frontiers | Research Topics

Assessing health-related quality of life of children

Edited by
Pei Wang, Shunping Li, Jing Wu and Nan Luo

Published in
Frontiers in Public Health
Frontiers in Pediatrics
Frontiers in Global Women's Health
Frontiers in Psychology

ISSN 1664-8714
ISBN 978-2-8325-4487-7
DOI 10.3389/978-2-8325-4487-7

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

Assessing health-related quality of life of children

Topic editors

Pei Wang — Fudan University, China

Shunping Li — Shandong University, China

Jing Wu — Tianjin University, China

Nan Luo — National University of Singapore, Singapore

Citation

Wang, P., Li, S., Wu, J., Luo, N., eds. (2024). *Assessing health-related quality of life of children*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-4487-7

Table of contents

frontiers | Frontiers in Pediatrics

TYPE Original Research
PUBLISHED 15 November 2022
DOI 10.3389/fped.2022.1045666

OPEN ACCESS

EDITED BY
Jing Wu,
Tianjin University, China

REVIEWED BY
Michał Ginszt,
Medical University of Lublin, Poland
Vadim Byvaltsev,
Irkutsk State Medical University, Russia

*CORRESPONDENCE
Milan Bajin
milanbajin@yahoo.com

SPECIALTY SECTION
This article was submitted to Children and Health, a section of the journal Frontiers in Pediatrics

RECEIVED 15 September 2022
ACCEPTED 27 October 2022
PUBLISHED 15 November 2022

CITATION
Bajin M, Kojić M, Romanov R and Ahmetović Z (2022) Neglected problem: Influence of school bag on lumbar segment in children. Front. Pediatr. 10:1045666. doi: 10.3389/fped.2022.1045666

Neglected problem: Influence of school bag on lumbar segment in children

Milan Bajin[1]*, Milan Kojić[2], Romana Romanov[1] and Zlatko Ahmetović[1]

[1]Faculty of Sport and Psychology, Educons University in Novi Sad, Novi Sad, Serbia, [2]Faculty of Sports and Physical Education, University of Belgrade, Belgrade, Serbia

Background and Objectives: School bag (SB) load causes significant changes in the height and symmetry of the intervertebral discs at each level of the spine from T12-L1 to L5-S1. This study aims to determine the change in the size of the lumbar segment angle at a particularly critical point L3-L4 of the spine in relation to the load of the average weight of SB in healthy male children (students) at standing and after 2-minute gait.

Methods: 47 boys, aged 12.2 ± 0.92 years, underwent photogrammetric measurements in the sagittal plane in statics and dynamics, walking on a laboratory treadmill. Measurements were repeated with the weight of SB with a constant load of 6,251 kg, which represents 13.78% of the average body weight of our sample. The lumbar angle (LA) connecting the point of the big toe, the lumbar point L3-L4 and the processus spinosus C7 was measured. In gait, LA was measured in the phases of the middle support and the initial contact of the heel.

Results: T-test of paired samples was used to estimate the change in LA at standing from 4.953° and walking phases from 6.295° to 7.332° in relation to the unloaded state, and the value of the effect size (ES) indicates that the impact of SB load is significant.

Conclusions: Cumulatively, microtraumas caused by SB load significantly affect the increase in intervertebral pressure at the L3-L4 point, which is susceptible to degenerative processes and which can be the cause of lumbar syndrome (LS). Preventive measures are needed in order to lighten SB in this population and introduce up to 10% of students' body weight into the safe zone.

KEYWORDS

photogrammetry, sagittal plane, lumbar angle, schoolchildren, standing posture, gait, backpack weight

Introduction

Researches, that have dealt with risk factors that affect the degenerative processes of the spine, suggest that this issue requires a complex analysis of their occurrence. In addition to genetic preconditions (1), anthropometric parameters (2) and microtrauma are the most common exogenous factors that can overload the spine during growth and development and thus have an impact on the etiology of degenerative processes (3–7). Exogenous factors, especially in the sensitive period of

growth and development, can have a negative impact on posture (8, 9). One of the most frequently mentioned factors that intrigues the scientific public and worries parents is the load conditioned by the weight of the school bag - school backpack (SB) (10–16). A number of studies have identified that students most often carry SB on both shoulders (83%) (17), and that SB weight ranged from 10%–22% of the total body weight of students (18–24). The load caused by SB leads to an increase in axial pressure on the spinal column and causes changes in the height and symmetry of the intervertebral discs of the T12-S1 level, but also changes in the angular values of the intervertebral relations (25). The size, frequency and duration of the load conditioned by SB, additionally increases the risk of creating postural changes in students during the period of rapid growth and development (26, 27).

Previous researches, that have dealt with the problem of SB load on the spinal column and the reaction in the form of changing the angle for the torso segment, analyzed different angles. They have dealt with the angles formed between the points trochanterion (femoris) and acromion and in relation to the reference horizontal (18, 19, 21, 28, 29), i.e., vertical line that intersects the great trochanter of the femur (30–32). A group of German authors examined the angle of inclination of the torso in relation to the horizontal line, where the torso is defined as a vertical line formed by the mean values of anatomical points on its upper (both acromion, jugular fossa and processus spinosus C7) and lower segment (upper sacrum, left and right iliac bone - anterior spina iliaca superior) (33). The angle of inclination of the torso that forms the anatomical points on the big toe, the intervertebral space between L3-L4 and the processus spinosus C7, was the subject of a study that considered the change in angle under different loads due to carrying different types of women's bags (female subjects, aged 19 to 37) (34). In accordance with all the above, the angle of the lumbar segment that defines the anatomical points of the big toe, the intervertebral space L3-L4 and the processus spinosus C7, can be studied by considering additional load in the form of SB and in the student population. The reasons are primarily related to the understanding of anatomical-physiological structure, and forces influencing the lumbar spine segment located on the disc between L3-L4, which also forms the peak of the lordotic curve and is the riskiest place for degenerative changes (35, 36). Intervertebral pressure in resting phases such as standing, sitting without support, as well as all types of lying, is highest in the L3-L4 region (37). Anatomically, L4 has the smallest prosessus transverus and relatively long iliotransversal ligaments that provide less support and stability (38). The multifidus muscle maintains lumbar lordosis by acting as a tendon that helps transmit certain axial compression forces on the disc to the anterior longitudinal ligament by shifting from the compression to the tensile load and further protects the discs by preventing unexpected movements such as twisting and bending (39). Regardless of the more significant role which the paraspinal muscles have in the protection of the L3-L4 segment, the greatest correlation of multifidus muscle atrophy is in that region, which additionally makes this point susceptible to degenerative changes (36).

The large weight of SB on the student's back changes the center of gravity of the body and conditions the adjustment to the posterior load (40). As the external load increases, the flexion of the torso increases significantly in response to the motor control mechanism to move the combined center of gravity of the body and the bag forward in order to maintain balance (21, 29, 40–43). The combination of external load and degree of torso flexion leads to increased pressure on the intervertebral discs (44–47).

This study aims to determine the change in the size of the angle of the lumbar segment of the spine in relation to the load of the average weight of SB in healthy male children (students) aged 10 to 13 years. In accordance with the reviewed literature, the hypothesis of this study is that the load of the SB significantly increases the flexion of the torso, and thus the compression of the vertebral discs. The flexion of the torso is related to the reduction of the lumbar angle formed by the examined anatomical points of the big toe, the lumbar points L3-L4 and the processus spinosus C7 (LA). The task of this research is to define the acute effect of SB load when standing and after 2-minute gait.

Materials and methods

Examinees

This study was a 2-factor repeated-measures design, with the participants serving as their own controls, a sample of 47 male students aged 10–13 years, mean age 12.2 ± 0.92 years, mean height 155.2 ± 8.0 cm and average body weight 45.3 ± 8.7 kg (**Table 1**). All students are of the same nationality, live in the same city and in the same municipality and attend the same school. Population and age were selected for this study to exclude factors of full demorphism in changes in the musculoskeletal system during the second sensitive period associated with the impact of accelerated growth and development (48). The phase of accelerated growth is a period

TABLE 1 Demographics of the subjects, male ($n = 47$).

Variables	Mean (AS)	Minimum	Maximum
Age, year	12.155 (0.9184)	9.9	13.6
Height, cm	155.162 (80.475)	139.00	180.70
Weight of subject, kg	45.336 (86.520)	30.10	70.20
Weight of school bag, kg	6.251 (1.0467)	2.3	8.1

in which posture irregularities often occur (49) and is accompanied by rapid bone growth and certain limitations of muscle elasticity, a characteristic primarily of postural muscles (50). The basic criteria for inclusion were that they always carry SB symmetrically over both shoulders, that they do not have any acute injuries and diseases and do not engage in any organized physical activity outside the school, except for physical education classes.

Study design and procedures

The experimental factor of this study represents the load in the form of SB, was defined on the basis of the weight which is conditioned by the content of SB, and in relation to the weekly school schedule (5 days a week, for 6 school lessons). The measured load of SB was 6,251 kg, that is, the load was an average 13.78% of the body weight of each subject. In addition to identifying the workload on a weekly basis, examinees were monitored from school entry to the classroom to identify how the workload was distributed, i.e., how respondents carried SB, and what time period was required to put the load down. It was found that the respondents, in 85% of cases, carried the load, SB, on both shoulders, and that on average it took about 120 s from entering the school to putting the load down (this time measure was used to assess variables in dynamic conditions).

The size of the lumbar segment (LA) angle at the L3-L4 level and its changes, caused by the additional SB load in statics and dynamics, were identified. The LA angle is defined by anatomical points: the tip of the big toe, the intervertebral space L3-L4 and the processus spinosus C7. Identification of LA size in relation to SB load in statics and dynamics was performed by photogrammetric measurement in the sport diagnostic laboratory. In the statics, the values of the angle, without and with the load SB, were identified. The values of angles in dynamics (without and with the load) were identified when walking in the phase of the middle support and the phase of initial contact of the heel, and at a speed of 4 km/h on the treadmill because it represents a comfortable, normal walking speed of children (51). In order to familiarize the subjects with the test in which angles are identified in dynamic conditions, they were exposed to a 5-minute walk (4 km/h) without load, prior to the measurement procedure (52).

The study was conducted following the ethical standards of the Declaration of Helsinki. The school and the parents have given the permission for their children's participation in this research. We have an Institutional Review Board approval document given by the Ethical Board of the Faculty of Sport and Psychology, Novi Sad, no. 423-2/2022, dated 12.05.2022.

Variables

The variables in this study were analyzed in the sagittal plane. The flexion of the torso was analyzed on the basis of the defined lumbar angle (LA) which forms the anatomical points of the big toe, the lumbar points L3-L4 and the processus spinosus C7 in: (1) Statics (a) without SB; and (b) with SB; (2) Dynamics in the phase of middle support with the right foot (a) without SB and (b) with SB; (3) Dynamics in the phase of middle support with the left foot (a) without SB and (b) with SB; (4) Dynamics in the phase of initial contact of the heel with the right foot (a) without SB and (b) with SB and (5) Dynamics in the phase of initial contact of the heel with the left foot (a) without SB and (b) with SB. The middle phase of the support was analyzed at the moment when the weight of the body was transferred to the standing leg, and the opposite, swinging leg, was off of the ground. The initial contact was analyzed at the moment when the heel of the front foot touched the ground (48, 53).

Tools

For the purposes of this research, two instruments were used: (1) The system (hardware and software) of the German company Contemplas was used for the estimated variables in statics. Templo software with 2D in 3 View protocol. The cameras used in the postural analysis are from the manufacturer "IDS Imaging", model UI164xLE-C, type uEye, Firmware V1.0, USB interface, max. resolution 1280 × 1024, frequency 50 Hz. (2) For the estimated variables in dynamics, on the treadmill, Templo - Software of the German company GMBH Contemplas with software module: General Motion Analysis. Two high-speed cameras were used - Basler full HD 240 Hz. The laboratory treadmill tape is Sprintex CALLIS Z6000, with a maximum speed of 22 km per hour with a possible elevation of 15%. The lamella technology of the laboratory treadmill belt, on which we performed the test, allows a natural step, provides the best ergonomic experience of walking and running with minimal impact on the joints (51, 52). This treadmill belt has no sliding speed, it has a true zero starting speed, which is very convenient for testing and measurements. In both procedures, retroreflective markers with a diameter of 19 mm were placed on the anatomical points of the subjects for imaging. Both systems are non-invasive and completely safe for testing. The drawing and calculation of angles was done with the help of Software, marking pre-marked anatomical points with the help of retroreflective markers. On the basis of previously defined and marked anatomical points, drawing angles with tools in the software was carried out by two diagnosticians independently of each other, and under the control of a third certified

photogrammetric expert who, in the case of differently plotted angles, with expert argumentation, made the final judgment on the position of the arms of the angle.

Data analysis

The obtained results were statistically analyzed in the program SPSS V.26. package. Descriptive statistics were applied, *t*-test of paired samples was used to identify differences in LA with and without SB in statics and dynamics. Calculate the value of Cohen's *d* and the effect size (ES) correlation, $r_Y\lambda$, using the *t*-test value for a between subjects *t*-test and the degrees of freedom: Cohen's $d = 2t / \sqrt{(df)}$; $r_Y\lambda = \sqrt{/[t^2 / (t^2 + df)]}$ (54).

Results

Table 2 shows the results of the *t*-test of paired samples which tested the difference in the size of the LA of the subjects, observed in the sagittal plane, and measured in statics. A statistically significant decrease in LA was found during SB load. The average decrease in LA was 4.953°, and the value of the effect size (ES = 0.634) indicates that the impact of the SB load is significant.

Tables 3, 4 show the results of the *t*-test of paired samples, which tested the difference in the size of LA of the subjects, observed in the sagittal plane, and measured in dynamics. In this test, LA was observed during the middle support with the right / left foot, i.e., in the phase of initial contact. For the phase of the middle support on the right foot (Table 3), a statistically significant decrease in LA was determined during SB load. The average decrease in LA was 6.295°, and the value of the effect size (ES = 0.706) indicates that the impact of the SB load is significant. The value of the change in LA at the phase of the middle support with the left foot (Table 3) indicates a statistically significant decrease in LA during SB load. The average decrease in LA was 6.664°, and the value of the effect size (ES = 0.792) indicates that the impact of the SB load is significant (Table 3).

Table 4 shows the values for LA with and without SB and in the initial contact phase for the right/left leg. Statistically significant values can be observed for both left and right leg. The average reduction of LA at the initial contact for the right leg was 7.332°, and the value of the effect size (ES = 0.782) indicates that the impact of the SB load is significant. Also, a statistically significant change of the value in LA with the load in the phase of initial contact for the left leg is also observed. The average decrease in LA was 7.106°, and the value of the effect size (ES = 0.790).

TABLE 2 Influence of SB load on LA size change in statics.

Variable	AS_1	AS_2	t	Df	*p*	ES
LA	167.238	162.285	8.926	46	0.000	0.634

AS_1, arithmetic mean for LA with SB; AS_2, arithmetic mean for LA without SB; t, t-test value; df, degrees of freedom; *p*, t-test significance; ES, effect size.

TABLE 3 Influence of SB load on the change in the size of LA at the phase of the middle support for the right/left leg.

Variable	AS_1	AS_2	t	df	*p*	ES
LA right support	159.272	152.977	10.514	46	0.000	0.706
LA left support	159.743	153.079	13.243	46	0.000	0.792

AS1, arithmetic mean for LA with SB; AS2, arithmetic mean for LA without SB; t, t-test value; df, degrees of freedom; *p*, t-test significance; ES, effect size.

TABLE 4 Influence of SB load on LA size change at initial contact phase for right / left foot.

Variable	AS_1	AS_2	t	df	*p*	ES
LA initial contact right	142.000	134.668	12.832	46	0.000	0.782
LA initial contact left	141.617	134.511	13.150	46	0.000	0.790

AS1, arithmetic mean for LA with SB; AS2, arithmetic mean for LA without SB; t, t-test value; df, degrees of freedom; *p*, t-test significance; ES, effect size.

Discussion

In this study, the effects of short-term SB load on the lumbar segment, i.e., the change in angle in the L3-L4 level of students aged 10–13 years, in statics and dynamics, were investigated. For all examinees, the average weight of SB (6,251 kg) was defined, which is coordinated with the weekly school schedule. The load of SB at the group level was 13,78% of the average body weight of the subjects. This value can also represent a limitation of the study, because the load is not defined in relation to body mass and for each subject individually. At the group level in accordance with the defined workload of SB, it was found that there were changes in LA in statics with load, that is, the measured angle was reduced by an average of 4,953° and the value of ES (0.634) indicates that the impact of the load is significant. This result is in accordance with previous studies of the impact of SB load, which ranged from 10%–20% of the average body weight of the subjects (20–23). In the studies that treated the problem of changing the angle of the torso under load, it was found that the forward inclination of the torso in relation to the vertical line, caused by the load SB, ranges from 3.02° to 6.8° (31, 55, 56).

In our study, the phases of the middle support and the initial contact for the left and right foot when walking on a treadmill in a time frame of 110–120 s, were also analyzed. The change of LA at SB load was also identified, and the

values for both tested phases ranged from 6.295° to 7.332°, while the ES value ranged from 0.706 to 0.792, which indicates that the effect of load for all four variables tested in dynamics was significant. Studies that have treated the problem of SB load (10%–20% of total body weight) in dynamic conditions and changes in the forward inclination of the of the torso in relation to the axial line, identified values ranging from 4.55° to as high as 19.80° (18, 19, 21, 28–30, 32, 33), which is also confirmed in this study. Differences in the range of angle changes in these studies can be explained by various experimental factors such as: anthropological differences in children (students) from different areas, gender, whether the dynamics were analyzed in the laboratory or open field, time spent in the dynamics after which the change in the inclination of the torso was analyzed, the weight and composition of the load filled SB (sand, weights, books, etc.) and the height at which their center of gravity was adjusted relative to the back of the subjects, etc. Therefore, the closest study to this one in relation to experimental factors is the study of Lee, Hong and Robinson on 15 boys, average age 10.36 years, with SB load of average weight of 10% of body weight, after 1 min gait on the treadmill, verified the change of the torso angle of 4.55° in relation to the torso angle without load SB (28). The results of studies in 10-year-old boys walking on a treadmill, where the average weight of SB was 15% and 20% of body weight, also identified a great influence on the change in the angle of the torso. In the same studies, when the load was 10% of the body weight of the subjects, i.e., without the SB load, there was no significant effect on the change in the angle of the torso (19, 21). The studies that have treated different weights, i.e., SB load on subjects aged 10–13 years, found that any load greater than 10% relative to average body weight has a large impact on torso inclination and LA change (18, 30).

The child's skeleton has a large amount of cartilage, significantly in the region of the spine as a prerequisite for growth. These cartilaginous regions consist of articular cartilage, epiphyses, and apophysis (24). Each of these structures is subject to different types of injury. Articular cartilage is susceptible to current stress, while the epiphyses and apophysis are more susceptible to recurrent microtrauma (57). Studies *in vivo* measuring intervertebral pressure, measured a pressure of about 0.3 MPa at the L3-L4 level (range 0.27–0.33 MPa) (37). When the torso was bent forward, an increase in intervertebral pressure of 2.5–3.6 times was found (45). The combination of intervertebral disc compression and torso flexion pre-induced by SB weight is a microtrauma that will cause an increase in intervertebral pressure (46, 47). Increased intervertebral pressure may be the cause of the accelerated process of disc degeneration (58, 59). In such a degenerated disc, the fluid level decreases, while the nucleus pulposus is not able to maintain uniform pressure on the adjacent fibrous ring and end plates (46, 47, 60). The fibrous ring consists of several nerve endings in its posterior part (61). The concentration of stress due to compression and flexion of the torso is maximal in the posterior part of the fibrous ring and irritates nerve endings that mechanically cause acute but can also be the cause of chronic lower back pain - lumbar syndrome (LS) (60, 62). LS is a major problem in the population of 40% of young people aged 9 to 18 worldwide (20, 63). The load of SB, which is greater than 10% of the total body weight (as is the case in this study), has a negative effect and contributes to the development of LS in students (20, 64). The fact that the painful condition in LS is associated with SB load was identified in 82% of students aged 11 to 14 years (65). These data are worrying because LS in youth plays an important role in the development of chronic LS in adulthood (20, 66). However, the negative impact of SB can be prevented, primarily by limiting the weight to a maximum of 10% in relation to the total body weight of the child, which can have a positive impact on the spine and thus on quality of life in later years (67). Also, it was found that the design of SB (55) and the acquisition of knowledge about its proper wearing (68, 69) can significantly mitigate its negative effects during statics and dynamics.

Conclusions

The innovativeness of this study is reflected in the fact that for the first time in children (students), the change of LA at the level of L3-L4 with and without SB load was analyzed by photometric method in statics and dynamics. The study found that the average weight of SB was 13.78% of the average body weight of the tested sample. The results showed that the influence of SB load on LA change is great both in statics and dynamics, which is in line with previous research where SB load is greater than 10% of body weight, and, as such, causes a negative effect in terms of degenerative changes. Slightly greater changes in LA in dynamics relative to statics point to differences in the motor control strategies children use to maintain balance. Cumulatively, the microtrauma caused by SB load significantly increases the intervertebral pressure in a very critical segment, L3-L4 (which is susceptible to degenerative processes) and which can be the cause of LS. Preventive measures are needed in order to lighten SB in this population and introduce up to 10% of the child's body weight to the safe zone, as well as designing the most anatomically functional SB and educating about its proper carrying.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by Ethical Board of the Faculty of Sport and Psychology, Novi Sad, no. 423-2/2022, dated 12.05.2022. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

MB, and RR.; methodology, MB, MK.; software, MB, MK validation, RR, ZA.; formal analysis, MB, RR; investigation, MB, MK; resources, MB.; data curation, MB MK.; writing—original draft preparation, MB; writing—review and editing, RR, ZA; visualization, MK.; supervision, RR, ZA.; project administration, RR.; funding acquisition, MB, ZA. All authors contributed to the article and approved the submitted version.

Acknowledgments

The authors thank the participants for their contribution to the study.

Conflict of interest

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

Publisher's note

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

References

1. Ala-Kokko L. Genetic risk factors for lumbar disc disease. *Ann Med*. (2002) 34:42–7. doi: 10.1080/078538902317338634

2. Videman T, Levalahti E, Battié MC. The effects of anthropometrics, lifting strength,and physical activities in disc degeneration. *Spine*. (2007) 32:1406–13. doi: 10.1097/BRS.0b013e31806011fa

3. Miller JA, Schmatz C, Schultz AB. Lumbar disc degeneration: correlation with age, sex, and spine level in 600 autopsy specimens. *Spine (Phila Pa 1976)*. (1988) 13:173–8. doi: 10.1097/00007632-198802000-00008

4. Videman T, Nurminen M, Troup JD. Volvo award in clinical sciences. Lumbar spinal pathology in cadaveric material in relation to history of back pain, occupation, and physical loading. *Spine (Phila Pa 1976)*. (1990) 15:728–40. doi: 10.1097/00007632-199008010-00002

5. Videman T, Battié MC. The influence of occupation on lumbar degeneration. *Spine*. (1999) 24:1164–8. doi: 10.1097/00007632-199906010-00020

6. Hassett G, Hart DJ, Manek NJ, Doyle DV, Spector TD. Risk factors for progression of lumbar spine disc degeneration. *Arthritis Rheum*. (2003) 48:3112–7. doi: 10.1002/art.11321

7. Macedo LG, Battié MC. The association between occupational loading and spine degeneration on imaging – a systematic review and meta-analysis. *BMC Musculoskelet Disord*. (2019) 20:1–15. doi: 10.1186/s12891-019-2835-2

8. O' Sullivan PB, Smith AJ, Beales DJ, Straker LM. Association of biopsychosocial factors with degree of slump in sitting posture and self-report of back pain in adolescents: a cross selectional-stydy. *Phys Ther*. (2011) 91:470–83. doi: 10.2522/ptj.20100160

9. Latalski M, Bylina J, Fatuga M, Repko M, Filipović M, Jarosz JM, et al. Risk factors of postural defects in children at school age. *Ann Agric Envrion Med*. (2013) 20:583–7.

10. Vitiello A, Pollard H. Backpack design: the use of ratings of perceived exertion (Borg scale) – a review. *Chiropractic J Australia*. (2002) 32:91–8.

11. Cottalorda J, Rahmani A, Diop M, Gautheron V, Ebermeyer E, Belli A. Influence of school bag carrying on gait kinetics. *J Pediatric Orthopaedics*. (2003) 12:357–64. doi: 10.1097/01.bpb.0000078270.58527

12. Cardon G, Balague F. Backpacks and spinal disorders in school children. *Eura Medicophys*. (2004) 40:15–20. PMID: 16030489

13. Negrini S, Politano E, Carabalona R, Tartarotti L, Marchetti M. The backpack load in schoolchildren: clinical and social importance, and efficacy of a community-based educational intervention. *Eur Medicophys*. (2004) 40:185–90.

14. Al-Hazzaa HM. How much load do Saudi school boys carry on their shoulder. *Saudi Med J*. (2006) 27:1567–71.

15. Chow D, Leung K, Holmes A. The effects of load carriage and bracing on the balance of schoolgirls with adolescent idiopathic scoliosis. *Eur Spine J*. (2007) 16:1351–8. doi: 10.1007/s00586-007-0333-y

16. Ismaila SO, Oriolowo KT. *Determination of safe backpack mass for students in tertiary institutions. Proceedings of the 2015 international conference on industrial engineering and operations managemen*, Dubai: Piscataway (2015). p. 2526–9.

17. Talbott NR, Bhattacharya A, Davis KG, Shukla R, Levin L. School backpacks: it's More than just a weight problem. *Work*. (2009) 34:481–94. doi: 10.3233/WOR-2009-0949

18. Pascoe DD, Pascoe DE, Wang YT, Shim DM, Kim CK. Influence of carrying book bags on gait cicle and posture youths. *Ergonomics*. (1997) 40:631–41. doi: 10.1080/001401397187928

19. Hong Y, Brueggemann GB. Changes in gait patterns in 10-year-old boys with increasing loadswhen walking on a treadmill. *Gait Posture*. (2000) 11:254–9. doi: 10.1016/S0966-6362(00)00055-2

20. Negrini S, Carabalona R. Backpack on ! Schoolchildren's Perceptions of load, association with back pain and factors determining the load. *Spine*. (2002) 27:187–95. doi: 10.1097/00007632-200201150-00014

21. Hong Y, Cheung C. Gait and posture responses to backpack load during level walkingin children. *Gait Posture*. (2003) 17:28–33. doi: 10.1016/S0966-6362(02)00050-4

22. Cavallo CM, Hlavaty TM, Tamase MGM. A pilot study for the development of a primaryprevention program: what is the average weight of a fourth graders backpack? *Work*. (2003) 20:137–58. Corpus ID: 35490437

23. Forjuoh SN, Schuchmann JA, Lane BL. Correlates of heavy backpack use by elementary school children. *Public Health*. (2004) 118:532–5. doi: 10.1016/j.puhe.2003.10.010

24. Vaghela NP, Parekh SK, Padsala D, Patel D. Effect of backpack loading on cervical and sagittal shoulder posture in standing and after dynamic activity in school going children. *J Family Med Prim Care*. (2019) 8:1076–81. doi: 10.4103/jfmpc.jfmpc_367_18

25. Neuschwander TB, Cutrone J, Macias BR, Cutrone S, Murthy G, Chambers H, et al. The effect of backpacks on the lumbar spine in children: a standing

magnetic resonance imaging study. *Spine*. (2010) 35:83–8. doi: 10.1097/BRS.0b013e3181b21a5d

26. Chan SC, Ferguson SJ, Gantenbein-Ritter B. The effects of dynamic loading on the intervertebral disc. *Eur Spine J*. (2011) 20:1796–812. doi: 10.1007/s00586-011-1827-1

27. Cosma G, Ilinca I, Rosu L, Burileanu A. Physical exercise and it s role in a correct postural alignment. *Discobolul Phys Educ Sport Kinetother J*. (2015) 1:58–64.

28. Li JX, Hong Y, Robinson PD. The effect of load carriage on movement kinematics and respiratory parameters in children during walking. *Eur J Appl Physiol*. (2003) 90:35–43. doi: 10.1007/s00421-003-0848-9

29. Li JX, Hong Y. Age difference in trunk kinematics during walking with different backpack weights in 6-to 12-year-old children. *Res Sports Med*. (2004) 12:135–42. doi: 10.1080/15438620490460486

30. Goodgold S, Mohr K, Samant A, Parke T, Burns T, Gardner L. Effects of backpack load and task demand on trunk forward lean: pilot findings on two boys. *Work*. (2002) 18:213–20.

31. Kistner F, Fiebert I, Roach K, Moore J. Postural compensations and subjective complaints due to backpack loads and wear time in schoolchildren. *Pediatr Phys Ther*. (2013) 25:15–24. doi: 10.1097/PEP.0b013e31827ab2f7

32. Zhou N, Yan J, Chen W, Hu C, Zhang X. Effects of daily used bags on biomechanics characteristic of elementary students. *Leather Footwear J*. (2016) 16:265–74. doi: 10.24264/lfj.16.4.2

33. Hell AK, Braunschweig L, Grages B, Brunner R, Romkes J. Einfluss des schulrucksackgewichtes bei grundschulkindern: gang, muskelaktivität, haltung und stabilität [the influence of backpack weight in school children: gait, muscle activity, posture and stability. *Orthopade*. (2021) 50:446–54. doi: 10.1007/s00132-020-04047-8

34. Lyu S, LaBat KL. Effects of natural posture imbalance on posture deviation caused by load carriage. *Int J Ind Ergon*. (2016) 56:115–23. doi: 10.1016/j.ergon.2016.09.006

35. Pritchett JW, Bortel DT. Degenerative symptomatic lumbar scoliosis. *Spine*. (1993) 18:700–3. doi: 10.1097/00007632-199305000-00004

36. Sun D, Liu P, Cheng J, Ma Z, Liu J, Qin T. Correlation between intervertebral disc degeneration, paraspinal muscle atrophy, and lumbar facet joints degeneration in patients with lumbar disc herniation. *Musculoskelet Dis*. (2017) 18:1–7. doi: 10.1186/s12891-016-1361-8

37. Dreischarf M, Shirazi-Adl A, Arjmand N, Rohlmann A, Schmidt H. Estimation of loads on human lumbar spine: a review of in vivo and computational model studies. *J Biomech*. (2016) 49:833–45. doi: 10.1016/j.jbiomech.2015.12.038

38. Kirkaldy-Willis WH, Farfan HF. Instability of the lumbar spine. *Clin Orthop*. (1982) 165:110–23. doi: 10.1097/00003086-198205000-00015

39. Kader DF, Wardlaw D, Smith FW. Correlation between the MRI changes in the lumbar multifidus muscles and leg pain. *Clin Radiol*. (2000) 55:145–9. doi: 10.1053/crad.1999.0340

40. Grimmer K, Dansie B, Milanese S, Pirunsan U, Trott P. Adolescent standing postural response to backpack loads: a randomized controlled experimental study. *BMC Musculoskelet Dis*. (2002) 3:1–10. doi: 10.1186/1471-2474-3-10

41. Motmans RREE, Tomlow S, Vissers D. Trunk muscle activity in different modes of carrying schoolbags. *Ergonomics*. (2006) 49:127–38. doi: 10.1080/00140130500435066

42. Fiolkowski P, Horodyski M, Bishop M, Williams M, Stylianou L. Changes in gait kinematics and posture with the use of a front pack. *Ergonomics*. (2006) 49:885–94. doi: 10.1080/00140130600667444

43. Chen YL, Nguyen HT, Chen Y. Influence of school bag loads and carrying methods on body strain among young Male students. *Int J Ind Ergon*. (2021) 82:1–9. doi: 10.1016/j.ergon.2021.103095

44. Rohlmann A, Neller S, Claes L, Bergmann G, Wilke HJ. Influence of a follower load on intradiscal pressure and intersegmental rotation of the lumbar spine. *Spine*. (2001) 26:557–61. doi: 10.1097/00007632-200112150-00014

45. Takahashi I, Kikuchi S, Sato K, Sato N. Mechanical load of the lumbar spine during forward bending motion of the trunk–A biomechanical study. *Spine*. (2006) 31:18–23. doi: 10.1097/01.brs.0000192636.69129.fb

46. Zahaf S, Mansouri B. The effect of the posterior loading on the spine of a school child. *Adv Cancer Prev*. (2016) 1:1–19. doi: 10.4172/2472-0429.1000112

47. Li JQ, Kwong WH, Chan YL, Kawabata M. Comparison of in vivo intradiscal pressure between sitting and standing in human lumbar spine: a systematic review and meta-analysis. *Life (Basel, Switzerland)*. (2022) 12:1–18. doi: 10.3390/life12030457

48. Abernethy B, Kippers V, Hanrahan SJ. In: Schrag M, editor. Human kinetics. *Biophysical foundations of human movement*. Champaign, IL, United States: Human Kinetics (2013). p. 41–54.

49. Schlösser TP, Vincken KL, Rogers K, Castelein RM, Shah SA. Natural sagittal spino-pelvic alignment in boys and girls before, at and after the adolescent growth spurt. *Eur Spine J*. (2015) 24:1158–67. doi: 10.1007/s00586-014-3536-z

50. Upasani VV, Hedequist DJ, Hresko MT, Karlin LI, Emans JB, Glotzbecker MP. Spinal deformity progression after posterior segmental instrumentation and fusion for idiopathic scoliosis. *J Child Orthop*. (2015) 9:29–37. doi: 10.1007/s11832-015-0632-5

51. Romkes J, Bracht-Schweizer K. The effects of walking speed on upper body kinematics during gait in healthy subjects. *Gait Posture*. (2017) 54:304–10. doi: 10.1016/j.gaitpost.2017.03.025

52. Leardini A, Sawacha Z, Paolini G, Ingrosso S, Nativo R, Benedetti MG. A new anatomically based protocol for gait analysis in children. *Gait Posture*. (2007) 26:560–71. doi: 10.1016/j.gaitpost.2006.12.018

53. Pirker W, Katzenschlager R. Gait disorders in adults and the elderly. *Wien Klin Wochenschr*. (2016) 129:81–95. doi: 10.1007/s00508-016-1096-4

54. Becker LA. Effect size calculators. 2000, Retrieved from https://lbecker.uccs.edu/

55. Mackie HW, Legg SJ, Beadle J, Hedderle D. Comparison of four different backpacks intended for school use. *Appl Ergon*. (2003) 34:257–64. doi: 10.1016/S0003-6870(03)00034-6

56. Brackley HM, Stevenson JM, Selinger JC. Effect of backpack load placement on posture and spinal curvature in prepubescent children. *Work*. (2009) 32:351–60. doi: 10.3233/WOR-2009-0833

57. Whittfield JK, Legg SJ, Hedderley DL. The weight and use of schoolbags in New Zealand secondary schools. *Ergonomics*. (2001) 44:819–24. doi: 10.1080/00140130117881

58. Luoma K, Riihimäki H, Luukkonen R, Raininko R, Viikari-Juntura E, Lamminen A. Low back pain in relation to lumbar disc degeneration. *Spine*. (2000) 25:487–92. doi: 10.1097/00007632-200002150-00016

59. Zhang S, Hu B, Liu W, Wang P, Lv X, Chen S, et al. The role of structure and function changes of sensory nervous system in intervertebral disc-related low back pain. *Osteoarthritis Cartilage*. (2021) 29:17–27. doi: 10.1016/j.joca.2020.09.002

60. Adams MA, Roughley PJ. What is intervertebral disc degeneration, and what causes it? *Spine*. (2006) 31:2151–61. doi: 10.1097/01.brs.0000231761.73859.2c

61. Raj Arjun SI, Chandrashekara CV, Parth G. Fracture strength estimation of L3-L4 intervertebral disc using FEA. *Vibroeng Procedia*. (2019) 27:67–72. doi: 10.21595/vp.2019.20976

62. Mackie HW, Legg SJ. Postural and subjective responses to realistic schoolbag carriage. *Ergonomics*. (2008) 51:217–31. doi: 10.1080/00140130701565588

63. Calvo-Muñoz I, Gómez-Conesa A, Sánchez-Meca J. Prevalence of low back pain in children and adolescents: a meta-analysis. *BMC Pediatr*. (2013) 13:1–12. doi: 10.1186/1471-2431-13-14

64. Nicolet T, Mannion AF, Heini P, Cedraschi C, Balague F. No kidding: low back pain and type of container influence adolescents' perception of load heaviness. *Eur Spine J*. (2014) 23:794–9. doi: 10.1007/s00586-014-3213-2

65. Shymon SJ, Yaszay B, Dwek JR, Proudfoot JA, Donohue M, Hargens AR. Altered disc compression in children with idiopathic low back pain: an upright magnetic resonance imaging backpack study. *Spine*. (2014) 39:243–8. doi: 10.1097/BRS.0000000000000114

66. Jones GT, Macfarlane GJ. Epidemiology of low back pain in children and adolescents. *Arch Dis Child*. (2005) 90:312–6. doi: 10.1136/adc.2004.056812

67. Steele E, Bialocerkowski A, Grimmer K. The postural effects of load carriage on young people – a systematic review. *BMS Musculosekelet Disord*. (2003) 4:1–7. doi: 10.1186/1471-2474-4-1

68. Feingold AJ, Jacobs K. The effect of education on backpack wearing and posture in a middleschool population. *Work*. (2002) 18:287–94. PMID: 12441569

69. Brzek A, Plinta R. Exemplification of movement patterns and their influence on body posture in younger school-age children on the basis of an authorial program "I take care of my spine". *Med (Baltimore)*. (2016) 95:1–11. doi: 10.1097/MD.0000000000002855

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 17 November 2022
DOI 10.3389/fpubh.2022.1047617

OPEN ACCESS

EDITED BY
Jing Wu,
Tianjin University, China

REVIEWED BY
Ananda Vayaravel Cassinadane,
Sri Venkateswaraa College of Paramedical Sciences, India
Shalini Bassi,
Public Health Foundation of India, India

*CORRESPONDENCE
S. M. Mahbubur Rashid
mahbubur.rashid@du.ac.bd

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

RECEIVED 18 September 2022
ACCEPTED 31 October 2022
PUBLISHED 17 November 2022

CITATION
Akter F, Rashid SMM, Alam N, Lipi N, Qayum MO, Nurunnahar M and Mannan A (2022) Knowledge, attitude and practice of diabetes among secondary school-going children in Bangladesh.
Front. Public Health 10:1047617.
doi: 10.3389/fpubh.2022.1047617

Knowledge, attitude and practice of diabetes among secondary school-going children in Bangladesh

Farhana Akter[1,2], S. M. Mahbubur Rashid[3]*, Nazmul Alam[4], Nasrin Lipi[5], Md. Omar Qayum[6], Mehejabin Nurunnahar[6] and Adnan Mannan[2,7]

[1]Department of Endocrinology, Chittagong Medical College, Chattogram, Bangladesh, [2]Disease Biology & Molecular Epidemiology Research Group, Chattogram, Bangladesh, [3]Department of Genetic Engineering & Biotechnology, University of Dhaka, Dhaka, Bangladesh, [4]Department of Public Health, Asian University for Women, Chittagong, Bangladesh, [5]Institute of Statistical Research and Training, University of Dhaka, Dhaka, Bangladesh, [6]Department of Epidemiology, Institute of Epidemiology, Disease Control & Research (IEDCR), Dhaka, Bangladesh, [7]Department of Genetic Engineering & Biotechnology, University of Chittagong, Chattogram, Bangladesh

Diabetes Mellitus (DM) is a global public health concern. DM has been increasing alarmingly among the young people and childhood-onset has now become an emerging issue worldwide. Unlike other chronic diseases, DM requires constant and active attention of the patients, sometimes of their family members for successful management of this disease. Knowledge, attitude, and practices make significant differences at the population level, which largely depend on socioeconomic status, area of residence, level of education, and other socio-demographic attributes. A descriptive cross-sectional study was carried out among secondary school students in grades 6 through 10 attending schools in 18 districts of Bangladesh. A total of 2009 students were enrolled for the study from the selected schools and madrasas from Bengali, English, and Arabic medium. The majority of responders (79.34%) reported that they had heard of diabetes, however, only 45% knew that diabetes can cause blood glucose levels to rise. Among different therapeutic options, only 15% of students had heard the name of metformin, while 56.2% were familiar with insulin. English medium students were significantly more likely to have good knowledge than the Bangla medium students (19.31%, OR: 1. 44; 95% CI: 1.07, 1.95). Respondents of semi-urban (19.66%, OR: 1.7; 95% CI: 1.21, 2.36) and urban (18.47%, OR: 1.48; 95% CI: 1.17, 1.86) areas were more likely to have good knowledge than rural areas. Of the respondents, 20.61% did not know of any diabetic treatment options. Only a small percentage of students (11%) attended any diabetes education classes and 10.7% of students followed any recommended diabetes prevention methods. Knowledge, attitude, and practice related to diabetes were found to have been influenced by socioeconomic factors, societal practices, and behavioral patterns. This is the first nationwide research study in Bangladesh among secondary school students to study knowledge, attitudes, and practices related to diabetes. With a focus on Bangladesh's youthful population, this study sought to provide an informational framework that can be supportive for effective intervention to increase knowledge of diabetes and its implications.

KEYWORDS

diabetes, school-going children, Bangladesh, awareness, healthy lifestyle

Introduction

Diabetes Mellitus (DM) is one of the major non-communicable diseases that has a significant association with increasing morbidity and mortality globally (1). According to the World Health Organization, DM is a silent global epidemic affecting mainly older adults; however, DM among the young population has been increasing consistently (1, 2). The prevalence of DM has been rising more rapidly in low- and middle-income countries than in high-income countries (1, 3, 4). Approximately 537 million adults (20–79 years) live with diabetes; and by 2045, this will rise to 783 million (1). The International Diabetes Federation estimated about 13.1 million people have diabetes in Bangladesh with a prevalence of 14.2% in adults and almost 5.7 million people have undetectable diabetes (5). The number is suspected to be doubled by 2045. Moreover, the prevalence of type-2 diabetes is increasing at an alarming rate among adults in both the rural and urban settings of the country (6, 7).

Besides adult Diabetes, childhood-onset DM has now become a burning issue worldwide (8–10). A recently published study conducted on US youths aged 10 to 19 showed that 22.3 and 13.8 per 100,000 youths were affected with Type 1 and Type 2 diabetes, respectively, in 2015 (8). The rates of incident diabetes cases have been increasing by about 1.9% for type 1 diabetes (T1DM) and 4.8% for type 2 diabetes (T2DM). Asian youths have shown the highest rate of increased prevalence of T2DM. The prevalence of impaired fasting glucose and type 2 diabetes among Bangladeshi youths aged 10–18 is 3.4 and 1.8%, respectively, which can be considered as very high (11).

Several factors have been attributed to the rising prevalence of diabetes including urbanization, adopting a western lifestyle, inadequate physical activity, and excess calorie intake (12). Both micro and macrovascular complications like cardiovascular diseases, retinopathy, nephropathy, and neuropathy have been known as long-term consequences of DM (13). This will significantly deteriorate patients' quality of life and creates a significant financial burden at the national and individual level. Prevention and self-care are the cornerstones of diabetes management as evidenced by various studies and populations (14–16). Unlike other chronic diseases, DM requires the constant, active participation of the patients, sometimes of their family members for successful management of this disease (15). Several studies have reported that knowledge, attitude, and practices at the population level that make significant differences largely depend on socioeconomic status, residing area, level of education, and other socio-demographic attributes (17, 18). Knowledge of diabetes mellitus at an early stage of life can assist in early diagnosis with subsequent minimizing of diabetes-related life-threatening complications (19–23). So, to prevent the upward trend of increasing young diabetic patients, it is necessary to know how much awareness has been built among people, especially children about different factors like lifestyle modification, dietary readjustment, treatment, medication adherence, physical activity, etc. Lack of appropriate perception about the disease results in poor compliance. However, this type of study has not been done yet on school-going children in Bangladesh. To build a strategy to create awareness, it is important to know to what extent the young people of this country are concerned about these health issues.

This is the first nationwide research study in Bangladesh among secondary school students to evaluate knowledge, attitudes, and practices related to diabetes. This study aimed to help guide effective intervention to raise awareness about diabetes and its consequences, especially among the young generation in Bangladesh, and contribute to much-needed healthy lifestyle adoption.

Materials and methods

Study design, participants and settings

Between January 2021 and January 2022, a descriptive cross-sectional study was carried out among high school students in grades 6 through 10 attending schools from Bengali, English and Arabic medium. For a better understanding of the real KAP scenarios of DM and geographic representation of Bangladesh, a total of 18 districts, drawn from all 8 administrative divisions of Bangladesh: Dhaka, Chattogram, Sylhet, Khulna, Rajshahi, Mymensingh, Barisal, and Rangpur were included in the study. A total of 18 districts were selected purposively, eight administrative districts of all divisions, top 7 densely populated districts to represent densely populated districts from census data and 3 indigenous districts were chosen to represent the indigenous population and hill tracts area. Using census data from the Bangladesh Bureau of Statistics (BBS) and nine of the education boards of Bangladesh, a total of 78 schools and madrasa were selected randomly from a list of all eligible and currently functioning schools. The schools and madrasas having the highest number of students from each district were selected. Selected schools follow both Bangla and English medium curricula while madrasas follow Arabic curriculum. English medium schools are mostly private, and students studying there mostly belong to the economically well-off families. A total of 2009 students were enrolled for the study from the selected schools and madrasas, ranging in grade level from sixth to tenth.

