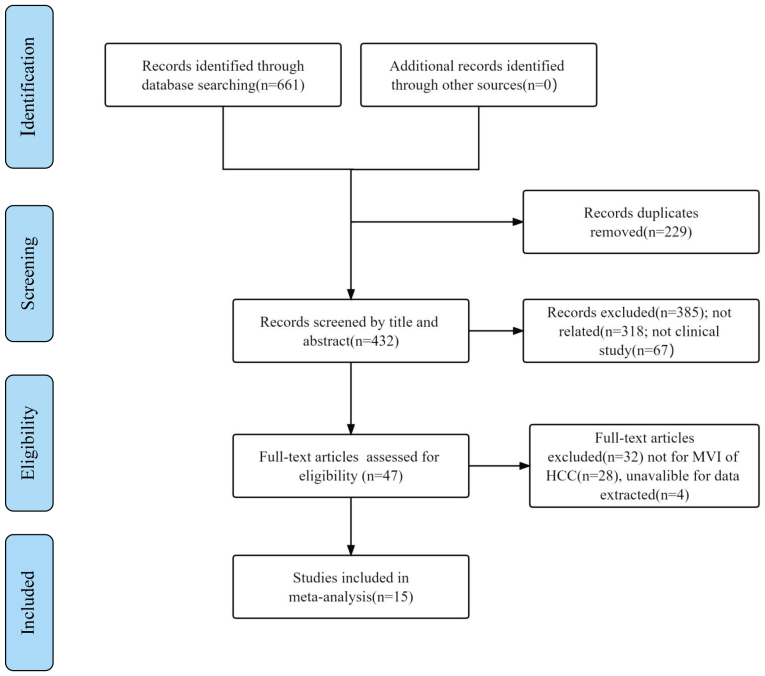


FIGURE 1
Included study selection process for this meta-analysis.

**TABLE 1** Characteristics of included studies in the meta-analysis.

| Author | Year | Country | Study design | Imaging-to-surgery interval | Tumor size (cm), mean (range) | Tumor number | Patient number (all) | (Male/female) | MVI (+) | MVI (-) | MRI parameters | | Radiomics software | Gold standard | Data | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | Contrast media | Field strength (T) | | | TP | FP | FN | TN |
| Feng (18) | 2019 | China | Re | Within 1 month | 4.3 (2.7, 6.0) | 50 | 50 | 46/4 | 20 | 30 | HSCM | 3.0 | A.K | Histology | 18 | 7 | 2 | 23 |
| Zhang R (19) | 2019 | China | Re | Within 1 month | MVI(+)<br>5.13 (1.4–10.2), MVI(-)<br>4.00 (0.8–9.7) | 73 | 73 | 64/9 | 26 | 47 | Other | 3.0 | MATLAB | Histology | 21 | 15 | 5 | 32 |
| Chong (20) | 2018 | China | Re | Within 1 month | within 5.0 | 106 | 106 | 88/18 | 30 | 76 | Other | 1.5 | Python | Histology | 28 | 11 | 2 | 65 |
| Zhu YJ (21) | 2019 | China | Re | 15 days (range, 7–35 days) | MVI(+)3.82 ±0.88, MVI(-) 3.21 ± 0.94 | 99 | 99 | 32/54 | 37 | 62 | Other | 3.0 | Omni-Kinetics | Histology | 28 | 17 | 9 | 45 |
| Willson G (22) | 2020 | USA | Re | Within 3 months | 4.5 (2.3–6) | 36 | 36 | 32/4 | 22 | 14 | Other | 1.5 or 3.0 | TexRAD | Histology | 15 | 5 | 7 | 9 |
| Zhang Y (23) | 2021 | China | Re | Within a week | MVI(+)4.00 (2.73–5.00), MVI(-) 3.20 (2.00–5.00) | 59 | 59 | 50/9 | 34 | 25 | NA | 3.0 | A.K | Histology | 26 | 5 | 8 | 20 |
| Nebbia (24) | 2020 | USA | Re | Within a week | MVI(+)3.45, MVI(-) 3.84 | 99 | 99 | 83/16 | 61 | 38 | Other | 1.5 | Python | Histology | 49 | 8 | 12 | 30 |
| Chen Y (25) | 2020 | China | Re | Within 2 weeks | NA | 81 | 81 | NA | 33 | 48 | HSCM | 3.0 | Python | Histology | 26 | 0 | 7 | 48 |
| Dai (26) | 2020 | China | Re | Within a month | MVI(+) 5.54 ± 2.68 (2.3–11.3), MVI(-) 4.49 ± 2.12(1.4–9.2) | 69 | 69 | 65/4 | 29 | 40 | Other | 3.0 | MATLAB | Histology | 27 | 7 | 2 | 33 |
| Meng (27) | 2021 | China | Re | Within a month | 3.4 (2.4–4.7) | 102 | 102 | 84/18 | 31 | 71 | NA | 3.0 | Python | Histology | 16 | 9 | 15 | 62 |
| Yang Y (28) | 2021 | China | Re | Within a month | NA | 53 | 53 | 40/13 | 26 | 27 | HSCM | 1.5 or 3.0 | Python | Histology | 14 | 1 | 12 | 26 |
| Qu C (29) | 2022 | China | Re | Within a month | MVI(+) 2.98 ± 1.13, MVI(-)2.94 ± 1.04 | 53 | 53 | 45/8 | 24 | 29 | Other | 3.0 | Python | Histology | 17 | 12 | 7 | 17 |
| Jiang T (30) | 2022 | China | Re | NA | MVI(+) 5.70 ± 3.97, MVI(-) 3.91 ± 1.92 | 21 | 21 | 17/4 | 10 | 11 | HSCM | 3.0 | R software | Histology | 9 | 1 | 1 | 10 |
| Gao L (31) | 2022 | China | Re | Within a month | NA | 35 | 35 | 29/6 | 19 | 16 | HSCM | 3.0 | R software | Histology | 17 | 4 | 2 | 12 |
| Tian Y (32) | 2022 | China | Re | Within a month | Within 3.0 | 45 | 45 | 35/10 | 13 | 32 | HSCM | 3.0 | R software | Histology | 11 | 14 | 2 | 18 |