Data collection procedure and tools

A pre-tested, structured close-ended questionnaire was used for data collection. The questionnaire consisted of five sections: socio-demographic characteristics; lifestyle-related practices; knowledge of diabetes, attitude toward diabetes,

and practices on diabetes issues. The knowledge, attitude and practice level of the students were assessed based on the lifestyle modification/behavioral risk factors/choices of diabetes management included in the questionnaire. A semi-structured questionnaire was used to collect data by face-to-face interviews with trained interviewers having study background on life science or public health. This included family status, obesity, anxiety, salt intake, lack of exercise and physical activity, high fat diet, use of soft drinks, fast food consumption, symptoms of diabetes, and treatment of diabetes. The questionnaire was developed with input from study investigators with specialties in the field of epidemiology, endocrinology, public health and pediatrics. A pilot testing of the questionnaire was conducted among the children of Noakhali and Manikgonj district, the districts not included the main study. Feedback from pilot test has been incorporated for finalization of the questionnaire.

After informing and obtaining written permission from the school authorities and legal guardians, the purpose of the study was explained to the students. Parental assent was taken after explaining the study objectives and purpose of the study. The study tool was converted into the Bengali language for easy understanding of the questions. Anonymity and confidentiality were highly maintained. The collected data were entered into an excel spreadsheet.

Statistical analysis

Survey data was imported through Microsoft Excel and analyzed via R version 3.5.2. Knowledge, attitude, and practice scores of the respondents were calculated from 15 knowledge questions, 6 attitude questions, and 9 practice questions. Scores were estimated using a score of 1 for each right answer and 0 for each wrong answer. Maximum score of knowledge, attitude and practice was 13, 6, and 8, respectively. Correct answer to 80% of questions is considered as the cutoff point to differentiate between good and poor knowledge, positive and negative attitude, good and bad practice (24).

Descriptive statistics were employed to describe socio demographic characteristics of the respondents and to demonstrate students' knowledge, attitude and practice toward diabetes. Pearson chi-square test was used to test the association of socio-demographic characteristics with the knowledge and practice status of the students. Multiple logistic regression analyses were done to estimate univariate and multivariable Odds ratio and 95% confidence interval to determine the predictors of knowledge level, where dependent variables were dichotomized as Good Knowledge and Poor Knowledge and predictor variables included socio-demographic variables including age, gender, medium of education, family income and father's and mother's education.

TABLE 1 Socio-demographic characteristics of the study participants ($N = 2,009$).

Characteristics		Number of participants	Percentage (%)
Gender	Male	970	48.28
	Female	1,039	51.71
Medium of education	Bangla	1,386	68.98
	English	264	13.14
	Madrasha	359	17.87
Place of residence	Rural	931	46.31
	Semi-Urban	239	11.89
	Urban	839	41.76
Ethnic origin	Indigenous	525	26.13
	Non-Indigenous	1,484	73.86
Father's education	None	162	8.06
	Primary	504	25.08
	Secondary	738	36.73
	Graduation	605	30.11
Mother's education	None	170	8.46
	Primary	576	28.67
	Secondary	794	39.52
	Graduation	469	23.34
Monthly family income (in taka)	≤10,000	485	24.14
	10,001–20,000	413	20.55
	20,001–30,000	460	22.89
	30,001–50,000	359	17.87
	≥50,001	292	14.53

Results

Socio-demographics

In total, 2009 students took part in the survey; 970 of them (48.28%) were boys and 1,039 (51.71%) were girls. Bangla was the language of instruction for the majority of students (68.98%), followed by Madrasha (17.87%) and English (13.14). The survey covered 26.13% indigenous and 73% non-indigenous students. In comparison to rural (46.31%) and urban (41.76%) areas, a smaller share (11.89%) of the sample population came from semi-urban areas. Most study participants had a family income of ≤100 USD per month. In terms of their parents' education, the majority had only completed primary school (Table 1).

Knowledge

The majority of responders (79.34%) reported that they had heard of diabetes, however, only 45% knew that diabetes can cause blood glucose levels to rise. Only 15% of students had

10.3389/fpubh.2022.1047617

TABLE 2 Frequency distribution of the responses to knowledge questions.

Questions	Response	Frequency	Percentage
Heard of diabetes			
	Yes	1,594	79.34
	No	415	20.65
What increases in blood when diabetes occurs?			
	Glucose	911	45.34
	Other	422	21.00
	Don't Know	676	33.64
Know about types of diabetes			
	Yes	512	25.48
	No	1,497	74.51
Diabetes increase the risk of other disease			
	Yes	1,180	58.73
	No	829	41.27
Diabetes, is a hereditary disease?			
	Yes	1,164	57.93
	No	845	42.06
Diabetes is a life -long disease			
	Yes	782	38.92
	No	1,227	61.07
Heard of Insulin?			
	Yes	1,130	56.24
	No	879	43.75
Diabetes reduces life expectancy			
	Yes	934	46.49
	No	1,075	53.51
Have you heard the name of "Metformin"?			
	Yes	314	15.62
	No	1,695	84.37
Symptoms of diabetes			
	Frequent Urination	1,011	50.32
	Always feel tired	382	19.01
	Increasing weight	247	12.29
	Tired after meal	120	5.97
	Don't know	495	24.63
Risk factors of diabetes			
	Got it from family	552	27.47
	Eating Excess carbohydrates or sugary food	1,116	55.55
	Overweight	532	26.48
	Don't know	510	25.38

heard the name of metformin, while 56.2% were familiar with insulin. A whopping 58% of students believed that diabetes was a hereditary condition and 46.49% of students thought that it could reduce life expectancy (Table 2).

TABLE 3 Frequency distribution of the responses to attitude questions.

Questions	Response	Frequency	Percentage
Is it possible to control diabetes?			
	Yes	558	27.77
	No	899	44.74
	Don't know	552	27.47
What should a diabetic patient do to control diabetes?			
	Balanced diet	190	9.45
	Exercise/Walking	599	29.81
	Both	897	44.64
	Don't know	323	16.07
How can diabetes be treated?			
	Regular exercise	1,167	58.08
	Balanced diet	404	20.11
	Regular checkup	329	16.37
	Insulin	857	42.65
	Metformin	269	13.39
	Paracetamol	88	4.38
	Don't know	414	20.61
Long term use of Insulin intake increases complications			
	Agree	234	11.64
	Disagree	342	17.02
	Don't know	1,433	71.33
Controlled eating habits and regular exercise greatly reduce the risk of diabetes			
	Agree	1,251	62.27
	Disagree	79	3.93
	Don't know	679	33.79

In multiple response analysis, 24.63% of respondents did not know any symptoms of diabetes, whereas 50.32% of respondents reported that frequent urination might be a sign of diabetes and 19.01% answered feeling tired all the time as a symptom. Regarding the risk factors for diabetes, 27.47% of respondents reported that the disease could occur from family heredity. Overconsumption of carbs or sugary meals can raise the risk of diabetes reported by 55.55% of the participants. Furthermore, being overweight was cited as a risk factor for diabetes by 26.48% of students, while 25.38% reported that they were unaware of any diabetes risk factors.

Overall, a higher proportion of girls were found to have good knowledge of diabetes than boys (18.57% vs. 14.53%). Interestingly, the difference in knowledge status is insignificant between mediums of education. Conversely, there were statistically significant differences in knowledge level by place of residence, ethnicity and monthly family income (Table 3).

From the logistic regression analysis, we found that for 1 year change of age there were no significant changes in students' knowledge level regarding diabetes. English medium students

were significantly more likely to have good knowledge than the Bangla medium students (19.31%, OR: 1. 44; CI: 1.07, 1.95). Respondent's place of residence was significantly associated with their knowledge regarding diabetes. Respondents of semi-urban (19.66%, OR: 1.7; CI: 1.21, 2.36) and urban (18.47%, OR: 1.48; CI: 1.17, 1.86) areas were more likely to have good knowledge than rural areas. Monthly family income level had a significant effect on students' knowledge. The likelihood of having good knowledge was more among students whose monthly family income was more than 500 USD. After adjusting for known covariates and confounding factors, we found no significant association between demographic characteristics and student knowledge about diabetes except the student's place of residence. Students of semi-urban areas were more likely to have good knowledge than those in rural areas (19.66%, OR: 1.62; CI: 1.14, 2.31) (Table 4).

TABLE 4 Frequency distribution of the responses to practice questions.

Questions	Response	Frequency	Percentage
How often do you eat Fruit?			
	Everyday	651	32.41
	More than twice in a week	649	32.31
	Once in a week	424	21.10
	Once in a month	254	12.64
	Never	31	1.54
How often do you eat Vegetable?			
	Everyday	1,260	62.72
	More than twice in a week	541	26.92
	Once in a week	139	6.91
	Once in a month	43	2.14
	Never	26	1.29
How often do you eat Soft Drinks?			
	Everyday	264	13.14
	More than twice in a week	267	13.29
	Once in a week	574	28.57
	Once in a month	694	34.54
	Never	210	10.45
How often in a month do you eat junk/ fast food?			
	≤10 days	1,667	82.97
	More than 10 days	342	17.02
Have you ever attended an educational class on diabetes?			
	Yes	221	11.00
	No	1,788	88.99
Do you follow any measures to prevent diabetes?			
	Yes	215	10.70
	No	1,225	60.97
	Don't know	569	28.32
Do you follow any specific eating habits to avoid diabetic?			
	Yes	652	32.45
	No	1,357	67.54
Do you exercise/walk/run in a typical day?			
	Yes	1,742	86.71
	No	218	10.85
Do you play outdoor games in a day (cricket/football/others)?			
	Yes	1,440	71.67
	No	569	28.32

Attitude

All about 28% of respondents reported that it might be possible to control diabetes; 62.27% of them reported controlled diet and regular exercise significantly lower the chance of developing diabetes (Table 5). A frequent checkup is important for managing diabetes as reported by 16.37% of the students and 20.61% of students did not know of any diabetic treatment options. Of the participants, 42.65% reported that insulin could treat diabetes, while the majority of students (71.33%) were unaware that long-term insulin use could raise difficulties.

Practice

To prevent diabetes, 32.41% of students reported that they eat fruits every day (Table 6). Most of the students reported doing physical activities daily (86.71%) and playing outdoor games (71.67%). Only a small percentage of students (11%) attended any diabetes education classes and 10.7% of students followed any recommended diabetes prevention methods.

Practices related to diabetes prevention were not found to be significantly different by gender, parents' educational levels, or place of living.

Determinants of diabetes knowledge

It was found from logistic regression analysis that age was not a significant predictor of students' knowledge level regarding diabetes (Table 4). English medium students were significantly more likely to have good knowledge than the Bangla medium students (19.31%, OR: 1. 44; CI: 1.07, 1.95) as found from an unadjusted model; however, this finishing was not significant after adjustment with covariates. Respondent's place of residence was significantly associated with their knowledge regarding diabetes. Respondents from semi-urban areas (19.66%, OR: 1.7; CI: 1.21, 2.36) and urban (18.47%, OR: 1.48; CI: 1.17, 1.86) areas were more likely to have good knowledge than those living in rural areas. Monthly family income level had a significant effect on students' knowledge. The likelihood of having good knowledge was more among students whose monthly family income was more than USD 500, however, after adjusting

10.3389/fpubh.2022.1047617

TABLE 5 Relationship among knowledge and practice with socio demographic variables.

Variables	Knowledge			Practice		
	Good *n* (%)	Poor *n* (%)	χ2 (*p*-value)	Good *n* (%)	Poor *n* (%)	χ2 (*p*-value)
Gender						
Male	141 (14.53)	829 (85.46)	5.62 (0.017)	104 (10.87)	852 (86.95)	1.98 (−0.1591)
Female	193 (18.57)	846 (81.41)		131 (13.04	873 (89.12	
Medium of education						
Bangla	230 (16.59)	1,156 (83.41)	2.28 (−0.319)	146 (10.79)	1,206 (89.20)	5.97 (−0.0504)
English	51 (19.31)	213 (80.68)		37 (14.12)	225 (85.87)	
Madrasha	53 (14.76)	306 (85.24)		52 (15.02)	294 (84.97)	
Place of residence						
Rural	132 (14.17)	799 (85.82)	7.68 (−0.022)	99 (10.93)	806 (89.06)	1.76 (−0.414)
Semi-Urban	47 (19.66)	192 (80.33)		30 (12.82)	204 (87.17)	
Urban	155 (18.47)	684 (81.53)		106 (12.91)	715 (87.08)	
Father's education						
None	37 (24.46)	125 (77.16)	12.1 (−0.007)	19 (11.94)	140 (88.05)	1.16 (−0.763)
Primary	81 (16.07)	423 (83.92)		55 (11.24)	434 (88.75)	
Secondary	158 (21.41)	580 (78.59)		82 (11.53)	629 (88.46)	
Graduation	148 (24.46)	457 (75.37)		79 (13.14)	522 (86.85)	
Mother's education						
None	38 (22.35)	132 (77.64)	8.29 (−0.04)	19 (11.72)	143 (88.27)	3.7 (−0.295)
Primary	101 (17.53)	475 (82.46)		64 (11.53)	491 (88.46)	
Secondary	169 (21.28)	625 (78.71)		85 (10.89)	695 (89.10)	
Graduation	116 (24.73)	353 (75.26)		67 (14.47)	396 (85.53)	
Monthly family income						
≤10,000	61 (12.57)	424 (87.42)	13.17 (0.011)	50 (10.50)	426 (89.49)	1.75 (−0.781)
10,001–20,000	65 (15.74)	348 (84.26)		49 (12.09)	356 (87.90)	
20,001–30,000	82 (17.82)	378 (82.17)		58 (13.33)	377 (86.67)	
30,001–50,000	61 (16.99)	298 (83.00)		43 (12.07)	313 (87.92)	
≥50,001	65 (22.26)	227 (77.73)		35 (12.15)	253 (87.84)	

for known covariates and confounding factors, we found no significant association between demographic characteristics and student knowledge about diabetes except for students' place of residence. Students of the semi-urban areas were more likely to have good knowledge than rural areas (Adjusted OR: 1.63; CI: 1.14–2.31).

Discussion

This study aimed to study knowledge, attitudes and practices related to diabetes among secondary school students. Our findings revealed that a large proportion of students have overall poor knowledge and low awareness about diabetes, which is a major health issue of our time. This study included young population of diverse socio-demographic background to include indigenous students and others who opted for the private English medium schools. It was also highlighted that our school education systems rarely focus on heath literacy as it was found that only a small percentage of students' reporting having had attended session diabetes. Due to the rapid increase of diabetes cases worldwide; primary prevention, early diagnosis, and educational preventive measures are now being prioritized (13). It is suggested that the management of such health issues has been shown to be benefited greatly from people's education (17). Understanding diabetes and increasing general population awareness of the condition at a young age can help with early identification which ultimately can lessen life-threatening consequences that are linked to the disease (19, 20).

Overall participation rate was good in this study and the male-female ratio of the participants represents the current population scenario of Bangladesh (25). The study finds an average to low level of understanding of DM among young students in Bangladesh. While most of the participants knew about diabetes as a disease, but only half of them knew about the association of glucose and insulin with diabetes. Nearly half

10.3389/fpubh.2022.1047617

TABLE 6 Results of multiple logistic regression on factors determining level of diabetes knowledge.

Variables		Knowledge			
		Unadjusted		Adjusted	
		OR (95% CI)	*P*-value	OR (95% CI)	*P*-value
Age		0.97 (0.92–1.02)	0.248	0.96 (0.91–1.01)	0.1747
Gender	Female	Ref		Ref	
	Male	0.82 (0.66–1.01)	0.0676	0.81 (0.64–1.00)	0.0553
Medium of education	Bangla	Ref		Ref	
	English	1.45 (1.07–1.95)	0.0156*	1.18 (0.81–1.70)	0.3808
	Madrasha	0.89 (0.66–1.19)	0.4242	0.91 (0.65–1.26)	0.5698
Place of residence	Rural	Ref		Ref	
	Semi-Urban	1.70 (1.21–2.36)	0.0018**	1.63 (1.14–2.31)	0.0067*
	Urban	1.48 (1.17–1.86)	0.00103**	1.25 (0.92–1.69)	0.1553
Monthly family income	<10,000	Ref		Ref	
	10,000–20,000	1.33 (0.95–1.86)	0.0923	0.79 (0.56–1.11)	0.1697
	20,000–30,000	1.21 (0.87–1.68)	0.253	1.18 (0.77–1.81)	0.4515
	30,000–50,000	1.34 (0.95–1.9)	0.0923	0.88 (0.63–1.23)	0.4514
	>50,000	1.89 (1.34–2.68)	<0.001***	0.88 (0.6–1.28)	0.5029
Father's education	None	Ref		Ref	
	Graduation	1.09 (0.73–1.67)	0.668	0.96 (0.51–1.83)	0.8934
	Primary	0.65 (0.42–1.01)	0.0508	0.72 (0.43–1.22)	0.2164
	Secondary	0.92 (0.62–1.4)	0.6891	1 (0.58–1.74)	0.9987
Mother's education	None	Ref		Ref	
	Graduation	1.14 (0.76–1.75)	0.534	0.8 (0.42–1.55)	0.5033
	Primary	0.74 (0.49–1.34)	0.157	0.89 (0.54–1.5)	0.665
	Secondary	0.94 (0.64–1.41)	0.758	0.85 (0.49–1.47)	0.549

*Significant with $p < 0.05$.
**Significant with $p < 0.005$.
***Significant with $p < 0.001$.

of the students were found to have awareness of the negative effect of taking excess carbohydrates in their meals, however, only one-fourth of the participants were able to recognize the association of obesity with diabetes which is very low compared to the findings of studies conducted on adolescents of Kuwait and Nigeria (12, 20). This finding might coincide with the fact that diabetes in the adult population in Bangladesh who are also obese or overweight has been reported to be high (26, 27).

In general, the school children exhibited poor knowledge of the signs and symptoms of diabetes, for example, only half of the participants knew excessive thirst and increased urination as a sign of diabetes. A similar knowledge pattern was observed in Nigerian adolescents but young students in Kuwait were found quite aware of the diabetes symptoms (12, 20). The fact that young students have such a limited understanding of the signs and symptoms of diabetes needs further attention, as early detection of the condition and fast treatment are both possible with early awareness of the symptoms (24–26).

It is rather concerning that nearly about three-fourths of participants were not aware that diabetes can be controlled through physical exercise, regulated diet intake and recommended medication including insulin when needed, which is very low in contrast with another study (23).

The study identified inadequate awareness programs running in the country that can contribute to raising awareness of diabetes and its associated management system, particularly among the young population in Bangladesh. Only 11% of the participants reported attending any educational programs on diabetes leaving a huge majority of the participants out of any kind of formal awareness-raising programs that can prevent diabetes in the future. However, the study found a satisfactory level of practices that ultimately can help in reducing the risk of diabetes in the future. About two-thirds of the participants used to have fruits in their meals regularly. Almost a universal level of the participants reported having vegetables in their meals at least twice a week.

Knowledge, attitude, and practice related to diabetes were found to have been influenced by socioeconomic factors, societal practices, and behavioral patterns (20). This study supports the mainstream findings since the regression model suggests that

odds of having good knowledge about diabetes was significantly higher among students who belong to high-income families, those who live in urban or semi-urban areas, and studies in schools whose medium of instruction is English. However, this study did not find any significant differences in diabetes-related KAP status by their gender, medium of education, location, parents' education as well as family income.

This study was exclusive in that it focused on students in grades 6 through 10 who were between the ages of 9 and 16 and represented all geographical divisions of Bangladesh, encompassing rural, suburban, and urban areas. The study also includes students of indigenous origin. This is the first study in Bangladesh conducted to date among school children that evaluate the KAP of DM. The study also had a few limitations. Purposive sampling was used for selection of schools and the students as participants. Although the study area covered all eight administrative divisions in Bangladesh and study participants representing diverse socioeconomic status, however, precaution should be practiced for generalizability of the study finding. Data were collected through interview of school children without using particular scale, so collected data may suffer from recall bias or social desirability bias.

Conclusion

The study finds an average to low level of knowledge and awareness of DM among secondary school students in Bangladesh, while most of them knew about diabetes itself but they lack some basic knowledge related to signs and symptoms, prevention and control measures. There was a substantial lack of accessible awareness-raising programs for the young population that may be attributable to overall poor KAP among school children. This study highly recommends initiating large-scale country-wide awareness-raising events where young-age people can participate and acquire knowledge and skills to shape up their plans in fighting chronic diseases like diabetes in the future.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by Ethical Review Committee (CMC/PG/2020/122) of Chittagong Medical College, Chattogram, Bangladesh. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

FA, SR, MQ, NA, and AM conceived and designed the experiments. FA and AM coordinated the study. NL, MN, AM, and NA analyzed the data and drafted the manuscript. FA, MQ, MN, NA, and AM revised the manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgments

Authors would like to thank the authority of participating schools and madrashas, Disease Biology and Molecular Epidemiology Research Group (dBme) Chittagong, and Chittagong University Research and Higher Study Society (CURHS) for their support during the study.

Conflict of interest

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

Publisher's note

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

References

1. Sun H, Saeedi P, Karuranga S, Pinkepank M, Ogurtsova K, Duncan BB, et al. IDF Diabetes Atlas: global, regional and country-level diabetes prevalence estimates for 2021 and projections for (2045). *Diabetes Res Clin Pract.* (2022) 183:109119. doi: 10.1016/j.diabres.2021.109119

2. World Health Organization. *Diabetes: Key Facts 2020*. (2021). Available online at: https://www.who.int/news-room/fact-sheets/detail/diabetes.

3. Ewen M, Joosse H-J, Beran D, Laing R. Insulin prices, availability and affordability in 13 low-income and middle-income countries. *BMJ Glob Health.* (2019) 4:e001410. doi: 10.1136/bmjgh-2019-001410

4. Haque M, Islam T, Rahman NAA, McKimm J, Abdullah A, Dhingra S. Strengthening primary health-C are services to help prevent and control long-term (chronic) non-communicable diseases in low-and middle-income countries. *Risk Manag Healthc Policy.* (2020) 13:409. doi: 10.2147/RMHP.S239074

5. Federation ID. *IDF Diabetes Atlas Tenth edition*. 10. Magliano D, editor. ISBN: 978-2-930229-98-0 (2021). Available online at: https://diabetesatlas.org/atlas/tenth-edition/ (accessed July 19, 2022).

6. Biswas T, Islam A, Rawal L, Islam S. Increasing prevalence of diabetes in Bangladesh: a scoping review. *Public Health.* (2016) 138:4–11. doi: 10.1016/j.puhe.2016.03.025

7. Talukder A, Hossain M. Prevalence of diabetes mellitus and its associated factors in Bangladesh: application of two-level logistic regression model. *Sci Rep.* (2020) 10:1–7. doi: 10.1038/s41598-020-66084-9

8. Divers J, Mayer-Davis EJ, Lawrence JM, Isom S, Dabelea D, Dolan L, et al. Trends in incidence of type 1 and type 2 diabetes among youths—selected counties and Indian reservations, United States, 2002–2015. *Morbid Mortal Weekly Rep.* (2020) 69:161. doi: 10.15585/mmwr.mm6906a3

9. Kao K-T, Islam N, Fox DA, Amed S. Incidence trends of diabetic ketoacidosis in children and adolescents with type 1 diabetes in British Columbia, Canada. *J Pediatr.* (2020) 221:165–73. e2. doi: 10.1016/j.jpeds.2020.02.069

10. Candler T, Mahmoud O, Lynn R, Majbar A, Barrett T, Shield J. Continuing rise of type 2 diabetes incidence in children and young people in the UK. *Diabetic Med.* (2018) 35:737–44. doi: 10.1111/dme.13609

11. Sayeed MA, Rahman MM, Fayzunnessa N, Khanam PA, Begum T, Mahtab H, et al. Childhood diabetes in a Bangladeshi population. *J Diabetes Mellitus.* (2013) 3:33. doi: 10.4236/jdm.2013.31006

12. Unadike B, Chineye S. Knowledge, awareness, and impact of diabetes among adolescents in Uyo, Nigeria. *Afr J Diabetes Med.* (2009) 3:12–14. Available online at: https://www.africanjournalofdiabetesmedicine.com/abstract/knowledge-awareness-and-impact-of-diabetes-among-adolescents-in-uyo-nigeria-76887.html

13. Mohan D, Raj D, Shanthirani C, Datta M, Unwin N, Kapur A, et al. Awareness and knowledge of diabetes in Chennai-the Chennai urban rural epidemiology study [CURES-9]. *J Assoc Phys India.* (2005) 53:283–7. Available online at: https://www.japi.org/u28464b4/awareness-and-knowledge-of-diabetes-in-chennai-the-chennai-urban-rural-epidemiology-study-cures-9

14. Chatterjee S, Davies MJ, Heller S, Speight J, Snoek FJ, Khunti K. Diabetes structured self-management education programmes: a narrative review and current innovations. *Lancet Diabetes Endocrinol.* (2018) 6:130–42. doi: 10.1016/S2213-8587(17)30239-5

15. Ross J, Stevenson FA, Dack C, Pal K, May CR, Michie S, et al. Health care professionals' views towards self-management and self-management education for people with type 2 diabetes. *BMJ Open.* (2019) 9:e029961. doi: 10.1136/bmjopen-2019-029961

16. Ansari RM, Harris M, Hosseinzadeh H, Zwar N. Healthcare professionals' perspectives of patients' experiences of the self-management of type 2 diabetes in the rural areas of pakistan: a qualitative analysis. *Int J Environ Res Public Health.* (2021) 18:9869. doi: 10.3390/ijerph18189869

17. Fatema K, Hossain S, Natasha K, Chowdhury HA, Akter J, Khan T, et al. Knowledge attitude and practice regarding diabetes mellitus among nondiabetic and diabetic study participants in Bangladesh. *BMC Public Health.* (2017) 17:1–10. doi: 10.1186/s12889-017-4285-9

18. Chawla SPS, Kaur S, Bharti A, Garg R, Kaur M, Soin D, et al. Impact of health education on knowledge, attitude, practices and glycemic control in type 2 diabetes mellitus. *J Fam Med Primary Care.* (2019) 8:261. doi: 10.4103/jfmpc.jfmpc_228_18

19. Alemayehu AM, Dagne H, Dagnew B. Knowledge and associated factors towards diabetes mellitus among adult non-diabetic community members of Gondar city, Ethiopia (2019). *PloS ONE.* (2020) 15:e0230880. doi: 10.1371/journal.pone.0230880

20. Al-Hussaini M, Mustafa S. Adolescents' knowledge and awareness of diabetes mellitus in Kuwait. *Alexandria J Med.* (2016) 52:61–6. doi: 10.1016/j.ajme.2015.04.001

21. Banerjee AT, Mahajan A, Mathur-Balendra A, Qureshi N, Teekah M, Yogaratnam S, et al. Impact of the South Asian Adolescent Diabetes Awareness Program (SAADAP) on diabetes knowledge, risk perception and health behaviour. *Health Educ J.* (2022) 81:96–108. doi: 10.1177/00178969211051054

22. Alenazi MA, Alenezi SH, Alhablani MN, Alanazi MAM, Alenazi WH, AlQahtani MS, et al. Knowledge and awareness of diabetes mellitus disease among high school students in King Abdulaziz Military City, Tabuk, Saudi Arabia. *Open Access Macedonian J Med Sci.* (2020) 8:91–7. doi: 10.3889/oamjms.2020.4194

23. Holder M, Ehehalt S. Significant reduction of ketoacidosis at diabetes onset in children and adolescents with type 1 diabetes—The Stuttgart Diabetes Awareness Campaign, Germany. *Pediatric Diabetes.* (2020) 21:1227–31. doi: 10.1111/pedi.13064

24. Paul A, Sikdar D, Hossain MM, Amin MR, Deeba F, Mahanta J, et al. Knowledge, attitudes, and practices toward the novel coronavirus among Bangladeshis: implications for mitigation measures. *PloS ONE.* (2020) 15:e0238492. doi: 10.1371/journal.pone.0238492

25. Molla MA-M. Women outnumber men for first time- finds latest population census. *The Daily Star* (2022, July 27).

26. Hossain MB, Khan MN, Oldroyd JC, Rana J, Magliago DJ, Chowdhury EK, et al. Prevalence of, and risk factors for, diabetes and prediabetes in Bangladesh: evidence from the national survey using a multilevel Poisson regression model with a robust variance. *PLoS Glob Public Health.* (2022) 2:e0000461. doi: 10.1371/journal.pgph.0000461

27. Akter J, Shahjahan M, Hossain S, Chowdhury HA, Ahmed KR, Fatema K, et al. Determinants of overweight and obesity among Bangladeshi diabetic women of reproductive age. *BMC Res Notes.* (2014) 7:1–6. doi: 10.1186/1756-0500-7-513

frontiers | Frontiers in Pediatrics

TYPE Original Research
PUBLISHED 29 November 2022
DOI 10.3389/fped.2022.1051432

OPEN ACCESS

EDITED BY
Jing Wu,
Tianjin University, China

REVIEWED BY
Arlette Setiawan,
Universitas Padjadjaran, Indonesia
Hsin-Yu Lai,
Harvard University, United States

*CORRESPONDENCE
Dongchuan Yu
dcyu@seu.edu.cn

SPECIALTY SECTION
This article was submitted to Children and Health, a section of the journal Frontiers in Pediatrics

RECEIVED 22 September 2022
ACCEPTED 03 November 2022
PUBLISHED 29 November 2022

CITATION
Wang H, Liu F, Dong Y and Yu D (2022) Features of eye movements during rapid automatized naming in Chinese children aged 7–11 years. Front. Pediatr. 10:1051432. doi: 10.3389/fped.2022.1051432

Features of eye movements during rapid automatized naming in Chinese children aged 7–11 years

Hongan Wang[1], Fulin Liu[1], Yuhong Dong[2] and Dongchuan Yu[1,2,3]*

[1]Key Laboratory of Child Development and Learning Science of Ministry of Education, School of Biological Science and Medical Engineering, Southeast University, Nanjing, China, [2]Henan Provincial Medical Key Lab of Language Rehabilitation for Children, Sanmenxia Center Hospital, Sanmenxia, China, [3]Henan Provincial Medical Key Lab of Child Developmental Behavior and Learning, The Third Affiliated Hospital of Zhengzhou University, Zhengzhou, China

Rapid Automatized Naming (RAN) tests have been well-documented to predict reading abilities as well as a variety of neurobiological disorders (e.g., developmental dyslexia). Traditional measures of RAN tests only take into account the naming time and accuracy and cannot reflect temporal-spatial features during RAN tests. Although the eye tracking approach appears to be a promising tool for characterizing the essential temporal-spatial characteristics of RAN tests, no research has been conducted to investigate whether and how gender, age, and task-type alter those characteristics. Additionally, no study has examined eye movements during a Chinese adaptation of RAN in order to expand the applicability of RAN to developmental dyslexia in Chinese. To address the concerns stated above, this article recruited 408 children (206 males, aged 7–11 years) and adopted eight measures to quantify features of eye movements during a Chinese adaptation of RAN. Findings showed that: (1) eight eye-movement measures had the main effects of task-type and age, but only five of them had the main effect of gender (in particular, females outperformed males); (2) RAN abilities observed by eight eye-movement measures initially developed quickly before the age of 9, and then entered a relatively sluggish development phase; (3) non-alphanumeric RAN tasks generally required higher mental load (implying more fixation counts, saccade counts, and regression counts, smaller average saccade amplitude, fixation duration fluctuation and saccade amplitude fluctuation, and longer average fixation duration and total time of naming) than alphanumeric ones; (4) there were significant correlations between total time of naming (a widely-used behavioral parameter) and other eye-movement measures; and (5) there were significant correlation between eight eye-movement measures and three attention-related skills observed from a number cancellation task. The current study might offer some perspectives on the understanding of normative data of eye movements during RAN in Chinese school-aged children, as well as the applications (e.g., developmental dyslexia) associated with RAN.

KEYWORDS

developmental dyslexia, rapid automatized naming, eye tracking, school - aged children, reading abilities

Introduction

Rapid automatized naming (RAN) tests (1–12) have been proposed to evaluate an individual's ability to read letters, digits, objects, or other visual stimuli such as colors or geometric shapes as quickly as possible. These RAN tasks were initially used in studies on reading behavior and child development in healthy children, but they have gradually become recognized as a popular and useful psychometric test for identifying a variety of cognitive and neurobiological abnormalities, including developmental dyslexia (13–15), specific language impairment (16), attention deficit/hyperactivity disorder (ADHD) (17), learning disabilities, and autism spectrum disorder (ASD) (18, 19). In particular, findings (16, 20–24) have demonstrated that the RAN deficits may more accurately describe the characteristics of developmental dyslexia than other cognitive skill deficits. The goal of the current study was to offer some new perspectives on how to understand RAN more fully.

Traditional measures of RAN tests only take into account the naming time and accuracy and cannot reflect dynamical temporal-spatial features during RAN tests. While eye tracking techniques can be used to monitor the focus points in sequence and record the essential ocular activities throughout visual cognitive processes. Therefore, eye tracking approach would be a promising tool to characterize the visual cognitive features of RAN. Only a small number of studies (19, 25) have explored the characteristics of eye movements during RAN tests, but none of them have examined whether and how gender, age, and task type alter those characteristics. Additionally, no study has been done to analyze eye movements during a Chinese adaptation of RAN in order to expand the applicability of RAN to developmental dyslexia in Chinese. This study sought to identify the patterns or characteristics of Chinese school-aged children's eye movements during a Chinese adaptation of RAN (including naming Chinese characters). It should be noted that normative data of eye movement during RAN must be established in order to identify developmental abnormalities connected to RAN from eye movement data. However, thus far, no study has been conducted to establish normative data of eye movements during RAN in school-aged children, especially in China.

Taken together, we used the eye tracking technique to measure the characteristics of RAN eye movements and identified RAN eye movement characteristics in Chinese children aged 7–11 years. We specifically examined the characteristics of eye movements in Chinese children between the ages of 7 and 11 while they named Chinese characters. This is the first time, as far as we are aware, that eye movement characteristics during RAN testing for Chinese children aged 7–11 have been reported. In order to determine how the gender, age, and task type affect the characteristics of eye movements during RAN, we recruited 408 kids (206 of them were boys, ages 7–11). The relationship between eye-movement characteristics was examined. It was also explored how features of eye movements connect to abilities associated with attention.

Materials and methods

The Southeast University Research Ethics Committee gave its approval to all study protocols and research techniques, ensuring that they adhered to the World Medical Association's Declaration of Helsinki regarding the use of humans in testing. All participating children's parents gave their informed consent, and each participant gave their verbal consent. After finishing the research, each kid was given a toy that was appropriate for their age.

Study design and participants

The current study was conducted in Sanmenxia, Henan Province, China, between September 2021 to March 2022. According to the districts' rankings of GDP per person in 2020, the districts of Sanmenxia were divided into three levels, i.e., strong economic level (>90,000 RMB), medium economic level (70,000–90,000 RMB), weak economic level (<70,000 RMB). In order to prevent bias in sample selection, we randomly selected a district with medium economic level and randomly selected an ordinary primary school locally from the district. This primary school included 1,387 kids (aged 7–11). According to the sequence number chosen randomly, a coding number was given to each child who was recruited. We only invited kids whose coding numbers with 3, 6 or 9 in the ones digit were to participate in our experiments.

Exclusion criteria were as follows: (a) abnormal hearing functioning (i.e., hearing threshold levels bigger than 25 dB HL) and vision functioning (i.e., naked or corrected monocular visual acuities below than 1.0); (b) significant sensory or motor impairment; (c) a history of previous neurological or psychiatric disorders; (d) IQ score lower than 85 or bigger than 115; (e) children who had repeated a grade; and (f) incomplete measure data.

By steps above, a total of 408 children (206 males) attended the current experiments (see **Table 1** for detailed information). This survey complied with the sampling criteria since, according to Weeks' work (26), the sample size to be obtained in the event of a 95% confidence level and +5% accuracy was calculated to be $n = 384$.

TABLE 1 Demographic characteristics of participants.

Age groups	Males (*N*, %)	Total (*N*)	Age (years)
7-years children	33 (50.00)	66	7.55 ± 0.24
8-years children	45 (47.87)	94	8.46 ± 0.28
9-years children	45 (54.88)	82	9.43 ± 0.28
10-years children	31 (45.59)	68	10.41 ± 0.30
11-years children	52 (53.06)	98	11.50 ± 0.28
Total	496 (50.49)	408	N/A

Experimental tasks

RAN tasks

We employed a Chinese adaptation of RAN (C-RAN) (27) that substituted highly-frequently-used Chinese characters for English letters. The C-RAN paradigm in this study consisted of four tasks: Task N-number (i.e., naming of numbers), Task N-character (i.e., naming of Chinese characters), Task N-object (i.e., naming of objects), and Task N-color (i.e., naming of colors). While, N-object and N-color were non-alphanumeric RAN tasks; N-number and N-character were alphanumeric RAN tasks. For each task, a 5×10 matrix of objects was presented, in which each matrix used five repetitions of each of the ten different objects with the order pseudo-randomized.

The subject was situated between 60 and 90 cm away from the 21.5 in. TFT LCD monitor (with $1{,}920 \times 1{,}080$ resolution) displaying the stimuli for each C-RAN task. Eye movements were recorded using a Tobii 4C eye tracker (90 Hz; Tobii Technology AB, Danderyd, Sweden), which was calibrated using a standard 9-point grid. For each RAN task, participants were instructed to name the stimuli (numbers, Chinese characters, colors, or objects) as quickly and accurately as possible in a left-to-right and down fashion.

Number cancellation test

As a second experimental task, a number cancellation test (NCT) (28) was used to gauge a participant's attention-related skills. The participant was given a standard B5-sized piece of paper with a list of numbers structured into 26 rows and 40 columns. The participant, who was given a Digital Pen (with an integrated smart mini-camera), was required to find the number "3" (the targeted number) and then draw a circle on it, but ignore all other numbers (distractors), as quickly as possible within 2 min. The Digital Pen's technical advantage was the use of a smart mini-camera, which was designed to measure temporal-spatial features from the perspective of handwriting kinematics, such as pre-movement time (initiating), movement time (moving pen to a stimulus), drawing time (completing a cancellation), circumference of a drawn curve, real-time spatial positions (trajectory) of drawing, and drawing time sequence. It should be noted (28) that temporal-spatial features may outperform traditional NCT measures.

Quality control

To ensure the consistency and fidelity of administration of evaluation tools, a senior expert with professional experience more than 8 years carried out measures for all participating kids. The senior expert had training in administration of all tools used in this study.

Measures

Measures of eye movements during RAN

This study took into account six traditional eye-movement measures, including fixation counts, saccade counts, regression counts, average fixation duration, average saccade amplitude, and total time of naming, to assess eye movements during RAN tasks. Two novel measures, namely fixation length fluctuation and saccade amplitude fluctuation, were proposed to indicate the dynamic change of attentional maintenance and switching during RAN. The fixation duration fluctuation was defined as follows:

$$F = \frac{1}{n-1}\sum_{i=1}^{n-1} |duration(i+1) - duration(i)| \quad (1)$$

where $duration(i)$ is the time length of the i-th fixation.

While, the saccade amplitude fluctuation was defined as follows:

$$F = \frac{1}{n-2}\sum_{i=1}^{n-2} |distance(i+1) - distance(i)| \quad (2)$$

where $distance(i)$ is the Euclidean distance between the $(i+1)$-th and i-th fixation. In summary, this article will examine eight eye-movement measures (parameters).

Measures of a number cancellation test

Several temporal-spatial parameters can be measured with the Digital Pen (which has an integrated smart mini-camera) during NCT (28). In this study, we selected only three parameters (28) to evaluate individuals' performance during the NCT. Those parameters were defined as follows.

(1) Speed of cognitive processing (SpC) was defined as:

$$SpC = M\sum_{i=1}^{N} R_i \quad (3)$$

where M was the amount of numbers in one row (here $M = 40$); N was the total number of rows to be

circled; $R_i = 1$ represented the case if any number in the i-th row has been circled; and $R_i = 0$ represented the case if no number in the i-th row had been circled.

(2) Selective attention (SA) was defined as:

$$SA = \frac{1}{T}\frac{m-\omega}{m+o} \times \mathrm{SpC} \quad (4)$$

where O was the amount of omitted targets; ω was the number of distractors being circled; and m was the total amount of targets that should be circled; T was the task time (here $T = 120$); SpC was defined by Equation (1).

(3) Averaged time of circlings (ATC) was defined as:

$$ATC = \frac{1}{n} \times \sum_{i=1}^{n} t_i \quad (5)$$

where n was the amount of numbers being circled; and t_i was the time to circle the i-th number.

Statistical analysis

We performed a three-factor (gender, age, and task-type) ANOVA for each of the eight eye-movement measures, where age (7–11 years) and gender (male vs. female) were inter-group factors; and task-type (Task N-number, Task N-character, Task N-object, and Task N-color) was an intra-group factor. After confirming that our data (i.e., eight eye-movement measures) failed to pass the normality test and variance homogeneity test, we conducted a number of nonparametric ANOVA procedures [i.e., the Aligned Rank Transform (ART) procedures] using ARTool software package (29, 30). Traditional nonparametric statistical tests (like the Kruskal-Wallis test, Mann-Whitney U test, Friedman test, or Wilcoxon signed-rank test) are one-way tests and only permit the analysis of one factor at a time, whereas ARTool can be used to analyze multiple factors nonparametrically (29, 30). Additionally, for *post-hoc* multiple comparisons, we utilized the nonparametric Wilcoxon rank-sum test with the "FDR" approach to control the false discovery rate. The effect size was determined by the parameter r (low effect: $0.1 \leq r < 0.3$; medium effect: $0.3 \leq r < 0.5$; efficient response: $r \geq 0.5$).

Pearson's correlation method was used to calculate the correlation among eight eye-movement measures, as well as the correlation between eight eye-movement features and three parameters of the number cancellation test (28). All statistical analysis above was conducted with R language (version 4.0.2), and the significance level α was set at 0.05.

Results

General information of participants

The current study investigated a total of 408 children, including 206 males and 202 females. The ratio of males to females was 1.01:1 and the participants were split up into 5 age groups, see **Table 1** for detailed information. We verified that there was no significant gender difference ($\chi^2 = 1.81$, $p = 0.77$).

Main effects analysis

We performed a three-factor (gender, age, and task-type) ANOVA and *post-hoc* multiple comparisons for each of the eight eye-movement measurements. **Figures 1–8** summarized our results, which were listed as follows.

Fixation counts

Main effects

The main effects of task-type, age and gender were significant (Task-type: $F = 942.32$, $p < 1 \times 10^{-4}$, $\eta^2 = 0.506$; Gender: $F = 6.61$, $p = 0.01$, $\eta^2 = 0.009$; Age: $F = 15.15$, $p < 1 \times 10^{-4}$, $\eta^2 = 0.08$).

Interaction effects

There was a significant interaction effect between gender and task-type ($F = 6.83$, $p < 1 \times 10^{-4}$, $\eta^2 = 0.007$).

Post-hoc *multiple comparisons*

Figure 1 summarized the results of *post-hoc* multiple comparisons and indicated that: (1) For males or females (see **Figure 1A**), task N-number required less fixation counts than other three tasks (p's $< 1 \times 10^{-4}$, adjusted); task N-character required less fixation counts than tasks N-object and N-color (p's $< 1 \times 10^{-4}$, adjusted); task N-object required less fixation counts than tasks N-color (p's $< 1 \times 10^{-4}$, adjusted); (2) As shown in **Figure 1A**, males required more fixation counts in task N-color than females ($w = 25032.0$, $p < 0.001$, $r = 0.18$), but there was no significant gender difference in other tasks (p's > 0.05, adjusted); (3) As shown in **Figure 1B**, children aged 7 years required more fixation counts than children aged 9–11 years (p's < 0.01, adjusted); children aged 8 years required more fixation counts than children aged 10–11 years (p's < 0.01, adjusted); children aged 9 or 10 years required more fixation counts than children aged 11 years (p's < 0.05, adjusted); but there were no significant differences in fixation counts between children 7 and 8 years, between children 8 and 9 years, and between children 9 and 10 years (p's > 0.05, adjusted).

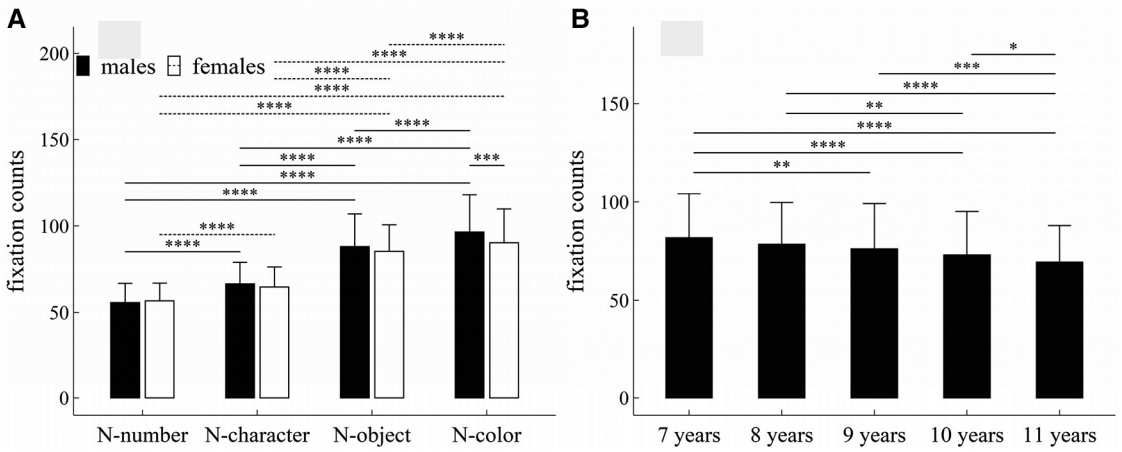


FIGURE 1
Influence of task-type, gender and age on fixation counts: (**A**) interactive effect of gender and task-type; (**B**) influence of age. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 1 \times 10^{-4}$.