HSCM, hepatocyte-specific contrast media; NA, not attended; Re, retrospective; A.K, Artificial Intelligent Kit software.

TABLE 2 Radiomic characteristics of included studies in the meta-analysis.

| Author | Investigated area | Segmentation method | Feature extraction | Radiomic feature categories | Machine-learning method for feature selection | Number of selected features | AUC of radiomic model with the best performance | AUC of radiomic–clinical model |
|---|---|---|---|---|---|---|---|---|
| Feng (18) | VOI: tumor | Manual delineation | 1,044 radiomic features | Gray-level histogram, texture analysis, wavelet features | LASSO/LR | 10 radiomic features | Training 0.850, Validation 0.833 | NA |
| Zhang R (19) | ROI: tumor and surrounding tissue | Manual delineation | 484 radiomic features | Intensity features, texture features, wavelet features | mRMR/LR | mRMR features | Training 0.784, Validation 0.820 | Training 0.753, Validation 0.729 |
| Chong (20) | VOI: tumor | Manual delineation | 854 radiomic features | Shape, size, intensity, and texture features | LASSO/RF, LR | 4 subsets of radiomic features | Training 0.999, Validation 0.918 | Training 0.798, Validation 0.725 |
| Zhu YJ (21) | VOI: tumor | Manual delineation | 58 texture features | Texture features | LR/texture analysis | 10, 12 texture features AP, PP | Training 0.765, Validation 0.773 | Training 0.810, Validation 0.794 |
| Willson G (22) | ROI: largest cross section | manual drawn | 6 type texture features | Texture features | NA/LR | NA | 0.83 | NA |
| Zhang Y (23) | VOI: tumor | Manual segmentation | 396 radiomic features | GLCM, GLSZM, RLM, formfactor, haralick features | LASSO/LR | 6 subsets of radiomic features | Training 0.889, Validation 0.822 | Training 0.901, Validation 0.840 |
| Nebbia (24) | VOI: tumor and margin | Manual segmentation | 100 radiomic features | Shape features, first-order features, texture features | LASSO/ SVM, decision trees, LR | NA | 0.808 | NA |
| Chen Y (25) | VOI: tumor | Manual segmentation | 1,395 radiomic features | First-order features, texture features, high-order features | LASSO/ SVM, XGBoost, LR | 6 subsets of radiomic features | Training 1.00, Validation 0.842 | NA |
| Dai (26) | ROI: axial slice | Manual segmentation | 167 radiomic features | Shape features, intensity features, texture features | mRMR, LASSO/RF, SVM, LR | 68 radiomic features | 0.792 | NA |
| Meng (27) | VOI: tumor | Manually drawn | 10,304 radiomic features | Shape features, first-order features, high-order features | LASSO/LR | 2,114 radiomic features | 0.804 | 0.872 |
| Yang Y (28) | VOI: tumor and margin | Manual segmentation | 851 radiomic features | First-order features, shape features, texture features, wavelet-transformed features | LASSO/ mRMR | NA | Training 0.896, Validation 0.788 | Training 0.932, Validation 0.917 |
| Qu C (29) | VOI: tumor and margin | Manual segmentation | 874 radiomic features | Shape, first-order statistics, GLCM, GLRLM, GLSZM, GLDM | RFE algorithm | 560 radiomic feature | Training 0.89, Validation 0.66 | Training 0.90, Validation 0.70 |
| Jiang T (30) | ROI: largest cross section | Manual segmentation | 1,967 radiomic features | Shapes, first-order statistics, filter-transformed features, GLCM, GLSZM, GLDM, GLCM | LASSO/ least absolute shrinkage | 11 radiomic features | Training 0.807, Validation 0.835 | NA |
| Gao L (31) | VOI: tumor and margin | Manual segmentation | 107 radiomic features | Shape-based characteristics, first-order statistics, textural features | LR, SVC, RFC, adaboost | NA | Training 0.823, Validation 0.740 | Training 0.915, Validation 0.868 |
| Tian Y (32) | VOI: tumor and margin | Manual segmentation | 1,561 radiomic features | Shape-based features, first-order statistics features, GLCM, GLRLM, GLSZM, GLDM | LASSO/ least absolute shrinkage | 43 radiomic features | Training 0.842, Validation 0.800 | Training 0.934, Validation 0.889 |

NA, not available; ROI, region of interest; VOI, volume of interest; LASSO, least absolute shrinkage and selection operator; GLCM, gray-level co-occurrence matrix; GLSZM, gray-level size zone matrix; LR, logistic regression; SVM, support vector machine; RLM, run length matrix; mRMR, minimum redundancy maximum relevance; GLRLM, gray-level run length matrix; GLDM, gray-level dependence matrix; RFE, recursive feature elimination; SVC, support vector classifier; RFC, random forest classifier.

TABLE 3 Results of the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) quality assessment of included studies.

| Study | Risk of bias | | | Applicability concerns | | | |
|---|---|---|---|---|---|---|---|
| | Patient selection | Index test | Reference standard | Flow and timing | Patient selection | Index test | Reference standard |
| Feng (18) | + | + | + | + | + | + | + |
| Zhang.R (19) | + | + | + | + | + | + | + |
| Chong (20) | + | + | + | + | + | + | + |
| Zhu YJ (21) | + | + | + | + | + | + | + |
| Willson G (22) | + | ? | + | + | + | + | + |
| Zhang Y (23) | + | ? | + | + | + | + | + |
| Nebbia (24) | + | + | + | + | + | + | + |
| Chen Y (25) | + | + | + | + | + | ? | + |
| Dai (26) | + | + | + | + | + | ? | + |
| Meng (27) | + | + | + | + | + | + | + |
| Yang Y (28) | + | + | + | + | + | + | + |
| Qu C (29) | + | + | + | + | + | + | + |
| Jiang T (30) | + | + | + | + | + | ? | + |
| Gao L (31) | + | ? | + | + | + | + | + |
| Tian Y (32) | + | + | + | + | + | + | + |

+: Low risk; -: High risk;?: Unclear risk.
QUADAS, Quality Assessment of Diagnostic Accuracy Studies.

0.89). These findings indicated that radiomics-based MRI has a high diagnostic performance for evaluating MVI in HCC.

## Exploration of heterogeneity

Heterogeneity was tested using Cochran-Q and $I^2$. In Figure 3, the *P*-value of the Cochran-Q test was 0.00 ($P < 0.05$), and $I^2$ was 61.12% in pooled sensitivity. Additionally, the *P*-value of the Cochran-Q test was 0.00 ($P < 0.05$), and $I^2$ was 71.58% in pooled specificity. These results indicated that there was significant heterogeneity in pooled sensitivity and specificity among the included studies.

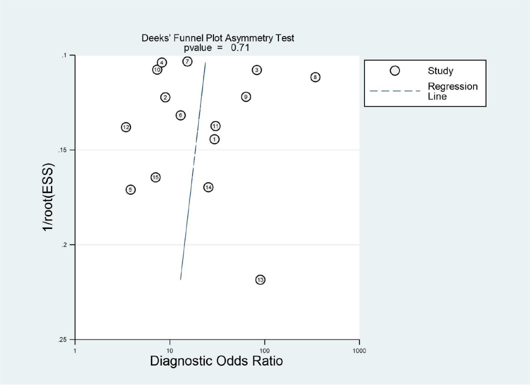


FIGURE 2
Deeks' funnel plot to test publication bias.

The result of sensitivity analysis showed that the bivariate model was moderately robust in goodness-of-fit and bivariate normality analyses (Supplemental Figure S2A, B). Influence analysis and outlier detection identified two outliers (Supplemental Figure S2C, D). After we excluded these outliers, the overall results did not change significantly, which suggested that the results of this study were statistically reliable.

Subgroup analysis was performed by comparing included studies with different variables. Six studies using HSCM had equivalent sensitivity (0.737 *vs.* 0.729) and specificity (0.816 *vs.* 0.820) compared to nine studies using the other. There were 11 studies with the LASSO method that had high sensitivity (0.775 *vs.* 0.620) and high specificity (0.842 *vs.* 0.765) than other methods. There were 11 studies using the investigated area of VOI that had equivalent sensitivity (0.731 *vs.* 0.730) and low specificity (0.814 *vs.* 0.844) than those studies with ROI. The imaging-to-surgery interval of 15 days had higher sensitivity (0.823 *vs.* 0.682) and slightly low specificity (0.790 *vs.* 0.837) than the others. The details of the subgroup analysis are shown in Table 4 and Figures 5A–D.