Saccade counts

Main effects

The main effects of task-type, age and gender were significant (Task-type: $F = 506.55$, $p < 1 \times 10^{-4}$, $\eta^2 = 0.36$; Gender: $F = 6.61$, $p = 0.01$, $\eta^2 = 0.009$; Age: $F = 7.63$, $p < 1 \times 10^{-4}$, $\eta^2 = 0.04$).

Interaction effects

There were significant interaction effects between gender and task-type ($F = 4.59$, $p < 0.001$, $\eta^2 = 0.005$), and between age and gender ($F = 2.63$, $p = 0.034$, $\eta^2 = 0.02$).

Post-hoc *multiple comparisons*

Figure 2 summarized the results of *post-hoc* multiple comparisons and indicated that: (1) For males or females (see **Figure 2A**), task N-number required less saccade counts than other three tasks (p's $< 1 \times 10^{-4}$, adjusted); task N-character required less saccade counts than tasks N-object and N-color (p's $< 1 \times 10^{-4}$, adjusted); task N-object required less saccade counts than tasks N-color (p's $< 1 \times 10^{-4}$, adjusted); (2) As shown in **Figure 2A**, males required more saccade counts in tasks N-character and N-color than females (N-character: $w = 23145.0$, $p = 0.049$, $r = 0.10$; N-color: $w = 24290.0$, $p = 0.003$, $r = 0.15$), but there was no significant gender difference in other tasks (p's > 0.05, adjusted); (3) As shown in **Figure 2B**, males aged 7 years required more saccade counts than males aged 8–11 years (p's < 0.01, adjusted); males aged 8 years required more saccade counts than males aged 11 years ($w = 21976.0$, $p = 0.006$, $r = 0.15$, adjusted); but there were no significant

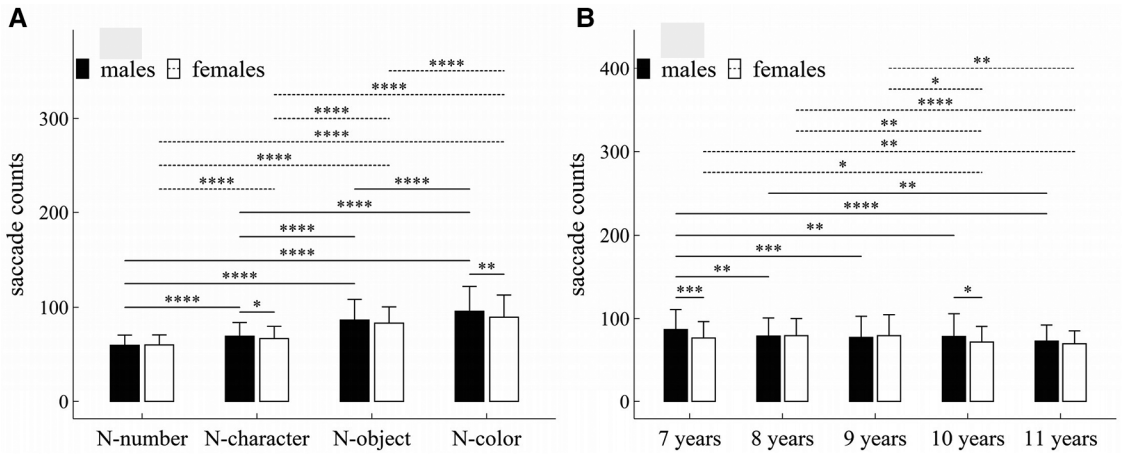


FIGURE 2
Influence of task-type, gender and age on saccade counts: (**A**) interactive effect of gender and task-type; (**B**) interactive effect of age and task-type. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 1 \times 10^{-4}$.

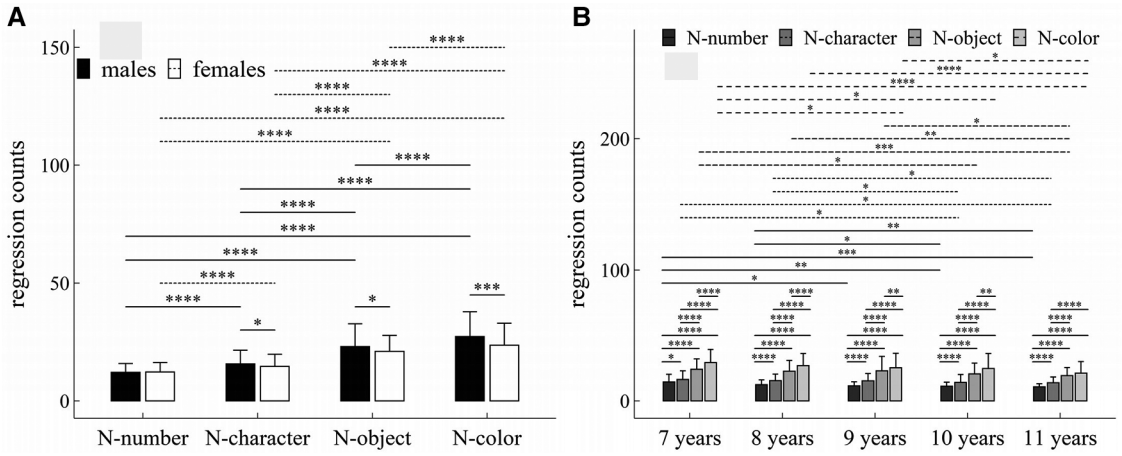


FIGURE 3
Influence of task-type, gender and age on regression counts: (A) interactive effect of gender and task-type; (B) interactive effect of age and task-type. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 1 \times 10^{-4}$.

differences in saccade counts among males aged 8–10 years, and among males aged 9–11 years (p's > 0.05, adjusted); (4) As shown in **Figure 2B**, females aged 7, 8 or 9 years required more saccade counts than females aged 10–11 years (p's < 0.05, adjusted); but there were no significant differences in saccade counts between females aged 7 and 8 years, between females aged 8 and 9 years, and between females aged 10 and 11 years, (p's > 0.05, adjusted); (5) As shown in **Figure 2B**, males aged 7 years required more saccade counts than females aged 7 ($w = 10905.0$, $p < 0.001$, $r = 0.22$); and males aged 10 years required more saccade counts than females aged 10 ($w = 10592.0$, $p = 0.028$, $r = 0.13$).

Regression counts

Main effects

The main effects of gender, age and task-type were significant (Task-type: $F = 407.45$, $p < 1 \times 10^{-4}$, $\eta^2 = 0.35$; Gender: $F = 13.27$, $p < 1 \times 10^{-4}$, $\eta^2 = 0.02$; Age: $F = 10.48$, $p < 1 \times 10^{-4}$, $\eta^2 = 0.05$).

Interaction effects

There were significant interaction effects between age and task-type ($F = 2.35$, $p = 0.01$, $\eta^2 = 0.01$), and between gender and task-type ($F = 7.78$, $p < 1 \times 10^{-4}$, $\eta^2 = 0.01$).

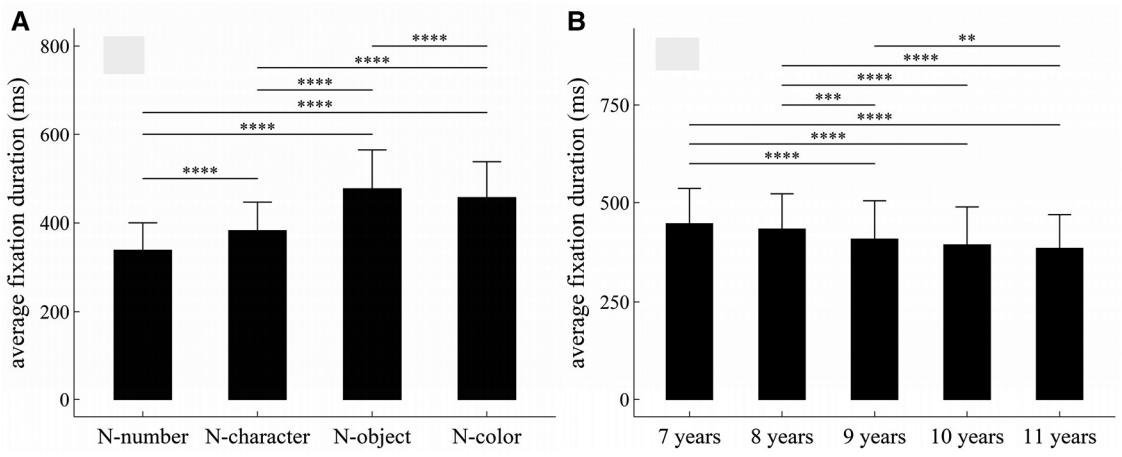


FIGURE 4
Influence of task-type, gender and age on average fixation duration: (A) influence of task-type; (B) influence of age. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 1 \times 10^{-4}$.

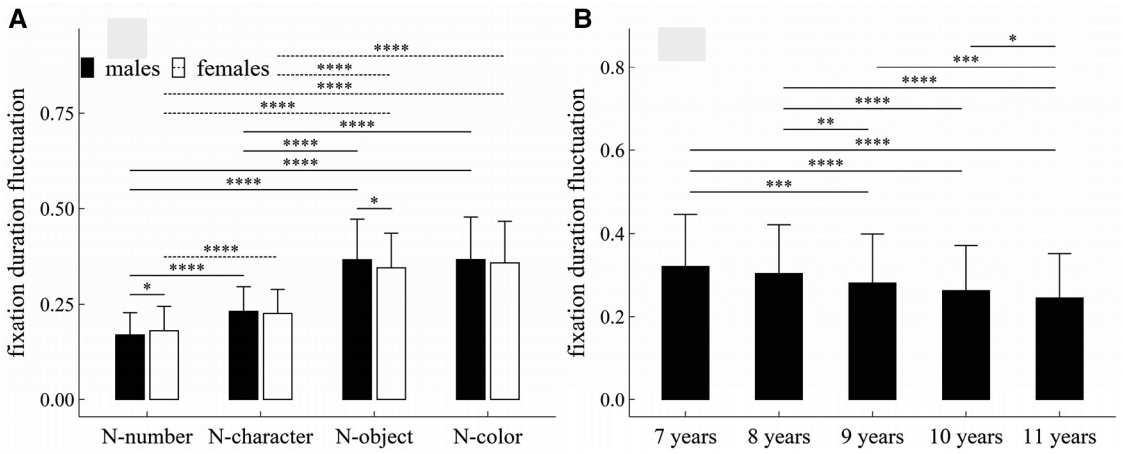


FIGURE 5
Influence of task-type, gender and age on fixation duration fluctuation: (A) interactive effect of gender and task-type; (B) influence of age. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 1 \times 10^{-4}$.

Post-hoc *multiple comparisons*

Figure 3 summarized the results of *post-hoc* multiple comparisons and indicated that: (1) For males or females (see **Figure 3A**), task N-number required more regression counts than other three tasks (p's $< 1 \times 10^{-4}$, adjusted); task N-character required more regression counts than tasks N-object and N-color (p's $< 1 \times 10^{-4}$, adjusted), task N-object required more regression counts than task N-color (p's $< 1 \times 10^{-4}$, adjusted); (2) As shown in **Figure 3A**, males required more regression counts in task N-character ($w = 23194.5$, $p = 0.044$, $r = 0.10$), task N-object ($w = 23145.5$, $p = 0.049$, $r = 0.097$), and task N-color ($w = 25164.5$, $p < 0.001$, $r = 0.18$) than females, but there was no significant gender difference in regression counts in task N-number ($w = 20148.0$, $p = 0.58$, $r = 0.03$); (3) For children aged 7–10 years (see **Figure 3B**), task N-number required less regression counts than other three tasks (p's $<$ 0.05, adjusted); task N-character required less regression counts than tasks N-object and N-color (p's $< 1 \times 10^{-4}$, adjusted), task N-object required less regression counts than task N-color (p's < 0.01, adjusted); (4) For children aged 11 years (see **Figure 3B**), task N-number required less regression counts than other three tasks (p's $< 1 \times 10^{-4}$, adjusted); task N-character required less regression counts than tasks N-object and N-color (p's $< 1 \times 10^{-4}$, adjusted); but there was no significant difference in regression counts between tasks N-object and N-color ($w = 1593.5$, $p = 0.09$, $r = 0.09$); (5) For task

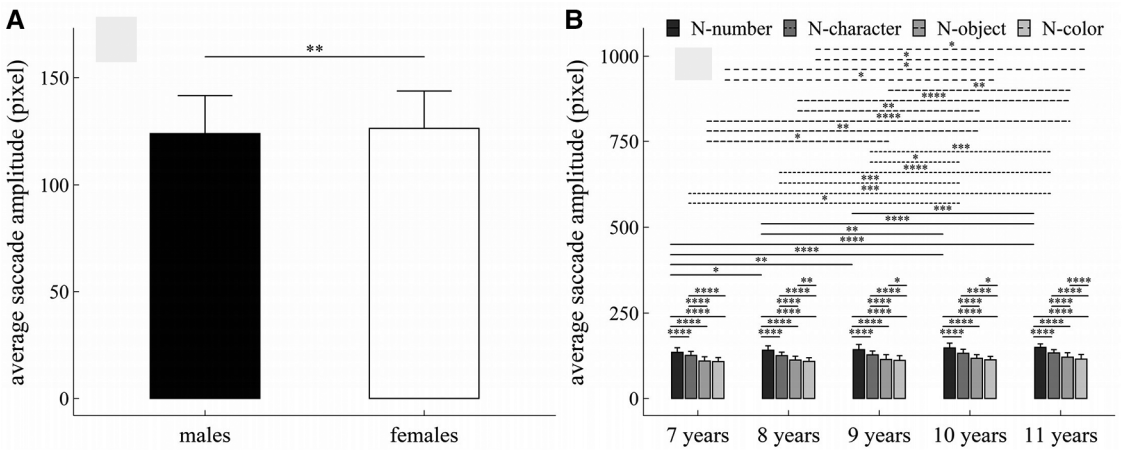


FIGURE 6
Influence of task-type, gender and age on average saccade amplitude: (A) influence of gender; (B) interactive effect of age and task-type. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$.

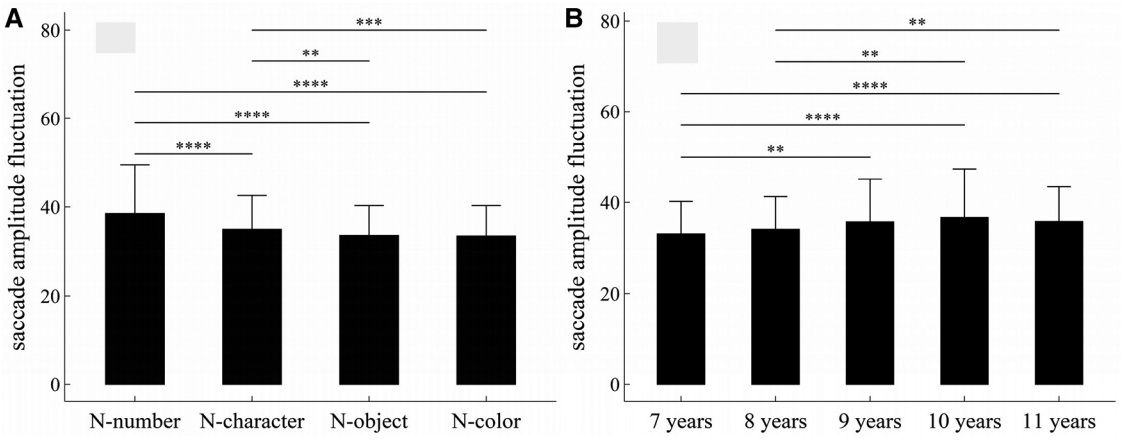


FIGURE 7
Influence of task-type, gender and age on saccade amplitude fluctuation: (A) influence of task-type; (B) influence of age. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 1 \times 10^{-4}$.

N-number (see **Figure 3B**), children aged 7 years required more regression counts than children aged 9–11 years (p's < 0.05, adjusted); children aged 8 years required more regression counts than children aged 10–11 years (p's < 0.05, adjusted); but there were no significant differences between children 7 and 8 years, and among children aged 9–11years (p's > 0.05, adjusted); (6) For task N-character (see **Figure 3B**), children aged 7 years required more regression counts than children aged 10–11 years (p's < 0.05, adjusted); children aged 8 years required more regression counts than children aged 10–11 years (p's < 0.05, adjusted); but there were no significant differences in regression counts between children aged 7 and 8 years, and among children aged 9–11years (p's > 0.05, adjusted); (7) For task N-object (see **Figure 3B**), children aged 7 years required more regression counts than children aged 10–11 years (p's < 0.05, adjusted); children aged 8 or 9 years required more regression counts than children aged 11 years (p's < 0.05, adjusted); but there were no significant differences in regression counts between children aged 7 and 8 years, between children aged 8 and 9 years, and between children aged 8 and 10 years, and between children aged 10 and 11 years (p's > 0.05, adjusted); (8) For task N-color (see **Figure 3B**), children aged 7 years required more regression counts than children aged 9–11 years (p's < 0.05, adjusted); children aged 8 or 9 years required more regression counts than children aged 11 years (p's < 0.05, adjusted); but there were no significant differences in regression counts between children aged 7 and 8 years, between children aged 8 and 9

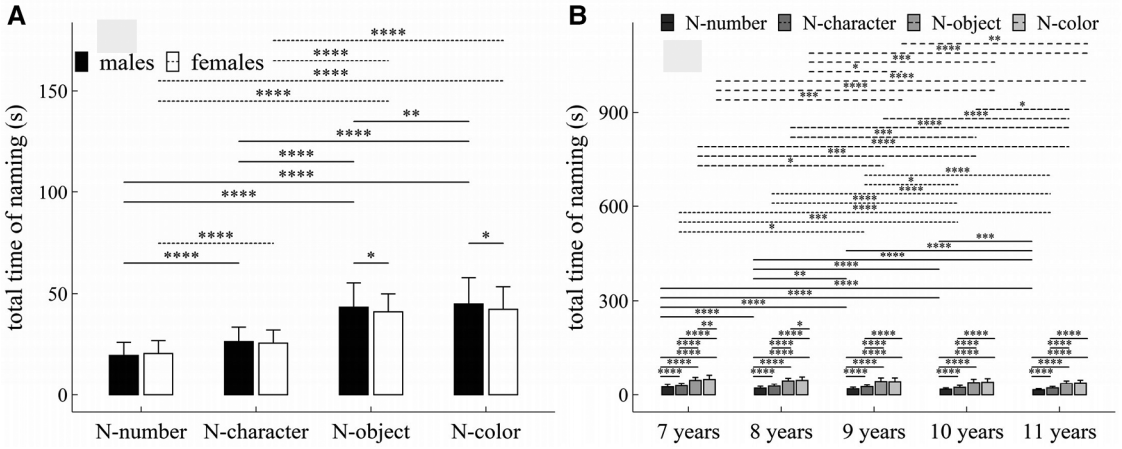


FIGURE 8
Influence of task-type, gender and age on total time of naming: (A) interactive effect of gender and task-type; (B) interactive effect of age and task-type. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 1 \times 10^{-4}$.

years, and between children aged 10 and 11 years (p's > 0.05, adjusted).

Average fixation duration

Main effects

The main effects of task-type and age were significant (Task-type: $F = 738.31$, $p < 1 \times 10^{-4}$, $\eta^2 = 0.37$; Age: $F = 15.58$, $p < 1 \times 10^{-4}$, $\eta^2 = 0.096$).

Interaction effects

There was no significant interaction effect among gender, age and task-type (p's > 0.05).

Post-hoc *multiple comparisons*

Figure 4 summarized the results of *post-hoc* multiple comparisons and indicated that: (1) As shown in Figure 4A, task N-number required shorter average fixation duration than other three tasks (p's $< 1 \times 10^{-4}$, adjusted), task N-character required shorter average fixation duration than tasks N-object and N-color (p's $< 1 \times 10^{-4}$, adjusted), task N-object required longer average fixation duration than task N-color ($w = 57045$, $p < 1 \times 10^{-4}$, $r = 0.12$, adjusted); (2) As shown in Figure 4B, children aged 7 or 8 years required longer average fixation duration than children aged 9–11 years (p's < 0.001, adjusted); children aged 9 years required longer average fixation duration than children aged 11 years (p's < 0.01, adjusted); but there were no significant differences in average fixation duration between children 7 and 8 years, between children 9 and 10 years, and between children 10 and 11 years (p's > 0.05, adjusted).

Fixation duration fluctuation

Main effects

The main effects of task-type and age were significant (Task-type: $F = 902.10$, $p < 1 \times 10^{-4}$, $\eta^2 = 0.495$; Age: $F = 19.71$, $p < 1 \times 10^{-4}$, $\eta^2 = 0.10$).

Interaction effects

There was a significant interaction effect between gender and task-type ($F = 4.53$, $p = 0.007$, $\eta^2 = 0.005$).

Post-hoc *multiple comparisons*

Figure 5 summarized the results of *post-hoc* multiple comparisons and indicated that: (1) For males or females (see Figure 5A), task N-number required smaller fixation duration fluctuation than other three tasks (p's $< 1 \times 10^{-4}$, adjusted), task N-character required smaller fixation duration fluctuation than tasks N-object and N-color (p's $< 1 \times 10^{-4}$, adjusted), but there were no significant differences in fixation duration fluctuation between task N-object and task N-color (p's > 0.05, adjusted); (2) As shown in Figure 5A, males required smaller fixation duration fluctuation in task N-number ($w = 18399.0$, $p = 0.043$, $r = 0.10$), but required larger fixation duration fluctuation in task N-object ($w = 23553$, $p = 0.021$, $r = 0.114$); there were no significant gender difference in tasks N-character and N-color (p's > 0.05, adjusted); (3) As shown in Figure 5B, children aged 7 or 8 years required larger fixation duration fluctuation than children aged 9–11 years (p's < 0.01, adjusted); children aged 9 years required larger fixation duration fluctuation than children aged 11 years ($w = 75325.5$, $p < 0.001$, $r = 0.15$, adjusted); children aged 10 years required larger fixation duration fluctuation than children aged 11 years ($w = 58458.5$, $p = 0.043$, $r = 0.08$, adjusted); but there were no significant differences in fixation duration fluctuation between children 7 and 8 years, and between children 9 and 10 years (p's > 0.05, adjusted).

Average saccade amplitude

Main effects

The main effects of gender, age and task-type were significant (Task-type: $F = 1024.49$, $p < 1 \times 10^{-4}$, $\eta^2 = 0.52$; Gender: $F = 9.88$, $p = 0.002$, $\eta^2 = 0.014$; Age: $F = 16.29$, $p < 1 \times 10^{-4}$, $\eta^2 = 0.087$).

Interaction effects

There were significant interaction effects between age and task-type ($F = 1.96$, $p = 0.03$, $\eta^2 = 0.008$).

Post-hoc *multiple comparisons*

Figure 6 summarized the results of *post-hoc* multiple comparisons and indicated that: (1) As shown in Figure 6A, females required larger average saccade amplitude than males in all RAN tasks ($w = 304618$, $p = 0.003$, $r = 0.07$); (2) For children aged 8–11 years (see Figure 6B), task N-number required larger average saccade amplitude than other three tasks (p's $< 1 \times 10^{-4}$, adjusted); task N-character required larger average saccade amplitude than tasks N-object and N-color (p's $< 1 \times 10^{-4}$, adjusted), task N-object required larger average saccade amplitude than task N-color (p's < 0.05, adjusted); (3) For children aged 7 years (see Figure 6B), task N-number required larger average saccade amplitude than other three tasks (p's $< 1 \times 10^{-4}$, adjusted); task N-character required larger average saccade amplitude than tasks N-object and N-color (p's $< 1 \times 10^{-4}$, adjusted); but there was no significant difference in average saccade amplitude between tasks N-object and N-color ($w = 1387$, $p = 0.07$, $r = 0.06$, adjusted); (4) For task N-number (see Figure 6B), children aged 7 years required smaller average saccade amplitude than children aged 8–11 years (p's < 0.05, adjusted); children aged 8 years required smaller average saccade amplitude than children aged 10–11 years (p's < 0.01, adjusted); children aged 9 years required smaller average saccade amplitude than children aged 11 years ($w = 2709$, $p < 0.001$, $r = 0.28$, adjusted); but there were no significant differences in average saccade amplitude between children 8 and 9 years, between children 9 and 10 years, and between children 10 and 11 years (p's > 0.05, adjusted); (5) For task N-character (see Figure 6B),

children aged 7 years required smaller average saccade amplitude than children aged 10–11 years (p's < 0.05, adjusted); children aged 8 years required smaller average saccade amplitude than children aged 10–11 years (p's < 0.001, adjusted); children aged 9 years required smaller average saccade amplitude than children aged 10–11 years (p's < 0.05, adjusted); but there were no significant differences in average saccade amplitude between children 7 and 8 years, between children 8 and 9 years, and between children 10 and 11 years (p's > 0.05, adjusted); (6) For task N-object (see **Figure 6B**), children aged 7 years required smaller average saccade amplitude than children aged 9–11 years (p's < 0.05, adjusted); children aged 8 years required smaller average saccade amplitude than children aged 10–11 years (p's < 0.01, adjusted); children aged 9 years required smaller average saccade amplitude than children aged 11 years ($w = 2955$, $p = 0.002$, $r = 0.23$, adjusted); but there were no significant differences in average saccade amplitude between children 7 and 8 years, between children 8 and 9 years, and between children 10 and 11 years (p's > 0.05, adjusted); (7) For task N-color (see **Figure 6B**), children aged 7 years required smaller average saccade amplitude than children aged 10–11 years (p's < 0.05, adjusted); children aged 8 years required smaller average saccade amplitude than children aged 10–11 years (p's < 0.05, adjusted); but there were no significant differences in average saccade amplitude between children 7 and 8 years, between children 8 and 9 years, between children 9 and 10 years, and between children 10 and 11 years (p's > 0.05, adjusted).

Saccade amplitude fluctuation

Main effects

The main effects of task-type and age were significant (Task-type: $F = 44.90$, $p < 1 \times 10^{-4}$, $\eta^2 = 0.055$; Age: $F = 4.50$, $p = 0.001$, $\eta^2 = 0.02$).

Interaction effects

There was no significant interaction effect among gender, age and task-type (p's > 0.05).

Post-hoc *multiple comparisons*

Figure 7 summarized the results of *post-hoc* multiple comparisons and indicated that: (1) As shown in **Figure 7A**, task N-number required larger saccade amplitude fluctuation than other three tasks (p's < 1×10^{-4}, adjusted); task N-character required larger saccade amplitude fluctuation than tasks N-object and N-color (p's < 0.01, adjusted); but there was no significant difference in saccade amplitude fluctuation between tasks N-object and N-color ($w = 43107$, $p = 0.50$, $r = 0.01$, adjusted); (2) As shown in **Figure 7B**, children aged 7 years required smaller saccade amplitude fluctuation than children aged 9–11 years (p's < 0.01, adjusted); children aged 8 years required smaller saccade amplitude fluctuation than children aged 10–11 years (p's < 0.01, adjusted); but there were no significant differences in saccade amplitude fluctuation between children aged 7 and 8 years, between children aged 8 and 9 years, and among children aged 9–11years (p's > 0.05, adjusted).

Total time of naming

Main effects

The main effects of gender, age and task-type were significant (Task-type: $F = 1440.33$, $p < 1 \times 10^{-4}$, $\eta^2 = 0.59$; Gender: $F = 4.74$, $p = 0.03$, $\eta^2 = 0.007$; Age: $F = 27.82$, $p < 1 \times 10^{-4}$, $\eta^2 = 0.14$).

Interaction effects

There were significant interaction effects between age and task-type ($F = 2.37$, $p = 0.01$, $\eta^2 = 0.009$), and between gender and task-type ($F = 7.04$, $p < 1 \times 10^{-4}$, $\eta^2 = 0.007$).

Post-hoc *multiple comparisons*

Figure 8 summarized the results of *post-hoc* multiple comparisons and indicated that: (1) For males (see **Figure 8A**), task N-number required shorter total time of naming than other three tasks (p's < 1×10^{-4}, adjusted); task N-character required shorter total time of naming than tasks N-object and N-color (p's < 1×10^{-4}, adjusted); and task N-object required shorter total time of naming than task N-color ($p < 0.01$, adjusted); (2) For females (see **Figure 8A**), task N-number required shorter total time of naming than other three tasks (p's < 1×10^{-4}, adjusted); task N-character required shorter total time of naming than tasks N-object and N-color (p's < 1×10^{-4}, adjusted); but there was no significant difference in total time of naming between tasks N-object and N-color ($w = 8845.0$, $p = 0.09$, $r = 0.03$, adjusted); (3) As shown in **Figure 8B**, males required longer total time of naming in tasks N-object and N-color than females (p's < 0.05), but there were no significant gender difference in total time of naming in tasks N-number and N-character (p's > 0.05); (4) For children aged 7–8 years (see **Figure 8B**), task N-number required shorter total time of naming than other three tasks (p's < 1×10^{-4}, adjusted), task N-character required shorter total time of naming than that tasks N-object and N-color (p's < 1×10^{-4}, adjusted), task N-object required shorter total time of naming task N-color (p's < 0.05, adjusted); (5) For children aged 9–11 years (see **Figure 8B**), task N-number required shorter total time of naming than other three tasks (p's < 1×10^{-4}, adjusted), task N-character required shorter total time of naming than that tasks N-object and N-color (p's < 1×10^{-4}, adjusted), but there was no significant difference in total time of naming between tasks N-object and N-color (p's > 0.05, adjusted); (6) For task N-number (see **Figure 8B**), children aged 7 years required longer total time of naming than children aged 8–11 years (p's < 1×10^{-4}, adjusted); children aged 8 years required longer total time of naming than children aged 9–11 years (p's < 0.01, adjusted);

children aged 9 or 10 years required longer total time of naming than children aged 11 years (p's < 0.001, adjusted); but there was no significant difference between children aged 9 and 10 years (p's > 0.05, adjusted); (7) For task N-character (see Figure 8B), children aged 7 years required longer total time of naming than children aged 9–11 years (p's < 0.05, adjusted); children aged 8 years required longer total time of naming than children aged 10–11 years (p's < 1×10^{-4}, adjusted); children aged 9 years required longer total time of naming than children aged 10–11 years (p's < 0.05, adjusted); but there were no significant differences between children aged 7 and 8 years, between children aged 8 and 9 years, and between children aged 9 and 10 years (p's > 0.05, adjusted); (8) For task N-object (see Figure 8B), children aged 7 years required longer total time of naming than children aged 9–11 years (p's < 0.05, adjusted); children aged 8 years required longer total time of naming than children aged 10–11 years (p's < 0.001, adjusted); children aged 9 or 10 years required longer total time of naming than children aged 11 years (p's < 0.05, adjusted); but there were no significant differences between children aged 7 and 8 years, between children aged 8 and 9 years, and between children aged 9 and 10 years (p's > 0.05, adjusted); (9) For task N-color (see Figure 8B), children aged 7 years required longer total time of naming than children aged 9–11 years (p's < 0.001, adjusted); children aged 8 years required longer total time of naming than children aged 9–11 years (p's < 0.05, adjusted); children aged 9 years required longer total time of naming than children aged 11 years ($w = 5068.5$, $p = 0.004$, $r = 0.22$, adjusted); but there were no significant differences between children aged 7 and 8 years, between children aged 9 and 10 years, and between children aged 10 and 11 years (p's > 0.05, adjusted).

Correlation among eye-movement measures

If there was a significant correlation between two of the eight eye movement measurements, there were $C_8^2 = 28$ possible pairs of two measures, and 28 different correlation coefficients between each pair of two measures. We artificially defined the sequence number for each of the 28 correlation coefficients to better highlight how they differ across the four RAN challenges (see Table 2 for the matching relationship between the sequence number and a correlation coefficient). Figure 9 provided a summary of our findings and displayed the correlation coefficients across four RAN tasks. Pairs having weak correlation in a RAN task were not plotted in Figure 9.

It can be easily seen from Figure 9 that: (i) There were significant positive correlation in 12 pairs of measures across all RAN tasks; (ii) There were significant negative correlation in 10 pairs of measures across all RAN tasks; (iii) There were no significant correlation in one pair (No. 27, corresponding to the pair of regression counts and saccade amplitude fluctuation) across all RAN tasks; and (iv) There were significant correlation in two pairs (Nos. 8 and 11, corresponding to the pair of average fixation duration and fixation counts, and the pair of average fixation duration and regression counts, respectively) in tasks N-number and N-character, but there were no significant correlation in other tasks.

As noted, the tasks N-number and N-character were alphanumeric RAN tasks, whereas the tasks N-object and N-color were non-alphanumeric ones. It would be interesting to see if there were any correlation differences between alphanumeric and non-alphanumeric RAN tasks. Our findings (see Figure 9) showed that while the correlations between 6 pairs of measures (i.e., Nos. 5, 7, 16, 21, and 28 in Table 2) in alphanumeric RAN tasks may be higher than those in non-alphanumeric RAN tasks, the correlations between 12 pairs of measures (i.e., Nos. 1, 2, 3, 4, 8, 11, 14, 17, 18, 19, and 26 in Table 2) may be lower in alphanumeric RAN tasks than in non-alphanumeric RAN tasks.

Some important correlations between pairs of measures were listed as follows: (i) The correlations (corresponding to Nos. 1, 2, 4, 5 and 6 in Figure 9) between total time of naming and five measures were positive (r between 0.56 and 0.87); the correlation (corresponding to No. 3 Figure 9) between total time of naming and one measures was negative (r between −0.68 and −0.58); and the correlation (corresponding to No. 7 in Figure 9)

TABLE 2 Sequence number of 28 correlation coefficients between two measures.

No.	Correlation coefficient	No.	Correlation coefficient	No.	Correlation coefficient	No.	Correlation coefficient
1	R (TTN, AFD)	8	R (AFD, FC)	15	R (FC, SC)	22	R (ASA, SAF)
2	R (TTN, FC)	9	R (AFD, ASA)	16	R (FC, RC)	23	R (SC, RC)
3	R (TTN, ASA)	10	R (AFD, SC)	17	R (FC, FDF)	24	R (SC, FDF)
4	R (TTN, SC)	11	R (AFD, RC)	18	R (FC, SAF)	25	R (SC, SAF)
5	R (TTN, RC)	12	R (AFD, FDF)	19	R (ASA, SC),	26	R (RC, FDF)
6	R (TTN, FDF)	13	R (AFD, SAF)	20	R (ASA, RC)	27	R (RC, SAF)
7	R (TTN, SAF)	14	R (FC, ASA)	21	R (ASA, FDF)	28	R (FDF, SAF)

R(x,y) is the correlation coefficient between measures x and y; FC, fixation counts; SC, saccade counts; RC, regression counts; AFD, average fixation duration; FDF, fixation duration fluctuation; ASA, average saccade amplitude; SAF, saccade amplitude fluctuation; TTN, total time of naming.

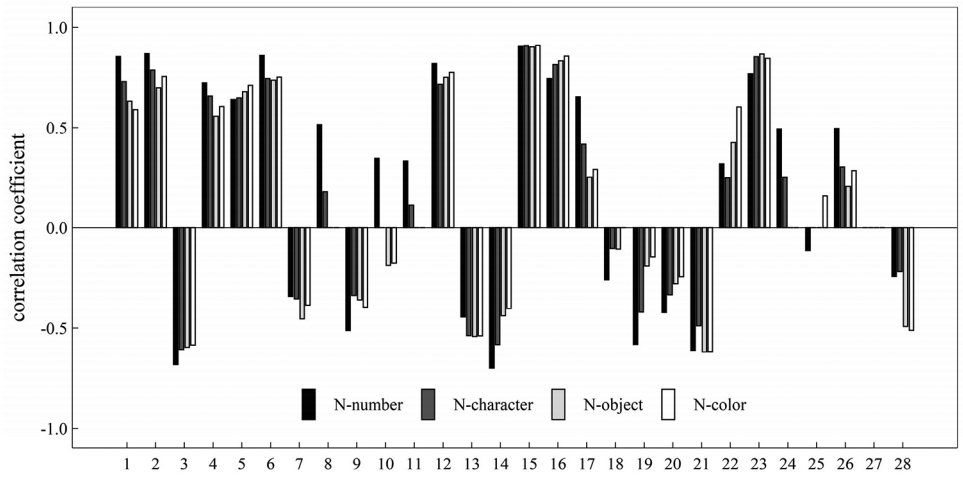


FIGURE 9
Plot showed correlation coefficients between two eye-movement measures during four tasks, in which the sequence number was defined in Table 2.

between total time of naming and saccade amplitude fluctuation (r between −0.45 and −0.34); (ii) The correlations (corresponding to Nos. 15, 16 and 23 in **Figure 9**) among three count-related measures (i.e., fixation counts, saccade counts, regression counts) were significantly positive (r between 0.74 and 0.91); (iii) The correlation (corresponding to No. 28 in **Figure 9**) between two fluctuation-related measures (i.e., fixation duration fluctuation, saccade amplitude fluctuation) was significantly negative in all tasks; and the correlation coefficients in alphanumeric RAN tasks ($r = -0.24$ and $r = -0.22$) were higher than that in non-alphanumeric tasks ($r = -0.49$ and $r = -0.5$); (iv) The correlation (corresponding to No. 12 in **Figure 9**) between two fixation-duration-related measures (i.e., average fixation duration and fixation duration fluctuation) was significantly positive in all tasks (r between 0.72 and 0.82); and the correlation coefficient in task N-number was higher than other tasks; and (v) The correlation (corresponding to No. 22 in **Figure 9**) between two saccade-amplitude-related measures (i.e., average saccade amplitude and saccade amplitude fluctuation) was significantly positive in all tasks (r between 0.25 and 0.60); and the correlation coefficients between them in alphanumeric RAN tasks ($r = 0.32$ and $r = 0.25$) were lower than that in non-alphanumeric counterparts ($r = 0.42$ and $r = 0.60$).

Correlation between eye movements and attention-related skills

We asked the participating kids to complete the RAN tasks and the NCT test (28) together in order to demonstrate the relationship between RAN skills and attention-related abilities. Eight eye-movement measures were selected for RAN tasks, while three attention-related skills indicated for the NCT test were derived using Equations (3–5). The relationship between the measures of the two tasks was then examined. **Tables 3–6** summarized our results and showed that: (i) Significant correlations were found between three attention-related skills and eight eye-movement measures (p's < 0.05); (ii) In general, two parameters (i.e., speed of cognitive processing and selective attention) significantly correlated negatively with six eye-movement measures (i.e., fixation counts, saccade counts, regression counts, average fixation duration, fixation duration fluctuation, and total time of naming) (p's < 0.05), but

TABLE 3 Correlation between number-cancellation-related skills and eye-movement measures in task N-naming.

Eye-movement measures	Number cancellation test		
	Speed of cognitive processing	Selective attention	Averaged time of circlings
Fixation counts	−0.25****	−0.26****	0.19***
Saccade counts	−0.19***	−0.18***	0.14**
Regression counts	−0.24****	−0.26****	0.18***
Average fixation duration	−0.22****	−0.28****	0.17****
Fixation duration fluctuation	−0.28****	−0.33****	0.27****
Average saccade amplitude	0.20****	0.27****	−0.18***
Saccade amplitude fluctuation	0.09	0.11*	−0.11*
Total time of naming	−0.34****	−0.38****	0.25****

*p < 0.05.
**p < 0.01.
***p < 0.001.
****p < 0.0001.

TABLE 4 Correlation between number-cancellation-related skills and eye-movement measures in task N-character.

Eye-movement measures	Number cancellation test		
	Speed of cognitive processing	Selective attention	Averaged time of circlings
Fixation counts	−0.17***	−0.23****	0.16***
Saccade counts	−0.10	−0.16**	0.09
Regression counts	−0.10*	−0.16**	0.14**
Average fixation duration	−0.34****	−0.37****	0.28****
Fixation duration fluctuation	−0.31****	−0.36****	0.31****
Average saccade amplitude	0.32****	0.36****	−0.27****
Saccade amplitude fluctuation	0.13*	0.10*	−0.09
Total time of naming	−0.32****	−0.39****	0.30****

*p<0.05; **p<0.01; ***p<0.001; ****p<0.0001.

TABLE 6 Correlation between number-cancellation-related skills and eye-movement measures in task N-color.

Eye-movement measures	Number cancellation test		
	Speed of cognitive processing	Selective attention	Averaged time of circlings
Fixation counts	0.20****	−0.24****	0.17***
Saccade counts	−0.14**	−0.19***	0.10*
Regression counts	−0.21****	−0.28****	0.14**
Average fixation duration	−0.23****	−0.26****	0.18***
Fixation duration fluctuation	−0.23****	−0.27****	0.19***
Average saccade amplitude	0.17***	0.20****	−0.16**
Saccade amplitude fluctuation	0.10*	0.11*	−0.13*
Total time of naming	−0.30****	−0.36****	0.24****

*p<0.05; **p<0.01; ***p<0.001; ****p<0.0001.

significantly correlated positively with two other eye-movement measures (i.e., average saccade amplitude and saccade amplitude fluctuation) (p's < 0.05); (iii) Basically, averaged time of circlings significantly negatively correlated with two eye-movement measures (i.e., average saccade amplitude and saccade amplitude fluctuation) (p's < 0.05), but significantly positively correlated with six eye-movement measures (i.e., fixation counts, saccade counts, regression counts, average fixation duration, fixation duration fluctuation, total time of naming) (p's < 0.05).

TABLE 5 Correlation between number-cancellation-related skills and eye-movement measures in task N-object.

Eye-movement measures	Number cancellation test		
	Speed of cognitive processing	Selective attention	Averaged time of circlings
Fixation counts	−0.26****	−0.26****	0.19***
Saccade counts	−0.19***	−0.18***	0.14**
Regression counts	−0.24*	−0.26****	0.19***
Average fixation duration	−0.22****	−0.28****	0.17***
Fixation duration fluctuation	−0.26****	−0.32****	0.15**
Average saccade amplitude	0.20****	0.27****	−0.18***
Saccade amplitude fluctuation	0.13**	0.19***	−0.09
Total time of naming	−0.34****	−0.38****	0.25****

*p<0.05; **p<0.01; ***p<0.001; ****p<0.0001.

Discussion

RAN tests have been utilized extensively in the examination of numerous cognitive and neurobiological disorders (13–19), as well as in studies on reading behavior and child development in healthy children (2–8). This article aimed to offer some new perspectives on RAN understanding by the eye tracking method, which is able to capture the visual and cognitive characteristics of RAN. To resolve issues with earlier research, this article involved measures of eye movements during RAN tests for Chinese children aged 7–11 years. This study's primary goal was to determine how age, gender, and task type affected measures of eye movements made by Chinese children aged 7–11 during RAN. Eight parameters (including two fluctuation parameters proposed in this article) were designed to measure eye movements.

First of all, our findings showed that all eye-movement measures had the main effects of task-type and age, but only five of them (i.e., fixation counts, saccade counts, regression counts, average saccade amplitude, total time of naming) had the main effect of gender. Additionally, three measures (i.e., saccade counts, regression counts, and total time of naming) had interaction effects between task-type and gender, and between age and gender; two measures (i.e., fixation counts and fixation duration fluctuation) had an interaction effect between gender and task-type, only; one measure (i.e., average saccade amplitude) had an interaction effect between age and gender, only.

Our results confirm that almost all eye-movement measures initially develop quickly before the age of 9, and then enter a relatively sluggish development phase without a ceiling or floor impact for kids between the ages of 7 and 11. In

particular, children aged 7–11 years can be roughly separated into three stages for the development of RAN-related skills: 7–8 years old, 9–10 years old, and 11 years old. This is in line with the widely-used three-stage model in Chinese educational practices, where the low stage corresponds to Grades 1–2 (corresponding to children aged 7–8), the middle stage corresponds to Grades 3–4 (corresponding to children aged 9–10), and the high stage corresponds to Grades 5–6 (corresponding to children aged 11–12 years old).

It is well established that females have a faster cognitive and social development up to the end of adolescence than males of the same age. Our results can likewise be used to support this conclusion. Indeed, we demonstrated that females outperformed males in RAN tasks as measured by five eye-movement parameters: fixation counts, saccade counts, regression counts, average saccade amplitude, and total time of naming. Additionally, the fixation duration fluctuation in tasks N-number and N-object was significantly different between males and females, as shown in **Figure 5A**. We also discovered that there was no gender difference in saccade amplitude fluctuation. Furthermore, almost all measures (except average fixation duration) had an interaction effect between gender and task-type, or between gender and age. These findings might offer some new perspectives on how gender affects children's cognitive development and neurological and psychiatric disorders connected to RAN.