## Evaluation of clinical utility

The clinical utility of radiomics-based MRI was evaluated by using the likelihood ratio to simulate a Fagan nomogram. The results are shown in Figure 6. With a 20% pretest probability of MVI, the posttest probabilities of MVI and given positive and negative results of radiomics-based MRI are 51% and 6%, respectively. The Fagan nomogram revealed that the posttest probability increased by 31% in

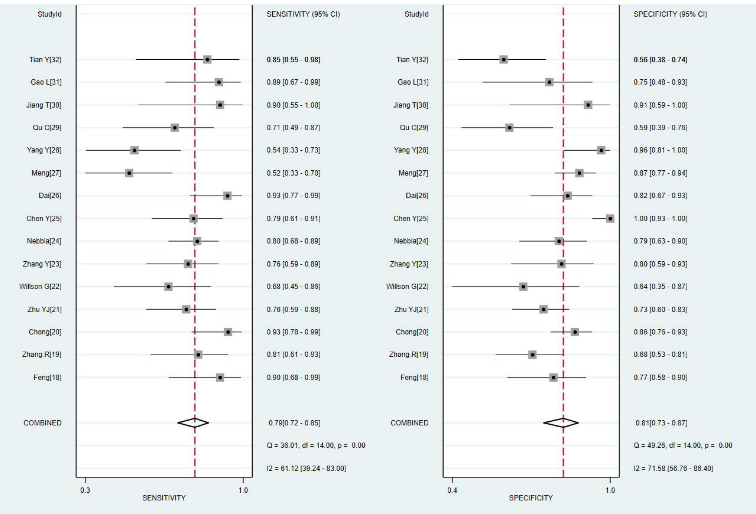


FIGURE 3
Coupled forest plots of the sensitivity and specificity of radiomics-based microvascular invasion (MRI) for predicting the MVI of hepatocellular carcinoma (HCC).

positive pretest patients but decreased by 14% in patients with a negative pretest, indicating that radiomics based-MRI was useful in clinical practice.

## Discussion

MVI is defined as the presence of cancer cells in the portal vein, hepatic vein, or a large capsular vessel of the surrounding hepatic tissue lined by the endothelium, which is visible on microscopy (12). MVI is recognized as the strongest independent predictor of the early recurrence and poor prognosis of HCC (8–10). Previous studies found that some conventional imaging features, such as the tumor margin, size, number, capsule, shape, apparent diffusion coefficient values, and enhancement pattern, may contribute to the diagnosis of MVI before surgery (33). However, imaging features have some limitations, such as the fact that the reviews of medical images rely on subjective experience. The quantitative radiomics features can reflect the microscopic pathological changes of HCC by extracting features from the overall level of the tumor on the basis of conventional imaging images and evaluating the internal heterogeneity of the tumor (34, 35). Several previous similar studies have demonstrated that radiomics has high accuracy in evaluating the MVI in HCC; however, all of these studies analyzed CT-, MRI-, and ultrasound-based radiomics (13–15). This meta-analysis demonstrates that radiomics-based MRI has high diagnostic performance for predicting the MVI of HCC and can be used as a reliable and quantitative method for the non-invasive diagnosis of MVI in clinical practice. MRI can provide better soft-tissue resolution, multiparameters, and more stable features for assessing tumor heterogeneity.

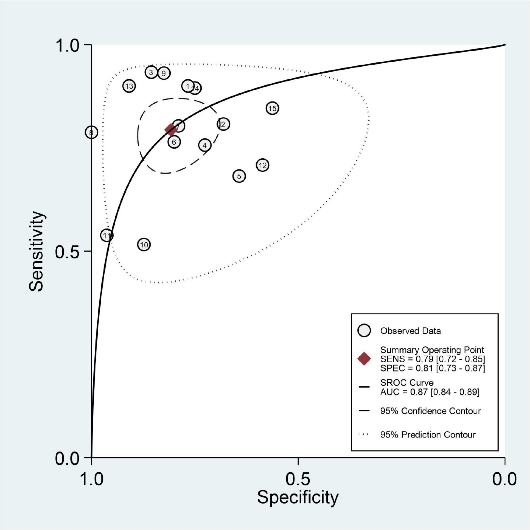


FIGURE 4
Summary receiver operating characteristic curve to evaluate the MVI of HCC.

However, obvious heterogeneity between included studies was noted. HSCM gadoxetate disodium was proven effective to assess the presence of MVI. The study demonstrated that the specificity of the hepatobiliary phase of gadolinium ethoxybenzyl diethylenetriamine pentaacetic acid Gd-EOB-DTPA-enhanced MRI combined with tumor margins and low signal intensity around the tumor to predict MVI is as high as 92.5% (36), but the contrast agent is expensive and not widely used in clinical practice. Subgroup analysis found that different contrast media (HSCM and others), the investigated area, and the method for selection were not the factors of significant heterogeneity. Furthermore, different imaging-to-surgery interval times have different. Therefore, the procedure and method should be standardized by conducting further research.

This study still has some limitations: (1) MRI scanning parameters (including the scanner machine model, field strength, and radiomics software) have not yet been unified; external datasets and different MRI scanning parameters are necessary for confirming the prediction value of the radiomics model. (2) Only English literatures of studies were included, which may result in applicable

TABLE 4 Results of subgroup analysis.

| Variate | Studies (n) | Sensitivity (95% CI) | Specificity (95% CI) | PLR | NLR | DOR |
|---|---|---|---|---|---|---|
| Contrast media | | | | | | |
| HSCM | 6 | 0.737 (0.547–0.867) | 0.816 (0.715–0.888) | 4.118 (2.513–6.748) | 0.182 (0.059–0.563) | 17.769 (5.572–56.079) |
| Other | 9 | 0.736 (0.654–0.804) | 0.824 (0.735–0.887) | 3.885 (2.626–5.747) | 0.320 (0.239–0.430) | 14.027 (8.626–25.021) |
| Method for the selection of radiomic features | | | | | | |
| LASSO | 11 | 0.775 (0.678–0.849) | 0.842 (0.770–0.895) | 4.719 (3.307–6.734) | 0.182 (0.099–0.335) | 23.092 (12.505–42.642) |
| Other methods | 4 | 0.620 (0.533–0.700) | 0.765 (0.586–0.865) | 2.625 (1.690–4.079) | 0.484 (0.344–0.681) | 6.042 (3.440–10.611) |
| Investigated area | | | | | | |
| VOI | 11 | 0.731 (0.630–0.812) | 0.814 (0.747–0.866) | 3.862 (2.681–5.213) | 0.228 (0.123–0.421) | 14.566 (8.007–26.498) |
| ROI | 4 | 0.730 (0.581–0.840) | 0.844 (0.625–0.946) | 4.684 (1.765–12.435) | 0.387 (0.246–0.609) | 16.222 (4.113–63.984) |
| Imaging-to-surgery interval | | | | | | |
| Within 15 days | 4 | 0.823 (0.636–0.925) | 0.790 (0.701–0.858) | 3.816 (2.573–5.659) | 0.145 (0.045–0.470) | 15.291(6.250–37.411) |
| Other | 10 | 0.682 (0.596–0.757) | 0.837 (0.747–0.899) | 4.035 (2.631–6.188) | 0.327 (0.199–0.538) | 13.491(6.699–27.168) |

PLR, positive likelihood ratio; NLR, negative likelihood ratio; DOR, diagnostic odds ratio.

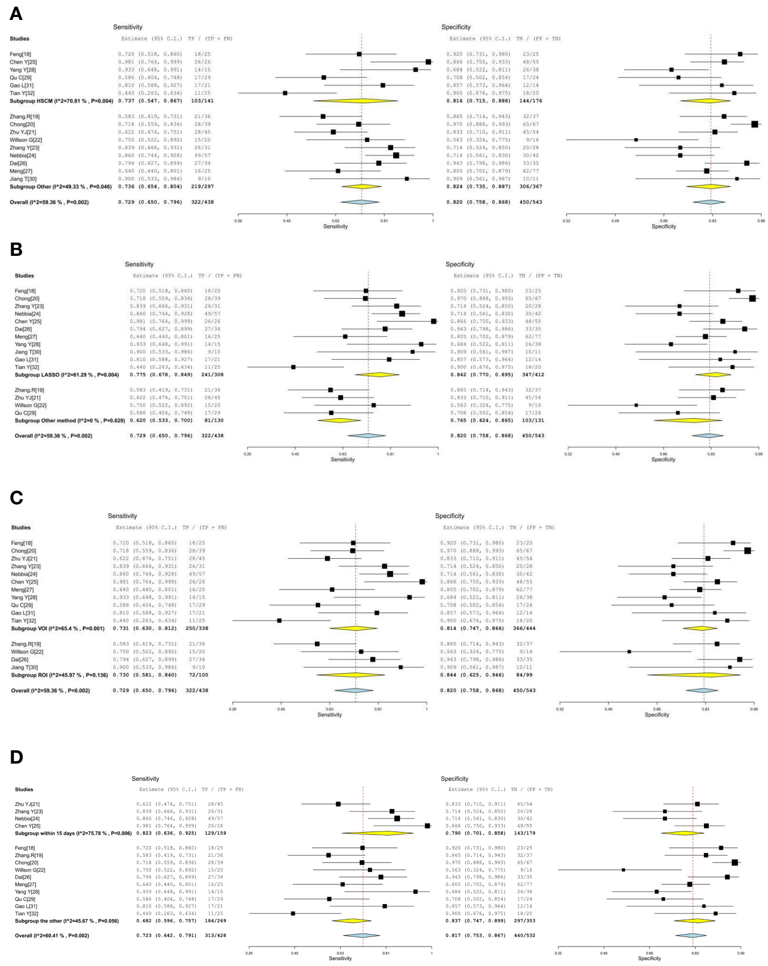


FIGURE 5
**(A, B)** The forest plots of subgroup analysis. **(C, D)** The forest plots of subgroup analysis.