RAN tasks can generally be grouped into two categories, i.e., alphanumeric and non-alphanumeric RAN tasks. Naming of numbers, letters, words, or Chinese characters are examples of alphanumeric RAN tasks, whereas naming of colors or objects are examples of non-alphanumeric RAN tasks. Alphanumeric RAN might necessitate primarily phonological processing because the associated linguistic codes of these stimuli are easily available at the surface level (31). While non-alphanumeric RAN appears to necessitate extra steps and need conceptual processing to establish meaning and then the appropriate name code, before phonological processing results in articulating a response (31). This suggests that non-alphanumeric RAN is more complex than alphanumeric RAN and hence generally takes more mental effort. This inference may be supported by our findings that in comparison to alphanumeric RAN tasks, non-alphanumeric RAN tasks generally required more fixation counts, saccade counts, regression counts, smaller average saccade amplitude, smaller fixation duration fluctuation and saccade amplitude fluctuation, and longer average fixation duration and total time of naming. This suggests that eight parameters used in this article may inherit key features of eye movements during RAN tasks, and more importantly, may be able to distinguish between alphanumeric and non-alphanumeric RAN.

The fluctuation, defined as the sum of absolute first-order differences in Equation (1) or (2), is taken from an earlier work (32), which can be applied to measure the oscillation, variability, and unpredictability of time series in a variety of nonlinear physical systems. According to the work of Paulson (33), eye movement can be considered as a self-similar nonlinear dynamic process, and thus can be measured using the fluctuation. The eye-movement changes associated with age or task-type may be observed by both fluctuation parameters, as shown in **Figures 5**, **7**. Additionally, the gender differences in tasks N-number and N-object may be reflected in the fixation duration fluctuation, but not in the saccade amplitude fluctuation.

Rather than emphasizing the superiority of the fluctuation definition, this study attempted to employ fluctuation as a supplement to the original measures of eye movements. In fact, we also examined how well the standard deviation performed as a definition of fluctuation and discovered that it had a strong correlation (r's ≥ 0.85) with our technique (i.e., the sum of absolute first-order differences) for all RAN tasks. It can be easily seen from **Figure 9** that the correlations among eight measures of eye movements may be positive, negative and null. Some correlations between two measures may decrease with the task complexity, but some of them may increase with the task complexity. While, two pairs of measures had significant correlations in tasks N-number and N-character, only, but not in other tasks. The principle of assessing total time of naming remains the same, even though it was measured in this study using an eye tracking method that differed from the traditional behavioral methodology. It should be noted that total time of naming was well-documented in traditional RAN measures. It was fascinating to see if there were any relationships between total time of naming and other RAN measures. Our findings showed that there were highly correlated between total time of naming and other six measures ($|r|$ between 0.56 and 0.87); and there was moderately correlated between total time of naming and saccade amplitude fluctuation (r between −0.45 and −0.34). This partially supports the effectiveness and feasibility of the eye-movement measures suggested in the current article.

The correlations among three count-related measures (i.e., fixation counts, saccade counts, regression counts) were significantly positive (r between 0.74 and 0.91). This implies that internal consistency among three count-related measures was high. Additionally, the correlation between two fixation-duration-related measures (i.e., average fixation duration and fixation duration fluctuation) was highly positive in all tasks (r between 0.72 and 0.82).

Findings (see **Tables 4**–**6**) showed that there were moderate correlations between eight eye-movement measures and three attention-related skills (p's < 0.05). This supports the weak correlation but strong dependence between RAN and NCT tasks. The NCT was applied to measure attention-related abilities, which involved cognitive skills in selective and sustained attention, motor inhibition, visuospatial search, planning, organizing, psychomotor speed, intact visual-

perception abilities, fine motor coordination, and sensory motor integration (28). On the other hand, both RAN and NCT might share several visual and cognitive neural circuits because they both need a similar "visual scanning" processing. Additionally, NCT and RAN are associated with "writing" and "reading", respectively. Hence, it is hypothesized that RAN, in combination with NCT, may bring some new insights into the understanding of developmental dyslexia and learning disabilities (34).

It is natural to adopt a Chinese adaptation of RAN in the understanding of developmental dyslexia in Chinese. The difference between both the original RAN and Chinese version is due to the features of Chinese characters: (i) Chinese characters not only have shape and sound attributes like English letters, but also represent meaning; (ii) Chinese characters have no clear form-to-sound conversion rules, so readers need to remember the pronunciation of Chinese characters; and (iii) The visual complexity of Chinese characters are much higher than that of English letters. Consequently, compared with the original RAN, the Chinese adaptation may have higher cognitive complexity, and thus activate a wider range of brain regions (35, 36). To extend the application of RAN to developmental dyslexia in Chinese, we suggested a Chinese adaptation of RAN (i.e., the C-RAN) by substituting Chinese characters (highly-frequently used) for English letters. We expect that the C-RAN should be more suitable in the evaluation of developmental dyslexia in Chinese than the original RAN.

Future study will explore a few issues. First, we will explore gender differences in eye-movement measures taken during RAN throughout the lifespan (especially for children older than 11-years). Second, we will examine the possibility that the apparent gender disparities in RAN-related neurological and mental disorders are due to age-related changes in RAN skills. Third, we will detect whether the diagnostic criteria for RAN-related neurological and psychiatric disorders might be biased or poorly specified for one gender and/or grade group. Finally, we will seek a new definition of fluctuation with increased performance.

Conclusion

This article investigated eight measures of eye movements during RAN tests for Chinese children aged 7–11 years. First of all, we showed that all eye-movement measures had the main effects of task-type and age, but only five of them had the main effect of gender, with interaction effect between task-type and gender or/and between age and gender. Second, we found that almost all eye-movement measures initially developed quickly before the age of 9, and then entered a relatively sluggish development phase. Third, we confirmed that non-alphanumeric RAN tasks generally required higher mental load than alphanumeric ones. Fourth, we showed that there were significant correlations between total time of naming and other eye-movement measures. Finally, we found significant relationships between eight eye-movement measures and three attention-related skills. Because eye tracking is a fundamental tool in psychological research, the technique suggested has the potential to be used in a wide range of applications (e.g., developmental dyslexia).

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by The Research Ethics Committee at Southeast University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

DY and HW developed the idea for the study. HW, FL and YD collected the data. HW and DY did the analyses. DY wrote the paper. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the National Natural Science Foundation of China under Grant Nos. 62073077 and 61673113, and by the Guangdong Key Project in "Development of new tools for diagnosis and treatment of Autism " (Grant No. 2018B030335001).

Acknowledgment

The authors would like to thank all participating children.

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. Hjetland HN, Brinchmann EI, Scherer R, Melby-Lervåg M. Preschool predictors of later reading comprehension ability: a systematic review. *Campbell Syst Rev.* (2017) 13(1):1–155. doi: 10.4073/csr.2017.14

2. Kail R, Hall LK, Caskey BJ. Processing speed, exposure to print, and naming speed. *Appl Psycholinguist.* (1999) 20:303–14. doi: 10.1017/S0142716499002076

3. Stainthorp RW, Stuart M, Powell D, Quinlan P, Garwood H. Visual processing deficits in children with slow RAN performance. *Sci Stud Read.* (2010) 14:266–92. doi: 10.1080/10888431003724070

4. Georgiou G, Parilla R, Cui Y, Papadopoulos TC. Why is rapid automatized naming related to reading? *J Exp Child Psychol.* (2013) 115:218–25. doi: 10.1016/j.jecp.2012.10.015

5. Bowers PG, Wolf M. Theoretical links among naming speed, precise timing mechanisms and orthographic skill in dyslexia. *Read Writ.* (1993) 5:69–85. doi: 10.1007/BF01026919

6. Decker SL, Roberts AM, Englund JA. Cognitive predictors of rapid picture naming. *Learn Individ Differ.* (2013) 25:141–9. doi: 10.1016/j.lindif.2013.03.009

7. Powell D, Stainthorp R, Stuart M. Deficits in orthographic knowledge in children poor at rapid automatized naming (RAN) tasks? *Sci Stud Read.* (2014) 18:192–207. doi: 10.1080/10888438.2013.862249

8. Wiig EH, Zureich P, Chan HNH. A clinical rationale for assessing rapid automatized naming in children with language disorders. *J Learn Disabil.* (2000) 33(4):359–74. doi: 10.1177/002221940003300407

9. Akhand O, Rizzo JR, Rucker JC, Hasanaj L, Balcer LJ. History and future directions of vision testing in head trauma. *J Neuroophthalmol.* (2019) 39(1):68–81. doi: 10.1097/WNO.0000000000000726

10. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (5th ed.).* Arlington: American Psychiatric Publishing (2013).

11. Araújo S, Huettig F, Meyer A. What underlies the deficit in rapid automatized naming (RAN) in adults with dyslexia? Evidence from eye movements. *Sci Stud Read.* (2021) 25(6):534–49. doi: 10.1080/10888438.2020.1867863

12. Ullman MT, Earle FS, Walenski M, Janacsek K. The neurocognition of developmental disorders of language. *Annu Rev Psychol.* (2020) 71:389. doi: 10.1146/annurev-psych-122216-011555

13. Goswami U. Sensory theories of developmental dyslexia: three challenges for research. *Nat Rev Neurosci.* (2015) 16(1):43–54. doi: 10.1038/nrn3836

14. Åvall M, Wolff U, Gustafsson JE. Rapid automatized naming in a developmental perspective between ages 4 and 10. *Dyslexia.* (2019) 25(4):360–73. doi: 10.1002/dys.1631

15. Georgiou G, Parrila R. What mechanism underlies the rapid automatized naming-reading relation? *J Exp Child Psychol.* (2020) 194:104840. doi: 10.1016/j.jecp.2020.104840

16. Snowling MJ, Melby-Lervag M. Oral language deficits in familial dyslexia: a meta-analysis and review. *Psychol Bull.* (2016) 142(5):498–545. doi: 10.1037/bul0000037

17. Tannock R, Martinussen R, Frijters J. Naming speed performance and stimulant effects indicate effortful, semantic processing deficits. *J Abnorm Child Psychol.* (2000) 28(3):237. doi: 10.1023/A:1005192220001

18. Zhao J, Chen S, Tong X, Yi L. Advantage in character recognition among Chinese preschool children with autism spectrum disorder. *J Autism Dev Disord.* (2019) 49(12):4929–40. doi: 10.1007/s10803-019-04202-x

19. Hogan-Brown AL, Hoedemaker RS, Gordon PC, Losh M. Eye-voice span during rapid automatized naming: evidence of reduced automaticity in individuals with autism spectrum disorder and their siblings. *J Neurodev Disord.* (2014) 6(1):33. doi: 10.1186/1866-1955-6-33

20. Araújo S, Faísca L. A meta-analytic review of naming-speed deficits in developmental dyslexia. *Sci Stud Read.* (2019) 23(5):349–68. doi: 10.1080/10888438.2019.1572758

21. Georgiou G, Parrila R, Kirby JR, Stephenson K. Rapid naming components and their relationship with phonological awareness, orthographic knowledge, speed of processing, and different reading outcomes. *Sci Stud Read.* (2008) 12(4):325–50. doi: 10.1080/10888430802378518

22. Mcweeny S, Choi SJ, Choe J, Latourrette A, Norton E. Rapid automatized naming (RAN) as a kindergarten predictor of future reading: a systematic review and meta-analysis (2021).

23. Norton ES, Wolf M. Rapid automatized naming (RAN) and reading fluency: implications for understanding and treatment of reading disabilities. *Annu Rev Psychol.* (2012) 63(1):427–52. doi: 10.1146/annurev-psych-120710-100431

24. Wolf M, Bowers PG, Biddle K. Naming-speed processes, timing, and reading: a conceptual review. *J Learn Disabil.* (2000) 33(4):387–407. doi: 10.1177/002221940003300409

25. Jones MW, Obregon M, Kelly ML, Branigan HP. Elucidating the component processes involved in dyslexic and non-dyslexic reading fluency: an eye-tracking study. *Cognition.* (2008) 109(3):389–407. doi: 10.1016/j.cognition.2008.10.005

26. Weeks S, Atlas A. Clinical audit for occupational therapy intervention for children with autism spectrum disorder: sampling steps and sample size calculation. *BMC Res Notes.* (2015) 8:282. doi: 10.1186/s13104-015-1247-0

27. Wang H, Liu F, Dong Y, Yu D. Entropy of eye movement during rapid automatized naming. *Front Hum Neurosci.* (2022) 16:945406. doi: 10.3389/fnhum.2022.945406

28. Xie Y, Wang H, Chen Y, Liu F, Yao M, Zhang L, et al. Establishing normative data for the number cancelation test among children in kindergartens and primary schools in China. *Front Psychiatry.* (2022) 13:788825. doi: 10.3389/fpsyt.2022.788825

29. Wobbrock JO, Findlater L, Gergle D, Higgins JJ. *The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. Proceedings of the ACM conference on human factors in computing systems (CHI '11)*; May 7–12, 2011; Vancouver, British Columbia. New York: ACM Press (2011). p. 143–6.

30. Elkin LA, Kay M, Higgins JJ, Wobbrock JO. *An aligned rank transform procedure for multifactor contrast tests. Proceedings of the ACM symposium on user interface software and technology (UIST '21)*; October 10–14, 2021; Virtual Event. New York: ACM Press (2021). p. 754–68.

31. Donker M, Kroesbergen E, Slot E, Van Viersen S, De Bree E. Alphanumeric and non-alphanumeric rapid automatized naming in children with reading and/or spelling difficulties and mathematical difficulties. *Learn Individ Differ.* (2016) 47:80–7. doi: 10.1016/j.lindif.2015.12.011

32. Liu Y, Zhou W, Yuan Q, Chen S. Automatic seizure detection using wavelet transform and SVM in long-term intracranial EEG. *IEEE Trans Neural Syst Rehabilitation Eng.* (2012) 20(6):749–55. doi: 10.1109/TNSRE.2012.2206054

33. Paulson EJ. Viewing eye movements during reading through the lens of chaos theory: how reading is like the weather. *Read Res Q.* (2005) 40(3):338–58. doi: 10.1598/RRQ.40.3.3

34. Benjamin Z, Black LI, Maenner MJ, Schieve LA, Danielson ML, Bitsko RH, et al. Prevalence and trends of developmental disabilities among children in the United States: 2009-2017. *Pediatrics.* (2019) 144(4):e20190811. doi: 10.1542/peds.2019-0811

35. Liao CH, Deng CP, Hamilton J, Lee CSC, Wei W, Georgiou GK. The role of rapid naming in reading development and dyslexia in Chinese. *J Exp Child Psychol.* (2015) 130:106–22. doi: 10.1016/j.jecp.2014.10.002

36. Peng P, Wang CC, Tao S, Sun CY. The deficit profiles of Chinese children with reading difficulties: a meta-analysis. *Educ Psychol Rev.* (2017) 29(3):513–64. doi: 10.1007/s10648-016-9366-2

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 07 December 2022
DOI 10.3389/fpubh.2022.1054023

OPEN ACCESS

EDITED BY
Pei Wang,
Fudan University, China

REVIEWED BY
Simone Battaglia,
University of Turin, Italy
Zhishan Hu,
Shanghai Jiao Tong University, China

*CORRESPONDENCE
Qun Ye
qun.ye@zjnu.edu.cn

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

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

RECEIVED 26 September 2022
ACCEPTED 18 November 2022
PUBLISHED 07 December 2022

CITATION
Cao Z, Huang Y, Song X and Ye Q (2022) Development and validation of children's mind wandering scales. *Front. Public Health* 10:1054023. doi: 10.3389/fpubh.2022.1054023

Development and validation of children's mind wandering scales

Zhiwei Cao[1,2†], Ying Huang[1,3†], Xiaolan Song[1,3] and Qun Ye[1,3*]

[1]Department of Psychology, College of Teacher Education, Zhejiang Normal University, Jinhua, China, [2]School of Pharmacy & School of Biological and Food Engineering, Changzhou University, Changzhou, China, [3]Key Laboratory of Intelligent Education Technology and Application of Zhejiang Province, Zhejiang Normal University, Jinhua, China

Introduction: Mind wandering is generally considered an endogenous mental state that arises spontaneously, which is one of the most common experiences of consciousness and typically occurs at a significant cost to mental health and behavioral performance. Previous studies have shown that mind wandering appears to be a stable trait and can be assessed reliably in adults. Surprisingly little, however, is known about how to measure the frequency of mind wandering in children, given that children can accurately introspect their experiences. The present studies aimed to develop the Frequency of Children's Mind Wandering Scale (CMWS-F) and the Context of Children's Mind Wandering Scale (CMWS-C) to assess the frequency of mind wandering and contexts in which mind wandering occurs for children aged 8 to 11 years.

Methods: The exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were used to develop the CMWS-F and CMWS-C. To further assess the validity of the scales, we compared the scores in CMWS-F/CMWS-C and the frequencies of probe-caught mind wandering in the typical tasks.

Results: In study 1a, the EFA (n = 292) and CFA (n = 346) showed that attentional failure and spontaneous thinking were the two main dimensions of CMWS-F. In study 1b, contexts about mind wandering in children could be divided into high-demand and low-demand contexts using EFA (n = 258) and CFA (n = 347). Study 2 showed moderate positive correlations between the frequencies of probe-caught mind wandering in the tasks and the scores in the scales.

Discussion: The results showed that scores on the two scales could predict the performance on the experimental tasks and further demonstrated empirical validity of the CMWS-F and CMWS-C scales. Taken together, the results of the current studies provided preliminary evidence for the validity and reliability of CMWS-F and CMWS-C in children, which can be used as a reference to balance its downsides and productive aspects of mind wandering.

KEYWORDS

mind wandering, task-unrelated thoughts, self-generated thoughts, attentional failure, spontaneous thinking, context

Introduction

Mind wandering emphasizes a state of consciousness that arises spontaneously in the waking state, when the content of an individual's consciousness is not determined by subjective will but is occupied by endogenous mental representations; this state can occur either in goal-directed tasks or in resting states (1). Adult studies have found that the frequency of mind wandering in daily life is as high as 46.9%, and at least 30% of the sample answered that mind wandering occurs in almost all activities (2). Furthermore, the frequency of mind wandering in children has also been found to reach 20 to 33% (3). Mind wandering as a trait has significant externalities in individual differences (1). For instance, individuals with a low frequency of mind wandering show that they are focused and have good control over their thinking activities, while individuals with a high frequency show that they are often distracted during the task and cannot maintain coherent thinking, thus affecting the performance of the task. In addition, the association between the frequency of mind wandering and individuals with cognitive or emotional disorders is also receiving increasing attention (4). Thus, it is of great theoretical and practical importance to study mind wandering from the perspective of individual difference (5).

Mind wandering has been extensively studied in adults (6–8), but relatively limited research has been conducted in children (3, 9–13). In fact, mind wandering has been shown to be reliably measured in children (9), and the study of its functions will help to understand the role it plays in children's cognitive and social development. In the present paper, we aim to investigate the characteristics of children's mind wandering by developing psychometric scales (Study 1a and Study 1b) and applying experience sampling method in the laboratory tasks (Study 2). In the section that follow, we review recent advances on theoretical basis and related scales of mind wandering.

Current Concerns Theory tells that once a goal is established, it becomes a current concern event (14, 15), and the individual's cognitive system maintains a high level of accessibility to environmental cues that make it easy to accomplish the goal, and facilitates behavior by enhancing the accessibility of goal-related stimuli. This accessibility will continue until the goal is achieved or abandoned. Assuming that an individual's goals have a hierarchical structure, mind wandering may arise spontaneously because alternative implicit goals are automatically activated, only that the individual fails to realize it at the time. 1 thus argue that mind wandering can be incorporated into the executive control model of attention, treating the occurrence of mind wandering as goal-driven processing that simply shifts executive control from the task at hand to the processing of internal goals, and arguing that this process is cognitively resource-intensive, which further leads to the Perceptual Decoupling Hypothesis (1, 16, 17). Several studies of event-related potentials (ERP) provide support for this hypothesis. Some studies found that participants' P300 amplitude during self-reported mind wandering was reduced relative to the events in the task (6, 18–20). Because the P300 can be treated as an index of executive resources, the decreased P300 amplitude during mind wandering indicates that executive resources have been withdrawn, at least partly, from the primary task and are presumably directed toward task-unrelated thoughts (21).

However, executive failure hypothesis argues that mind wandering arises as a result of executive control failure and that this process is not cognitively resource-intensive (22–24). First, it is the default state to constantly assess the gap between the ideal state and the present state and thus to continuously generate spontaneous thoughts beyond the level of awareness. Second, mind wandering arises during the task when controlled processing is not sufficient to handle the interference generated by spontaneous thought. Studies have found that the default mode network (DMN) and executive control network (ECN) are associated with mind wandering in adults. For instance, neuroimaging studies have shown that the DMN and ECN are activated during mind wandering, and the neural activations in both networks are strongest when subjects lack meta-awareness (25–27). Similarly, evidences suggest that mind wandering in children is related to specific executive functions like inhibition and set shifting/switching (3 which could be due to dysregulation of specific brain regions, like DMN. Recent studies on healthy individuals have shown how dysregulation of action control and inhibition capacity impairs task performance that involves inhibition of actions, indicating that children may have difficulty in switching off the DMN during some tedious tasks which require focused attention and inhibit task-unrelated thoughts (28, 29).

As can be seen from the above theories, researchers have accepted that mind wandering was determined by a combination of automatically generated thoughts in response to environmental or mental cues and the ability of the executive control system to deal with disturbing thoughts (1, 24), but most of the existing scales have been designed based on one perspective only, either focusing on the individual's ability to generate spontaneous thoughts or on the individual's ability to execute control during the task, thus lacking comprehensiveness. Study1a aims to explore the dimension of the generation of children's mind wandering, and we hypothesize that both the spontaneity of individual thinking and the executive control processes are involved from the perspective of scale development.

Based on the opposing hypotheses that executive control inhibits the occurrence of mind wandering and executive control supports the continuity of mind wandering, several studies have shown that individuals with high working memory capacity (WMC) report less mind wandering in the sustained attention response task (SART) (22, 30), reading task (23, 31), and memory span task (32–34) relative to individuals with low WMC, and that individuals with high WMC also report less

mind wandering when they are more focused on the task using experience sampling in everyday life (35). In contrast, some studies have shown that individuals with high WMC report more mind wandering than those with low WMC in low cognitive demand visual search tasks and respiratory perception tasks (36), and similar findings have been found in choice reaction time tasks (37), where individuals with high WMC report more mind wandering about future directions. More interestingly, in these studies, when the nature of the task was divided in a different way, WMC did not predict the frequency of mind wandering, e.g., in the above-mentioned experience sampling method (35), there was no association between WMC and mind wandering frequency when the task was divided in terms of whether it was challenging or the level of effort exerted. In addition, no association was found between WMC and mind wandering frequency in high cognitive demand visual search task (36). Indeed, researchers have suggested that individuals' executive control not only inhibits the generation of mind wanderings, but also flexibly adjusts the frequency of mind wanderings according to the demands of the task load (38). Thus, Study 1b aims to explore the dimension of the generation of children's mind wandering from the context of cognitive demand.

Researchers have largely acknowledged that differences in the content and frequency of mind wandering often reflect inter-individual differences in trait level, yet there still lacks validated scale to measure the frequency of mind wandering in children. Through extensive reading of relevant literature, we found that the existing measurement tools have the following drawbacks.

First, some studies have used concepts such as daydreaming and task-unrelated thinking, which are close to mind wandering, as alternatives to study the characteristics of mind wandering in terms of form, content, and frequency. Such studies often change the connotation and extension of the concept of mind wandering, which not only lacks face validity but also makes the survey less relevant [e.g., Imaginal Processes Inventory Questionnaire, (39)].

Second, attentional failure and endogenous spontaneity as the two main features of mind wandering are often taken in specific research contexts. Some scales focus only on the interference of mind wandering with specific tasks, emphasizing the characteristics of mind wandering that are not controlled by the subjective will of individuals. These instruments include Attention Related Cognitive Errors scale (40), Cognitive Failure Questionnaire (41) and Memory Failure Scale (42), which are often used to indirectly account for the frequency of mind wandering. This type of scale focuses on the consequences of attention lapses and correlates well with behavioral task performance, and is often used as a tool to illustrate the undesirable consequences of mind wandering. However, not all mind wandering is unhelpful, and not all lapses are necessarily due to mind wandering, and researchers has also found positive implications for the existence of mind wandering (43, 44). The other part of the scale focuses on the content of spontaneously generated mental activities, such as resting state thought activities. These instruments include Automatic Thought Questionnaire (45) and Resting State Questionnaire (46). On the one hand, these scales were developed based on the content of mind activity, not just the frequency of occurrence. On the other hand, these scales focused on the spontaneity of mind activity, ignoring the important role played by attentional control.

Last but not the least, some researchers have used Mindful Attention Awareness Scale (MAAS), developed by Brown and Ryan (47), to indirectly explain the individual differences on mind wandering, based on the logic that mindfulness and mind wandering are opposing concepts. Although the close conceptual relationship between mindfulness and wind wandering, the strength of their association has been surprisingly low (48). In fact, mindfulness is considered to be a general personality trait, but mind wandering is a much more transient and fluctuating phenomenon during an ongoing task. Thus, the MAAS can be used as a reference to study an individual's level of mind wandering, but is not a substitute.

To sum up, there is no scale developed for children on the frequency of mind wandering. However, elementary school is a critical period for good behavior and a foundation for future learning development, so it is necessary to accurately screen children with excessive mind wandering and lack of mind wandering. For children with excessive mind wandering, early intervention would improve academic performance and quality of life; for children with a lack of mind wandering, a package of counseling exercises would help improve individual attention, and social and emotional functioning.

Study 1a development of the frequency of children's mind wandering scale (CMWS-F)

Item generation

We first invited 22 elementary school students for interviews, which included conceptual understanding of mind wandering and frequency of mind wandering. After that, a total of 21 relevant items were generated by combining previous relevant scales, expert opinions, and interview contents.

In order to determine the validity of the generated items, two pilot tests were conducted. The total sample included 220 children, with 94 children in the first polit test (36 boys and 58 girls) and 126 in the second test (59 boys and 67 girls). All participants are elementary school students from grades 3 to 6 in Zhejiang Province, China. The first pilot analysis revealed that there were problems such as inappropriate reverse questions, and low item-total correlations for some items. After the modification, we decided not to apply the reverse questions

and replaced the abstract concepts with life-like terms. To explore the validity of the changes, the revised scale was tested for the second time. The item analysis was good and resulted in a preliminary scale. Finally, a total of 22 items were identified (including one lie detection item, e.g., please check option 4 on this question), and all response options were on a 5-point Likert response scale (1 = strongly disagree, 2 = disagree, 3 = unsure, 4 = agree, and 5 = strongly agree).

Exploratory factor analysis (EFA)

Participants

The sample consisted of 332 children, of which 40 children were excluded due to randomly check and failed the lie detection item. The remaining 292 participants' data were analyzed (151 boys and 141 girls, $M = 10.70$ years old, $SD = 1.30$, 24.3% of third graders, 28.1% of fourth graders, 22.6% of fifth graders, 25.0% of sixth graders).

Results

The structure of CMWS-F was analyzed using principal component analysis (PCA) and Promax rotation method, and the specific statistical processing was implemented with SPSS 10.0. We first conducted item analysis to assess the item appropriateness within the scale. Inappropriate items were eliminated based on the critical ratio (CR $\geq$ 3), the item-total correlation ($r \geq 0.3$). One item was removed first as it did not meet above criteria. One main step before conducting EFA was evaluating the data appropriateness for Factor analysis using Kaiser Meyer-Olkin (KMO) and the Bartlett test of sphericity. The KMO value was 0.926, indicating that the sample size was adequate for factor analysis. The Bartlett test was significant ($p < 0.001$), supporting the argument that the data were appropriate for conducting EFA. Following the EFA, eight items were removed from the scale as their factor loadings fell below 0.45. The scree plot and eigenvalues < 1.0 were evaluated to determine the number of factors in this scale. In the end, two factors were extracted and the two factors accounted for 56.32% of the total variance. The two factors respectively accounted for 42.42 and 13.91% of the variance. Factor loadings ranged from 0.59 to 0.84 (see details in Table 1).

Factor 1 consisted of six items in which children frequently "absent-minded" during the process that required them to continuously devote attentional resources on the task at hand. We named this factor "attentional failure." In contrast, Factor 2 consisted of 6 items that include involuntary thoughts unrelated to the current activity, such as reflections on past events, plans for the future, or even "whimsical thoughts," which we named "spontaneous thinking." The correlation coefficient between the two factors was 0.442, indicating the homogeneity and relative independence of the content measured by each dimension.

Confirmatory factor analysis (CFA)

Participants

The sample consisted of 346 children (155 boys, 190 girls, 1 other; $M = 10.92$ years old, $SD = 1.20$, 21.4% of third graders, 25.1% of fourth graders, 35.3% of fifth graders, 18.2% of sixth graders).

Results

Considering that mind wandering might differ in frequency for boys and girls, we conducted a series of independent-sample t-tests and found that there were no gender differences for each dimension in our scales ($p > 0.05$, Table 2).

We then used the confirmatory factor analysis (CFA) and test-retest reliability to evaluate the scale. CFA is applied as a way to test the construct's dimensionality and to confirm the factor structure that emerged in the EFA. CFA was conducted using AMOS 4.0, and maximum likelihood estimation (MLE) was used to examine the fit of the model to the data. Evaluation of different indices has been suggested to check the fit of the model. The result revealed that the model provided an acceptable fit to the data, and the factorial structure of children's mind wandering frequency was confirmed ($\chi^2/df = 3.14$, GFI = 0.922, CFI = 0.919, TLI = 0.899, RMSEA = 0.079). All factor loadings were significant at $p < 0.05$.

Test-retest reliability

Test-retest reliability reflects the stability and consistency of the scale across time. To test the degree of consistency of the scale, we assessed test-retest reliability, and the interval between the retests was 2 weeks. A total of 99 participants completed the scale twice (41 boys and 58 girls, $M = 10.54$ years old, $SD = 1.08$, 25.3% of third graders, 37.4% of fourth graders, 37.4% of fifth graders). The correlation analysis showed that the correlation coefficient of the attentional failure dimension was 0.76 ($p < 0.01$), the correlation coefficient of the spontaneous thinking dimension was 0.74 ($p < 0.01$), and the correlation coefficient of the total score was 0.80 ($p < 0.01$). The results showed that this scale had high test-retest reliability over time.

Discussion

In this study, the construct validity of the scale was verified in detail in two aspects: overall model fit, and intrinsic structure of the model, and it can be concluded that the two-factor structure of the Frequency of Children's Mind Wandering Scale was supported by the data, and the test-retest reliability of the scale is robust.

TABLE 1 Factor loadings of the CMWS-F.

Items	Factor loadings	
	F1 (Attentional failure)	F2 (Spontaneous thinking)
1. I often get distracted when I'm doing something, and I think about other things without realizing it.	**0.807**	0.010
2. During class, I often lose focus and think about other things.	**0.790**	−0.078
3. I would try very hard to listen to the lecture or concentrate on something, but often I would still be thinking about something unrelated to the class.	**0.763**	0.000
4. While thinking about the problem, I often suddenly think of other things that are unrelated to the problem.	**0.742**	0.083
5. I often can't help but think about other interesting things when I'm doing my homework.	**0.735**	−0.004
6. I often do one thing but can't help thinking about another.	**0.616**	0.113
7. I am a person with a lot of ideas, and often suddenly appear some interesting ideas.	−0.225	**0.840**
8. My mind often comes up with one different thought (idea) after another.	0.013	**0.778**
9. I often can't help but make a lot of associations.	0.007	**0.775**
10. I always think about many things without realizing it.	0.219	**0.621**
11. I often have a bee in my bonnet and come up with a lot of ideas without realizing it.	0.141	**0.595**
12. Ideas often pop up in my head for no apparent reason.	0.223	**0.587**

CMWS-F, Frequency of Children's Mind Wandering Scale. The bold values indicate the strongest loading for each item.

TABLE 2 Gender comparisons for CMWS-F/CMWS-C scores in study 1.

Scale	Boys (M ± SD)	Girls (M ± SD)	t	df	Cohen'd
CMWS-F					
Attentional failure	2.69 ± 0.96	2.66 ± 0.86	0.352	343	0.03
Spontaneous thinking	3.35 ± 0.89	3.33 ± 0.83	0.262	343	0.02
Total score	3.02 ± 0.81	2.99 ± 0.72	0.356	343	0.04
CMWS-C					
High-demand	2.45 ± 0.87	2.40 ± 0.85	0.614	340	0.06
Low-demand	3.34 ± 0.83	3.34 ± 0.79	−0.072	340	0.00
Total score	2.90 ± 0.72	2.87 ± 0.67	0.338	340	0.04

CMWS-F, Frequency of Children's Mind Wandering Scale; CMWS-C, Context of Children's Mind Wandering Scale.

Early scholars believed that mind wandering arises simply as a result of failures of attentional control, and thus individual differences in attentional control could predict the frequency of mind wandering. According to this conjecture, adolescents have significantly weaker attentional control than adults, and thus adolescents should experience more mind wandering than adults. Young adolescents, however, did not report more mind wandering than adults (49). Study 1a suggests that mind wandering is the result of both attentional failure and spontaneous thinking, reflecting the individual's executive control and spontaneity of thought, respectively, which we should consider when predicting the frequency of mind wandering. It has been found that the generation of mind wandering is complex in different contexts (50), and the weighting of attentional failure and spontaneous thinking is constantly changing. For this reason, we designed Study 1b to develop a contextual scale for the occurrence of children's mind wandering and to explore how the frequency of children's mind wandering changes in different contexts.

Study 1b development of the context of children's mind wandering scale (CMWS-C)

The development of the Context of Children's Mind Wandering Scale followed the same four steps of interview and pilot test, exploratory factor analysis, confirmatory factor analysis, and test-retest analysis, each of which was strictly implemented in accordance with psychometric requirements.

Item generation

The whole process of items generation was similar to the development of the CMWS-F. The total sample included 218 children, with a total of 92 children in the first pilot test and 126 in the second test. Through two rounds of test, a total of 22 items (including one lie detection item) were formed for the initial scale.

Exploratory factor analysis (EFA)

Participants

The sample consisted of 332 children, of which 74 children were excluded due to randomly check and failed the lie detection item. The remaining 258 participants' data were analyzed (134 boys and 124 girls, $M = 10.71$ years old, $SD = 1.28$, 25.6% of third graders, 26.7% of fourth graders, 22.1% of fifth graders, 25.6% of sixth graders).

Results

The structure of the CMWS-C was analyzed using PCA and Varimax rotation method, and the specific statistical processing was implemented using SPSS 10.0. Without the lie detection item, we conducted EFA on 21 items. The KMO value was 0.870, indicating that the sample size was adequate for factor analysis. The Bartlett test was significant ($p < 0.001$), supporting the argument that the data were appropriate for conducting EFA. Following the EFA, nine items were removed from the scale as their factor loadings fell below 0.45. Based on the results and scree plot, two factors were extracted. The results showed that the two factors accounted for 51.57% of the total variance. The two factors respectively accounted for 27.98 and 23.59% of the variance. Factor loadings ranged from 0.59 to 0.79 (see details in Table 3).

Factor 1 consisted of 6 items across which children's mind wandering invaded in contexts such as reading and taking classes, and thus we named this factor "high-demand." In contrast, Factor 2 consisted of 6 items across which children's mind wandering happened in contexts such as riding in the car and going for a walk, and this is usually a context where children are prone to mind wandering, and we named this factor "low-demand."

Confirmatory factor analysis (CFA)

Participants

The sample consisted of 347 children (155 boys,191 girls, 1 other; $M = 10.92$ years old, $SD = 1.21$, 21.6% of third graders, 25.1% of fourth graders, 35.2% of fifth graders, 18.2% of sixth graders).

Results

The result revealed that the model provided an acceptable fit to the data, and the factorial structure of CMWS-C was confirmed ($\chi^2/df = 2.428$, GFI = 0.940, CFI = 0.929, TLI = 0.911, RMSEA = 0.064).

Test-retest reliability

A total of 100 participants completed the scale twice (42 boys and 58 girls, $M = 10.52$ years old, $SD = 1.09$, 26.0% of third graders, 37.0% of fourth graders, 37.0% of fifth graders), and the test-retest analysis showed that the correlation coefficient of the dimension of high-demand context was 0.61 ($p < 0.01$); the correlation coefficient of the dimension of low-demand context was 0.62 ($p < 0.01$); and the correlation coefficient of the total score was 0.64 ($p < 0.01$). The results showed that this scale had a relatively high test-retest reliability over time. In addition, a matched samples t-test showed that children reported significantly more mind wandering in low-demand contexts ($M = 3.40$, $SD = 0.84$) than in high contexts ($M = 2.43$, $SD = 0.85$), $t = 18.68$, $p < 0.001$.

Relationship between the CMWS-F and the CMWS-C

After collapsing the above data, a total of 343 children completed both CMWS-F and CMWS-C (21.3% of third graders, 25.4% of fourth graders, 35.3% of fifth graders, 18.1% of sixth graders). To further examine the effects of attentional failure and spontaneous thinking on the frequency of mind wandering in different contexts, the factors of CMWS-C were used as dependent variables and the factors of CMWS-F as predictor variables in regression analysis. As seen in Table 4, both attentional failure ($\beta = 0.194$, $p < 0.01$) and spontaneous thinking ($\beta = 0.514$, $p < 0.001$) were valid predictors in the low-demand context, and spontaneous thinking had a greater weight in the regression model. In contrast, only the factor of

TABLE 3 Factor loadings of the CMWS-C.

Items	Factor loadings	
	Factor 1(high-demand)	Factor 2(low-demand)
1. When I'm doing my homework, I would often be miles away.	**0.791**	0.006
2. Even though I'm concentrating on one thing, I still often get distracted and think about other things.	**0.774**	0.120
3. I often get distracted during class and think about other things.	**0.754**	0.091
4. When I am in the hobby class, I'm often miles away and think about something else.	**0.737**	0.271
5. During the exam, I am sometimes interrupted by other ideas that pop up.	**0.622**	0.082
6. I often get distracted and think of other things while reading.	**0.586**	0.205
7. On the way home by car, I often can't help but think about many things.	0.009	**0.735**
8. When I take a leisurely walk, I often can't help but think about other things.	0.109	**0.725**
9. When I sit alone in the chair resting, I can't help but think of one thing after another.	0.050	**0.684**
10. Before sleeping at night (when you're lying in the bed ready to sleep), things come to my mind.	0.164	**0.631**
11. When I'm brushing my teeth or taking a shower, I can't help but think of other things.	0.181	**0.628**
12. When I finish the textbook that the teacher asks us to finish before the specified time, I can't help but think about something else.	0.464	**0.599**

CMWS-C, Frequency of Children's Mind Wandering Scale. The bold values indicate the strongest loading for each item.

TABLE 4 Multiple regression analysis of the CMWS-F on CMWS-C.

Dependent variable	Predictor variables	*R*	R^2	*F*	β (Standardized)	*t*
Low-demand context	regression model	0.632	0.400	113.319***		
	Attentional failure				0.194	4.025**
	Spontaneous thinking				0.514	10.653***
High-demand context	regression model	0.788	0.620	277.701***		
	Attentional failure				0.782	20.365***
	Spontaneous thinking				0.012	0.313

**$p < 0.01$, ***$p < 0.001$. CMWS-F, Frequency of Children's Mind Wandering Scale; CMWS-C, Context of Children's Mind Wandering Scale.

attentional failure ($\beta = 0.782$, $p < 0.001$) was a valid predictor in the high-demand context.

Discussion

Study 1b shows that the occurrence of children's mind wandering can be broadly classified into two types of contexts: low-demand contexts, which are commonly understood to be prone contexts in which individuals often let go of conscious control and allow their thoughts to drift; and high-demand contexts, in which individuals are occasionally distracted by inner thoughts. The competing relationship between mind wandering and task load has been found in adult research, i.e., mind wandering occurs less when the current task places a higher demand on an individual's cognitive resource, and conversely, mind wandering occurs more frequently. Study 1b illustrated by means of a scale that children did report significantly more frequent mind wandering in low-demand contexts than in high-demand contexts.

In combination with Study 1a and Study 1b, we investigated the psychometric properties of the CMWS-F and CMWS-C. However, it is the first time to develop a scale for children's mind wandering, and it is difficult to find an authoritative method as a

calibration for this study. For this reason, we designed Study 2 to support the validity of the scales with more objective behavioral experiments. On the one hand, we can explore the characteristics of children's mind wandering in the laboratory tasks, and on the other hand, we can explore the relationship on subjective reports of mind wandering in the tasks and in our scales.

Study 2 probe-caught mind wandering in the experimental tasks

Participants

A power analysis was performed on children's data from a similar protocol to estimate the appropriate sample size (12, 13). With an $\alpha = 0.05$ and power = 80%, we required at least 82 participants to obtain a medium-sized effect (two-tailed, $r = 0.3$) for the validity of the scales in each laboratory task using the experience sampling method. Additional children were recruited in case of potential withdrawal from our tasks (e.g., inability to perform the computerized task). Finally, the total sample consisted of 365 children (all participants were from grades 4–6, 95 children in Breathing task, 105 children in Sustained Attention Response Task, 94 children in Vigilance task and 71 children in one-back task).

Materials and procedure

In this study, we selected four typical tasks with "thought probe" inserted to study mind wandering.

In the Breathing task, participants needed to wear headphones and sit in front of a computer screen, close their eyes and pay attention to the movement of breath in and out. In addition, a "ding" sound would come from the headphones from time to time, and children needed to open their eyes to answer the questions that appeared on the screen. The question was as follows: what were you thinking before the sound appeared? The answer options were to think about something related to counting breaths (please press "F") and to think about something unrelated to counting breaths (please press "J"). Following the judgment, they were prompted to close their eyes and count breaths again. During the task, the program presented 8 times of thought probes, and the time interval between two question pages was 25, 35, 45, and 55s. A 40s buffer was set at the beginning of the program, and the entire program lasted approximately 6 min.

In the Vigilance task, there was only one stimulus "+" presented for 0.5s, and the inter-stimuli interval was 1.5, 2.5, 3.5, or 4.5s. Participants were instructed to press a button whenever they saw the "+." This task lasted 12 min and collected 30 times of thought probes on mind wandering.

In the Sustained Attention Response Task (SART), the stimuli were divided into two categories: white numbers 1–9 (non-targets) and red crosses (targets). Each stimulus presented for 0.5s and the inter-stimulus interval was 2s. Participants were asked to press the space bar to the numbers as fast and accurately as possible and to withhold the response when the red cross was presented. The ratio of the appearance of non-targets to targets was 8:1 (non-targets = 240, targets = 30), and the program would randomly probe 30 times on mind wandering. The entire program was approximately 11 min.

In the one-back task, participants were asked to wear headphones and sit in front of a computer screen. Two stimuli, black numbers 1–9 (non-target) and green question mark (target), were presented at random order, with 2s for non-target and 3s for target on screen, followed by 1s inter-stimulus interval. Participants were asked to determine the parity of the previous number when the green question mark appeared, pressing "F" for odd numbers and "J" for even numbers. Non-target to target ratio was approximately 7:1 (non-targets = 190, targets = 24). Similarly, the program would randomly appear 6 times on mind wandering, with a "ding" in the headset, asking the participants about their state of consciousness at that time. The entire program was approximately 11 min.

It is important to note that we used thought-sampling method to obtain participants' current state of consciousness. To help participants familiar with the procedures, we presented examples of different options to ensure all of them understood the question before the formal tests. At the end, the participants were asked to complete the CMWS-F and the CMWS-C.

Results

We first calculated the accuracy of one-back task, and the children showed reasonable task performance ($M_{one-back} = 0.80$, $SD_{one-back} = 0.22$). We then evaluated how this task performance was affected by probe-caught mind wandering across subjects, and results showed a significant negative correlation between the task accuracy and the frequency of probe-caught mind wandering (calculated as the proportion of task-unrelated thoughts in probes, $M = 0.37$, $SD = 0.30$; $r_{one-back} = -0.277$, $p < 0.05$), indicating that the participants who reported more mind wandering were associated with poor performance during a high-demand task. In addition to the breathing task (no response required), we also calculated the average coefficient of variability (CV = SD/mean) for the reaction time in our SART and Vigilance tasks ($M_{\text{SART}} = 0.34$, $SD_{\text{SART}} = 0.22$, $M_{\text{Vigilance}} = 0.38$, $SD_{\text{Vigilance}} = 0.22$), which is consistent with our previous children's study (13). Taken together, these results suggested that children are able to complete these computerized tasks.

Next, we considered the correspondence between the state-level (probe-caught) and the trait-level mind wandering by

TABLE 5 Correlations between the frequencies of probe-caught mind wandering in tasks and the scores of the CMWS-F/CMWS-C in Study 2 (descriptive statistics included).

Task Type (probe-caught mind wandering)	TUT probability ($M \pm SD$)	CMWS-F ($M \pm SD$)		CMWS-C ($M \pm SD$)	
		Attentional failure	Spontaneous thinking	High-demand	Low-demand
Breathing task	0.43 ± 0.26	15.57 ± 5.08 (0.316**)	19.17 ± 5.58 (0.180)	13.86 ± 5.81 (0.236*)	19.75 ± 5.37 (0.101)
SART	0.17 ± 0.14	14.65 ± 4.90 (0.318**)	18.68 ± 5.70 (0.197*)	13.30 ± 5.71 (0.180)	19.35 ± 5.40 (0.164)
Vigilance task	0.14 ± 0.12	14.73 ± 5.01 (0.284**)	19.13 ± 5.67 (0.138)	13.70 ± 6.01 (0.186)	19.47 ± 5.08 (0.184)

*$p < 0.05$, **$p < 0.01$. The data in the parenttheses represented the correlation coefficient between the scores on CMWS-F/CMWS-C and probe-caght mind wandering in tasks. CMWS-F, Frequency of Children's Mind Wandering Scale; CMWS-C, Context of Children's Mind Wandering Scale. The probability of task-unrelated thought (TUT) was calculated as the proportion of task-unrelated thoughts in probes.

calculating the correlation between the frequencies of mind wandering in laboratory tasks and the scores on CMWS-F/CMWS-C. As expected, the results showed moderate positive correlations between the frequencies of probe-caught mind wandering in the three tasks and the scores in the scales, especially in the dimension of attentional failure of CMWS-F, with correlation coefficients ranging from 0.284 to 0.316 (Table 5).

Discussion

Early scholars often wonder when children are able to correctly understand and report conscious state. Studies of young children aged 4 to 13 years suggested that the uncontrollability of comprehension awareness did not mature until 8 or 9 years of age or even later (51–53). Here, we show that children are able to correctly report their mind wandering, at least, at the age of 8 years old. Mind wandering has been shown in numerous studies in adults to be associated with sustained attention capacity (54–57), working memory (23, 30), and intelligence tests (32). Usually when individuals report more mind wandering during the task, the worse the individual performs on the task, and vice versa. Mrazek and colleagues found that the higher the frequency of mind wandering during adolescent reading, the worse the reading comprehension scores, and concluded that students aged 11–13 years already have the ability to correctly report their state of consciousness (5). In all three different task conditions, we derived an agreement between mind wandering frequencies reported in the task and the mind wandering frequencies measured by the scale, confirming the criterion validity of the CMWS-F/CMWS-C.

Study 1a showed that the production of mind wandering was influenced by both attentional failure and spontaneous thinking, while study 1b showed that the contexts in which mind wandering occurred could be broadly classified into two categories: low-demand and high-demand, and the weights of the effects of attentional failure and spontaneous thinking on the frequency of mind wandering changed continuously in different contexts, constituting a dynamic and complex relational model. This would explain why the use of working memory capacity to predict mind wandering frequency under different experimental task conditions has been controversial (17).

General discussion

This is the first study to develop the CMWS-F and CMWS-C scales to assess the frequency of mind wandering for children. Following the steps of standard scale development (item generation and selection, EFA, CFA and test-retest analysis), we showed scales with good reliability and validity. Study 2 further provided evidence for the criterion validity of the CMWS-F and CMWS-C scales within four laboratory tasks, which showed moderate consistency between the frequencies of probe-caught mind wandering in the tasks and the scores in the scales.

Attentional failure and spontaneous thinking are two main factors that caused mind wandering in the CMWS-F. Attentional failure occurs when an individual's attention shifts from the current task to something unrelated, which arises as a result of executive control failure (22–24). Executive control is closely related to the main characteristic of mind wandering: disengagement of the external task/environment (58). Children's capacity for attentional control is relatively immature throughout childhood and adolescence (59, 60), so their inhibitory control capacity is so weak that they are not able to focus on the current task long time. Therefore, it's not hard to explain why inhibitory control capacity is a significant predictor of children's mind wandering frequency (3).

The second factor, spontaneous thinking, means that individual's mentation is occupied by implicit goals, such as memories from the past and plans for the future [also called unintentional mind wandering, (61)]. On the one hand,

studies have shown that spontaneous thinking is correlated with depression, stress and anxiety disorder (62, 63). In terms of personality, Spontaneous thinking has negative correlations with agreeableness, and extraversion (64). Thus, this may also bring negative consequences, like emotional problems and learning difficulties, for children. Therefore, if children score too high on the CMWS-F for spontaneous thinking, parents should pay attention to whether the children have emotional and interpersonal problems. On the other hand, this does not mean spontaneous thinking is useless for children. Studies have shown that mind wandering has a prospective bias (65–68), the content of mind wandering about future planning is significantly more than that about past, suggesting that mind wandering may contribute to the ability of future planning. Indeed, Ye et al. (12) found that there is a significant forward-looking bias of mind wandering among children. All of these suggest that only by correctly understanding the causes and the characteristics of children's mind wandering, can we give full play to the positive role of mind wandering and reasonably explore how to restrain children with high-frequency mind wandering and improve children with low-frequency mind wandering.

The CMWS-C showed that the contexts mind wandering often occurs in children can be divided into two categories: low-demand context and high-demand context. For children, the low-demand context is the leisure context like playing games and having a walk; the high-demand context is the task context like doing homework and reading books. Mind wandering can be beneficial or detrimental depending on the flexibility of cognitive resources in the specific context (58). In easy tasks that only demand low cognitive resources, individuals can tolerate longer delays in waiting for rewards because of mind wandering (69). However, in high cognitively demanding tasks, studies have shown that mind wandering could be detrimental for individual's performance (18, 70–72). This reveals us to the children's mind wandering treatment should also be flexible, in low-demand context, children can be relatively unfettered when mind wandering occurs, while in high-demand context, children must be required to pay attention to the current task.

In study 2, we also found some evidence for CMWS-F and CMWS-C in experimental tasks. There was a significant correlation between the score of attentional failure dimension in CMWS-F and three tasks ($p < 0.01$), but no significant correlation on the spontaneous thinking. As expected, attentional failure and spontaneous thinking occur in different contexts. Because the participants were asked to complete the laboratory tasks during the experiment, they were actually in high-demand context. Just like our definitions of the two dimensions of CMWS-F, attentional failure occurs more during the task, while spontaneous thinking occurs more when the individual is free or in low-demand context. This can also explain why the scale score is significantly or nearly significantly correlated with the three tasks under the high-demand dimension in CMWS-C ($p = 0.07$).

As mentioned above, mind wandering has both benefits and costs (43), so it should be viewed dialectically in the field of education. Under some conditions, mind wandering will promote students' creativity (73). But we have to admit that in many educational contexts, mind wandering could lead some negative influence. The higher the mind wandering frequency, the worse the test scores were (74). Besides, if students fail to pay attention to the classroom or what they are learning, this may impede their chances of acquiring important knowledge or skills. The frequency of mind wandering mediates the relation between children's ratings of topic interest and learning scores (9), so improving children's interest in what they have learned is a good choice to counteract the downsides of mind wandering.

Previous studies focused on the influence of divided attention on children from the cognitive process perspective (75), ignoring the spontaneous characteristics of consciousness. Here we showed that spontaneous thinking and attentional failure are the main causes of mind wandering, and discussed the influence of context on children's mind wandering, which will be enlightening for future studies on children's learning difficulties like ADHD (76, 77). Furthermore, meta-awareness hypothesis showed that individuals who are more aware of their current mental activity could have a lower frequency of mind wandering (1, 5). Considering the uncontrollability of children's mental states, one direction for future study is to enhance children's metacognition to indirectly prevent excessive mind wandering. Similarly, mindfulness training, including practices that enhance awareness of thoughts, may modulate the occurrence of mind wandering, supported by bottom-up and top-down neural mechanisms (78).

Finally, it is not certain whether the structure of the CMWS-F/CMWS-C are different in varied cultures. Therefore, more cross-culture studies are needed to explore children's mind wandering. In conclusion, we developed and validated the CMWS-F/CMWS-C scales to explore the underlying causes and contexts of mind wandering in children. The results showed that children's mind wandering was mainly caused by attentional failure and spontaneous thinking, and the contexts could be divided into high-demand context and low-demand context. We have to admit that although our study provided a convenient measurement tool for future studies of mind wandering in children, more research is needed to complement and enrich CMWS-F/CMWS-C.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of Zhejiang Normal University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

XS and QY designed the research. ZC and QY collected the data and performed the statistical analysis. YH and QY wrote the first draft of the manuscript. All authors contributed to the research and approved the final version of the manuscript.

Funding

This research was supported by grants from National Natural Science Foundation of China (32200912), Zhejiang Provincial Social Sciences Foundation (No. 78871), Open Research Fund of Psychology Department of Zhejiang Normal University, and Open Research Fund of College of Teacher Education of Zhejiang Normal University (No. jykf22034) to QY.

Acknowledgments

We thank all the participants for their contribution to this study, and we thank the staff of the Xiaoshan Yinhe Experimental Primary School for their help with the data collection.

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. Smallwood J. The restless mind. *Psychol Bull.* (2006) 132:946–58. doi: 10.1037/0033-2909.132.6.946

2. Killingsworth MA. A wandering mind is an unhappy mind. *Science.* (2010) 330:932. doi: 10.1126/science.1192439

3. Keulers EHH. Mind wandering in children: examining task-unrelated thoughts in computerized tasks and a classroom lesson, and the association with different executive functions. *J Exp Child Psychol.* (2019) 179:276–90. doi: 10.1016/j.jecp.2018.11.013

4. Smallwood J, O'Connor RC, Sudbery MV. Mind-wandering and dysphoria *Cogn Emot.* (2007 21:816–42. doi: 10.1080/02699930600911531

5. Mrazek MD, Phillips DT, Franklin MS, Broadway JM. Young and restless: validation of the Mind-Wandering Questionnaire (MWQ) reveals disruptive impact of mind-wandering for youth. *Front Psychol.* (2013) 4:560. doi: 10.3389/fpsyg.2013.00560

6. Maillet D, Yu L, Hasher L. Age-related differences in the impact of mind-wandering and visual distraction on performance in a go/no-go task. *Psychol Aging.* (2020) 35:627–38. doi: 10.1037/pag0000409

7. Moran CN, McGovern DP, Warren G, Gralaigh RO, Kenney JPM, Smeaton A, et al. Young and restless, old and focused: age-differences in mind-wandering frequency and phenomenology. *Psychol Aging.* (2021) 36:252–67. doi: 10.1037/pag0000526

8. Smith AP, Brosowsky N, Murray S, Daniel R, Meier ME. Fixation, flexibility, and creativity: The dynamics of mind wandering. *J Exp Psychol Hum Percept Perform.* (2022) 48:689–710. doi: 10.1037/xhp0001012

9. Cherry J, McCormack T. The link between mind wandering and learning in children. *J Exp Child Psychol.* (2022) 217:105367. doi: 10.1016/j.jecp.2021.105367

10. Jones PR. Sit still and pay attention: using the wii balance-board to detect lapses in concentration in children during psychophysical testing. *Behav Res Methods.* (2019) 51:28–39. doi: 10.3758/s13428-018-1045-4

11. McCormack T, Burns P, O'Connor P, Jaroslawska A. Do children and adolescents have a future-oriented bias? a developmental study of spontaneous and cued past and future thinking. *Psychol Res.* (2019) 83:774–87. doi: 10.1007/s00426-018-1077-5

12. Ye Q, Song X, Zhang Y. Children's mental time travel during mind wandering. *Front Psychol.* (2014) 5:927. doi: 10.3389/fpsyg.2014.00927

13. Zhang Y, Song X, Ye Q. Children with positive attitudes towards mind-wandering provide invalid subjective reports of mind-wandering during an experimental task. *Conscious Cogn.* (2015) 35:136–42. doi: 10.1016/j.concog.2015.05.006

14. Klinger E. Daydreaming and fantasizing: thought flow and motivation. In: Markman KD, Klein WMP, Suhr J, editors. *Handbook of Imagination and Mental Simulation.* New York, NY: Psychology Press (2009). p. 225–39.

15. Klinger E. Thought flow: Properties and mechanisms underlying shifts in content. In: Singer JA, Salovey P, editors. *At Play in the Fields of Consciousness: Essays in the Honour of Jerome L. Singer.* Mahwah, NJ: Erlbaum (1999). p. 29–50.

16. Schooler JW, Smallwood J, Christoff K, Handy TC, Reichle ED. Meta-awareness, perceptual decoupling and the wandering mind. *Trends Cogn Sci.* (2011) 15:319–26. doi: 10.1016/j.tics.2011.05.006

17. Smallwood J. Distinguishing how from why the mind wanders: a process-occurrence framework for self-generated mental activity. *Psychol Bull.* (2013) 139:519–35. doi: 10.1037/a0030010

18. Baldwin CL, Roberts DM, Barragan D, Lee JD, Lerner N. Detecting and quantifying mind wandering during simulated driving. *Front Hum Neurosci.* (2017) 11:406. doi: 10.3389/fnhum.2017.00406

19. Barron E, Riby LM, Greer J. Absorbed in thought: the effect of mind wandering on the processing of relevant and irrelevant events. *Psychol Sci.* (2011) 22:596–601. doi: 10.1177/0956797611404083

20. Smallwood J, Beach E, Schooler JW. Going AWOL in the brain: mind wandering reduces cortical analysis of external events. *J Cogn Neurosci.* (2008) 20:458–69. doi: 10.1162/jocn.2008.20037

21. Smallwood J. Why the global availability of mind wandering necessitates resource competition: reply to McVay and Kane (2010). *Psychol Bull.* (2010) 136:202–7. doi: 10.1037/a0018673

22. McVay JC. Conducting the train of thought: working memory capacity, goal neglect, and mind wandering in an executive-control task. *J Exp Psychol Learn Mem Cogn.* (2009) 35:196–204. doi: 10.1037/a0014104

23. McVay JC. Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *J Exp Psychol Gen.* (2012) 141:302–20. doi: 10.1037/a0025250

24. McVay JC. Does mind wandering reflect executive function or executive failure? Comment on Smallwood and Schooler (2006) and Watkins (2008). *Psychol Bull.* (2010) 136:188–97. doi: 10.1037/a0018298

25. Andrews-Hanna JR, Reidler JS, Huang C. Evidence for the default network's role in spontaneous cognition. *J Neurophysiol.* (2010) 104:322–35. doi: 10.1152/jn.00830.2009

26. Christoff K, Gordon AM, Smallwood J, Smith R. Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proc Natl Acad Sci U.S.A.* (2009) 106:8719–24. doi: 10.1073/pnas.0900234106

27. Fox KC, Spreng RN, Ellamil M, Andrews-Hanna JR. The wandering brain: meta-analysis of functional neuroimaging studies of mind-wandering and related spontaneous thought processes. *Neuroimage.* (2015) 111:611–21. doi: 10.1016/j.neuroimage.2015.02.039

28. Battaglia S, Cardellicchio P, Fazio D, Nazzi C, Fracasso CA, et al. The influence of vicarious fear-learning in "infecting" reactive action inhibition. *Front Behav Neurosci.* (2022) 16:946263. doi: 10.3389/fnbeh.2022.946263

29. Battaglia S, Cardellicchio P, Fazio D, Nazzi C, Fracasso CA, et al. Stopping in (e)motion: Reactive action inhibition when facing valence-independent emotional stimuli. *Front Behav Neurosci.* (2022) 16:998714. doi: 10.3389/fnbeh.2022.998714

30. McVay JC. Drifting from slow to "D'oh!": working memory capacity and mind wandering predict extreme reaction times and executive control errors. *J Exp Psychol Learn Mem Cogn.* (2012) 38:525–49. doi: 10.1037/a0025896

31. Unsworth N. Mind wandering and reading comprehension: examining the roles of working memory capacity, interest, motivation, and topic experience. *J Exp Psychol Learn Mem Cogn.* (2013) 39:832–42. doi: 10.1037/a0029669

32. Mrazek MD, Smallwood J, Franklin MS, Chin JM, Baird B. The role of mind-wandering in measurements of general aptitude. *J Exp Psychol Gen.* (2012) 141:788–98. doi: 10.1037/a0027968

33. Robison MK, Unsworth N. Cognitive and contextual correlates of spontaneous and deliberate mind-wandering. *J Exp Psychol Learn Mem Cogn.* (2018) 44:444. doi: 10.1037/xlm0000444

34. Unsworth N. Working memory capacity and sustained attention: a cognitive-energetic perspective. *J Exp Psychol Learn Mem Cogn.* (2020) 46:77–103. doi: 10.1037/xlm0000712

35. Kane MJ, Brown LH, McVay JC, Silvia PJ, Myin-Germeys I. For whom the mind wanders, and when: an experience-sampling study of working memory and executive control in daily life. *Psychol Sci.* (2007) 18:614–21. doi: 10.1111/j.1467-9280.2007.01948.x

36. Levinson DB, Smallwood J. The persistence of thought: evidence for a role of working memory in the maintenance of task-unrelated thinking. *Psychol Sci.* (2012) 23:375–80. doi: 10.1177/0956797611431465

37. Baird B, Smallwood J. Back to the future: autobiographical planning and the functionality of mind-wandering. *Conscious Cogn.* (2011) 20:1604–11. doi: 10.1016/j.concog.2011.08.007

38. Rummel J. Controlling the stream of thought: working memory capacity predicts adjustment of mind-wandering to situational demands. *Psychon Bull Rev.* (2014) 21:1309–15. doi: 10.3758/s13423-013-0580-3

39. Singer J, Antrobus J. Day dreaming, imaginal processes, and personality: A normative study. In: Sheehan P, editor. *The Function and Nature of Imagery.* New York, NY: Academic Press (1972). p. 175–202.

40. Cheyne JA, Carriere JS. Absent-mindedness: lapses of conscious awareness and everyday cognitive failures. *Conscious Cogn.* (2006) 15:578–92. doi: 10.1016/j.concog.2005.11.009

41. Broadbent DE, Cooper PF, FitzGerald P, Parkes KR. The Cognitive Failures Questionnaire (CFQ) and its correlates. *Br J Clin Psychol.* (1982) 21:1–16. doi: 10.1111/j.2044-8260.1982.tb01421.x

42. Carriere JS, Cheyne JA. Everyday attention lapses and memory failures: the affective consequences of mindlessness. *Conscious Cogn.* (2008) 17:835–47. doi: 10.1016/j.concog.2007.04.008

43. Mooneyham BW. The costs and benefits of mind-wandering: a review. *Can J Exp Psychol.* (2013) 67:11–8. doi: 10.1037/a0031569

44. Pachai AA, Acai A, LoGiudice AB, Kim JA. The mind that wanders: challenges and potential benefits of mind wandering in education. *JoSoTL.* (2016) 2:134–46. doi: 10.1037/stl0000060

45. Hollon SD. Cognitive self-statements in depression: Development of an automatic thoughts questionnaire. *Cognit Ther Res.* (1980) 4:383–95. doi: 10.1007/BF01178214

46. Delamillieure P, Doucet G, Mazoyer B, Turbelin MR, Delcroix N, Mellet E, et al. The resting state questionnaire: An introspective questionnaire for evaluation of inner experience during the conscious resting state. *Brain Res Bull.* (2010) 81:565–73. doi: 10.1016/j.brainresbull.2009.11.014

47. Brown KW. The benefits of being present: mindfulness and its role in psychological well-being. *J Pers Soc Psychol.* (2003) 84:822–48. doi: 10.1037/0022-3514.84.4.822

48. Belardi A, Chaieb L, Rey-Mermet A, Mormann F, Rothen N, Fell J, et al. On the relationship between mind wandering and mindfulness. *Sci Rep.* (2022) 12:7755. doi: 10.1038/s41598-022-11594-x

49. Stawarczyk D, Majerus S, Catale C. Relationships between mind-wandering and attentional control abilities in young adults and adolescents. *Acta Psychol (Amst).* (2014) 148:25–36. doi: 10.1016/j.actpsy.2014.01.007

50. Kane MJ, Gross GM, Chun CA, Smeekens BA, Meier ME, Silvia PJ, et al. For whom the mind wanders, and when, varies across laboratory and daily-life settings. *Psychol Sci.* (2017) 28:1271–89. doi: 10.1177/0956797617706086

51. Flavell JH, Green FL. Children's understanding of the stream of consciousness. *Child Dev.* (1993) 64:387–98. doi: 10.2307/1131257

52. Flavell JH, Green FL. The mind has a mind of its own: developing knowledge about mental uncontrollability. *Cogn Dev.* (1998) 13:127–38. doi: 10.1016/S0885-2014(98)90024-7

53. Flavell JH, Green FL. Development of children's awareness of their own thoughts. *J Cogn Dev.* (2000) 1:97–112. doi: 10.1207/S15327647JCD0101N_10

54. Dixon P. Construction, integration, and mind wandering in reading. *Can J Exp Psychol.* (2013) 67:1–10. doi: 10.1037/a0031234

55. Feng S, D'Mello S. Mind wandering while reading easy and difficult texts. *Psychon Bull Rev.* (2013) 20:586–92. doi: 10.3758/s13423-012-0367-y

56. Franklin MS, Mooneyham BW, Baird B. Thinking one thing, saying another: the behavioral correlates of mind-wandering while reading aloud. *Psychon Bull Rev.* (2014) 21:205–10. doi: 10.3758/s13423-013-0468-2

57. Franklin MS, Smallwood J. Catching the mind in flight: using behavioral indices to detect mindless reading in real time. *Psychon Bull Rev.* (2011) 18:992–7. doi: 10.3758/s13423-011-0109-6

58. Smallwood J. The science of mind wandering: empirically navigating the stream of consciousness. *Annu Rev Psychol.* (2015) 66:487–518. doi: 10.1146/annurev-psych-010814-015331

59. Keulers EH, Goulas A, Jolles J. Maturation of task-induced brain activation and long range functional connectivity in adolescence revealed by multivariate pattern classification. *Neuroimage.* (2012) 60:1250–65. doi: 10.1016/j.neuroimage.2011.12.079

60. Schleepen TM. A longitudinal study of semantic grouping strategy use in 6-11-year-old children: investigating developmental phases, the role of working memory, and strategy transfer. *J Genet Psychol.* (2014) 175:451–71. doi: 10.1080/00221325.2014.958126

61. Seli P, Risko EF, Smilek D. Mind-wandering with and without intention. *Trends Cogn Sci.* (2016) 20:605–17. doi: 10.1016/j.tics.2016.05.010

62. Marcusson-Clavertz D. Psychometric properties of the spontaneous and deliberate mind wandering scales. *Eur J Psychol Assess.* (2018) 35:878–90. doi: 10.1027/1015-5759/a000470

63. Seli P, Beaty RE, Marty-Dugas J. Depression, anxiety, and stress and the distinction between intentional and unintentional mind wandering. *Psychol Conscious: Theory, Res, Pract.* (2019) 6:163–70. doi: 10.1037/cns0000182

64. Carciofo R. Deliberate and spontaneous mind wandering in Chinese students: Associations with mindfulness, affect, personality, and life satisfaction. *Pers Indiv Differ.* (2021) 180:110982. doi: 10.1016/j.paid.2021.110982

65. Smallwood J, Nind L. When is your head at? an exploration of the factors associated with the temporal focus of the wandering mind. *Conscious Cogn.* (2009) 18:118–25. doi: 10.1016/j.concog.2008.11.004

66. Smallwood J, Schooler JW, Turk DJ, Cunningham SJ, Burns P. Self-reflection and the temporal focus of the wandering mind. *Conscious Cogn.* (2011) 20:1120–6. doi: 10.1016/j.concog.2010.12.017

67. Song X. Mind wandering in Chinese daily lives—an experience sampling study. *PLoS ONE.* (2012) 7:e44423. doi: 10.1371/journal.pone.0044423

68. Stawarczyk D, Cassol H. Phenomenology of future-oriented mind-wandering episodes. *Front Psychol.* (2013) 4:425. doi: 10.3389/fpsyg.2013.00425

69. Smallwood J, Ruby FJ. Letting go of the present: mind-wandering is associated with reduced delay discounting. *Conscious Cogn.* (2013) 22:1–7. doi: 10.1016/j.concog.2012.10.007

70. Randall JG, Oswald FL. Mind-wandering, cognition, and performance: a theory-driven meta-analysis of attention regulation. *Psychol Bull.* (2014) 140:1411–31. doi: 10.1037/a0037428

71. Unsworth N. The influence of lapses of attention on working memory capacity. *Mem Cognit.* (2016) 44:188–96. doi: 10.3758/s13421-015-0560-0

72. Yanko MR. Driving with the wandering mind: the effect that mind-wandering has on driving performance. *Hum Factors.* (2014) 56:260–9. doi: 10.1177/0018720813495280

73. Baird B, Smallwood J, Mrazek MD, Kam JW, Franklin MS. Inspired by distraction: mind wandering facilitates creative incubation. *Psychol Sci.* (2012) 23:1117–22. doi: 10.1177/0956797612446024

74. Risko EF, Anderson N, Sarwal A, Engelhardt M. Everyday attention: variation in mind wandering and memory in a lecture. *Appl Cogn Psychol.* (2012) 26:234–42. doi: 10.1002/acp.1814

75. Savage R, Cornish K, Manly T. Cognitive processes in children's reading and attention: the role of working memory, divided attention, and response inhibition. *Br J Psychol.* (2006) 97:365–85. doi: 10.1348/000712605X 81370

76. Alperin BR, Christoff K, Mills C. More than off-task: increased freely-moving thought in ADHD. *Conscious Cogn.* (2021) 93:103156. doi: 10.1016/j.concog.2021.103156

77. Frick MA, Asherson P. Mind-wandering in children with and without ADHD. *Br J Clin Psychol.* (2020) 59:208–23. doi: 10.1111/bjc. 12241

78. Kaunhoven RJ. How does mindfulness modulate self-regulation in pre-adolescent children? an integrative neurocognitive review. *Neurosci Biobehav Rev.* (2017) 74:163–84. doi: 10.1016/j.neubiorev.2017.01.007

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 20 December 2022
DOI 10.3389/fpubh.2022.1054931

OPEN ACCESS

EDITED BY
Pei Wang,
Fudan University, China

REVIEWED BY
Xiaoyu Xi,
China Pharmaceutical University, China
Shunping Li,
Shandong University, China

*CORRESPONDENCE
Qi Kang
kangqi@shdrc.org
Chunlin Jin
jinchunlin@shdrc.org

†These authors have contributed equally to this work

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

RECEIVED 27 September 2022
ACCEPTED 11 November 2022
PUBLISHED 20 December 2022

CITATION
Hu J, Zhu L, Bao H, Liu Y, Xing H, Kang Q and Jin C (2022) Utility estimations of different health states of patients with type I, II, and III spinal muscular atrophy in China: A mixed approach study with patient and proxy-reported data. *Front. Public Health* 10:1054931. doi: 10.3389/fpubh.2022.1054931

Utility estimations of different health states of patients with type I, II, and III spinal muscular atrophy in China: A mixed approach study with patient and proxy-reported data

Jiahao Hu[1†], Lin Zhu[1†], Han Bao[2], Yuhan Liu[1], Huanping Xing[3], Qi Kang[1*] and Chunlin Jin[1*]

[1]Shanghai Health Development Research Centre (Shanghai Medical Information Centre), Shanghai, China, [2]Institute of Pharmaceutical Economics, Sun Yat-sen University, Guangzhou, China, [3]Meier Advocacy & Support Centre for SMA, Beijing, China

Introduction: Spinal muscular atrophy (SMA) is a rare autosomal-recessive neuromuscular disease. Health state utility values (HSUV) are used in health economic evaluation regarding the desirability of health outcomes such as a certain health state or change in health states over time. There is no utility data of Chinese patients with SMA.

Materials and methods: Vignettes were developed for 10 pediatric neurologists to value the utility of Chinese patients with Type I SMA. A mixed patient/proxy derived approach using EQ-5D-Y-3L, EQ-5D-3L, and CHU9D was adopted to estimate the HSUV data of patients with Type II and III SMA, including 112 patients and 301 caregivers.

Result: The utility of Type I SMA patients ranged from 0.19 to 0.72 with the health state improved from "permanent ventilation" to "walking". The utility of children patients with Type II and III SMA derived from EQ-5D-Y-3L ranged from 0.33 to 0.82 while that derived from CHU9D ranged from 0.46 to 0.75. The utility of adult patients with Type II and III SMA measured by EQ-5D-3L ranged from 0.30 to 0.83.

Conclusion: The better health states the patients with SMA were in, the higher were the HSUV. The utilities derived from population with different age and disease subtypes were not statistically different when patients with SMA were in the same health states. We recommend further studies on the Chinese specific value set for EQ-5D-Y-3L and other PBMs for children to derive more robust utility data.

KEYWORDS

spinal muscular atrophy, vignette approach, CHU9D, EQ-5D-Y, EQ-5D-3L, utility

Introduction

Spinal muscular atrophy (SMA) is a rare autosomal-recessive neuromuscular disease, which primarily affects children and is the leading genetic cause of infant death as well (1). It is caused by genetic deletions or mutations of the survival motor neuron (SMN) 1 gene (2), which results in progressive proximal muscle weakness, atrophy and, in the most severe types, paralysis (3). SMN2, a related gene, produces insufficient levels of stable SMN protein to compensate for the SMN1 deficiency, and the number of SMN2 copies that an individual carries is generally inversely proportional to the severity of the disease (4). The estimated worldwide incidence of SMA is 1 in 10,000 live births (1). However, the prevalence of SMA carrier is estimated to be 2.0% in China, which may be on a slow upward trend (5). SMA is divided into five clinical subtypes (0, I, II, III, IV) based on patient age at onset of disease symptoms and the highest motor function achieved (6). Type 0 is the most severe subtype with onset in utero and death before 6 months of age (7). Type I (infantile-onset) SMA is a common subtype in living patients which usually presents before 6 months of age with a life expectancy of less than 2 years. Without intervention, patients with Type I SMA are never able to sit independently (8). Type II usually onsets between 6 and 18 months of age. Patient with Type II could sit independently but never be able to walk (9). Type III normally onsets after age 18 months, with which patients may acquire independent ambulation, while some may lose the ability to walk in adulthood owing to the progressive nature of the disease (10). Type II and III SMA are known as later-onset SMA as well. Type IV SMA (adult-onset) is the rarest subtype and has the lowest morbidity and mortality, which generally occurs after the age of 20 (6).

The treatment of SMA involves multidisciplinary inputs from neurologists, respiratory specialists, gastroenterologists, geneticists, palliative care physicians, orthopedic surgeons, and physical therapists, etc., (11, 12). After 2016, the introduction of disease-modifying treatments (DMTs) brought new management options for patients with SMA. There are currently two approved DMTs for SMA in China. Nusinerisen was approved in China in 2019, is an antisense oligonucleotide that modifies SMN2 pre-messenger RNA splicing to increase functional SMN production. It needs to be intrathecally administered to patients (13). Risdiplam is a daily orally administered SMN2 splicing modifier that is distributed centrally and peripherally and increases SMN production. Risdiplam was approved in the China in 2021 for patients with SMA aged more than 2 months (14).

Though there are available DMTs in China, the high price raised concerns from the Chinese SMA community and payers, for example, the National Healthcare Security Administration (NHSA), which is responsible for reimbursement decisions in China. Health economic evaluation has been a necessary step in the annual appraisal procedure of National Reimbursed Drug List since 2019 in China.

There is currently no published health economic evaluation study on DMTs for SMA in the Chinese context. The published health economic evaluation studies from other countries used a discrete-event Markov model structure. Type I SMA was modeled separately, while Type II and III were often modeled together (15–21). Thus, outcome measurements for patients with Type I and Type II/III SMA, respectively, are needed in health economic evaluations. Quality-adjusted life-years (QALYs), which measures both the survival and health-related quality of life, is the most recommended and widely used outcome measurement in health economic evaluation (22). QALYs are calculated with health state utility values (HSUV), which could be derived with direct or indirect approaches. The direct approaches include methods like standard gambling (SG), time trade-off (TTO) and discrete choice experiment (DCE), while the indirect approach includes using generic preference-based measures (PBMs) such as EQ-5D (including EQ-5D adult and youth versions) and SF-6D. The indirect approach is recommended based on China Guidelines for Pharmacoeconomic Evaluations (23). It further recommends that EQ-5D-3L or EQ-5D-5L be used to derive utility for adults and EQ-5D-Y-3L be used to derive utility for children and adolescents (23). If there is no EQ-5D or other PBM instrument data collected from patients, mapping algorithm from disease-specific/generic HRQoL data to a generic PBM is acceptable. For example, an algorithm has been used to map EQ-5D-Y-3L utility scores from Pediatric Quality of Life Inventory (PedsQL) generic core scales (24).

There have been a few studies reporting the utility of patients with SMA published with different study designs and study locations. Chambers et al. surveyed 40 patient-caregiver pairs with Type I, II and III SMA in Australia using EQ-5D-Y-3L, while EQ-5D-3L Australian value sets were used as a proxy as EQ-5D-Y-3L value set was unavailable (25). Lloyd et al. used vignette approach by interviewing five clinical experts in SMA using EQ-5D-Y-3L and Pediatric Quality of Life (PedsQL) Neuromuscular Module to estimate the utility of health states describing different severities of Type I and II SMA (26). López-Bastida et al. reported proxy-reported utility of SMA patients derived from EQ-5D-3L (27). Malone et al. used PedsQL data collected in clinical trial and mapped it onto EQ-5D-Y-3L for patients with Type I SMA (20). Lo et al. designed a DCE to estimate the health utility for SMA treatment outcomes from a general public sample (28). Only one study reported the quality of life of Chinese children with SMA and their caregivers while PedsQL Neuromuscular Module and Family Impact Module were applied to the patients and their caregivers (29). Thus, no utility data of Chinese patients with SMA is currently available for inputs in economic evaluations.

The objective of this study was using a mixed approach of novel vignette approach and traditional questionnaire survey to estimate the utility of Chinese patients with SMA in different health states. It was the first such study with a largest number of SMA patients and would help methodological development in relevant areas in non-western countries. With validated methods and measurements, the trustworthy health utility data would help generate health economics evidence and further support NHSA's relevant reimbursement decision-makings, from which the Chinese SMA community would benefit.

Materials and methods

Study design

This study used a mixed approach of qualitative and quantitative methods. Patients with Type I SMA was extremely difficult to recruit due to the short life expectancy and a low prevalence. Furthermore, the patients with Type I SMA would never be able to sit independently, it was highly unlikely to recruit patients with better motor functions than sitting. Patients at such a young age do not have the literacy to complete a generic PBMs and no generic PBMs (except for HUI3, which did not have a Chinese value set) have been validated in infants younger than 2 years old. Thus, a vignette approach was adopted to measure the utility of Type I SMA patients. For Type II and III SMA patients, a mixed patient or caregiver-derived approach was adopted.

The whole process of collecting utility data was done through January to March 2022. The study design was approved by the Medical Ethics Committee of Shanghai Health Development Research Center (No. 2022004).

Vignette approach for type I SMA

Defining health states for type I SMA

The health states of Type I SMA included four motor function milestones: "not sitting", "sitting", "standing" and "walking", defined by Hammersmith Infant Neurological Exam-Part 2 (HINE-2), together with "permanent ventilation" (30).

Vignette development

A vignette is a description of the impact of a medical condition that is valued in a preference elicitation task to obtain a utility estimate. Vignettes may describe a medical condition, its treatment, and its impact on various domains of HRQoL (31). Lloyd et al. had published a study to estimate the utility of patients with Type I and Type II SMA (26). The example case study description was referred to and forward translated from English to Chinese by JH and ZL. The differences between the two translated versions were discussed by the authors to form the first Chinese case history description of "not sitting" state. The first version was further backward translated to English to ensure that the Chinese contents were accurately expressed, and the language was precise and easy to understand. The backward translated version was compared with the original English version to find the differences and further modify until all authors approved. The second version of Chinese case history description of "not sitting" state was thus formed.

Two pediatric neurologists with over 10 years' experiences in diagnosis and treatment of SMA were interviewed to validate the second Chinese case history description of "not sitting" state to ensure the content and expression was appropriate and easy to understand. Comments raised in the interviews were discussed and revised by the authors. In addition, questions about the differences of symptoms and functional limitations between health states were interviewed to develop the other case history descriptions of the resting health states. When all the Chinese health states descriptions had been developed, the two pediatric neurologists were interviewed again to test the feasibility, comprehensibility, and acceptability of the questions all the case history descriptions.

Valuation

Ten senior pediatric neurologists with over 10 years' experiences in diagnosis and treatment of SMA from hospitals in different cities/municipals were chosen to participate in valuation. The digital version of all the case history descriptions were sent to the participants through email. The participants were asked to read each case history description and to draw on their clinical experience to try to imagine how patients would be affected by that severity of SMA. After finishing reading all the case history descriptions, the neurologists were asked to complete a proxy assessment for all the health states using EQ-5D-Y-3L Chinese version.

Mixed patient/proxy-derived approach for types II and III SMA

Defining health states for types II and III SMA

Seven discrete health states of Type II and III SMA were divided, including "not sitting", "sitting supported", "sitting unsupported", "standing supported", "standing unsupported", "walking supported" and "walking unsupported", with the increasing motor function. Adding "supported or unsupported" was to further explore the potential differences in utility (30). Health states related to sitting and standing were defined by the scores of Items 9 and 25 in Motor Function Measurement-32 items (MFM-32), respectively (32). The two highest levels of motor function about walking were defined as keeping with the highest independent mobility in Hammersmith Functional Motor Scale Expanded (HFMSE) (33).

TABLE 1 Type I SMA health states utility values.

Health states	Permanent ventilation	Not sitting	Sitting	Standing	Walking
		Mean (SD)			
EQ-5D-Y-3L health utility	0.19 (0.06)	0.29 (0.08)	0.39 (0.19)	0.60 (0.21)	0.72 (0.21)

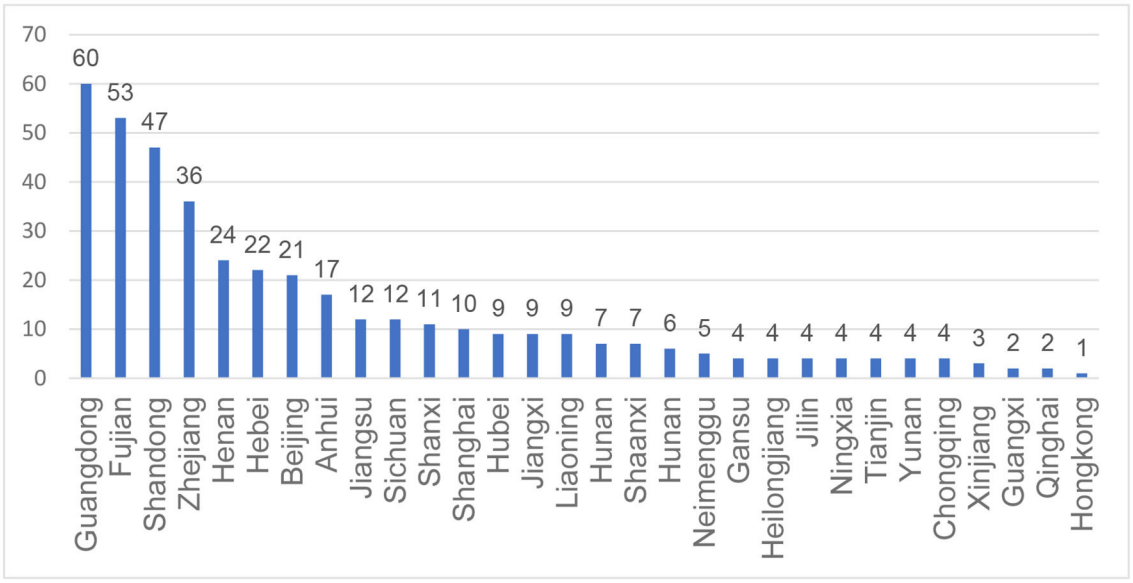


FIGURE 1
Geographical distribution of participated SMA patients.

Patient recruitment

A patient/caregiver derived approach was applied for Types II and III SMA patients. The patient recruitment was *via* snowballing sampling with the help of Meier Advocacy & Support Center for SMA, a Chinse patient advocacy group focusing on SMA, which had registered over 1,500 patients with different motor functions. Meier Advocacy & Support Group was responsible for patient recruitments through group chats, where only patients with definitive diagnosis and their caregivers could join. Participants were sent a link to the online survey (registered on www.wjx.cn, an online survey platform), which contained an electronic informed consent forms where participants had to complete a tick box indicating consent to participate prior to survey completion. The inclusion criterion was that patients aged over 4 who had been diagnosed with Types II or III SMA and were willing to participate in. The exclusion criterion was that the patients were diagnosed with other life-threatening diseases or in other conditions that were not suitable to participate in the study.

Questionnaire design

The questionnaire used in this phase included three parts: patients' characteristics, simplified motor function measurements and HRQoL measurements.

The first part collected information regarding patients' demographic variables (gender, age) and disease conditions (including disease subtypes, number of SMN2 copies and motor function status).

The second part adapted the abovementioned items and their answers from the MFM-32 and HFMSE into 7 corresponding statements. For example, respondents were asked the following question "which of the following description represents the best movements patients could do at present?". If the statement "patient could sit with no upper limb support, basically maintain for more than 5 seconds and then hold hands touching for 5 sec" was chosen, that meant patients were in the health state of "sitting without support".

The third part comprised two children friendly PBMs to collect utility data. EQ-5D-Y-3L and Child Health Utility Index-9 Dimensions (CHU9D) were used. EQ-5D-Y-3L were chosen as

it was a widely used children-friendly PBM, while CHU9D was the only available pediatric-specific PBMs with a Chinese version and the affiliated value set for children and adolescents in China. EQ-5D-Y was designed for children aged from 8 to 15 years old and a proxy version for children aged 4 to 7 was also available (34). CHU9D was developed originally for children aged 7–17 while a proxy version could be used for children aged 5–7 (35). Thus, participated patients aged 5–7 were asked to complete the proxy version of EQ-5D-Y-3L and CHU9D by their main caregivers, while patients aged 8–15 could choose to use self-reported or proxy-reported version of both PBMs. The patients over 15 were asked to finish EQ-5D-3L.

Data analysis

Based on the Chinese Pharmacoeconomics Evaluation Guideline, EQ-5D health states should be directly converted into China-specific utilities using Chinese value sets (23). Since there were no EQ-5D-Y-3L Chinese value set, an EQ-5D-3L Chinese value set were used as proxy (utility ranged from 0.1702 to 1.0000) in circumstances where EQ-5D-Y-3L were applied (36). A CHU9D China value set were used to score CHU9D data (utility ranged from −0.0855 to 1.0000) (37). Greater utility value indicated better health status in both value sets. Descriptive statistics were applied to describe the characteristics of the patients. The HSUV data of different states derived from two age subgroups were summarized as mean (standard deviations, SD) and median (interquartile range, IQR). The intraclass correlation coefficient (ICC) was calculated to assess the level of agreement between the CHU9D and the EQ-5D-Y-3L when used in patients under 16. An ICC below 0.75 implies poor to moderate agreement, whereas an ICC above 0.75 implies good agreement (38). Mann Whitney-U test was applied to explore the differences of utility of patients with similar motor functions but with different disease subtypes. Mann Whitney-U test was chosen as utility was not normally distributed and the homogeneity of variance was unknown. SPSS (IBM Corp. Released 2019. IBM SPSS Statistics for Windows, Version 26.0. Armonk, NY) was used for data analysis.

Results

HSUV of patients with type I SMA

All the 10 participated neurologists completed the valuation tasks. The mean utility values Type I SMA health states assessed by EQ-5D-Y-3L are shown in Table 1. The utility values of different health states range from 0.19 (the worst state "Permanent Ventilation") to 0.72 (the best state "Walking"). The utility values increase with the improvements in motor functions.

HSUV of patients with type II and III SMA

Characteristics of the patients

A total of 413 patients/caregivers participated in the study. Four hundred and twelve respondents were from 29 provinces in mainland China and one was from Hongkong, no patients from Guizhou or Xizang province participated into this study. The geographical distribution of the participants was shown in Figure 1. A total of 112 patients completed the survey by themselves while the rest results were reported by the caregivers. The age of the participants ranges from 1 to 53. The mean (SD) and median (IQR) age is 13.46 (10.10) and 9 (13.00), respectively. Most patients (48.67%) had 3 SMN2 copies while 130 patients were unclear about the number of SMN2 copies. Most SMA patients (30.27%) were in the "sitting unsupported" state, followed by the "walking unsupported" state (23.49%). Only 5.57, 5.81, and 4.84% of the patients were in the "standing supported", "standing unsupported" and "walking supported"

TABLE 2 Characteristics of the participated SMA patients ($N = 413$).

Variable	*N* (%)
Age (years old)	
<16	288 (69.73)
≥16	125 (30.27)
Gender	
Male	219 (53.02)
Female	194 (46.97)
Respondent	
Patient	112 (27.11)
Caregiver	301 (72.89)
Disease subtype	
Type II	254 (61.50)
Type III	159 (38.50)
Number of SMN2 Copies	
1	4 (0.97)
2	29 (7.02)
3	201 (48.67)
4	44 (10.65)
More than 4	5 (1.21)
Unclear	130 (31.48)
Motor function	
Not able to sit	50 (12.10)
Sitting supported	74 (17.92)
Sitting unsupported	125 (30.27)
Standing supported	23 (5.57)
Standing unsupported	24 (5.81)
Walking supported	20 (4.84)
Walking unsupported	97 (23.49)

TABLE 3 Response frequencies (%) for the EQ-5D-Y-3L ($N = 288$).

Dimensions	No problems	Some problems	Extreme problems/unable to do
Mobility	29 (10.07)	39 (13.54)	220 (76.39)
Self-care	27 (9.38)	56 (19.44)	205 (71.18)
Usual activities	18 (6.25)	72 (25.00)	198 (68.75)
Pain/discomfort	160 (55.56)	118 (40.97)	10 (3.47)
Anxiety/depression	150 (52.08)	128 (44.44)	10 (3.47)

TABLE 4 Response frequencies (%) for the CHU9D ($N = 288$).