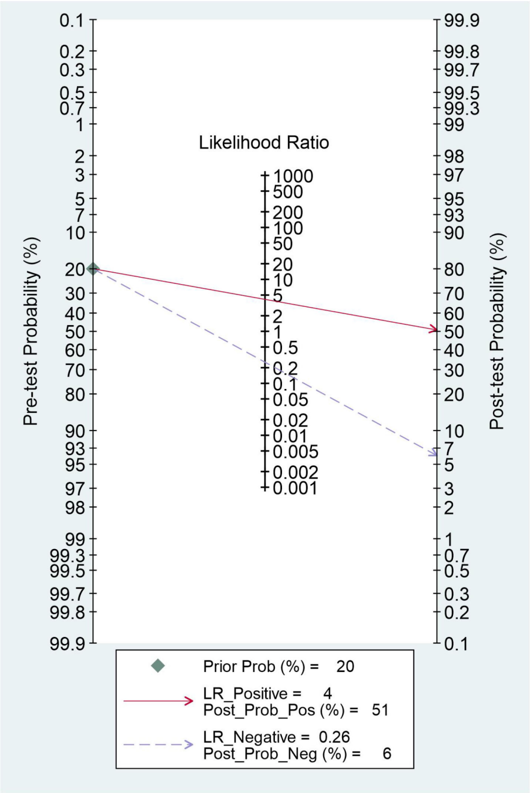


FIGURE 6
Fagan nomogram for the elucidation of post-test probabilities with a pre-test probability.

articles not being included in the review. (3) There was great heterogeneity in pooled estimates between the included studies. All of these factors may reduce the reliability of the results of this study. In the future, a large number of unified and standardized prospective studies are still needed to confirm the value of radiomics based-MRI in predicting the MVI of HCC.

## Conclusion

In conclusion, this study demonstrated that radiomics based on MRI has high accuracy for predicting MVI in HCC, and it can be used as a reliable method to predict the presence of MVI in HCC before surgery in clinical applications.

## Data availability statement

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

## Author contributions

GL and WY have contributed equally to this work and share first Authorship. GL and WY completed manuscript together. GL collected validation group data. SL and MX processed the data and the statistics. ML gave the support of everything we need. All authors did literature researches. All authors contributed to the article and approved the submitted version.

## Conflict of interest

Author ML was employed by West China-Frontier PharmaTech Co., Ltd.

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

## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.960944/full#supplementary-material

## References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2018) 68:394–424. doi: 10.3322/caac.21492

2. Forner A, Reig M, Bruix J. Hepatocellular carcinoma. *Lancet* (2018) 391 (10127):1301–14. doi: 10.1016/S0140-6736(18)30010-2

3. Bakr S, Echegaray S, Shah R, Kamaya A, Louie J, Napel S, et al. Noninvasive radiomics signature based on quantitative analysis of computed tomography images as a surrogate for microvascular invasion in hepatocellular carcinoma: a pilot study. *J Med Imaging (Bellingham)* (2017) 4:41303. doi: 10.1117/1.JMI.4.4.041303

4. Zheng Y, Cai Q, Peng L, Sun S, Wang S, Zhou J. Related factors of hepatocellular carcinoma recurrence associated with hyperglycemia after liver transplantation. *Transplant Proc* (2020) 53:177–92. doi: 10.1016/j.transproceed.2020.10.027

5. Mazzaferro V, Sposito C, Zhou J, Pinna AD, De Carlis L, Fan J, et al. Metroticket 2.0 model for analysis of competing risks of death after liver transplantation for hepatocellular carcinoma. *Gastroenterology.* (2018) 154(1):128–39. doi: 10.1053/j.gastro.2017.09.025

6. Kardashian A, Florman SS, Haydel B, Ruiz RM, Klintmalm GB, Lee DD, et al. Liver transplantation outcomes in a U.S. multicenter cohort of 789 patients with hepatocellular carcinoma presenting beyond Milan criteria. *Hepatology.* (2020) 72(6):2014–28. doi: 10.1002/hep.31210

7. Vogel A, Cervantes A, Chau I, Daniele B, Llovet JM, Meyer T, et al. Hepatocellular carcinoma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* (2019) 30(5):871–3. doi: 10.1093/annonc/mdy308

8. Rodríguez-Perálvarez M, Luong TV, Andreana L, Meyer T, Dhillon AP, Burroughs AK. A systematic review of microvascular invasion in hepatocellular carcinoma: Diagnostic and prognostic variability. *Ann Surg Oncol* (2013) 20:325–39. doi: 10.1245/s10434-012-2513-1

9. Imamura H, Matsuyama Y, Tanaka E, Ohkubo T, Hasegawa K, Miyagawa S, et al. Risk factors contributing to early and late phase intrahepatic recurrence of hepatocellular

carcinoma after hepatectomy. *J Hepatol* (2003) 38:200–7. doi: 10.1016/s0168-8278(02)00360-4

10. Du M, Chen L, Zhao J, Tian F, Zeng H, Tan Y, et al. Microvascular invasion (MVI) is a poorer prognostic predictor for small hepatocellular carcinoma. *BMC Cancer* (2014) 14:38. doi: 10.1186/1471-2407-14-38