Dimensions	Not/No. problems/any	A little bit/a few problems/most	A bit/some problem/some	Quite/many problems/ a few	Very/can't do/NO
Worried	101 (35.07)	67 (23.26)	33 (11.46)	44 (15.28)	43 (14.93)
Sad	143 (49.65)	3 (0.72)	5,820.14)	62 (21.53)	22 (7.64)
Pain	176 (61.11)	70 (24.31)	20 (6.94)	20 (6.94)	2 (0.69)
Tired	81 (28.13)	108 (37.50)	48 (16.67)	33 (11.46)	18 (6.25)
Annoyed	186 (64.6)	53 (18.40)	23 (7.99)	14 (4.86)	12 (4.17)
Schoolwork/ homework	139 (48.26)	42 (14.58)	29 (10.07)	17 (5.90)	61 (21.18)
Sleep	176 (61.11)	60 (20.83)	31 (10.76)	18 (6.25)	3 (1.04)
Daily routine	43 (14.93)	54 (18.75)	38 (13.19)	51 (17.70)	102 (35.42)
Able to join in activities	13 (4.51)	18 (6.25)	31 (10.76)	98 (34.03)	128 (44.44)

TABLE 5 HSUV derived from Type II and III SMA patients under 16 using EQ-5D-Y-3L and CHU9D.

Health state		EQ-5D-Y-3L utility		CHU9D utility	
	N (%)	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)
Not sitting	32 (11.11)	0.33 (0.07)	0.32 (0.04)	0.46 (0.22)	0.45 (0.36)
Sitting supported	52 (18.05)	0.37 (0.09)	0.36 (0.06)	0.55 (0.21)	0.55 (0.33)
Sitting Unsupported	95 (32.99)	0.39 (0.10)	0.36 (0.06)	0.61 (0.20)	0.62 (0.31)
Standing supported	18 (6.25)	0.47 (0.18)	0.39 (0.25)	0.65 (0.20)	0.71 (0.32)
Standing unsupported	19 (6.60)	0.50 (0.17)	0.40 (0.26)	0.65 (0.22)	0.70 (0.28)
Walking supported	12 (4.17)	0.53 (0.18)	0.46 (0.26)	0.70 (0.21)	0.60 (0.22)
Walking unsupported	60 (20.83)	0.82 (0.16)	0.85 (0.14)	0.75 (0.18)	0.82 (0.23)
Overall*	288 (100)	0.49 (0.21)	0.39 (0.30)	0.62 (0.21)	0.66 (0.33)

* ICC was 0.443, $p < 0.01$.

states. The basic information about demographics and diagnosis of the patients is summarized in Table 2.

HSUV of type II and III SMA patients younger than 16

Two hundred and eighty-eight patients/caregivers used EQ-5D-Y-3L and CHU9D to derive utility value. A total of 76.39%, 71.18% and 68.75% of the the patients/caregivers reported extreme problems or unable to do in three dimensions, which are mobility, self-care, and usual activities, when completing EQ-5D-Y-3L. However, less problems were reported in the pain/discomfort and anxiety/depression dimension. When completing CHU9D, the participants similarly reported many problems/can't do in the "daily routine" and "able to join in activities" dimension, which takes 53.12% and 78.47%. respectively. The response frequencies for EQ-5D-Y-3L and CHU9D are shown in Tables 3, 4.

The average utility values of the patients under 16 with Type II and III SMA is 0.49 and 0.62 derived from EQ-5D-Y-3L and CHU-9D. With motor functions increased from "Not Sitting" to "Walking unsupported", the average EQ-5D-Y-3L utility increased from 0.33 to 0.82, while the average CHU9D increased from 0.46 to 0.75. Notably, the CHU9D utility of "Standing supported" and "Standing unsupported" is the same, which is 0.65. ICC suggested that EQ-5D-Y and CHU9D utility indicated a moderate level of agreement. The HSUV of different health states are shown in Table 5.

HUSV of type II and III SMA patients above 16

One hundred and twenty-five patients/caregivers used EQ-5D-3L to derive utility value. Similar to patients under 16, 68.00%, 44.00%, and 54.40% of the patients above 16 reported extreme problems or unable to do in three dimensions, which are mobility, self-care, and usual activities. However, only 5 patients (4.00%) and 15 patients (12.00%) report extreme problems in the "Pain/Discomfort" and "Anxiety/Depression" dimensions, respectively. The response frequencies were shown in Table 6.

The average utility values of the patients above 16 with Type II and III SMA is 0.56. With motor functions increased from "Not Sitting" to "Walking unsupported", the average EQ-5D-3L utility increased from 0.30 to 0.83. The mean utility of "sitting unsupported" and "standing supported" is the same, which is 0.46. The HSUV of different health states are shown in Table 7.

TABLE 6 Response frequencies (%) for the EQ-5D-3L ($N = 125$).

Dimensions	No problems	Some problems	Extreme problems/ unable to do
Mobility	11 (8.80)	29 (23.20)	85 (68.00)
Self-care	38 (30.40)	32 (25.60)	55 (44.00)
Usual activities	11 (8.80)	46 (36.80)	68 (54.40)
Pain/Discomfort	53 (42.40)	67 (53.60)	5 (4.00)
Anxiety/Depression	36 (28.80)	74 (59.20)	15 (12.00)

TABLE 7 HSUV derived from Type II and III SMA patients above 16 using EQ-5D-3L ($N = 125$).

Health States		EQ-5D-3L utility	
	N (%)	Mean (SD)	Median (IQR)
Not sitting	18 (14.40)	0.30 (0.71)	0.32 (0.06)
Sitting supported	22 (17.60)	0.42 (0.17)	0.35 (0.26)
Sitting unsupported	30 (24.00)	0.46 (0.16)	0.38 (0.28)
Standing supported	5 (4.00)	0.46 (0.13)	0.39 (0.24)
Standing unsupported	5 (4.00)	0.52 (0.11)	0.57 (0.21)
Walking supported	8 (8.00)	0.70 (0.11)	0.70 (0.24)
Walking unsupported	37 (29.60)	0.83 (0.14)	0.85 (0.14)
Overall	125 (100.00)	0.56 (0.24)	0.57 (0.46)

Comparison of utility values of patients with different subtypes and age groups

Mann-Whitney U test was used to compare the utility values of patient of similar health states with Type I, Type II and III SMA, for example, by combing "Sitting supported" and "Sitting unsupported" in Type II and III together. The same test was applied to the utility value derived from patients under and above 16 years old as well. No statistical difference was found in both analyses. The utility difference between different subtypes were shown in Table 8 and the utility difference between patients under and above 16 were shown in Table 9.

Discussion

The present study used vignette approach and mixed patient/caregiver derived approach to estimate the utility of Chinese patients with Type I, II and III SMA in different health states. As Type I, II and III SMA are the three main subtypes of SMA (39), and we recruited patients nationwide, the study could reflect the utility preference of the patients with SMA in China. Generally, we found that utility values of patients with SMA were much lower than the Chinese population norm (0.985) (40). The utility values increased with the improvements of patients' motor functions, no matter the disease subtypes or the PBMs used to measure. This finding is generally consistent with the previous studies which reported the utility of SMA patients in different health states (19, 26, 28). However, "sit/walk unaided" was the best health state and "sit/walk with assistance" was the second-best health state in Lloyd et al. (26) and Biogen's submission to National Institute for Health and Care Excellence (NICE) (19), which resulted no difference in utility for "standing with assistance" and "walking with assistance", as well as "standing unaided" and "walking unaided" in the abovementioned two studies, respectively.

The utility values of patients in similar health states with Type I and Type II/III were slightly different but no statistical difference was found, especially the utility value of "sitting" state from Type I and Type II/III was the same. This finding may prove the validity of the vignette approach to some extent. Another possible reason was that, based on the feedbacks from pediatric neurologists participated in the valuation task, four of them claimed that they had never met a Type I SMA patient who could stand or walk, thus they brought in their experiences in treating Type II SMA patients. Furthermore, the clinicians may be familiar with several pediatric patients with SMA, which may confound their assessments of an individual case in the

TABLE 8 The difference between utility values of Type I, Type II and III SMA patients.

	Not sitting	Sitting	Standing	Walking
Type I	0.29 (0.08)	0.39 (0.19)	0.60 (0.21)	0.72 (0.21)
Type II/III	0.33 (0.33)	0.39 (0.94)	0.49 (0.18)	0.69 (0.25)
p value	0.09	0.57	0.15	0.46

TABLE 9 The difference between utility values of Type II and III SMA patients under and above 16.

	Not sitting	Sitting supported	Sitting unsupported	Standing supported	Standing unsupported	Walking supported	Walking unsupported
EQ-5D-Y-3L	0.33 (0.72)	0.37 (0.09)	0.39 (0.10)	0.47 (0.18)	0.50 (0.17)	0.53 (0.18)	0.82 (0.16)
EQ-5D-3L	0.30 (0.71)	0.42 (0.17)	0.46 (0.16)	0.46 (0.13)	0.52 (0.11)	0.70 (0.11)	0.83 (0.14)
p value	0.34	0.83	0.28	0.69	0.22	0.39	0.74

valuation task, which could be reason why NICE prefer utility data obtained from the person who acts as their caregiver rather than healthcare professionals (41).

We used two different PBMs for children and adolescents, which were CHU9D and EQ-5D-Y-3L to estimate the utility of adolescent patients with SMA. The EQ-5D-Y-3L utility value was on average 0.13 lower than the CHU9D utility value and the result showed CHU9D and EQ-5D-Y-3L had a moderate agreement in SMA patients. Similar results have been proven in a previous study in the population of cerebral palsy (42). We suggest that these two instruments may not be used inter-changeably to measure and value utility among children and adolescents with SMA. Though CHU9D provided another option to measure utility values in adolescents, further research would be necessary.

The utility data derived from children and adult patients with Type II and III SMA using the same value set was not statistically significant. The possible reason may be the same value set was used to derive utility. Though HTA bodies may consider adult value set acceptable for EQ-5D-Y-3L because of the lack of a validated value set, this goes against EuroQoL's recommendations (34). Japan was the only Asian country which had EQ-5D-3L and EQ-5D-Y-3L value set at the same time. The EQ-5D-Y-3L population norms for Japanese children and adolescents were lower than the EQ-5D-3L population norms for Japanese general population (43, 44), which indicated that using a proxy value set may result in the utility of adolescent patients with SMA being overestimated.

Our study has some strengths. To our best knowledge, this was the first study to estimate the utility of Chinese patients with SMA and we included patients of all the health states based on nature history of SMA. Compared with the similar studies involving patients with SMA or other rare diseases, this study had a considerably big sample size. We adopted different approaches and different reliable and validated instruments to comprehensively estimate the utility to support relevant health economic evaluations and reimbursement decisions in China as we strongly recommend that local utility data should be used to support Chinese reimbursement-decisions. Given that China is different from developed countries in socio-economic status and medical practices, it is of great significance to measure the utility of Chinse patients with SMA and further fulfill their unmet needs by valuing utility of patients with different motor functions.

Our study has several limitations worthy to notice as well. First, though vignettes in this study were developed rigorously following the good practices (31), its intrinsic methodological weakness couldn't be ignored that it may omit details that are meaningful to some patients. Furthermore, the pediatric neurologists valuated the HSUV of type I SMA patients using EQ-5D-Y proxy version. Although the applicable population age of 4–7 years is inconsistent with the age of younger patients, it is the most appropriate instrument available. Second, the patients were not evenly distributed in all the health states. The number of patients in "Standing supported", "Standing unsupported" and "Walking supported" were significantly less than that of the other four disease states, which was due to the nature history of the disease as introduced before and the extremely high price of the DMTs. Thirdly, we adopted a mixed patient/proxy-derived utilities for patients with SMA, which may result in unintentional bias from proxy assessments. For example, parents and caregivers usually have knowledge of one child with SMA, which may skew their interpretation of a particular health state (45). Fourthly, the using of Chinese value set for EQ-5D-3L in EQ-5D-Y-3L is a consequent limitation due to the lack of Chinese specific value set for EQ-5D-Y-3L. Finally, the influencing factors of the utility values were not fully analyzed, as our main objective was to report the utility value of different health states among Chinese patients with SMA.

Conclusion

To conclude, the better health states the patients with SMA were in, the higher were the health utilities. The utilities derived from population with different age and disease subtypes were not statistically different when patients with SMA were in the same health states. The utility generated from this study could be used in future health economic evaluations of intervention indicated for SMA. We also recommend further studies on the Chinese specific value set for EQ-5D-Y-3L and other PBMs for children to derive more robust utility data.

Data availability statement

The original contributions of this study and important findings are included in the text, further inquiries can be directed to the corresponding authors.

Author contributions

QK, CJ, and JH designed the survey questionnaire. HX and QK managed and coordinated the patient survey. JH and LZ developed the vignettes, was in charge of collecting survey data, and quality control. YL and HB analyzed and interpreted the survey data. JH, LZ, and YL contributed to draft the manuscript. All authors read and approved the final manuscript.

Funding

The study was funded by Shanghai Municipal Health Commission (Grant No. 20214Y0304) and Shanghai Roche Pharmaceuticals Ltd., (Grant No. 2021035B). The authors declare that this study received funding from Shanghai Roche Pharmaceuticals Ltd. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

Acknowledgments

We would like to thank all participants of the survey, including the SMA patients and their family members in China, for their generosity in sharing their lives and experiences *via* the survey. We would also like to thank all the clinicians participated in the study for their inputs. Dr. Andrew Lloyd is greatly appreciated for his kindness and expertise in inspiring the study design through personal communication.

Conflict of interest

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

Publisher's note

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

References

1. Verhaart IEC, Robertson A, Wilson IJ, Aartsma-Rus A, Cameron S, Jones CC, et al. Prevalence, incidence and carrier frequency of 5q-linked spinal muscular atrophy - a literature review. *Orphanet J Rare Dis.* (2017) 12:1–15. doi: 10.1186/s13023-017-0671-8

2. Lefebvre S, Bürglen L, Reboullet S, Clermont O, Burlet P, Viollet L, et al. Identification and characterization of a spinal muscular atrophy-determining gene. *Cell.* (1995) 80:155–65. doi: 10.1016/0092-8674(95)90460-3

3. Farrar MA, Park SB, Vucic S, Carey KA, Turner BJ, Gillingwater TH, et al. Emerging therapies and challenges in spinal muscular atrophy. *Ann Neurol.* (2017) 81:355–68. doi: 10.1002/ana.24864

4. Ahmad S, Bhatia K, Kannan A, Gangwani L. Molecular mechanisms of neurodegeneration in spinal muscular atrophy. *J Exp Neurosci.* (2016) 10:39–49. doi: 10.4137/JEN.S33122

5. Li C, Geng Y, Zhu X, Zhang L, Hong Z, Guo X, et al. The prevalence of spinal muscular atrophy carrier in China: evidences from epidemiological surveys. *Medicine.* (2020) 99:e18975. doi: 10.1097/MD.0000000000018975

6. Russman BS. Spinal muscular atrophy: Clinical classification and disease heterogeneity. *J Child Neurol.* (2007) 22:946–51. doi: 10.1177/0883073805673

7. Al Dakhoul S. Very severe spinal muscular atrophy (Type 0) avicenna. *J Med.* (2017) 07:32–3. doi: 10.4103/2231-0770.197512

8. Oskoui M, Levy G, Garland CJ, Gray JM, O'Hagen J, De Vivo DC, et al. The changing natural history of spinal muscular atrophy type 1. *Neurology.* (2007) 69:1931–6. doi: 10.1212/01.wnl.0000290830.40544.b9

9. Messina S, Pane M, De Rose P, Vasta I, Sorleti D, Aloysius A, et al. Feeding problems and malnutrition in spinal muscular atrophy type II. *Neuromuscul Disord.* (2008) 18:389–93. doi: 10.1016/j.nmd.2008.02.008

10. Zerres K, Rudnik-Schöneborn S, Forrest E, Lusakowska A, Borkowska J, Hausmanowa-Petrusewicz I, et al. collaborative study on the natural history of childhood and juvenile onset proximal spinal muscular atrophy (type II and III SMA): 569 patients. *J Neurol Sci.* (1997) 146:67–72. doi: 10.1016/S0022-510X(96)00284-5

11. Mercuri E, Finkel RS, Muntoni F, Wirth B, Montes J, Main M, et al. Diagnosis and management of spinal muscular atrophy: part 1: recommendations for diagnosis, rehabilitation, orthopedic and nutritional care. *Neuromuscul Disord.* (2018) 28:103–15. doi: 10.1016/j.nmd.2017.11.005

12. Finkel RS, Mercuri E, Meyer OH, Simonds AK, Schroth MK, Graham RJ, et al. Diagnosis and management of spinal muscular atrophy: part 2: pulmonary and acute care; medications, supplements and

immunizations; other organ systems; and ethics. *Neuromuscul Disord.* (2018) 28:197–207. doi: 10.1016/j.nmd.2017.11.004

13. Finkel RS, Mercuri E, Darras BT, Connolly AM, Kuntz NL, Kirschner J, et al. Nusinersen versus sham control in infantile-onset spinal muscular atrophy. *N Engl J Med.* (2017) 377:1723–32. doi: 10.1056/NEJMoa1702752

14. Ratni H, Ebeling M, Baird J, Bendels S, Bylund J, Chen KS, et al. Discovery of risdiplam, a selective survival of motor neuron-2 (SMN2) gene splicing modifier for the treatment of spinal muscular atrophy (SMA). *J Med Chem.* (2018) 61:6501–17. doi: 10.1021/acs.jmedchem.8b00741

15. NICE. *Risdiplam for Treating Spinal Muscular Atrophy in Children and Adults [ID1631]*. (2021). Available online at: https://www.nice.org.uk/consultations/1349/1/recommendations (accessed September 26, 2022).

16. Institute for Clinical and Economic Review. *Spinraza® and Zolgensma® for Spinal Muscular Atrophy: Effectiveness and Value - Final Evidence Report*, Vol. 2019. (2019). Available online at: https://icer.org/assessment/spinal-muscular-atrophy-2019/ (accessed September 26, 2022).

17. Zuluaga-Sanchez S, Teynor M, Knight C, Thompson R, Lundqvist T, Ekelund M, et al. Cost effectiveness of nusinersen in the treatment of patients with infantile-onset and later-onset spinal muscular atrophy in Sweden. *Pharmacoeconomics.* (2019) 37:845–65. doi: 10.1007/s40273-019-00769-6

18. Thokala P, Stevenson M, Kumar VM, Ren S, Ellis AG, Chapman RH. Cost effectiveness of nusinersen for patients with infantile-onset spinal muscular atrophy in US. *Cost Eff Resour Alloc.* (2020) 18:1–12. doi: 10.1186/s12962-020-00234-8

19. National Institute for Health and Care Excellence. *Single Technology Appraisal-Nusinersen for Treating Spinal Muscular Atrophy*. London (2019).

20. Malone DC, Dean R, Arjunji R, Jensen I, Cyr P, Miller B, et al. Cost-effectiveness analysis of using onasemnogene abeparvocec (AVXS-101) in spinal muscular atrophy type 1 patients. *J Mark Access Heal Policy.* (2019) 7:1601484. doi: 10.1080/20016689.2019.1601484

21. Canadian Agency for Drugs and Technologies in Health (CADTH*). Pharmacoeconomic Review Report - NUSINERSEN.* (2020). Available online at: https://cadth.ca/sites/default/files/cdr/pharmacoeconomic/sr0618-takhzyro-pharmacoeconomic-review-report.pdf (accessed September 26, 2022).

22. Neumann, Peter J; Goldie, Sue J; Weinstein MC. Preference-based measures in economic evaluation in health care. *Annu Rev Public Health.* (2000) (21):587–611. doi: 10.1146/annurev.publhealth.21.1.587

23. Chinese Pharmacoeconomics Evaluation Guideline Editing Group. C*hinese Pharmacoeconomics Evaluation Guideline, 2019 Version*. (2019). Available online at: http://proa28198.pic40.websiteonline.cn/upload/0ikw.pdf (accessed September 26, 2022).

24. Khan KA, Petrou S, Rivero-Arias O, Walters SJ, Boyle SE. Mapping EQ-5D utility scores from the PedsQL^{TM} generic core scales. *Pharmacoeconomics.* (2014) 32:693–706. doi: 10.1007/s40273-014-0153-y

25. Chambers GM, Settumba SN, Carey KA, Cairns A, Menezes MP, Ryan M, et al. Prenusinersen economic and health-related quality of life burden of spinal muscular atrophy. *Neurology.* (2020) 95:1–10. doi: 10.1212/WNL.0000000000009715

26. Lloyd AJ, Thompson R, Gallop K, Teynor M. Estimation of the quality of life benefits associated with treatment for spinal muscular atrophy. *Clin Outcomes Res.* (2019) 11:615–22. doi: 10.2147/CEOR.S214084

27. López-Bastida J, Peña-Longobardo LM, Aranda-Reneo I, Tizzano E, Sefton M, Oliva-Moreno J. Social/economic costs and health-related quality of life in patients with spinal muscular atrophy (SMA) in Spain. *Orphanet J Rare Dis.* (2017) 12:1–7. doi: 10.1186/s13023-017-0695-0

28. Lo SH, Gorni K, Sutherland CS, Martí Y, Lloyd A, Paracha N. Preferences and utilities for treatment attributes in type 2 and non-ambulatory type 3 spinal muscular atrophy in the United Kingdom. *Pharmacoeconomics.* (2022) 40:91–102. doi: 10.1007/s40273-021-01092-9

29. Yao M, Ma Y, Qian R, Xia Y, Yuan C, Bai G, et al. Quality of life of children with spinal muscular atrophy and their caregivers from the perspective of caregivers: a Chinese cross-sectional study. *Orphanet J Rare Dis.* (2021) 16:1–13. doi: 10.1186/s13023-020-01638-8

30. Paracha N, Gorni K, Hudson P. Pro62 spinal muscular atrophy: development of natural history models for disease subtypes. *Value Health.* (2020) 23:S339. doi: 10.1016/j.jval.2020.04.1284

31. Matza LS, Stewart KD, Lloyd AJ, Rowen D, Brazier JE. Vignette-Based Utilities: Usefulness, Limitations, and Methodological Recommendations. *Value Heal.* (2021) 24:812–21. doi: 10.1016/j.jval.2020.12.017

32. Trundell D, Le Scouiller S, Le Goff L, Gorni K, Vuillerot C. Assessment of the validity and reliability of the 32-item motor function measure in individuals with type 2 or non-ambulant type 3 spinal muscular atrophy. *PLoS ONE.* (2020) 15:1–12. doi: 10.1371/journal.pone.0238786

33. Pera MC, Coratti G, Forcina N, Mazzone ES, Scoto M, Montes J, et al. Content validity and clinical meaningfulness of the HFMSE in spinal muscular atrophy. *BMC Neurol.* (2017) 17:1–10. doi: 10.1186/s12883-017-0790-9

34. EuroQol Research Foundation. *EQ-5D-Y User Guide*. (2020). Available online at: https://euroqol.org/publications/user-guides (accessed September 26, 2022).

35. Chen G, Flynn T, Stevens K, Brazier J, Huynh E, Sawyer M, et al. Assessing the health-related quality of life of Australian adolescents: an empirical comparison of the child health utility 9D and EQ-5D-Y instruments. *Value Health.* (2015) 18:432–8. doi: 10.1016/j.jval.2015.02.014

36. Zhuo L, Xu L, Ye J, Sun S, Zhang Y, Burstrom K, et al. Time trade-off value set for EQ-5D-3L based on a nationally representative Chinese population survey. *Value Health.* (2018) 21:1330–7. doi: 10.1016/j.jval.2018.04.1370

37. Chen G, Xu F, Huynh E, Wang Z, Stevens K, Ratcliffe J. Scoring the child health utility 9D instrument: estimation of a Chinese child and adolescent-specific tariff. *Qual Life Res.* (2019) 28:163–76. doi: 10.1007/s11136-018-2032-z

38. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educ Psychol Meas.* (1973) 33:613–9. doi: 10.1177/001316447303300309

39. Jones CC, Cook SF, Jarecki J, Belter L, Reyna SP, Staropoli J, et al. Spinal muscular atrophy (SMA) subtype concordance in siblings: findings from the cure SMA cohort. *J Neuromuscul Dis.* (2020) 7:33–40. doi: 10.3233/JND-190399

40. Yao Q, Liu C, Zhang Y, Xu L. Population norms for the EQ-5D-3L in China derived from the 2013 National Health Services Survey. *J Glob Health.* (2021) 11:1–15. doi: 10.7189/jogh.11.08001

41. National Institute for Health and Care Excellence. *NICE Health Technology Evaluations: The Manual*. London (2022) 1–181.

42. Ryan JM, McKay E, Anokye N, Noorkoiv M, Theis N, Lavelle G. Comparison of the CHU-9D and the EQ-5D-Y instruments in children and young people with cerebral palsy: a cross-sectional study. *BMJ Open.* (2020) 10:1–11. doi: 10.1136/bmjopen-2020-037089

43. Shiroiwa T, Fukuda T, Ikeda S, Igarashi A, Noto S, Saito S, et al. Japanese population norms for preference-based measures: EQ-5D-3L, EQ-5D-5L, and SF-6D. *Qual Life Res.* (2016) 25:707–19. doi: 10.1007/s11136-015-1108-2

44. Shiroiwa T, Fukuda T. EQ-5D-Y population norms for Japanese children and adolescents. *Pharmacoeconomics.* (2021) 39:1299–308. doi: 10.1007/s40273-021-01063-0

45. Sutherland CS, Hudson P, Mitchell S, Paracha N. Systematic literature review to identify utility values in patients with spinal muscular atrophy (SMA) and their caregivers. *Pharmacoeconomics.* (2022) 40:39–67. doi: 10.1007/s40273-021-01115-5

frontiers | Frontiers in Pediatrics

TYPE Original Research
PUBLISHED 23 December 2022
DOI 10.3389/fped.2022.1057256

OPEN ACCESS

EDITED BY
Nan Luo,
National University of Singapore, Singapore

REVIEWED BY
Ralph C.A. Rippe,
Leiden University, Netherlands
Pei Wang,
Fudan University, China

*CORRESPONDENCE
Jana Willems
jana.willems@uniklinik-freiburg.de

†These authors share last authorship

SPECIALTY SECTION
This article was submitted to Children and Health, a section of the journal Frontiers in Pediatrics

RECEIVED 29 September 2022
ACCEPTED 07 December 2022
PUBLISHED 23 December 2022

CITATION
Willems J, Bablok I, Sehlbrede M, Farin-Glattacker E and Langer T (2022) The German pediatric integrated care survey (PICS-D): Translation, adaptation, and psychometric testing.
Front. Pediatr. 10:1057256.
doi: 10.3389/fped.2022.1057256

The German pediatric integrated care survey (PICS-D): Translation, adaptation, and psychometric testing

Jana Willems[1]*, Isabella Bablok[1], Matthias Sehlbrede[1], Erik Farin-Glattacker[1†] and Thorsten Langer[2†]

[1]Section of Health Care Research and Rehabilitation Research, Faculty of Medicine and Medical Center, University of Freiburg, Freiburg, Germany, [2]Department of Neuropediatrics and Muscle Disorders, Center for Pediatrics, Faculty of Medicine, University of Freiburg, Freiburg, Germany

Background: Integrated care models aim to strengthen the collaboration between healthcare sectors to ensure a higher quality of care for children with chronic conditions. Assessing pediatric healthcare integration through families' experiences, therefore, is essential. Our study aimed to psychometrically test the PICS-D, the first German-language integrated care questionnaire, which is based on the Pediatric Integrated Care Survey (PICS) developed in the USA.
Methods: We examined construct validity using exploratory and confirmatory factor analyses (structural validity). Cronbach's α and McDonald's ω_t coefficients explored reliability. Cognitive interviews assessed wording comprehension and item appropriateness.
Results: PICS-D was completed by 204 caregivers of children with chronic conditions (women = 84%; mean age = 41.4 years). Factor analyses identified 3 factors: "Team quality & communication", "Family impact", and "Access to care". The 3-factorial structure showed a satisfying fit to confirmatory classical-test-theory-based models. Due to the insufficient reliability of the third factor, we recommend using only factors 1 and 2 for scoring the PICS-D.
Conclusion: The PICS-D is a 13-item questionnaire to assess family-reported experiences in pediatric care integration, which has good psychometric properties. It may be useful in guiding quality improvement efforts or measuring the impact of a care plan or care model.
Trial registration: German Clinical Trials Register (DRKS): DRKS00018778; Trial registration date 05. December 2019—Retrospectively registered; https://apps.who.int/trialsearch/Trial2.aspx?TrialID=DRKS00018778.

KEYWORDS

chronic condition, reported outcome measures, integrated care, psychometric properties, chronic disease

Abbreviations
CFA, confirmatory factor analysis; CFI, comparative fit index; EFA, exploratory factor analysis; FACIT, functional assessment of chronic illness therapy translation procedures and guidelines; IC, integrated care; KMO, kaiser-mayer-olkin coefficient; PICS, pediatric integrated care survey; PICS-D, german version of the pediatric integrated care survey; RMSEA, root mean square error of approximation; SMA-C+, name of the overall project in which the study took place ("SMA" for spinal muscular atrophy; "C" for CM; "+" for adding an IT component to CM); TLI, Tucker–Lewis index; Velicer's MAP, Velicer's minimum average partial.

Introduction

One of the most persistent problems in many healthcare systems is the fragmentation between sectors and disciplines, which particularly affects patients with chronic and complex conditions who rely on services from a variety of disciplines (1–6). Integrated care (IC) models aim to strengthen the collaboration between sectors to ensure a higher quality of care (2, 7). The definition of IC varies widely, focusing primarily on adult care (8, 9). Tailoring these definitions to the pediatric setting is essential, as children have different healthcare and social needs (4, 10). By integrating the child's health perspective, Antonelli et al. extend existing definitions and offer the framework upon which this study is based: "*Integrated Care is the seamless provision of healthcare services, from the perspective of the patient and family, across entire care continuum. It results from coordinating the efforts of all providers, irrespective of institutional, departmental, or community-based organizational boundaries.*" (11–13). This definition implies that patient and family experiences are relevant when assessing IC in pediatric settings (6, 14). To date, evidence on the effectiveness of existing IC models (especially in adult care) has been predominantly specific for certain conditions (4, 8, 15–17). Other pediatric questionnaires, such as the EMPATHIC-30, cover only single facets of care integration (e.g., family-centered care experiences) and are limited to individual pediatric disciplines (18, 19). Generic evaluations often fail to adequately define goals and outcomes, and provide too little information on the availability of established measurement instruments (8, 11–13, 20). The Pediatric Integrated Care Survey (PICS) is a questionnaire for the caregivers of chronically ill children which measures the family-reported experience of pediatric care integration in the US (21). In German-speaking countries, instruments to measure IC currently exist only in the field of adult care (22–26)—a fact highlighting the need for a validated questionnaire to assess how well integrated pediatric healthcare is for families.

Materials and methods

Overview of the pediatric integrated care survey (PICS)

Antonelli et al. at Boston Children's Hospital developed the PICS in 2013. They formed and guided focus groups and interviews with parents of chronically ill children receiving care from multiple medical and non-medical providers (11, 21). It consists of 19 validated (mainly Likert-scaled) core items that form five scales (e.g., Access to care, Family impact, and Team functioning/performance/quality/connectivity). An example item in the "Family impact" scale is as follows: "In the past 12 months, how often have your child's care team members talked with you about how healthcare decisions for your child will affect your whole family?". The six-point response scale ranges from "never" to "always". The authors propose a 12-month reference period to evaluate IC. There are supplementary items that have not been validated due to sample size limitations in the Ziniel et al. study, as well as accompanying demographic and healthcare status questions (21). The supplementary items are grouped into so-called "modules" and "supplementary question topics". Antonelli et al. recommend selecting items in the set of supplementary questions in addition to the validated core items, depending on the focus of use (11).

Research aims

This study is part of the exploratory, prospective, controlled, two-armed SMA-C + -study, developed as an IT-supported Case Management to improve the care of patients with spinal muscular atrophy I and II (27). In this comprehensive project, we use the German version of the PICS to evaluate the Case Management as an IC intervention. Furthermore, we evaluated the questionnaire's psychometric properties by testing construct validity (as factorial/structural validity) and internal consistency. We, therefore, included the set of core items that were already validated in the English version.

Pre-data collection methods

The original PICS items' translation followed the six steps in the "Functional Assessment of Chronic Illness Therapy Translation Procedures and Guidelines" (FACIT): First, two native German speakers translated the instrument independently into German. A third native German speaker reviewed the translations and produced a preliminary version. A native English speaker then translated the preliminary version back into English. In subsequent steps, methodological and bilingual experts evaluated and discussed the different versions. In the final revision, the translated instrument is checked orthographically by a bilingual expert (28).

We conducted cognitive semi-structured interviews to assess wording comprehension and the item appropriateness of the preliminary PICS-D from the caregiver's perspective (29). For this purpose, we recruited $N = 10$ caregivers of chronically ill children in the Clinic of Neuropediatrics and Muscle Disorders of the University Medical Center Freiburg. After seven interviews, we revised problematic items based on their feedback. We then conducted another three interviews.

Pilot testing

We used the PICS-D for a cross-sectional study of family-reported experiences with IC involving caregivers of children

with a chronic disease. Taking the original authors' approach, we identified patients with chronic conditions based on their care needs. Patients are categorized as having a chronic condition if at least one of the following criteria is met: regular use of medications, regular use of therapies, use of auxiliary aids, regular counseling, or psychotherapy. We recruited $N = 204$ caregivers Germany-wide *via* personal contact from the University Medical Center, Freiburg; the Clinic for Pediatrics I, University Hospital, Essen; and several social pediatric centers (SPC). SPCs are interdisciplinary outpatient clinics that provide multidisciplinary care for children with complex healthcare needs. The participating SPCs are located in Bochum, Bremen, Celle, Freiburg, Lübeck, Rotenburg (Wuemme) and, Stuttgart. Caregivers were approached individually and were given the questionnaire and an information leaflet. In addition, the patient organization "Children's Network Germany" posted an invitation to participate on its website. The "Children's Network" represents pediatric patient organizations in Germany. All participants received a 20 € voucher per completed questionnaire. Exclusion criteria for participating in the questionnaire were not having a chronically ill child, limited German proficiency, or a patient's or caregiver's refusal to participate. Data collection took place between March 2019 and May 2020.

Statistical analyses

The statistical methods applied in psychometrically testing the PICS rely on the procedure described in the original publication and are supplemented by analyzing additional properties (21). Antonelli et al. recommend a top/bottom box scoring (11). The top/bottom boxes are the highest and lowest ratings (most positive/negative rating) on a response scale. Therefore, the item is "transformed" into a binary variable, and the respondent receives a box score of 1 indicating that he/she checked the desirable answer (most positive/negative answer). In contrast, we did not dichotomize the items and considered the response scale as an interval-scaled variable. The original authors indicate that the recommended scoring method is primarily for communicating quality improvement to healthcare institutions/professions, which is not a primary endpoint of our study. We described sample characteristics for children and respondents (caregivers) and assessed distribution properties, structural validity, and internal consistency (Cronbach's α and McDonald's ω_t). We performed descriptive analyses of the sample using IBM SPSS® and applied factor analyses in R (30).

Due to the content adjustment of the PICS-D, we refrained from confirmatory testing of the factor structure of the original PICS and conducted two exploratory factor analyses (EFA; extraction method: principal axis factoring after direct oblimin rotation) to identify concepts underlying the IC experience in German pediatric settings. Kaiser-Mayer-Olkin (KMO) coefficient and Bartlett sphericity test were used to analyze the suitability of the data for EFA. We calculated different criteria to determine the number of retained factors [Kaiser-Guttman criterion, parallel analysis and Velicer's MAP (31)]. We excluded three items on goal setting and access to medical records (original items 13, 30, and 31) from the beginning because the response scale deviated from a Likert scale (e.g., yes/no response format). Therefore, we included 16 Likert-scaled experience items (instead of the 19 original items) in the first EFA and 13 Likert-scaled items in the second EFA (another three items were excluded during the analysis process). We recoded response options such as "I don't know." or "I have no concerns." as missing values. Because the first EFA was mainly intended to provide exploratory support for selecting items, only the second EFA's results are reported in detail below.

To further analyze the PICS-D structure, we conducted two confirmatory factor analyses (CFA) using confirmatory classical-test-theory-based models. In each case, the CFAs followed an EFA that had been already been performed. We used the first CFA to obtain statistical criteria to exclude further items. The second CFA tested the performance of the final factor structure. Global measures of fit indicate the discrepancy between the data and hypothetical model (CFI, TLI, RMSEA) (32, 33). Bad fit of models may indicate violation of unidimensionality (i.e., item not related to the underlying concept) or insufficient item reliability. To determine the best dimensional structure, we conducted the CFAs on each factor as well as the whole model. Furthermore, we explored each factor's internal consistency *via* Cronbach's α and McDonald's ω_t coefficient.

Results

Translation and adaptation

This section describes modifications to the original core items for assessing pediatric care integration in Germany. **Table 1** shows the original 19 core items in both German and English translations (note: the latter does not match the wording of the original PICS items due to adaptations to the German health care system). The assignment of the original items, as well as their respective numbering, is found in the PICS "User Manual" (11). Most of the item changes resulted from the need to linguistically simplifying the wording, or adding information (concrete examples) for clarification. To ensure that the originally intended meaning of an item's wording remained intact, we consulted with the original authors at regular intervals. Moreover, cognitive interviews showed that the wording "care team" created confusion between the concepts of a "fixed team", that works cohesively as a unit and a team that

10.3389/fped.2022.1057256

TABLE 1 Translation of the 19 core items of the PICS-D.

No. of item in original PICS	Item wording of the PICS-D	Translated item wording of the PICS-D
13**	Hatten alle Behandelnden Zugang zu den für sie wichtigen medizinischen Unterlagen (z.B. Arztbriefe)? Wie ist Ihre Einschätzung?	Did all medical providers have access to the medical records that were important to them (e.g., doctors' letters)? What is your opinion?
17*	Wie oft hatten Sie Schwierigkeiten bzgl. medizinischer oder sozialer Dienste (z.B. Therapien, Integrationshilfe etc.), weil es Wartelisten oder andere Probleme bei der Terminvergabe gab?	How often did you have difficulties getting medical or social services (e.g., therapies, integration assistance, etc.), because there were waiting lists or other problems getting appointments?
18*	Wie oft hatten Sie Schwierigkeiten bzgl. medizinischer oder sozialer Dienste (z.B. Therapien, Integrationshilfe etc.), weil Sie nicht wussten, wer der richtige Ansprechpartner ist?	How often did you have difficulties getting medical or social services (e.g., therapies, integration assistance, etc.), because you did not know whom to contact?
22*	Wie oft haben die Mitglieder des Versorgungsnetzes Ihnen Dinge so erklärt, dass Sie sie verstehen konnten?	How often have the care network members explained things to you in a way that you could understand?
23*	Wie oft hatten Sie das Gefühl, dass Behandlungsempfehlungen zwischen den Mitgliedern des Versorgungsnetzes weitergegeben wurden?	How often did you feel that treatment recommendations were passed between the members of the care network?
24*	Manchmal haben Eltern Bedenken zur Gesundheit und Versorgung Ihres Kindes. Wie oft ist es Ihnen leicht gefallen, den Mitgliedern des Versorgungsnetzes Ihre Bedenken mitzuteilen?	Sometimes parents have concerns about their child's health and care. How often did you feel comfortable to share your concerns with the care network members?
25*	Wie oft hatten Sie das Gefühl, dass Sie von den Mitgliedern des Versorgungsnetzes gehört wurden, wenn Sie etwas über die Gesundheit Ihres Kindes zu sagen hatten?	How often did you feel that the members of the care network listened to what you had to say about your child's health?
26*	Wie oft wurde mit Ihnen besprochen, wer für die verschiedenen Bereiche der Versorgung Ihres Kindes verantwortlich ist?	How often has someone explained to you who was responsible for different parts of your child's care?
27*	Wie oft hatten Sie das Gefühl, dass die Mitglieder des Versorgungsnetzes über alle Tests und Untersuchungen Ihres Kindes informiert waren, um unnötige Untersuchungen zu vermeiden?	How often did you feel that the care network members were aware of all tests and evaluations of your child in order to avoid unnecessary testing?
28*	Wie oft hatten Sie das Gefühl, dass die Mitglieder des Versorgungsnetzes ihre Aufgaben in der Versorgung Ihres Kindes erfüllten?	How often did you feel that the care network members were fulfilling their roles in your child's care?
29*	Wie oft hatten Sie das Gefühl, dass die Mitglieder des Versorgungsnetzes bei der Betreuung Ihres Kindes seine Gesamtsituation berücksichtigt haben, d.h. alle Bedürfnisse Ihres Kindes?	How often have you felt that care network members thought about the "big picture" when caring for your child, meaning dealing with all of your child's needs?
30**	Haben die Mitglieder des Versorgungsnetzes kurzfristige Behandlungsziele für Ihr Kind gesetzt, d.h. Ziele bis zu 6 Monaten in der Zukunft?	Have the care network members created short-term care goals, meaning goals up to 6 months in the future?
31**	Haben die Mitglieder des Versorgungsnetzes langfristige Behandlungsziele für Ihr Kind gesetzt, d.h. Ziele, die 6 Monate oder länger in der Zukunft liegen?	Have the care network members created long-term care goals, meaning goals that are 6 months or more in the future?
32*	Wie oft haben die Mitglieder des Versorgungsnetzes Sie als vollwertige/n Partner/in behandelt?	How often have the care network members treated you as a full partner?
33*	Wie oft haben die Mitglieder des Versorgungsnetzes mit Ihnen darüber gesprochen, wie sich Entscheidungen über die Versorgung Ihres Kindes auf Ihre ganze Familie auswirken werden?	How often have the care network members talked with you about how care decisions for your child will affect your whole family?
34*	Wie oft haben die Mitglieder des Versorgungsnetzes mit Ihnen über Belastungen gesprochen, die sich für Sie durch die Erkrankung Ihres Kindes ergeben können?	How often have members of the care network talked with you about burdens you may face as a result of your child's illness?
35*	Wie oft haben die Mitglieder des Versorgungsnetzes Situationen angesprochen, die es Ihnen schwer machen können, sich um die Gesundheit Ihres Kindes zu kümmern (z.B. Arbeit, eigene Krankheit etc.)?	How often have the care network members addressed situations that may make it difficult for you to care for your child's health (e.g., work, own illness, etc.)?
36*	Wie oft haben die Mitglieder des Versorgungsnetzes angeboten, mit Ihnen auf andere Weise als durch einen persönlichen Besuch zu kommunizieren (z.B. Telefon, E-Mail oder Skype), wenn keine körperliche Untersuchung erforderlich war?	How often have the care network members offered to communicate with you in ways other than an in-person visit, such as phone, email, skype, if no physical examination was necessary?
37*	Wie oft haben Ihnen die Mitglieder des Versorgungsnetzes die Möglichkeit angeboten, mit anderen betroffenen Familien in Kontakt zu treten?	How often have the members of the care network offered you the opportunity to connect with other affected families?

*Items scored from 1 to 6: 1 = "never"; 2 = "rarely"; 3 = "sometimes"; 4 = "usually"; 5 = "almost always"; 6 = "always".
**Additional items: not included in scoring, with two answer modalities: "yes" "no".

results from different responsibilities within the child's care. Caregivers reported that the care of their children had no "fixed team". That is why we added an instruction defining the care team as all individuals involved in caring for the chronically ill child in the past 12 months (called the "care network"). After finalizing the questionnaire, we discussed the items with the original authors *via* videoconference. The translated PICS-version contained 19 core items on experiences (Table 1), a demographic and descriptive section including 32 items, and one open-ended item.

Sample characteristics

204 caregivers of children with chronic conditions completed the PICS-D, 84.3% of participants were female, 14.7% male. Two-thirds of respondents were between 30 and 50 years old and married. Slightly more than half of the participants were employed either part-time, not employed or not capable of gainful employment. The vast majority of the respondents' chronically ill children were between 1 and 12 years old. In 47% of participants, 2–5 health care providers were involved in the child's care. For around one-third of respondents, their care network even consisted of 6–10 health care providers. Further characteristics are detailed in Table 2.

Structural validity

The KMO coefficient (KMO = .89) and Bartlett sphericity test ($\chi^2 = 1427.29$, $p < .001$) indicated that the data we collected are suitable for the EFA. Parallel analysis and Kaiser-Guttman criterion suggest a three-factor solution, MAP two factors. We compared both EFA models and selected the three-factor solution because of a substantial better fit (RMSA < .08, SRMR < .05) in comparison to the two-factor solution. The three factors solution explained 53% of variance with low factor loadings on item 24 (share concerns with care network) and item 36 (communicate in ways other than an in-person visit). Factor loadings for the two items were item 24: 0.23, 0.19, and 0.10 in factors 1, 2, and 3, respectively and item 36: 0.33, 0.26, and 0.03 in factor 1, 2, and 3, respectively. Cognitive interviews revealed that both items were poorly understood due to their length and complicated syntax and we therefore excluded these items from our analysis. After exclusion, we conducted a first CFA for factor 1 and 2 that revealed high residual correlation within factor 1 for the item 27 with the items 23 and 25. Based on statistical and clinical considerations, we thus eliminated that item and then repeated the EFA and CFA analysis with 13 items.

The three factors of the second EFA for the PICS-D items explained 58% of the variance. The items' standardized loadings were ≥0.53 on factor 1 ("Team quality &

TABLE 2 Child and caregiver descriptive characteristics.

Variable	%
Respondent gender (*n* = 202)	
Female	84.3
Male	14.7
Respondent age at questionnaire completion (*n* = 201)	
<30 years.	7.5
30–40 years.	40.8
40–50 years.	35.3
>50 years.	15.4
Respondent relationship to child (*n* = 203)	
Mother	80.4
Father	13.7
Other	5.5
Respondent education (*n* = 203)	
Primary school. secondary school and secondary modern	10.3
High school	6.4
Completed training	38.4
University degree (Bachelor. Master. Doctorate)	22.7
Other	22.3
Respondent family status (*n* = 202)	
Single	5.4
Married	77.7
Living in a steady partnership	5.4
Divorced. separated	10.4
Widowed	1.0
Respondent employment status (*n* = 201)	
Employee full-time	18.4
Employee part-time	41.8
Civil servant	6.0
Self-employed	8.5
Not gainfully employed or capable of gainful employment	15.4
Other	10.0
Child gender (*n* = 204)	
Female	49.5
Male	50.5
Child age at questionnaire completion (*n* = 202)	
<1 year.	2.5
1–3 years.	20.3

(*continued*)

TABLE 2 Continued

Variable	%
3–12 years.	44.1
12–18 years.	24.3
>18 years.	8.9
Number of healthcare providers involved in the care of the child ($n = 202$)	
2–5	47.0
6–10	38.1
11–15	10.4
>15	4.5
Health insurance ($n = 199$)	
Statutory	75.4
Statutory with additional insurance	12.1
Private	12.1
I do not know.	0.5
Primary language spoken at home ($n = 203$)	
German	85.2
Other	14.8

communication") for items 22, 23, 25, 26, 28, 29, and 32; ≥0.40 on factor 2 ("Family impact") for items 33, 34, 35, and 27; and ≥0.71 on factor 3 ("Access to care") for items 17 and 18. All factor loadings can be found in **Table 3**.

We tested the model fit of the congeneric model for factors 1 and 2. The congeneric model is the least restrictive model of classical test theory, in which all items belong to one latent dimension. The fit indices suggested a good fit between the model and the data for factor 1 (CFI = 0.96, TLI = 0.94, RMSEA = 0.10) and a reasonable fit for factor 2 (CFI = 0.99, TLI = 0.96, RMSEA = 0.12). Since factor 3 consists of only two items, we used a parallel model to test unidimensionality because less restrictive models require more items. The parallel model combines the characteristics of the congeneric model with the assumption that correlations and measurement errors should be the same above all items. The fit indices suggested a good fit between the model and data for factor 3 (CFI = 1.00, TLI = 1.00, RMSEA < 0.01). Examining the entire three-factorial model, the fit indices suggested a good fit between the model and the data (CFI = 0.95, TLI = 0.94, RMSEA = 0.07).

Looking at the content of the items loading on factor 1, we observed that the focus was primarily on the interaction between families and health care providers. Simultaneously, factor 1 included items on the quality of communication in the care network. Therefore, we named factor 1 "Team quality & communication". Items that loaded on factor 2 mainly addressed the impact of care (decisions) on families' daily lives. For this reason, we designated factor 2 "Family impact". Factor 3 consisted of two items that concerned the access to care, so we named this factor "Access to care".

Reliability

Cronbach's α was 0.89 and 0.84 in factors 1 and 2, respectively. Due to the absence of essential tau-equivalence in the model test, Cronbach's α is not interpretable as a reliability measure. We thus also report McDonald's ω_t. McDonald's ω_t was 0.93 and 0.88 in factors 1 and 2, respectively. Since factor 3 contains only two items, we could not calculate Cronbach's α or McDonald's ω_t. As the inter-item correlation with $r = 0.55$ can be used as a reliability value, this factor shows insufficient reliability.

Discussion

In this study, we developed and psychometrically tested a German version of the Pediatric Integrated Care Survey (PICS), the so-called PICS-D. We can provide evidence that the PICS-D is a reliable and valid instrument by which to assess the experiences of IC as reported by the caregivers of chronically ill children. It takes a multidimensional approach and explores several crucial aspects of IC, categorized into three factors. Factor 1 "Team quality & communication" reflects the perceived quality of care and the cooperation within the care network (between all persons involved in caring for the chronically ill child in the last 12 months at the time of reporting), e.g., exchange of treatment recommendations, fulfillment of tasks involved in the care (2). Moreover, factor 1 analyzes the experiences concerning the communication between healthcare professionals and caregivers and the degree of the caregiver's involvement in decisions concerning the child's health, e.g., the feeling of being treated as a full partner (6, 22). Factor 2 "Family impact" evaluates the influence of care decisions and interventions on the entire family, e.g., burdens arising from the child's illness, opportunities for networking with other affected families (34, 35). Factor 3 "Access to care" investigated key elements of access to care, e.g., difficulties coordinating appointments (4, 36). Due to the insufficient reliability of the third PICS-D factor, we recommend using only factors 1 and 2 for scoring the PICS-D. To address this issue, it is necessary to develop a new associated item pool to re-examine this factor *via* factor analysis. More than four items would be desirable to perform in-depth statistical analyses. To obtain a rudimentary comparison with the psychometric properties of the original PICS, we used Cronbach's α for the two respective largest factors (21). The first PICS-D factor's internal consistency ("Team quality & communication"; Cronbach's $\alpha = 0.89$) is slightly higher compared to the first original PICS factor ("Communication between health care

TABLE 3 Exploratory factor analysis of psychometrically tested, Likert-scaled PICS items.

No. of item in original PICS	Translated item wording of the PICS-D	Factor 1	Factor 2	Factor 3
22	How often have the care network members explained things to you in a way that you could understand?	*0.63*	−0.11	−0.03
23	How often did you feel that treatment recommendations were passed between the members of the care network?	*0.53*	0.22	0.01
25	How often did you feel that the members of the care network listened to what you had to say about your child's health?	*0.88*	−0.06	0.10
26	How often has someone explained to you who was responsible for different parts of your child's care?	*0.61*	0.28	0.00
28	How often did you feel that the care network members were fulfilling their roles in your child's care?	*0.78*	−0.05	−0.08
29	How often have you felt that care network members thought about the "big picture" when caring for your child, meaning dealing with all of your child's needs?	*0.61*	0.20	−0.12
32	How often have the care network members treated you as a full partner?	*0.68*	0.02	−0.11
33	How often have the care network members talked with you about how care decisions for your child will affect your whole family?	0.15	*0.70*	−0.12
34	How often have members of the care network talked with you about burdens you may face as a result of your child's illness?	0.01	*0.92*	−0.02
35	How often have the care network members addressed situations that may make it difficult for you to care for your child's health (e.g., work, own illness, etc.)?	−0.04	*0.91*	0.07
37	How often have the members of the care network offered you the opportunity to connect with other affected families?	0.05	*0.40*	−0.02
17	How often did you have difficulties getting medical or social services (e.g., therapies, integration assistance, etc.), because there were waiting lists or other problems getting appointments?	0.07	−0.04	*0.77*
18	How often did you have difficulties getting medical or social services (e.g., therapies, integration assistance, etc.), because you did not know who to contact?	−0.11	0.06	*0.71*

provider and parent"; Cronbach's $\alpha = 0.80$). In addition, the second PICS-D factor's internal consistency ("Family impact"; Cronbach's $\alpha = 0.84$) is higher compared to the second original PICS factor ("Family impact", Cronbach's $\alpha = 0.72$).

The project team reviewed the items excluded from the analyses at both methodological and content levels. They seem to be well captured by the other items' content in the factors. A possible explanation for misunderstanding an item's wording could be the translation: In German, items tend to be longer and less commonly used, possibly triggering different latent constructs. However, the recovered factor structure in PICS-D appears to be topically similar to the original PICS' factor structure (except for the excluded items) (11, 21). Factors 2 and 3 consist of identical items such as the "Access to care"- and "Family impact"-factors in the original PICS. Factor 1 represents a composite of the original factors "Communication between health care provider and parent" and "Team functioning/performance/quality/connectivity".

Several strengths must be underlined. As far as we know, the PICS-D is the first and currently only German instrument by which to assess the perceived IC quality reported by caregivers of chronically ill children. This study examined its psychometric properties using a relevant nationwide sample. Our project provides compelling support for the critical importance of the role of patients and families in the design of health care delivery models, and in the evaluation of the performance of those models (37). The PICS-D is an initial step toward evaluating possible IC interventions to improve care for chronically ill children and increase caregivers' involvement in the care process. We used rigorous pre- and post-data collection procedures to enable the PICS' transferability to the German healthcare context (extensive translation process; two phases of cognitive interviews with intermittent and final revisions of the PICS-D; multiple virtual meetings with original authors to ensure original meanings of items; combining methodological and qualitative considerations in excluding items, etc.).

Our study has some limitations. Our sample size is at the lower limit of feasibility for psychometric testing. However, through intensive data management, we were able to ensure that the number of missing values was kept to a minimum to guarantee the feasibility of all our statistical analyses. Our assessment of the questionnaire's additional psychometrically important properties (e.g., other facets of construct validity, criterion validity, test-retest reliability, etc.) requires further methodological efforts. It is therefore essential to replicate the

questionnaire's structure working with other samples, especially with regard to the factor structure, since factors with only two items are statistically problematic and should only interpreted with care (38, 39). The original PICS was developed in the US healthcare system context by relying on interviews and focus groups with patients and caregivers. The scanned literature contained various definitions of care integration (2, 4, 6, 8, 11–13, 20, 21, 37). Although these are similar in essence, we could not rule out that different latent IC facets are triggered when answering the PICS-D. To design a valid measurement instrument, it is essential to develop a profound theoretical basis. Further psychometric testing in other countries is needed to refine a theoretical IC framework in pediatric settings by considering different experiences in international health contexts. Our study may support the further development of just such a framework.

Conclusion

The PICS-D is the first German-language questionnaire to assess the integration of pediatric care. Thanks to this study's positive findings, we can recommend the use of the PICS-D. Ziniel and co-authors describe a broad range of applications for the original PICS (21). Accordingly, the PICS-D can also be used to identify gaps in care delivery, guide quality improvement efforts, or measure the impact of a care plan or care model. Used under consideration of its framework, the results generated from our study reveal benefits for researchers, decision-makers, and field practitioners alike.

Principal axis factoring after direct oblimin rotation. Item information was summarized in 3 factors: Team quality & communication (factor 1), Family impact (factor 2), and Access to care (factor 3). Values are standardized factor loading: a higher value indicates a strong correlation with the corresponding factor. Values greater or equal 0.40 are in italics.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by Ethik-Kommission, Albert-Ludwigs-Universität Freiburg. The patients/participants provided their written informed consent to participate in this study.

Author contributions

Initiation, conception, design of the research project was done by TL and EFG. Coordination of the research project was done by TL, EFG and JW. Material preparation, data collection and analysis were performed by all authors. The first draft of the manuscript was written by JW and all authors commented on previous versions of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This study is being funded by the Innovationsausschuss of the Gemeinsamer Bundesausschuss, Wegelystraße 8, 10623 Berlin (grant number: 01VSF18005). The funding body was not involved in any aspect of the study's design, in collecting study data, in writing the manuscript, or in the decision to submit this article for publication.

Acknowledgments

We are grateful for the open exchange and helpful comments from the original authors of the PICS survey Richard Antonelli and Sonja Ziniel throughout the study. We also would like to thank Michaela Ruf for her assistance and valuable discussions in planning and developing this study. We acknowledge support by the Open Access Publication Fund of the University of Freiburg.

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. Gerlach FM, Schaeffer D, Greiner W, Thürmann PA, Wille E, Haubitz M, et al. *Wettbewerb an der schnittstelle zwischen ambulanter und stationärer gesundheitsversorgung. Bericht des sachverständigenrates zur begutachtung der entwicklung im gesundheitswesen*. Bonn: NOMOS; 2012 (2012).

2. Sterly C, Hassler M. Integrierte versorgung. In: CH Thielscher, editors. *Medizinökonomie*. Wiesbaden: Gabler Verlag | Springer Fachmedien (2012). p. S.483–504.

3. Mühlbacher A. *Integrierte versorgung. Management und organisation*. Hans Huber: Bern u.a. (2002).

4. Wolfe I, Lemer C, Cass H. Integrated care: a solution for improving children's Health? *Arch Dis Child*. (2016) 101(11):992–7. doi: 10.1136/archdischild-2013-304442

5. World Health Organisation. *Integrated care models: An overview*. Geneva, Switzerland: World Health Organisation (2016). Verfügbar unter: www.euro.who.int/__data/assets/pdf_file/0005/322475/Integrated-caremodels- overview.pdf

6. Singer SJ, Burgers J, Friedberg M, Rosenthal MB, Leape L, Schneider E. Defining and measuring integrated patient care: promoting the next frontier in health care delivery. *Med Care Res Rev*. (2011) 68(1):112–27. doi: 10.1177/1077558710371485

7. Goodwin N. Understanding integrated care. *Int J Integr Care*. (2016) 16 (4):1–4. doi: 10.5334/ijic.2530. Verfügbar unter: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5354214/

8. Wolfe I, Satherley RM, Scotney E, Newham J, Lingam R. Integrated care models and child health: a meta-analysis. *Pediatrics*. (2020) 145(1):1–12. doi: 10.1542/peds.2018-3747. Verfügbar unter: https://pediatrics.aappublications.org/content/145/1/e20183747

9. Armitage GD, Suter E, Oelke ND, Adair CE. Health systems integration: state of the evidence. *Int J Integr Care*. (2009) 9(2):1–11. doi: 10.5334/ijic.316. Verfügbar unter: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2707589/

10. Schor EL. American Academy of pediatrics task force on the family. Family pediatrics: report of the task force on the family. *Pediatrics*. (2003) 111(6 Pt 2):1541–71.

11. Antonelli R. *Pediatric integrated care survey 1.0 user manual*. Boston: Boston Children's Hospital (2015).

12. Antonelli R, Mcallister J, Popp J. Making care coordination a critical component of the pediatric health system: a multidisciplinary framework (2009). Verfügbar unter: https://www.commonwealthfund.org/publications/fund-reports/2009/may/making-care-coordination-critical-component-pediatric-health

13. Turchi R, Antonelli R, Norwood K, Adams R, Brei T. Burke R. Patient-and family-centered care coordination: a framework for integrating care for children and youth across multiple systems. *Pediatrics*. (2014) 133:E1451–60.

14. Eikötter T, Greiner W. Instrumente zur messung der versorgungsqualität in der integrierten versorgung. *Gesundheitsökonomie Qualitätsmanagement*. (2008) 13(1):25–31. doi: 10.1055/s-2007-963333

15. Wolfe I. Health services for children with long-term conditions and non-communicable diseases. In: I Wolfe, M McKee, Herausgeber, editors. *European Child health services and systems: lessons without borders*. Maidenhead: Open University Press (2013). p. S.63–93.

16. Weatherly JN, Lägel R, Herausgeber. *Neue versorgungsansätze in der psychiatrie, neurologie und psychosomatik*. Berlin: Medizinisch Wissenschaftliche Verlagsgesellschaft (2009).

17. Asarnow JR, Rozenman M, Wiblin J, Zeltzer L. Integrated medical-behavioral care compared with usual primary care for child and adolescent behavioral health: a meta-analysis. *JAMA Pediatr*. (2015) 169(10):929–37. doi: 10.1001/jamapediatrics.2015.1141

18. Latour JM, Duivenvoorden HJ, Tibboel D, Hazelzet JA, EMPATHIC Study Group. The shortened EMpowerment of PArents in THe intensive care 30 questionnaire adequately measured parent satisfaction in pediatric intensive care units. *J Clin Epidemiol*. (2013) 66(9):1045–50. doi: 10.1016/j.jclinepi.2013.02.010

19. Girch A, Rippe RCA, Latour JM, Jönebratt Stocker M, Blendermann M, Hoffmann K. The German EMPATHIC-30 questionnaire showed reliability and convergent validity for use in an intermediary/general pediatric cardiology unit: a psychometric evaluation. *Front Cardiovasc Med*. (2022) 9:901260. doi: 10.3389/fcvm.2022.901260

20. Cohen E, Coller RJ. Evaluating integrated care for children: a clarion call or a call for clarity? *Pediatrics*. (2020) 145(1):1–2. doi: 10.1542/peds.2019-3282. Verfügbar unter: https://pediatrics.aappublications.org/content/145/1/e20193282

21. Ziniel SI, Rosenberg HN, Bach AM, Singer SJ, Antonelli RC. Validation of a parent-reported experience measure of integrated care. *Pediatrics*. (2016) 138 (6):1–11. doi: 10.1542/peds.2016-0676

22. Noest S, Ludt S, Klingenberg A, Glassen K, Heiss F. Ose D. Involving patients in detecting quality gaps in a fragmented healthcare system: development of a questionnaire for Patients' experiences across health care sectors (PEACS). *Int J Qual Health Care*. (2014) 26(3):240–9. doi: 10.1093/intqhc/mzu044

23. Hildebrandt H, Pimperl A, Schulte T, Hermann C, Riedel H. Schubert I. Triple aim—evaluation in der integrierten versorgung gesundes kinzigtal—gesundheitszustand, versorgungserleben und wirtschaftlichkeit. *Bundesgesundheitsbl*. (2015) 58(4):383–92. doi: 10.1007/s00103-015-2120-y

24. Klingenberg A, Bahrs O, Szecsenyi J. How do patients evaluate general practice? German results from the European project on patient evaluation of general practice care (EUROPEP). *Zeitschrift für ärztliche Fortbildung und Qualitätssicherung*. (1999) 93:437–45.

25. Braun S, Kreimeier S, Greiner W. Messung der patientenzufriedenheit in der integrierten versorgung—eine pilotstudie mit dem modifizierten ZAP-fragebogen. Zeitschrift für evidenz. *Fortbildung und Qualität im Gesundheitswesen*. (2010) 104 (2):106–12. doi: 10.1016/j.zefq.2009.12.001

26. Bethge M, Bartel S, Streibelt M, Lassahn C, Thren K. Verbesserte behandlungsqualität durch integrierte versorgung bei knie- und hüftgelenkersatz: ergebnisse einer kontrollierten studie. *Rehabilitation*. (2011) 50 (2):86–93. doi: 10.1055/s-0030-1265144

27. Willems J, Farin-Glattacker E, Langer T. Evaluation of a case management to support families with children diagnosed with spinal muscular atrophy—protocol of a controlled mixed-methods study. *Front Pediatr*. (2021) 9:614512. doi: 10.3389/fped.2021.614512

28. Eremenco SL, Cella D, Arnold BJ. A comprehensive method for the translation and cross-cultural validation of health status questionnaires. *Eval Health Prof. Juni*. (2005) 28(2):212–32. doi: 10.1177/0163278705275342

29. Willis GB. *Cognitive interviewing: a tool for improving questionnaire design*. California, USA: SAGE Publications (2004). 375 S.

30. R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing (2020). Verfügbar unter: https://www.R-project.org/

31. Velicer WF. Determining the number of components from the matrix of partial correlations. *Psychometrika*. (1976) 41:321–7. doi: 10.1007/BF02293557

32. Hooper D, Coughlan J, Mullen M. Structural equation modeling: guidelines for determining model fit. *Electron J Bus Res Methods*. (2007) 30:6.

33. Hu Lt, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling Multidiscip J*. (1999) 6(1):1–55. doi: 10.1080/10705519909540118

34. Hjorth E, Kreicbergs U, Sejersen T, Jeppesen J, Werlauff U, Rahbek J. Bereaved parents more satisfied with the care given to their child with severe spinal muscular atrophy than nonbereaved. *J Child Neurol*. (2019) 34(2):104–12. doi: 10.1177/0883073818811544

35. Mastellos N, Gunn L, Harris M, Majeed A, Car J, Pappas Y. Assessing patients' experience of integrated care: a survey of patient views in the north west London integrated care pilot. *Int J Integr Care*. (2014) 14:1–9. doi: 10.5334/ijic.1453

36. Altman L, Zurynski Y, Breen C, Hoffmann T, Woolfenden S. A qualitative study of health care providers' perceptions and experiences of working together to care for children with medical complexity (CMC). *BMC Health Serv Res*. (2018) 18 (1):70. doi: 10.1186/s12913-018-2857-8

37. Bautista MAC, Nurjono M, Lim YW, Dessers E, Vrijhoef HJ. Instruments measuring integrated care: a systematic review of measurement properties. *Milbank Q*. (2016) 94(4):862–917. doi: 10.1111/1468-0009.12233

38. Marsh HW, Hau KT, Balla JR, Grayson D. Is more ever too much? The number of indicators per factor in confirmatory factor analysis. *Multivariate Behav Res*. (1998) 33(2):181–220. doi: 10.1207/s15327906mbr3302_1

39. Yong A, Pearce S. A Beginner's guide to factor analysis: focusing on exploratory factor analysis. *Tutor Quant Methods Psychol*. (2013) 9:79–94. doi: 10.20982/tqmp.09.2.p079

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 23 December 2022
DOI 10.3389/fpubh.2022.1056967

OPEN ACCESS

EDITED BY
Pei Wang,
Fudan University, China

REVIEWED BY
Arlette Setiawan,
Universitas Padjadjaran, Indonesia
Valeria Di Onofrio,
University of Naples Parthenope, Italy

*CORRESPONDENCE
Dalia Al-Abdulrazzaq
dalia.alabdulrazzaq@ku.edu.kw

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

RECEIVED 29 September 2022
ACCEPTED 25 November 2022
PUBLISHED 23 December 2022

CITATION
Al-Abdulrazzaq D, Khalifa D, Alqaisi T, Al-Juailla F, Othman F, Qabazard S and Al-Kandari H (2022) Health-related quality of life of children and adolescents with type 1 diabetes during the COVID-19 pandemic in Kuwait.
Front. Public Health 10:1056967.
doi: 10.3389/fpubh.2022.1056967

Health-related quality of life of children and adolescents with type 1 diabetes during the COVID-19 pandemic in Kuwait

Dalia Al-Abdulrazzaq [1,2]*, Doaa Khalifa[2], Taiba Alqaisi[2], Fatima Al-Juailla[2], Fouzeyah Othman[2], Sarah Qabazard[2] and Hessa Al-Kandari[2,3]

[1]Department of Pediatrics, Faculty of Medicine, Kuwait University, Kuwait City, Kuwait, [2]Department of Population Health, Dasman Diabetes Institute, Kuwait City, Kuwait, [3]Department of Pediatrics, Al-Farwaniya Hospital, Kuwait City, Kuwait

Background: With the rapid transmission of COVID-19 globally, countries have implemented strict governmental measures and social distancing rules that aimed to minimize the spread of the virus. School closures, quarantine orders, and social isolation, coupled with a surge in family stress and lack of peer interactions, are probable causes of mental health complications and psychological symptoms faced by children. This study aims at comparing the HRQoL of children and adolescents with type 1 diabetes (T1D) and affected by COVID-19 infection (personal or familial) and those who were not affected by COVID-19.

Materials and methods: A random sample was selected from children and adolescents diagnosed with T1D at the six major governmental diabetes centers in Kuwait. To measure HRQoL, parent-proxy and self-reports from the Pediatric Quality of Life Inventory (PedsQL^{TM}) 3.0 Diabetes Module were used.

Results: A sample of 455 children and adolescents with T1D diabetes (44.6% male participants and 41.98% affected by COVID-19 infection) was studied. The total score of the HRQoL self-reports was significantly higher compared with parent-proxy reports (79.06 ± 15.19 vs. 73.79 ± 15.17, $p < 0.01$). Children reported significantly higher HRQoL scores in the "treatment I" domain and "worry" domain and lower scores in the "diabetes" symptoms domain, compared with their parents' reports ($p < 0.01$). In the COVID-19-affected group, a major difference was noticed between the total scores of children and parent-proxy reports (77.04 ± 15.81 vs. 72.80±14.90, $p = 0.047$). The affected children reported significantly lower scores in "diabetes" symptoms (59.50) ($p < 0.01$) and higher scores in the "treatment I" domain (81.43) than their parent-proxy reports (72.05) ($p < 0.01$).

Conclusion: This is the first report on the health-related quality of life of children with T1D in Kuwait during the COVID-19 era. Parents' or caregivers' experience of caring for their children was negative, as they worried, and reported poorer HRQoL compared with their children's perceptions. There is a need to empower healthcare professionals to support parents and caregivers

of children with chronic diseases such as T1D in promoting self-management, enhancing physical and psychological wellbeing, treatment adherence, and continuous health education during pandemics of any kind.

KEYWORDS

children, type 1 diabetes, health-related quality of life, PedsQL^{TM} Diabetes Module, COVID-19

Introduction

The COVID-19 pandemic has undoubtedly impacted all areas of the global population's daily lives, both physically and mentally, making it a worldwide healthcare disaster. With the rapid transmission of COVID-19 globally, countries have implemented strict governmental measures and social distancing rules that aimed to minimize the spread of the virus. School closures, quarantine orders, and social isolation, coupled with a surge in family stress and lack of peer interactions, are probable causes of mental health complications and psychological symptoms faced by children (1, 2). A recent meta-analysis using data from 80,879 youths globally, to investigate the pandemic's effect on mental health, reported a pooled prevalence estimate of clinically elevated depression and anxiety: 25.2% for children and 20.5% for adolescents. Compared with pre-pandemic estimates, the prevalence of depression and anxiety symptoms during COVID-19 has doubled (1). Mental health problems, particularly hyperactivity and peer problems, as well as psychosomatic complaints, such as sleeping problems and headaches, were also experienced by children during the pandemic (3).

Children with chronic illnesses faced additional challenges during the COVID-19 pandemic, including hindrance to accessing inpatient and outpatient clinical care. These limitations in care may have negatively impacted children with type 1 diabetes (T1D) since diabetes pediatric centers shifted to telemedicine (4). Families with children and adolescents with T1D were forced to change their approach to disease management as health services and delivery became significantly disrupted (4, 5). This may have resulted in delayed diabetes diagnosis, delayed insulin delivery, more episodes of diabetic ketoacidosis (DKA) and hyperglycemia, and admission to the pediatric intensive care unit (PICU); families also experienced fear of contracting COVID-19 infection while obtaining care (6). Furthermore, COVID-19 infection itself in individuals with preexisting diabetes has been associated with higher rates of hospitalization and admission to the intensive care unit (7, 8). These might indeed have a negative impact on an infected child's mental health and wellbeing. It might be also safe to assume that COVID-19 infection in a close family member might have a negative impact on the child, as the role of family dynamics and functioning influence diabetes management behavior and metabolic control (9).

Within and beyond the context of the pandemic, children with T1D face daily challenges, such as intensive therapeutic insulin regimes, dietary restrictions, and necessary physical activity (10). These children are also prone to the development of profound feelings of uncertainty, fear, or irritability that are a result of home confinement and disruptions in normal routine, which may ultimately influence their diabetes management (4, 11). These factors in combination make it imperative to address the emotional state of this population and assess their health-related quality of life (HRQoL) during the pandemic.

The assessment of HRQoL is crucial in clinical trials and healthcare as it measures the impact of diseases or disabilities on the mental, physical, and social dimensions of a patient's health (12, 13). Diabetes-specific HRQoL measurement instruments are critical for the assessment of symptoms and problems relevant to people living with diabetes, as well as the person's adherence to diabetes self-management (14, 15). Considering the obstacles of accessing healthcare and diabetes management as well as the emotional consequences resulting from the COVID-19 pandemic, it may be expected that pediatric patients with T1D have impaired HRQoL.

One objective of this study was to compare the HRQoL of children and adolescents (aged 2–18 years) with T1D and affected by COVID-19 infection (personal or familial) and those who were not affected by COVID-19. An additional objective is to measure the similarities and differences between the children's and parents' HRQoL reports during the COVID-19 pandemic.

2. Methods

2.1. Study design and subjects

Starting in January 2021, a random sample was selected from children and adolescents diagnosed with T1D between January 2011 and January 2020 at the six major governmental diabetes centers in Kuwait. Children aged 2–18 years who met the inclusion criteria: diabetes duration for at least 1 year, resident of Kuwait during the pandemic for at least 6 months, and Arabic-speaking parents and children, were included. Participants with other comorbidities (such as

developmental delay and chronic renal disease) apart from T1D and its complications, and parents who refused to participate were excluded.

A total of 656 families were approached; of these, 586 (89.3%) were eligible while 70 (10.6%) did not meet the inclusion criteria; among those eligible, 455 (77.6%) agreed to participate. Of the 455 families, 116 fathers, 334 mothers, and 3 legal guardians were interviewed. A total of 203 boys and 252 girls were enrolled, including 49 children aged 2–4 years, 74 children aged 5–7 years, 204 children aged 8–12 years, and 128 children aged 13–18 years at the time of interview; these age groups correspond with the HRQoL instrument's age categories.

2.2. Data collection

Interviews were performed from January to October 2021, after parental consent and child assent were obtained. Most families participated over the phone or *via* a web-based platform ($n = 433$), but 22 families were directly interviewed during clinical visits. The PedsQL^{TM} 3.0 Diabetes Module Arabic version, a validated questionnaire for children, was used to assess diabetes-specific HRQoL (16, 17). Before completion of the PedsQL^{TM} 3.0 forms, baseline data were retrieved from parents or caregivers regarding demographic characteristics, diabetes-related history, and COVID-19 infection history. Glycemic status was obtained from medical records in terms of glycated hemoglobin (HbA1c) within 3 months of the recruitment. Optimal control was defined as HbA1c of <7% as per the 2018 International Society of Pediatric and Adolescent Diabetes (ISPAD) guidelines, and HbA1C of above 9.0% was considered high risk (18). The study protocol was approved by the Standing Committee for Coordination of Health and Medical Research (Ethics Review Committee) at the Ministry of Health (MoH) of Kuwait (Reference No. 1588/2020).

2.3. PedsQL^{TM} 3.0 Diabetes module

The 28-item module is composed of five scales: (1) diabetes symptoms (11 items), (2) treatment barriers I (4 items), (3) treatment adherence II (7 items), (4) worry (3 items), and (5) communication (3 items). The module is designed for four age groups: parent-proxy reports for toddlers (2–4 years), parent-proxy and child self-reports for young children (5–7 years), school-age children (8–12 years), and teenagers (13–18 years). The PedsQL^{TM} items follow the five-point Likert scale from 0 (never) to 4 (almost always) for all age groups, except for young children (aged 5–7 years), which follows a three-point scale: 0 (not at all), 2 (sometimes), and 4 (a lot). To calculate the health-related quality of life (HRQoL), items were reverse-scored and linearly transformed on a scale from 0 to 100 (0 = 100, 1 = 75, 2 = 50, 3 = 25, and 4 = 0), and the sum of all items was divided over the number of items answered to calculate the total score. Higher scores indicate better HRQoL (19, 20).

The original PedsQL^{TM} 3.0 Diabetes module was translated into Arabic and validated on children with T1D and their parents/caregivers at Kuwait University, Kuwait, based on the linguistic validation guidelines of the PedsQL^{TM} Quality of Life Questionnaire and approved by MAPI Research Trust, on behalf of Dr. James W. Varni, the copyright owner of PedsQL^{TM} (17, 21).

2.4. Procedure

In preparation for the study, a comprehensive well-established data collection sheet was developed. Research assistants were trained as interviewers. Following patient identification, a pilot phase was implemented over a 1-month period to test the procedures with a convenient sample of 50 families. The interview process and data collection forms were adjusted after the pilot phase.

The interviews took 20–25 min. The survey consisted of five sections: personal information, diabetes history, COVID-19-related history, COVID-19 family history, and the PedsQL 3.0 questionnaires. The answers were documented by the interviewers based on parents'/children's responses.

To ensure proper enrollment, COVID-19 status in children was confirmed through the Pediatric COVID-19 Registry (PCR-Q8) in Kuwait (MoH Reference No. 2020/1233), a national registry that records all cases of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in children in the country since the start of the pandemic in February 2020.

2.5. Statistical analysis

Statistical analysis was conducted using STATA software 13.1 (STATA Corp, College Station, TX). Continuous variables were expressed as median and interquartile ranges (IQR), as they were non-normally distributed. The Kruskal–Wallis test was used to test for differences in the continuous variables. The HRQoL scores were expressed as mean and standard deviation (SD). Unpaired *t*-tests were used to compare total and dimension-specific HRQoL scores between groups (e.g., no COVID-19 infection vs. COVID-19 infection).

3. Results

A total of 455 children and adolescents [203 (44.6%) male participants and 252 (55.4%) female participants] with T1D were included. The baseline characteristics of study participants and COVID-19 infection status are shown in Table 1. Among the participants, 191 (41.98%) children were affected by COVID-19.

10.3389/fpubh.2022.1056967

TABLE 1 Baseline characteristics of children with T1D at HRQoL assessment.

	Total ($n = 455$)	Affected by COVID-19 ($n = 191$, 41.98%)	Unaffected by COVID-19 ($n = 264$, 58.02%)	*p*
	n (%)	*n (%)*	*n (%)*	
Age at interview				
2–4	49 (10.8)	15 (7.9)	34 (12.9)	0.359
5–7	74 (16.3)	31 (16.2)	43 (16.3)	
8–12	204 (44.8)	87 (45.5)	117 (44.3)	
13–18	128 (28.1)	58 (30.4)	70 (26.5)	
Median (IQR)	11.0 (7–13)	11.5 (8.3–13.5)	10.7 (7.2–13.1)	0.035*
Gender				
Male	203 (44.6)	75 (39.3)	128 (48.5)	0.051
Female	252 (55.4)	116 (60.7)	136 (51.5)	
Father education level				
None	0 (0.0)	0 (0.0)	0 (0.0)	0.461
Primary	5 (1.3)	2 (1.2)	3 (1.3)	
Intermediate	41 (10.5)	19 (11.7)	22 (9.6)	
Secondary	92 (23.5)	44 (27.2)	48 (20.7)	
Diploma	70 (17.9)	24 (14.8)	46 (20.1)	
University degree or above	183 (46.8)	73 (45.1)	110 (48.0)	
Mother education level				
None	3 (0.8)	2 (1.2)	1 (0.4)	0.932
Primary	2 (0.5)	1 (0.6)	1 (0.4)	
Intermediate	38 (9.6)	16 (9.6)	22 (9.6)	
Secondary	64 (16.2)	29 (17.5)	35 (15.3)	
Diploma	92 (23.3)	39 (23.5)	53 (23.1)	
University degree or above	196 (49.6)	79 (47.6)	117 (51.1)	
Total family income (KD)				
<1,000	104 (24.7)	35 (19.9)	69 (28.2)	0.031*
1,000–1,499	89 (21.1)	31 (17.6)	58 (23.7)	
1,500–1,999	54 (12.8)	23 (13.1)	31 (12.6)	
2,000 and above	105 (24.9)	56 (31.8)	49 (20.0)	
Refused to answer	69 (16.4)	31 (17.6)	38 (15.5)	
Household dynamics				
Dual-parent	383 (89.1)	164 (90.6)	219 (87.9)	0.384
Single-parent	47 (10.9)	17 (9.4)	30 (12.1)	
Diabetes duration				
≤4 years	203 (49.4)	82 (46.3)	121 (51.7)	0.280
>4 years	208 (50.6)	95 (53.7)	113 (48.3)	
Median (IQR)	4 (2–6)	4.21 (2.6–6.8)	3.8 (2.1–5.9)	0.106

(Continued)

TABLE 1 (Continued)

	Total ($n = 455$) *n (%)*	Affected by COVID-19 ($n = 191$, 41.98%) *n (%)*	Unaffected by COVID-19 ($n = 264$, 58.02%) *n (%)*	*p*
HbA1c				
≤7.0%	12 (5.0)	2 (1.9)	10 (7.3)	0.018*
7.0% ≤ 9.0%	65 (27.1)	22 (21.2)	43 (31.6)	
>9.0%	163 (67.9)	80 (76.9)	83 (61.0)	
Median (IQR)	9.7 (8.4–11.3)	10.1 (9.1–11.6)	9.4 (8.2–10.8)	<0.01*
Insulin regimen				
MDI	380 (85.6)	162 (85.7)	218 (85.5)	0.947
CSII	64 (14.4)	27 (14.3)	37 (14.5)	
T1D family history**				
Yes	108 (24.8)	45 (24.6)	63 (24.9)	0.941
No	328 (75.2)	138 (75.4)	190 (75.1)	

T1D, type 1 diabetes; IQR, interquartile range; KD, Kuwaiti Dinar; HbA1c, Hemoglobin A1c; MDI, multiple daily injections; CSII, continuous subcutaneous insulin infusion; HRQoL, Health-Related Quality of Life; COVID-19, coronavirus disease of 2019.

*$p < 0.05$.

**Defined as a participant having a first-degree relative diagnosed with Type 1 diabetes.

Children who had had COVID-19 were significantly older at the time of the study than those who had not (11.5 vs. 10.7 years, $p = 0.035$). The majority (31.8%) of the COVID-19-affected children were from families with incomes of 2,000 KD (equivalent to approximately €6,000) and above ($p = 0.031$). The majority (76.9%) of the children affected by COVID-19 had glycemic control in the high-risk category ($p = 0.018$), with a significantly higher HbA1c than those who had not been infected with COVID-19 [10.1 (9.1;11.6) vs. 9.4 (8.2;10.8), $p < 0.01$]. Children who had and had not been affected by COVID-19 were not different in terms of gender, parents' education, household dynamics, diabetes duration, insulin regimen, or family history of T1D. HRQoL was collected for 620 participants based on 165 child self-reports and 455 parent-proxy reports from the PedsQL™ 3.0 Diabetes module. A total of 260 participants were affected by COVID-19, of which 69 were children and 191 were parents. A total of 360 participants not affected by COVID-19 responded to the HRQoL questionnaire, of which 96 were children and 264 were parents.

Table 2 describes the self-reports and parent-proxy reports of total and domain-specific HRQoL scores of the whole study population. The mean ± SD total HRQoL score of the self-report was significantly higher compared with those of the parent-proxy reports (79.06 ± 15.19 vs. 73.79 ± 15.17, $p < 0.01$). Children reported significantly higher scores than their parents in the "treatment barriers I" and "worry" domains ($p < 0.01$). However, children reported a significantly lower score (62.59 ± 18.04) in the "diabetes" symptoms domain than their parent-proxy reports (70.06 ± 19.66, $p < 0.01$).

In the COVID-19-affected group, there was a significant difference between the total self-reported HRQoL (77.04) and the parent-proxy HRQoL score (72.80) ($p = 0.047$). The COVID-19-affected children reported a significantly lower HRQoL score in the "diabetes" symptoms domain in comparison to parent-proxy reports (59.50 vs. 68.30, $p < 0.01$). However, they reported a significantly higher score in the "treatment barriers I" domain (81.43) than their parent-proxy reports (72.05) ($p < 0.01$).

Similar to the COVID-19-affected group, children unaffected by COVID-19 reported higher total HRQoL scores (80.51) compared with their parents (74.50) ($p < 0.01$). Furthermore, parent-proxies in the unaffected COVID-19 group consistently reported significantly lower scores compared with self-reports in the "treatment barriers I" domain ($p = 0.039$). The children unaffected by COVID-19 also reported a significantly lower score in the "diabetes" symptoms domain than their parent-proxy reports, similar to the COVID-19 affected group (64.93 ± 18.05 vs. 71.34 ± 19.22, $p < 0.01$).

Table 3 compares the self-report and parent-proxy reports of the affected and unaffected groups (by COVID-19). There was no significant difference between the two groups with regard to self-reports in total or domain-specific scores. However, the parents of COVID-19-affected children reported a significantly lower score in the "treatment barriers I" domain than the unaffected group (72.05 ± 26.12 vs. 77.30 ± 23.67,

TABLE 2 HRQoL of self-reports and parent-proxy reports during the COVID-19 pandemic.

	Total			Affected by COVID-19			Unaffected by COVID-19		
	Self-report ($n = 165$)	Parent-proxy ($n = 455$)	*p*	Self-report ($n = 69$)	Parent-proxy ($n = 191$)	*p*	Self-report ($n = 96$)	Parent-proxy ($n = 264$)	*p*
Peds QL™ Diabetes Module domains									
	M ± SD	**M ± SD**		**M ± SD**	**M ± SD**		**M ± SD**	**M ± SD**	
Total	79.06 ± 15.19	73.79 ± 15.17	<0.01*	77.04 ± 15.81	72.80 ± 14.90	0.047*	80.51 ± 14.65	74.50 ± 15.36	<0.01*
Diabetes	62.59 ± 18.04	70.06 ± 19.66	<0.01*	59.50 ± 17.69	68.30 ± 20.16	<0.01*	64.93 ± 18.05	71.34 ± 19.22	<0.01*
Treatment I	82.28 ± 20.81	75.09 ± 24.84	<0.01*	81.43 ± 24.28	72.05 ± 26.12	<0.01*	82.92 ± 17.86	77.30 ± 23.67	0.039*
Treatment II	83.23 ± 19.66	83.99 ± 18.67	0.660	82.58 ± 16.94	82.08 ± 20.08	0.856	83.70 ± 21.49	85.38 ± 17.48	0.454
Worry	73.48 ± 33.28	61.98 ± 41.88	<0.01*	69.71 ± 36.63	65.20 ± 40.04	0.418	76.39 ± 30.35	59.65 ± 43.10	<0.01*
Communication	82.15 ± 26.43	78.64 ± 29.82	0.187	81.40 ± 26.74	76.63 ± 30.29	0.251	82.71 ± 26.33	80.06 ± 29.46	0.446

Peds QL™, The Pediatric Quality of Life Inventory (Peds QLTM) Diabetes Module 3.0.
*$p < 0.05$.
Higher scores indicate better HRQoL.
M, mean; SD, standard deviation; HRQoL, Health-Related Quality of Life; COVID-19, Coronavirus disease of 2019.

TABLE 3 Comparison of self-reports and parent-proxy reports for those affected vs. unaffected by COVID-19.

	Self-report			Parent-proxy report		
	Affected by COVID-19	Unaffected by COVID-19	*p-value*	Affected by COVID-19	Unaffected by COVID-19	*p-value*
Peds QL™ Diabetes Module domains						
	M ± SD	**M ± SD**		**M ± SD**	**M ± SD**	
Total	77.04 ± 15.81	80.51 ± 14.65	0.149	72.80 ± 14.90	74.50 ± 15.36	0.238
Diabetes	59.50 ± 17.69	64.93 ± 18.05	0.059	68.30 ± 20.16	71.34 ± 19.22	0.104
Treatment I	81.43 ± 24.28	82.92 ± 17.86	0.655	72.05 ± 26.12	77.30 ± 23.67	0.026*
Treatment II	82.58 ± 16.94	83.70 ± 21.49	0.720	82.08 ± 20.08	85.38 ± 17.48	0.068
Worry	69.71 ± 36.63	76.39 ± 30.35	0.218	65.20 ± 40.04	59.65 ± 43.10	0.164
Communication	81.40 ± 26.74	82.71 ± 26.33	0.757	76.63 ± 30.29	80.06 ± 29.46	0.232

Peds QL™, The Pediatric Quality of Life Inventory (Peds QLTM) Diabetes Module 3.0.
*$p < 0.05$.
Higher scores indicate better HRQoL.
M, mean; SD, standard deviation; HRQoL, Health-Related Quality of Life; COVID-19, coronavirus disease of 2019.

10.3389/fpubh.2022.1056967

$p = 0.026$). Although marginally significant ($p = 0.068$), parents in the COVID-19-affected group reported lower scores in the "treatment adherence II" domain compared with parents unaffected by COVID-19 (82.08 ± 20.08 vs. 85.38 ± 17.48, $p = 0.068$).

4. Discussion

This study aimed to evaluate the HRQoL of children and adolescents with T1D based on parent-proxy reports using the PedsQL™ 3.0 Diabetes module during the COVID-19 pandemic. This information may provide insights into how experiences of disease management during the COVID-19 pandemic may have been perceived differently by children and their parents. We further compared HRQoL reports from children and parent-proxies who had been affected with COVID-19 infection (personal or familial) with those unaffected.

We found that, regardless of COVID-19 infection status, children and adolescents with T1D reported higher scores (indicating better HRQoL) in the total HRQoL, "treatment barriers I," and "worry" domains, compared with their parents. Furthermore, children and adolescents in both groups reported lower scores in the "diabetes symptoms" domain, indicating poorer HRQoL. In those affected by COVID-19, parents reported significantly worse HRQoL in "treatment barriers I" and marginally worse HRQoL in the "treatment adherence II" domain, compared with parents whose children were unaffected by COVID-19 infection.

Regardless of being affected or unaffected by COVID-19, parents of children with T1D during the pandemic reported worse HRQoL than what children reported. This finding is consistent with other studies published before the pandemic on children who suffer from T1D and obesity in Kuwait, and children with diabetes in Saudi Arabia (16, 22, 23). In Kuwait, Abdul-Rasoul et al. found that parents of children with T1D reported worse HRQoL than their children (16). Studies in Italy and Norway also showed that parents of children with T1D perceived a lower total quality of life compared with their children (20, 24). This might be attributed to the burden of the disease, the complexity of management, and concern for their child's lifestyle and future (16, 23, 25). Our study demonstrated that such discrepancy between the parents' reports and their children's reports on HRQoL continues to exist during the pandemic. It is therefore warranted to suggest incorporating an assessment of HRQoL during a pandemic to provide support, if needed, by facilitating open communication between children and their families, identifying points of concern from both sides, and planning for joint intervention.

In both groups, parent and child perception of diabetes symptoms, treatment barriers and adherence, worry about sleep quality, anger management, and hypoglycemic episodes were significantly different, consistent with the findings from Kuwait, Saudi Arabia, and Italy (16, 23, 24). The discrepancy in several of the domains indicates the importance of collecting information from the viewpoint of both child and parent, to better understand the impact of disease on HRQoL. The reduced quality of life in the "diabetes symptoms" domain in the children's report may be attributed to more personal effects of diabetes on the child's experiences with the disease.