11. Wakai T, Shirai Y, Sakata J, Kaneko K, Cruz PV, Akazawa K, et al. Anatomic resection independently improves long-term survival in patients with T1-T2 hepatocellular carcinoma. *Ann Surg Oncol* (2007) 14:1356–65. doi: 10.1245/s10434-006-9318-z

12. Lei Z, Li J, Wu D, Xia Y, Wang Q, Si A, et al. Nomogram for preoperative estimation of microvascular invasion risk in hepatitis b virus-related hepatocellular carcinoma within the Milan criteria. *JAMA Surg* (2016) 151(4):356–63. doi: 10.1001/jamasurg.2015.4257

13. Zhong X, Long H, Su L, Zheng R, Wang W, Duan Y, et al. Radiomics models for preoperative prediction of microvascular invasion in hepatocellular carcinoma: A systematic review and meta-analysis. *Abdom Radiol (NY).* (2022) 47(6):2071–88. doi: 10.1007/s00261-022-03496-3

14. Li L, Wu C, Huang Y, Chen J, Ye D, Su Z, et al. Radiomics for the preoperative evaluation of microvascular invasion in hepatocellular carcinoma: A meta-analysis. *Front Oncol* (2022) 12:831996. doi: 10.3389/fonc.2022.831996

15. Wang Q, Li C, Zhang J, Hu X, Fan Y, Ma K, et al. Radiomics models for predicting microvascular invasion in hepatocellular carcinoma: A systematic review and radiomics quality score assessment. *Cancers (Basel).* (2021) 13(22):5864. doi: 10.3390/cancers13225864

16. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* (2011) 155(8):529–36. doi: 10.7326/0003-4819-155-8-201110180-00009

17. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* (2017) 14(12):749–62. doi: 10.1038/nrclinonc.2017.141

18. Feng ST, Jia Y, Liao B, Huang B, Zhou Q, Li X, et al. Preoperative prediction of microvascular invasion in hepatocellular cancer: a radiomics model using gd-EOB-DTPA-enhanced MRI. *Eur Radiol* (2019) 29(9):4648–59. doi: 10.1007/s00330-018-5935-8

19. Zhang R, Xu L, Wen X, Zhang J, Yang P, Zhang L, et al. A nomogram based on bi-regional radiomics features from multimodal magnetic resonance imaging for preoperative prediction of microvascular invasion in hepatocellular carcinoma. *Quant Imaging Med Surg* (2019) 9(9):1503–15. doi: 10.21037/qims.2019.09.07

20. Chong HH, Yang L, Sheng RF, Yu Y, Wu DJ, Rao XS, et al. Multi-scale and multi-parametric radiomics of gadoxetate disodium-enhanced MRI predicts microvascular invasion and outcome in patients with solitary hepatocellular carcinoma ≤ 5 cm. *Eur Radiol* (2021) 31(7):4824–38. doi: 10.1007/s00330-020-07601-2

21. Zhu YJ, Feng B, Wang S, Wang LM, Wu JF, Ma XH, et al. Model-based three-dimensional texture analysis of contrast-enhanced magnetic resonance imaging as a potential tool for preoperative prediction of microvascular invasion in hepatocellular carcinoma. *Oncol Lett* (2019) 18:720–32. doi: 10.3892/ol.2019.10378

22. Wilson GC, Cannella R, Fiorentini G, Shen C, Borhani A, Furlan A, et al. Texture analysis on preoperative contrast-enhanced magnetic resonance imaging identifies microvascular invasion in hepatocellular carcinoma. *HPB (Oxford)* (2020) 22:1622–30. doi: 10.1016/j.hpb.2020.03.001

23. Zhang Y, Shu Z, Ye Q, Chen J, Zhong J, Jiang H, et al. Preoperative prediction of microvascular invasion in hepatocellular carcinoma *via* multi-parametric MRI radiomics. *Front Oncol* (2021) 11:633596. doi: 10.3389/fonc.2021.633596

24. Nebbia G, Zhang Q, Arefan D, Zhao X, Wu S. Pre-operative microvascular invasion prediction using multi-parametric liver MRI radiomics. *J Digit Imaging.* (2020) 33(6):1376–86. doi: 10.1007/s10278-020-00353-x

25. Chen Y, Xia Y, Tolat PP, Long L, Jiang Z, Huang Z, et al. Comparison of conventional gadoxetate disodium-enhanced MRI features and radiomics signatures with machine learning for diagnosing microvascular invasion. *AJR Am J Roentgenol* (2021) 216:1–11. doi: 10.2214/AJR.20.23255

26. Dai H, Lu M, Huang B, Tang M, Pang T, Liao B, et al. Considerable effects of imaging sequences, feature extraction, feature selection, and classifiers on radiomics-based prediction of microvascular invasion in hepatocellular carcinoma using magnetic resonance imaging. *Quant Imaging Med Surg* (2021) 11(5):1836–53. doi: 10.21037/qims-20-218

27. Meng XP, Wang YC, Zhou JY, Yu Q, Lu CQ, Xia C, et al. Comparison of MRI and CT for the prediction of microvascular invasion in solitary hepatocellular carcinoma based on a non-radiomics and radiomics method: Which imaging modality is better? *J Magn Reson Imaging* (2021) 54(2):526–36. doi: 10.1002/jmri.27575

28. Yang Y, Fan W, Gu T, Yu L, Chen H, Lv Y, et al. Radiomic features of multi-ROI and multi-phase MRI for the prediction of microvascular invasion in solitary hepatocellular carcinoma. *Front Oncol* (2021) 11:756216. doi: 10.3389/fonc.2021.756216

29. Qu C, Wang Q, Li C, Xie Q, Cai P, Yan X, et al. A radiomics model based on gd-EOB-DTPA-Enhanced MRI for the prediction of microvascular invasion in solitary hepatocellular carcinoma ≤ 5 cm. *Front Oncol* (2022) 12:831795. doi: 10.3389/fonc.2022.831795

30. Jiang T, He S, Yang H, Dong Y, Yu T, Luo Y, et al. Multiparametric MRI-based radiomics for the prediction of microvascular invasion in hepatocellular carcinoma. *Acta Radiol* (2022) 30, 210336510. doi: 10.1177/02841851221080830

31. Gao L, Xiong M, Chen X, et al. Multi-region radiomic analysis based on multi-sequence MRI can preoperatively predict microvascular invasion in hepatocellular carcinoma. *Front Oncol* (2022) 12:818681. doi: 10.3389/fonc.2022.818681

32. Tian Y, Hua H, Peng Q, Zhang Z, Wang X, Han J, et al. Preoperative evaluation of gd-EOB-DTPA-Enhanced MRI radiomics-based nomogram in small solitary hepatocellular carcinoma (≤3 cm) with microvascular invasion: A two-center study. *J Magn Reson Imaging* (2022) 56(5):1459–1472. doi: 10.1002/jmri.28157

33. Suh YJ, Kim MJ, Choi JY, Park MS, Kim KW. Preoperative prediction of the microvascular invasion of hepatocellular carcinoma with diffusion-weighted imaging. *Liver Transpl* (2012) 18:1171–8. doi: 10.1002/lt.23502

34. Davnall F, Yip CS, Ljungqvist G, Selmi M, Ng F, Sanghera B, et al. Assessment of tumor heterogeneity: An emerging imaging tool for clinical practice? *Insights Imaging* (2012) 3:573–89. doi: 10.1007/s13244-012-0196-6

35. Lewis S, Hectors S, Taouli B. Radiomics of hepatocellular carcinoma. *Abdominal Radiol (New York)* (2021) 46:111–23. doi: 10.1007/s00261-019-02378-5

36. Lee S, Kim SH, Lee JE, Sinn DH, Park CK. Preoperative gadoxetic acid-enhanced MRI for predicting microvascular invasion in patients with single hepatocellular carcinoma. *J Hepatol* (2017) 67:526–34. doi: 10.1016/j.jhep.2017.04.024

# Frontiers in
# Oncology

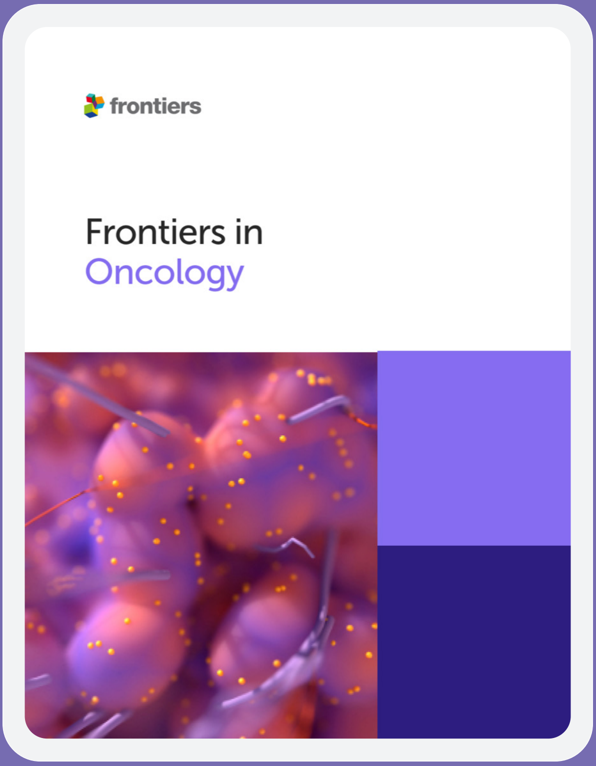
**Advances knowledge of carcinogenesis and tumor progression for better treatment and management**

The third most-cited oncology journal, which highlights research in carcinogenesis and tumor progression, bridging the gap between basic research and applications to imrpove diagnosis, therapeutics and management strategies.

---

## Discover the latest Research Topics

See more →



frontiers

Frontiers in
Oncology

**Frontiers**

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

**Contact us**

+41 (0)21 510 17 00
frontiersin.org/about/contact