This study also found that COVID-19-affected parents experienced worse HRQoL in the "treatment barrier" domain, compared with their unaffected counterparts. Treatment barriers include, for example, feeling pain when getting a finger prick or taking an insulin shot, feeling embarrassed about treatment, arguing with parents, or difficulty in caring for their diabetes. To this end, this is the first report that evaluates HRQoL between COVID-19-affected and unaffected children with T1D. Parents' added fear and worry regarding COVID-19 infection in the child or a family member adds to the challenges of managing diabetes and compliance with therapy during such stressful periods (26). The parental burden of management of a chronic condition such as T1D could also be amplified due to the lack of knowledge in relation to dealing with T1D during the pandemic, especially when infected with the infection itself (24–26). Infection control measures implemented by health authorities has led to limited interactions between children with T1D and their healthcare providers, which might indeed contribute to the lack of guidance during the pandemic period (27).

Our study is one of the first reports on HRQoL in children with T1D during the COVID-19 pandemic, regionally or internationally. To conduct the study, we used a validated translated questionnaire, with a disease-specific module, to assess the HRQoL of children and adolescents. There are a number of limitations in the current study. We can only generalize to Arabic speakers in Kuwait, as the data collection instruments were available in Arabic for the current study. The study was cross-sectional; therefore, it cannot provide information on changes in HRQoL over time, or before or after being affected by COVID-19.

5. Conclusion

This is the first report on the health-related quality of life of children with T1D in Kuwait during the COVID-19 era. Parents' or caregivers' experience of caring for their children was negative, as they worried, and reported poorer HRQoL compared with their children's perceptions. There is a need to empower healthcare professionals to support parents and caregivers of children with chronic diseases such as T1D in promoting self-management, enhancing physical and psychological wellbeing, treatment adherence, and continuous health education during pandemics of any kind.

Longitudinal studies on the effect of the COVID-19 pandemic on children and youth with T1D are necessary to enhance our understanding of the effects of the pandemic on disease management, glycemic control, day-to-day activities, and patients' wellbeing. Future studies could also add new insights by comparing the experiences of children and adolescents with and without diabetes.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by Ethical Committee at the Ministry of Health of Kuwait. Written/verbal consents were provided by legal guardians or next of kin.

Author contributions

DA-A was responsible for planning, designing, data managing, conducting the study, and writing the manuscript. TA, FO, and FA-J collected data and recruited patients. DK participated in directing the reported work, data management, critically reviewed the manuscript, and participated in the discussions. TA, DA-A, and DK carried out data analysis and interpretation of the results. TA, FO, SQ, and FA-J contributed to the writing of the manuscript. HA-K participated in the planning, data management, and conducting of the study. All authors contributed to the article and approved the submitted version.

Acknowledgments

The authors would like to acknowledge the children and their families who participated in this study. Furthermore, we would like to acknowledge Kuwait University and Dasman Diabetes Institute for the support provided to the investigators and authors during all phases of this project. We would also like to acknowledge the MAPI institute. Finally, we would like to acknowledge Solveig Argeseanu Cunningham for her expert review of the manuscript.

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. Racine N, McArthur BA, Cooke JE, Eirich R, Zhu J, Madigan S. Global prevalence of depressive and anxiety symptoms in children and adolescents during COVID-19: a meta-analysis. *JAMA Pediatr.* (2021) 175:1142–50. doi: 10.1001/jamapediatrics.2021.2482

2. Xie X, Xue Q, Zhou Y, Zhu K, Liu Q, Zhang J, et al. Mental health status among children in home confinement during the coronavirus disease 2019 outbreak in Hubei Province, China. *JAMA Pediatr.* (2020) 174:898–900. doi: 10.1001/jamapediatrics.2020.1619

3. Ravens-Sieberer U, Kaman A, Erhart M, Devine J, Schlack R, Otto C. Impact of the COVID-19 pandemic on quality of life and mental health in children and adolescents in Germany. *Eur Child Adolesc Psychiatry.* (2021) 2021:1–11. doi: 10.2139/ssrn.3721508

4. Danne T, Lanzinger S, de Bock M, Rhodes ET, Alonso GT, Barat P, et al. A worldwide perspective on COVID-19 and diabetes management in 22,820 children from the SWEET project: diabetic ketoacidosis rates increase and glycemic control is maintained. *Diabetes Technol Ther.* (2021) 23:632–41. doi: 10.1089/dia.2021.0110

5. Passanisi S, Pecoraro M, Pira F, Alibrandi A, Donia V, Lonia P, et al. Quarantine due to the COVID-19 pandemic from the perspective of pediatric patients with type 1 diabetes: a web-based survey. *Front Pediatr.* (2020) 2020:491. doi: 10.3389/fped.2020.00491

6. Al-Abdulrazzaq D, Alkandari A, Alhusaini F, Alenazi N, Gujral UP, Narayan KMV, et al. Higher rates of diabetic ketoacidosis and admission to the paediatric intensive care unit among newly diagnosed children with type 1 diabetes in Kuwait during the COVID-19 pandemic. *Diabetes Metab Res Rev.* (2022) 38:e3506. doi: 10.1002/dmrr.3506

7. McGurnaghan SJ, Weir A, Bishop J, Kennedy S, Blackbourn LA, McAllister DA, et al. Risks of and risk factors for COVID-19 disease in people with diabetes: a cohort study of the total population of Scotland. *Lancet Diab Endocrinol.* (2021) 9:82–93. doi: 10.1016/S2213-8587(20)30405-8

8. Diedisheim M, Dancoisne E, Gautier J-F, Larger E, Cosson E, Fève B, et al. Diabetes increases severe COVID-19 outcomes primarily in younger adults. *J Clin Endocrinol Metab.* (2021) 106:e3364–e8. doi: 10.1210/clinem/dgab393

9. Solowiejczyk J. The family approach to diabetes management: Theory into practice toward the development of a new paradigm. *Diab Spectrum.* (2004) 17:31–6. doi: 10.2337/diaspect.17.1.31

10. Control D, Group CTR. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *New Engl J Med.* (1993) 329:977–86. doi: 10.1056/NEJM199309303291401

11. Rochmah N, Faizi M, Hisbiyah Y, Triastuti IW, Wicaksono G, Endaryanto A. Quality of life differences in pre-and post-educational treatment in type 1 diabetes

mellitus during COVID-19. *Diab Metab Synd Obesity Targets Therapy*. (2021) 14:2905. doi: 10.2147/DMSO.S313575

12. Arab-Zozani M, Hashemi F, Safari H, Yousefi M, Ameri H. Health-related quality of life and its associated factors in COVID-19 patients. *Osong Public Health Res Perspect*. (2020) 11:296. doi: 10.24171/j.phrp.2020.11.5.05

13. Varni JW, Burwinkle TM, Jacobs JR, Gottschalk M, Kaufman F, Jones KL. The PedsQL^{TM} in type 1 and type 2 diabetes: reliability and validity of the Pediatric Quality of Life Inventory^{TM} generic core scales and type 1 diabetes module. *Diabetes Care*. (2003) 26:631–7. doi: 10.2337/diacare.26.3.631

14. Varni JW, Delamater AM, Hood KK, Driscoll KA, Wong JC, Adi S, et al. Diabetes management mediating effects between diabetes symptoms and health-related quality of life in adolescents and young adults with type 1 diabetes. *Pediatr Diabetes*. (2018) 19:1322–30. doi: 10.1111/pedi.12713

15. Rubin RR. Diabetes and quality of life. *Diabetes Spectr*. (2000) 13:21.

16. Abdul-Rasoul M, AlOtaibi F, Abdulla A, Rahme Z, AlShawaf F. Quality of life of children and adolescents with type 1 diabetes in Kuwait. *Medical principles and practice*. (2013) 22:379–84. doi: 10.1159/000347052

17. Abdul-Rasoul M, AlOtaibi F, AlMahdi M, AlKandari H. Reliability and validity of the Arabic version of the PedsQL^{TM} 4.0 generic ore scales and PedsQL^{TM} 3.0 diabetes module. *J Diab Mellit*. (2012) 2:301–7. doi: 10.4236/jdm.2012.23047

18. DiMeglio LA, Acerini CL, Codner E, Craig ME, Hofer SE, Pillay K, et al. ISPAD Clinical Practice Consensus Guidelines 2018: Glycemic control targets and glucose monitoring for children, adolescents, and young adults with diabetes. *Pediatr Diab*. (2018) 27:105–14. doi: 10.1111/pedi.12737

19. Varni JW, Seid M, Rode C. PedsQL: Measurement model for the pediatric quality of life inventory Version. *Med Care*. (1998) 4:126–39. doi: 10.1037/t65932-000

20. Kalyva E, Malakonaki E, Eiser C, Mamoulakis D. Health-related quality of life (HRQoL) of children with type 1 diabetes mellitus (T1DM): self and parental perceptions. *Pediatr Diabetes*. (2011) 12:34–40. doi: 10.1111/j.1399-5448.2010.00653.x

21. Varni JW. *Linguistic validation of the PedsQL—a Quality of Life Questionnaire*. Mapi Research Institute (2002). Available online at: https://www.pedsql.org/PedsQL-Linguistic-Validation-Guidelines.doc (accessed August, 2022).

22. Boodai SA, Reilly JJ. Health related quality of life of obese adolescents in Kuwait. *BMC Pediatr*. (2013) 13:1–5. doi: 10.1186/1471-2431-13-105

23. AlBuhairan F, Nasim M, Al Otaibi A, Shaheen NA, Al Jaser S, Al Alwan I. Health related quality of life and family impact of type 1 diabetes among adolescents in Saudi Arabia. *Diabetes Res Clin Pract*. (2016) 114:173–9. doi: 10.1016/j.diabres.2016.01.001

24. Nardi L, Zucchini S, D'Alberton F, Salardi S, Maltoni G, Bisacchi N, et al. Quality of life, psychological adjustment and metabolic control in youths with type 1 diabetes: a study with self-and parent-report questionnaires. *Pediatr Diabetes*. (2008) 9:496–503. doi: 10.1111/j.1399-5448.2008.00414.x

25. Graue M, Wentzel-Larsen T, Hanestad B, Båtsvik B, Søvik O. Measuring self-reported, health-related, quality of life in adolescents with type 1 diabetes using both generic and disease-specific instruments. *Acta Paediatr*. (2003) 92:1190–6. doi: 10.1111/j.1651-2227.2003.tb02483.x

26. Odeh R, Gharaibeh L, Daher A, Kussad S, Alassaf A. Caring for a child with type 1 diabetes during COVID-19 lockdown in a developing country: Challenges and parents' perspectives on the use of telemedicine. *Diabetes Res Clin Pract*. (2020) 168:108393. doi: 10.1016/j.diabres.2020.108393

27. Lazzerini M, Barbi E, Apicella A, Marchetti F, Cardinale F, Trobia G. Delayed access or provision of care in Italy resulting from fear of COVID-19. *Lancet Child Adolesc Health*. (2020) 4:e10–e1. doi: 10.1016/S2352-4642(20)30108-5

frontiers | Frontiers in Pediatrics

TYPE Original Research
PUBLISHED 30 December 2022
DOI 10.3389/fped.2022.1049942

OPEN ACCESS

EDITED BY
Jing Wu,
Tianjin University, China

REVIEWED BY
Marilena Trozzi,
Bambino Gesù Children's Hospital (IRCCS), Italy
Artur Rebelo,
University Hospital in Halle, Germany

*CORRESPONDENCE
Juliane Hardt
juliane.hardt@uni-wh.de

SPECIALTY SECTION
This article was submitted to Children and Health, a section of the journal Frontiers in Pediatrics

RECEIVED 21 September 2022
ACCEPTED 28 November 2022
PUBLISHED 30 December 2022

CITATION
Arens P, Hardt J, Angrick JC, Olze H and Coordes A (2022) Modified dindo-clavien system for registration of perioperative complications in children undergoing adenotonsillectomy.
Front. Pediatr. 10:1049942.
doi: 10.3389/fped.2022.1049942

Modified dindo-clavien system for registration of perioperative complications in children undergoing adenotonsillectomy

Philipp Arens[1], Juliane Hardt[2,3]*, Julie Charlotte Angrick[1], Heidi Olze[1] and Annekatrin Coordes[1]

[1]Department of Otorhinolaryngology, Charité—Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany, [2]Institute of Biometry and Clinical Epidemiology, Charité - Universitätsmedizin Berlin, Corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany, [3]Institute for Integrative Medicine, Department for Human Medicine, Faculty of Health, Witten/Herdecke University, Herdecke, Germany

Introduction: Surgical procedures in children are among the most commonly performed procedures in otolaryngology. Perioperative safety and documentation of complications are becoming increasingly important. This study investigates perioperative complications in a clinical cohort of children with adenotonsillar hyperplasia undergoing adenotonsillectomy using the standardized Dindo-Clavien reporting system.
Patients and Methods: Retrospective evaluation of 402 children who underwent adenotonsillectomy between 2009 and 2015. Patient parameters including all perioperative complications were investigated.
Results: In the study, 124 complications were found (106 mild, 16 severe). According to the Dindo-Clavien classification, 93 grade I, 15 grade II, 5 grade III, 11 grade IV and 0 grade V complications were documented. Complications were associated with additional diagnoses ($p = 0.001$), long-term medication intake ($p = 0.003$), duration of hospitalization ($p < 0.001$) and duration of surgery ($p < 0.001$), undergoing tonsillotomy ($p = 0.022$) or tonsillectomy ($p < 0.001$), differences in ASA score ($p = 0.005$) and differences in OSA-18 score ($p = 0.011$). Severe complications, classified as grade III and IV, were associated with premature birth ($p = 0.026$), additional diagnoses ($p = 0.017$), long-term medication intake ($p < 0.001$) and differences in ASA score ($p =< 0.001$).
Conclusion: The Dindo-Clavien classification is a standardized reporting system which can also be used for surgical procedures in children with adenotonsillar hyperplasia. The system shows associations with clinical parameters and thus can help to identify subgroups at risk of severe complications.

KEYWORDS

complication reporting, pediatrics, adenotonsillectomy, dindo-clavien classification, sleep apnea

Introduction

Surgical procedures in children with adenotonsillar hyperplasia are among the most commonly performed procedures in otolaryngology, head and neck surgery (1). The most frequent indications are recurrent or chronic inflammation of the tonsillar tissue (with or without otitis media) and obstructive sleep-apnea (OSA) (2). The safety of

these procedures is attracting increased attention. In children with OSA, there is often discussion about the need for an inpatient stay or admission to the intermediate care unit during the postoperative period (3–5). Furthermore, there is also a lack of standardized complication classifications in pediatric surgical patients, including children undergoing standardized surgery such as adenoidectomy (AT) and/or tonsillectomy/tonsillotomy (TE/TT).

Clavien et al. developed a classification system in 1992 for surgical complications in general surgery. Complications were classified into four grades depending on the therapy treating the complication (6). Dindo et al. modified the score and thereby focused on reporting permanently disabling and life-threatening complications (7).

Today, the modified Dindo-Clavien system is widely used to quantify patient outcomes in general surgery (8).

This study investigates perioperative complications in children with adenotonsillar hyperplasia undergoing AT and/or TE/TT using the standardized Dindo-Clavien reporting system.

Methods

Study design

The university hospital Charité—Universitätsmedizin Berlin is one of the largest children's clinic in Germany with children's intensive care units. Children with adenotonsillar hyperplasia who underwent AT and/or TE/TT at the Department of Otolaryngology at the Charité—Campus Virchow Klinikum between 2009 and 2015 were included in the study. The relevant data of all patients (aged 0–18) were extracted from the patients' files retrospectively. The study was conducted in accordance with the declarations of Helsinki and the guidelines of the Charité Berlin ethics committee (EA2/101/14).

All parents were routinely asked to complete the clotting questionnaire before surgery (9). If a coagulation disorder was suspected, patients were sent to the coagulation outpatient department.

All operations were performed under general anesthesia. Adenoidectomy was performed using the ring knife *via* a transoral approach under indirect mirror visualization. Tonsillectomies were performed using the cold dissection method and tonsillotomies with the monopolar needle. Bleeding lesions were gently coagulated with bipolar forceps. Patients received soft food postoperatively and pain medication as recommended. After adenoidectomy, the patients were monitored in hospital for one night and after tonsillotomy or tonsillectomy for three nights. Children with coagulation disorders were treated according to the recommendations of the coagulation department and were admitted for seven days. Antibiotic therapy was not prescribed prophylactically, except when local signs of inflammation or foetor ex ore were noticed.

Measurements and variables

Collected data of the patients included: age, sex, height, weight, body mass index, additional diagnoses, long-term medication, length of hospital stay, duration of surgery, previous general surgeries, previous surgeries of the upper respiratory tract, premature birth complications, OSA-18 score and ASA score. To capture patient comorbidity as completely as possible, "additional diagnoses" were collected. Any acute or chronic condition independent of the surgery-defining diagnosis that was reported by the patient or patient's family at the admission visit and noted in the patient's record was included. We counted as "long-term medication" all medications that the patient took permanently and not for acute reasons. This information was extracted from the medication schedules that were kept as admission medication in the patient's file. Any deviation from the standard of care or standard treatment course during hospital stay was considered a "complication". The entire inpatient stay period was evaluated.

Preoperatively, the parents filled in the ***OSA-18 Quality of Life Survey questionnaire*** (10). This questionnaire contains 18 quality of life items, which include sleep disorders, physical symptoms, mental complaints, daytime discomfort and patients' own fears. The OSA-18 total score including all 18 items ranges from 18 (no impact on quality of life) to 126 (major negative impact). A value >60 is considered abnormal in previous studies. The severity of the sleep-related breathing disorder recorded is divided into three categories: less likely (<60), possible (≥60–80) and high probability (>80) of the presence of OSA in a child. To evaluate a potential bias, patients with and without a completed OSA-18 questionnaire were compared.

The ***American Society of Anasthesiologists Physical Status Classification*** (ASA-PS) is an instrument for the preoperative determination of anesthetic risks. It records the physical condition of a patient and is divided into six categories. Group I describes healthy patients, group II is mild general disease, group III is severe general disease without performance limitations, group IV is severe, long-lasting general disease with long-lasting general disease with performance limitations, and Group V is a moribund patient who is expected to die within 24 h, with or without surgery. Group VI includes brain-dead patients (11).

Every complication was graded with the ***Dindo-Clavien-Score.*** The score classifies complications into five groups (see **Table 1**) (7). Every deviation from the standard treatment or the expected postoperative course was documented and regarded as a complication. Complications with grades III and IV were regarded as severe complications.

Statistics

In this exploratory study, the association of possible complicating or seriously complicating factors were

TABLE 1 Dindo-clavien classification of complications.

Dindo-Clavien classification	
Grade I	Any deviation from the normal postoperative course without the need for pharmacological treatment or surgical, endoscopic and radiological interventions. Permitted therapeutic regimens include: drugs as antiemetics, antipyretics, analgetics, diuretics and electrolytes and physiotherapy. This grade also includes wound infections opened at the bedside.
Grade II	Pharmacological treatment required with other medications than the permitted complications of grade I; including blood transfusions and total parenteral nutrition.
Grade III	Requiring surgical, endoscopic or radiological intervention
IIIa	Intervention not under general anesthesia
IIIb	Intervention under general anesthesia
Grade IV	Life-threatening complication (including CNS complications)* requiring ICU-management
IVa	Single organ dysfunction (including dialysis)
IVb	Multiorgan dysfunction
Grade V	Death of a patient

ICU, intensive care unit; CNS, central nervous system.

investigated and reported according to SAMPL-Guidelines (12). Data were presented according to the STROBE checklist (see **Supplementary Figure S1**). All statistical analyses were done with IBM SPSS version 25 (Statistical Package for the Social Sciences). Continuous variables were summarized with medians and interquartile ranges (IQR) and examined with the Mann-Whitney U test (as none of them were normally distributed). Categorical variables were expressed as numbers and percentages and examined using the Chi-squared test or Fisher's exact test. Group differences were systematically tested with significance tests in awareness that this procedure is associated with an inflation of type I error. As this is an exploratory study, significance tests are only used for detecting potential differences. Hence, *p*-values described as significant for two-sided *p*-values ≤ 0.05 are only given as an orientation and not to be interpreted as confirmatory.

Results

Description of the patients

Between 2009 and 2015, 402 children with adenotonsillar hyperplasia underwent surgery in the Department of Otolaryngology. Data collection was complete for all investigated factors, the OSA-18 Questionnaire only was completed with 354 children (88.1%). The **Supplementary Figure S2** shows the comparison of patients with vs. without completed questionnaires. 121 of 402 patients experienced one or more complications. In total, 124 complications were found. Three children had two independent complications (**Tables 2,3**). However, only the most severe complication was evaluated for statistical analysis (**Table 4**). 106 complications were wound- related and 15 were due to respiratory conditions. Thus, at least one complication occurred in 121 patients (30.1%). In 106 patients, the complication was mild (Dindo-Clavien grade I and II). In 15 patients a severe complication occurred (Dindo-Clavien grade III and IV). Grade V complications were not found.

The clinical data of the patients are summarized in **Table 4**. The median patient age was 55 months. 60.9% were male and 39.1% female. Adenoidectomies, tonsillotomies, tonsillectomies, and paracenteses were performed in 93.0%, 28.1%, 31.8% and 68.8% of patients respectively. 35.3% had undergone previous general operations and 19.7% previous operations of the upper airways. Other operations included surgery of the nasal conchae. The median inpatient duration (with IQR) was 5 (3–5) days, and the total range was 0–16 days. Nine children (2.2%) had coagulation disorders (e.g.,

TABLE 2 Number of total complications during hospital stay (3 patients had 2 independent complications).

Complications (Dindo-Clavien grade)	*n*
Wound	
• Wound infection or fever requiring antibiotic treatment and/or antipyretics (D-C I)	90
• Bleeding controlled with conservative intervention (D-C II)	11
• Bleeding controlled surgically (D-C III)	5
Airway	
• Perioperative respiratory problems requiring no specific therapy or ICU monitoring (D-C I)	1
• Perioperative respiratory problems requiring inhalative therapy (D-C II)	3
• Perioperative respiratory problems requiring re-intubation and/or ICU monitoring (D-C IV)	11
Others	
• Injury of the teeth (D-C I)	1
• Extravasate of the lower extremities (D-C I)	1
• Hypertensive derailment (D-C II)	1
Total complications, *n*	**124**

D-C, Dindo-Clavien grade; *n*, number.

TABLE 3 Total complications according to the dindo-clavien classification.

Complications according to Dindo-Clavien	*n* = 124
Grade I	93
Grade II	15
Grade III	5
Grade IV	11
Grade V	0
Mild complications (D-C I and II)	108
Severe complications (D-C III and IV)	16

D-C, Dindo-Clavien grade; *n*, number.

TABLE 4 Patient characteristics and statistical analyses.

Characteristics	Total	No complication	With complications			Comparison 1: without vs. all complications		Comparison 2: severe complications vs. all other patients	
			Simple	Severe	All	test	*p*-value	test	*p*-value
	402 (100%)	281 (69.9%)	106	15	121 (30.1%)				
Age (months), median (IQR)	**55.0 (37.8–83.0)**	56.0 (38.0–83.0)	53.5 (36.0–83.5)	36.0 (23.0–78.0)	52.0 (35.0–80.5)	U	0.521	U	0.133
Sex									
Male, *n* (%)	**245 (60.9)**	170 (60.5)	68 (64.2)	7 (46.7)	75 (62.0)	chi^2	0.866	chi^2	0.286
Female, *n* (%)	**157 (39.1)**	111 (39.5)	38 (35.8)	8 (53.3)	46 (38.0)				
Body height (cm), median (IQR)	**107.0 (97.0–124.0)**	108.0 (98.0–124.0)	105.5 (96.8–125.3)	97.0 (80.0–123.0)	105.0 (95.0–124.5)	U	0.632	U	0.132
Body weight (kg), median (IQR)	**18.0 (14.4–26.0)**	18.0 (14.5–26.0)	18.0 (14.5–27.0)	16.0 (10.5–22.5)	17.5 (14.0–26.5)	U	0.796	U	0.110
Body mass index, median (IQR)	**16.2 (14.7–18.5)**	16.3 (14.6–18.5)	16.1 (14.7–19.2)	16.1 (14.7–19.2)	16.1 (14.9–17.2)	U	0.651	U	0.766
Additional diagnoses *n* (%)	**119 (29.6)**	69 (24.6)	41 (38.7)	9 (60.0)	50 (41.3)	chi^2	**0.001**	Fisher	**0.017**
Long-term medication, *n* (%)	**51 (12.7)**	26 (9.3)	17 (16.0)	8 (53.3)	25 (20.7)	chi^2	**0.003**	Fisher	**<0.001**
Premature birth, *n* (%)	**32 (8.0)**	19 (6.8)	9 (8.5)	4 (26.7)	13 (10.7)	chi^2	0.249	Fisher	**0.026**
ASA score (score 1–6), median (IQR)	**1 (1–2)**	1 (1–2)	1 (1–2)	2 (1–3)	1 (1–2)	U	**0.005**	U	<0.001
1 (%)	**281 (69.9)**	208 (74.0)	68 (64.2)	5 (33.3)	73 (60.3)				
2 (%)	**99 (24.6)**	61 (21.7)	34 (32.1)	4 (26.7)	38 (31.4)				
3 (%)	**20 (5.0)**	11 (3.9)	4 (3.8)	5 (33.3)	9 (7.4)				
4 (%)	**2 (0.5)**	1 (0.4)	0 (0.0)	1 (6.7)	1 (0.8)				
Surgical procedures:									
Adenoidectomy, *n* (%)	**374 (93.0)**	264 (94.0)	97 (91.5)	13 (86.7)	110 (90.9)	chi^2	0.376	Fisher	0.281
Tonsillotomy, *n* (%)	**113 (28.1)**	69 (24.6)	39 (36.8)	5 (33.3)	44 (36.4)	chi^2	**0.022**	Fisher	0.770
Tonsillectomy, *n* (%)	**128 (31.8)**	69 (24.6)	54 (50.9)	5 (33.3)	59 (48.8)	chi^2	**<0.001**	Fisher	1,000
Paracentesis/tube, *n* (%)	**276 (68.8)**	194 (69.0)	71 (67.6)	11 (73.3)	82 (68.3)	chi^2	0.982	Fisher	0.785
Other surgeries, *n* (%)	**85 (21.1)**	64 (22.8)	19 (17.9)	2 (14.3)	21 (17.5)	chi^2	0.294	Fisher	0.743
Duration of surgery (min), median (IQR)	**23.0 (16.0–33.0)**	20.0 (15.0–30.0)	27.0 (18.0–36.3)	24.0 (15.0–35.0)	27.0 (18.0–36.0)	U	**<0.001**	U	0.719
Duration of hospital stay (days), median (IQR)	**5 (3–5)**	4 (3–5)	5 (5–6)	4 (3–6)	5 (5–6)	U	**<0.001**	U	0.542
Duration of extended stay (days), median (IQR)	**1.0 (0.8–1.0)**	1.0 (0.0–1.0)	1.0 (1.0–2.0)	1.0 (0.0–2.0)	1.0 (1.0–2.0)	U	**<0.001**	U	0.298
Proportion with extended stay, *n* (%)	**302 (75.1)**	202 (71.9)	89 (84.0)	11 (73.3)	100 (82.6)	chi^2	**0.031**	Fisher	1.000

(*continued*)

10.3389/fped.2022.1049942

TABLE 4 Continued

Characteristics	Total	No complication	With complications			Comparison 1: without vs. all complications		Comparison 2: severe complications vs. all other patients	
			Simple	Severe	All	test	*p*-value	test	*p*-value
	402 (100%)	281 (69.9%)	106	15	121 (30.1%)				
OSA-18 questionnaire score (18–126), median (IQR)	**56.5 (41.0–72.0)**	54.0 (38.5–69.0)	59.0 (46.0–73.0)	57.0 (48.0–80.0)	58.0 (46.5–74.5)	U	**0.011**	U	0.443
OSA risk scale (0–2; 0 = unlikely, 1 = possible, 2 = very likely), median (IQR)	**0 (0.0–1.0)**	0 (0.0–1.0)	0 (0.0–1.0)	0 (0.0–1.5)	0 (0.0–1.0)	U	0.113	U	0.607
OSA unlikely (%)	**208 (58.8)**	148 (61.4)	53 (53.0)	7 (53.8)	60 (53.1)				
OSA possible (%)	**90 (25.4)**	59 (24.5)	28 (28.0)	3 (20)	31 (27.4)	chi^2	0.172†	chi^2	0.937†
OSA very likely (%)	**56 (15.8)**	34 (14.1)	19 (19.0)	3 (20)	22 (19.5)	chi^2	0.257‡	Fisher	0.441‡

†OSA possible and OSA very likely vs. OSA unlikely.
‡OSA very likely vs. OSA unlikely and OSA possible.
IQR, interquartile range; *n*, number; OSA, obstructive sleep apnea; Fisher, Fisher's exact test; chi^2, Chi-squared test; *U*, Mann-Whitney U test for independent groups.

von Willebrand factor syndrome, hemophilia) and were monitored regularly for 7 days after surgery. 75.1% of patients had an extended stay [median 1.0 (IQR 0.8–1.0) days]. The ASA score was determined as one and two in 69.9% and 24.6%.

Which factors were associated with complications?

The presence of additional diagnoses ($p = 0.001$), long-term medication ($p = 0.003$), duration of surgery ($p < 0.001$), duration of hospital stay ($p < 0.001$), duration of extended stay ($p < 0.001$), differences in ASA score ($p = 0.005$), surgical procedures tonsillotomy ($p = 0.022$) and tonsillectomy ($p < 0.001$) and OSA-18 score ($p = 0.011$) were significantly associated with the occurrence of complications. 46.9% of the patients with complications had a possible or probable OSA compared with 38.6% without any complications. Age, sex, premature birth, height, weight, body mass index and previous operations were not associated with overall complications.

What overall factors are associated with severe complications?

Severe complications were associated with premature birth ($p = 0.026$), and also the presence of additional diagnoses ($p = 0.017$), long-term medication (<0.001) and differences in ASA score ($p < 0.001$). Even serious complications were not associated with age, gender, height, weight, BMI and pre-operations.

Discussion

This study investigated perioperative complications in children with adenotonsillar hyperplasia undergoing AT and/or TE/TT using the standardized Dindo-Clavien reporting system. In general surgery, the Dindo-Clavien classification system is widely used to quantify patient outcomes (8). In this study, the total complication rate per patient was 30.1% (121/402) and the rate of severe complications per patient was 3.7% (15/402). The total complication rate in this cohort was increased compared with the literature. The rate of severe complications (grade III and IV according to the Dindo-Clavien reporting system) was consistent with current literature (4, 13–18).

Various factors may have contributed to the high total complication rate. Firstly, the children in this cohort were less healthy compared to children with adenotonsillar hyperplasia undergoing AT in an ambulatory surgery center. 119 of 402 (29.6%) children had comorbidities, and a probable OSA did not count as independent comorbidity in our analysis (4, 14). Another aspect was the rigorous file review. Any deviation from the standard protocol was counted as a complication.

Therefore, 74,4% (90/121) children were scored as Dindo-Clavien grade I for children with wound infection or fever requiring antibiotic treatment or antipyretics. In daily routines, these mild complications are sometimes not regarded as relevant by physicians (**Table 2**). However, one of the advantages of standardized complication classification (e.g., Dindo-Clavien Score) is a clear definition of complications.

The grading of complications may help to identify relevant patient subgroups. In this study, we were able to identify subgroups at increased risk of complications (e.g., comorbidities, long-term medication, children with a higher preoperative ASA-score and children with an increased OSA-18 questionnaire score). The relationship between comorbidities and complications is well known in pediatric literature. Amoils et al. found a significantly increased risk of respiratory complications in children with cardiovascular disease or respiratory disease who underwent tonsillectomy or adenoidectomy (19). The use of long-term medication indicates concomitant diseases; therefore, this cannot be assumed as an independent risk factor. Additionally, concomitant diseases influence the ASA score. OSA itself is a known risk factor for perioperative complications (3). In this context, our findings are interesting because even though the OSA-18 is not able to detect an OSA with diagnostic certainty (20), the questionnaire might be helpful to identify children at higher risk of perioperative complications when preoperative polysomnography is not available. The standardized grading system of the Dindo-Clavien Score enables identification of children with serious complication risks (e.g., premature birth). Premature birth has been described as a risk factor for complications in children undergoing ATE with a diagnosis of OSA (19). Thus, a standardized tool like the Dindo-Clavien Score can help to identify subgroups at risk of severe complications in this special cohort of patients.

According to the criteria for complication reporting in surgical patients, defined by Martin et al., this study fulfilled nine out of ten criteria (21): definition of the data acquisition method, indication of the follow-up duration, definition of complications, registration of the mortality and morbidity rate, causes of death, indication of total complications and procedure-specific complications, severity grade used, length-of-stay data, and risk factors included in the analysis.

Limitations of this study from a methodological view are the size of the cohort and consequentially the low number of severe complications (grade III and IV).

Unfortunately, only 88.1% completed the OSA-18 questionnaire (see also **Supplementary Figure S2**). Children with completed questionnaires underwent tonsillectomy and tonsillotomy more frequently. These surgical procedures have a longer planned hospital stay and also the risk of complications is naturally higher. It can also be assumed that patients who undergo tonsillectomy as opposed to, for example, adenoidectomy alone are also more likely to suffer from OSAS. It can also be assumed that the parents of children with suspected OSAS are more willing to complete the OSA-18 questionnaire. Not fully explained is the fact that, contrary to expectations, the proportion of patients who unexpectedly stayed longer in the hospital is higher in the group of patients who did not complete the questionnaire. It is possible that risk underestimation took place beforehand.

A further limitation is the focus on the immediate postoperative period during the hospital stay. Outpatient postoperative course data were not included which may be helpful to evaluate the safety of surgical procedures as complications like bleeding occur later (22).

Conclusion

The Dindo-Clavien classification is a standardized reporting system which can also be used for surgical procedures in children with adenotonsillar hyperplasia. Treatment with antibiotics or additional antipyretics in case of suspected wound infection or fever are deviations from the standard protocol and therefore count as mild complications (grade I according to Dindo-Clavien Score). This system shows associations with clinical parameters and thus can help to identify subgroups at risk of severe complications.

Data availability statement

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

Author contributions

PA, JCA, AC: planed and organised study. PA, AC: prepared manuscript. JH: statistical analyis. HO: reviewed manuscript. 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

All claims expressed in this article are solely those of the authors and do not necessarily represent those of

their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fped.2022.1049942/full#supplementary-material.

References

1. Cullen KA, Hall MJ, Golosinskiy A. Ambulatory surgery in the United States, 2006. *Natl Health Stat Rep* (2009) 28(11):1–25. PMID: 19294964.

2. Parker NP, Walner DL. Trends in the indications for pediatric tonsillectomy or adenotonsillectomy. *Int J Pediatr Otorhinolaryngol*. (2011) 75:282–5. doi: 10.1016/j.ijporl.2010.11.019

3. Brown KA. Outcome, risk, and error and the child with obstructive sleep apnea. *Paediatr Anaesth*. (2011) 21:771–80. doi: 10.1111/j.1460-9592.2011.03597.x

4. Tweedie DJ, Bajaj Y, Ifeacho SN, Jonas NE, Jephson CG, Cochrane LA, et al. Peri-operative complications after adenotonsillectomy in a UK pediatric tertiary referral centre. *Int J Pediatr Otorhinolaryngol*. (2012) 76:809–15. doi: 10.1016/j.ijporl.2012.02.048

5. Baguley KE, Cheng AT, Castro C, Wainbergas N, Waters KA. Is day stay adenotonsillectomy safe in children with mild to moderate obstructive sleep apnoea? A retrospective review of 100 patients. *Int J Pediatr Otorhinolaryngol* (2014) 78:71–4. doi: 10.1016/j.ijporl.2013.10.050

6. Clavien PA, Sanabria JR, Strasberg SM. Proposed classification of complications of surgery with examples of utility in cholecystectomy. *Surg*. (1992) 111:518–26. PMID: 1598671.

7. Dindo D, Demartines N, Clavien P-A. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg*. (2004) 240:205–13. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1360123&tool=pmcentrez&rendertype=abstractAccessed October 8, 2012. doi: 10.1097/01.sla.0000133083.54934.ae

8. Clavien PA, Barkun J, De Oliveira ML, Vauthey JN, Dindo D, Schulick RD, et al. The clavien-dindo classification of surgical complications: five-year experience. *Ann Surg* (2009) 250:187–96. doi: 10.1097/SLA.0b013e3181b13ca2

9. Eberl W, Wendt I, Schroeder H-G. Preoperative coagulation screening prior to adenoidectomy and tonsillectomy. *Klin Padiatr*. (n.d.) 217:20–4. doi: 10.1055/s-2004-818789

10. Franco RA, Rosenfeld RM, Rao M. Quality of life for children with obstructive sleep apnea. *Otolaryngol Head Neck Surg*. (2000) 123:9–16. doi: 10.1067/mhn.2000.105254

11. ASA Physical Status Classification System American Society of Anesthesiologists (ASA), (n.d.). Available at: https://www.asahq.org/standards-and-guidelines/asa-physical-status-classification-system (Accessed November 29, 2019).

12. Lang TA, Altman DG. Statistical analyses and methods in the published literature: the SAMPL guidelines*. In: Moher D, Altman DG, Schulz KF, Simera I, Wager E, editors. *Guidel. Report. Heal. Res. A User's Man*. Oxford, UK: John Wiley & Sons, Ltd (2014). p. 264–74. doi: 10.1002/9781118715598.ch25

13. del-Río Camacho G, Martínez González M, Sanabria Brossart J, Gutiérrez Moreno E, Gómez García T, Troncoso Acevedo F. Post-operative complications following adenotonsillectomy in children with severe sleep apnea-hypopnea syndrome. Do they need to be admitted to an intensive care unit? *Acta Otorrinolaringológica Española*. (2014) 65(5):302–7. doi: 10.1016/j.otorri.2014.03.004

14. Wijayasingam G, Deutsch P, Jindal M. Day case adenotonsillectomy for paediatric obstructive sleep apnoea: a review of the evidence. *Eur Arch Oto-Rhino-Laryngol*. (2018) 275:2203–8. doi: 10.1007/s00405-018-5071-8

15. Sanders JC, King MA, Mitchell RB, Kelly JP. Perioperative complications of adenotonsillectomy in children with obstructive sleep apnea syndrome. *Anesth Analg*. (2006) 103:1115–21. doi: 10.1213/01.ane.0000244318.77377.67

16. Granell J, Gete P, Villafruela M, Bolaños C, Vicent JJA. Safety of outpatient tonsillectomy in children: a review of 6 years in a tertiary hospital experience. *Otolaryngol Head Neck Surg*. (2004) 131:383–7. doi: 10.1016/j.otohns.2004.03.027

17. Lavin JM, Shah RK. Postoperative complications in obese children undergoing adenotonsillectomy. *Int J Pediatr Otorhinolaryngol*. (2015) 79:1732–5. doi: 10.1016/j.ijporl.2015.07.038

18. Rodríguez-Catalán J, Fernández-Cantalejo Padial J, Rodríguez Rodríguez P, González Galán F, del-Río Camacho G. Postoperative complications after adenotonsillectomy in two paediatric groups: obstructive sleep apnoea syndrome and recurrent tonsillitis. *Acta Otorrinolaringol Esp (Engl Ed)*. (2020) 71(1):32–9. doi: 10.1016/j.otorri.2019.01.002

19. McGowan FX, Kenna MA, Fleming JA, O'Connor T. Adenotonsillectomy for upper airway obstruction carries increased risk in children with a history of prematurity. *Pediatr Pulmonol*. (1992) 13:222–6. doi: 10.1002/ppul.1950130409

20. Constantin E, Tewfik TL, Brouillette RT. Can the OSA-18 quality-of-life questionnaire detect obstructive sleep apnea in children? *Pediatr*. (2010) 125(1):e162–8. doi: 10.1542/peds.2009-0731

21. Martin RCG, Brennan MF, Jaques DP. Quality of complication reporting in the surgical literature. *Ann Surg* (2002):803–13. doi: 10.1097/00000658-200206000-00007. Ann Surg. 2002 Jun;235(6):803–13. doi: 10.1097/00000658-200206000-00007

22. Alsuhebani M, Walia H, Miller R, Elmaraghy C, Tumin D, Tobias JD, et al. Overnight inpatient admission and revisit rates after pediatric adenotonsillectomy. *Ther Clin Risk Manag*. (2019) 15:689–99. doi: 10.2147/TCRM.S185193

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 04 January 2023
DOI 10.3389/fpubh.2022.1069336

OPEN ACCESS

EDITED BY
Jing Wu,
Tianjin University, China

REVIEWED BY
Xiaoyu Xi,
China Pharmaceutical University, China
Zhihao Yang,
Erasmus Medical Center, Netherlands
Lei Dou,
Shandong University, China

*CORRESPONDENCE
Xiaoyan Zhang
✉ zhang.xiaoyan@scmc.com.cn
Jiaoyang Cai
✉ caijiaoyang@scmc.com.cn

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

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

RECEIVED 13 October 2022
ACCEPTED 05 December 2022
PUBLISHED 04 January 2023

CITATION
Wang W, Dong Y, Ji M, Zhang X and Cai J (2023) Health utility of children with acute lymphoblastic leukemia in China.
Front. Public Health 10:1069336.
doi: 10.3389/fpubh.2022.1069336

Health utility of children with acute lymphoblastic leukemia in China

Wei Wang[1†], Yaqi Dong[1†], Mingjing Ji[2], Xiaoyan Zhang[2*‡] and Jiaoyang Cai[2*‡]

[1]Key Lab of Health Technology Assessment, National Health Commission of the People's Republic of China (Fudan University), School of Public Health, Fudan University, Shanghai, China, [2]Key Laboratory of Pediatric Hematology and Oncology of China Ministry of Health, Department of Hematology and Oncology, Shanghai Children's Medical Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China

Background: Acute lymphoblastic leukemia is the most common cancer in children. As the 5-year survival rate has been improved to over 80%, more emphasis is now placed on reducing therapy toxicities and enhancing health-related quality of life (HRQoL) of patients during treatment. Our objective was to measure health utility of pediatric acute lymphoblastic leukemia (pALL) patients in China, examine utility weights of different treatment phases and influencing factors of health utility, as well as identify which aspects of HRQoL were most impaired.

Methods: A cross-sectional study was conducted in Shanghai Children's Medical Center (SCMC) Affiliated to Shanghai Jiao Tong University School of Medicine in China from April to November 2021. Primary caregivers of 247 patients completed the assessment by CHU9D-CHN and health utility scores were computed for all the patients and stratified by treatment phases. Various multivariable models were constructed and the best was chosen to identify independent factors associated with utility scores. Factors affecting the most impaired dimensions were also examined.

Results: The overall mean (SD) health utility score was 0.79 (±0.17) and significantly increased from induction (0.73 [±0.19], $P < 0.001$) to consolidation (0.74 [±0.18]), and to maintenance (0.82 [±0.16]). After adjusting for potentially influencing factors, utility scores in induction ($Beta = -0.086$, $P = 0.005$) and consolidation ($Beta = -0.074$, $P = 0.043$) were constantly lower than those during maintenance. In item-level analysis, lower age and induction phase were found to be significantly associated with high severity reported on the "school work/homework" dimension. Additionally, only the induction phase (vs. maintenance, OR = 2.24, $P = 0.016$) was independently associated with the high severity level reported on the "able to join in activities" dimension.

Conclusions: This is the first study that measured health utility of children with pALL in China. Mean health utility scores increased from induction to maintenance. These provided important utility estimates that help inform

future health economic models. The phrasing of "School work/homework" in CHU9D-CHN could be further improved. More efforts are needed to design and implement specific interventions targeting at the dimension "able to join in activities" for enhancing HRQoL of children with pALL in China.

KEYWORDS

acute lymphoblastic leukemia, children, quality of life, health utility, outcome research, treatment phrase

Introduction

Acute lymphoblastic leukemia is the most common cancer in children, accounting for about 26% of cancers diagnosed between the ages of 0 and 14 years (1). Because of advances in early diagnosis and treatment, the 5-year survival rate of pediatric acute lymphoblastic leukemia (pALL) can be over 80% (2). While new types of treatment are being tested in clinical trials such as targeted therapy and immunotherapy, chemotherapy remains the standard treatment for newly diagnosed pALL patients (3, 4). The chemotherapy usually has three phases–induction, consolidation and maintenance, reflecting a reliance on sequential multidrug regimens of varying intensity to avoid development of resistance: Remission induction is the first phase of chemotherapy, lasting 4 to 6 weeks. Patients are usually admitted to a hospital for initial treatment and workup; Consolidation aims to eradicate the submicroscopic residual disease that remains after a complete remission is obtained; Maintenance is the final and longest phase with a much less intensive regimen than the prior chemotherapy (3, 5). Each phase may vary in exact length and intensity among different treatment protocols, but they all shared the same principle of chemotherapy for pALL: to provide phased and long-term standardized treatment with intensive chemotherapy in the early phase, followed by less intensive treatment (4). Today, more emphasis has been placed on reducing the toxicities of chemotherapy, focusing on the health status and quality of life of patients during the treatment.

Health utility is a quantitative measurement of health-related quality of life (HRQoL) and describes an individual's preference for a health state on a scale from 0 to 1, where 0 represents death and 1 indicates full health status with negative values assigned to states worse than death (6). It has been an essential element in the value frameworks of the American Society of Clinical Oncology (7), and also a crucial part of cost utility analysis which is now strongly recommended for clinical comprehensive evaluation of drugs and pharmacoeconomic evaluation according to two recent Chinese guidelines (8, 9). In health economic evaluation (HEE) models for pALL (10, 11), treatment phases were commonly used to define different health states. For example, Health Quality Ontario (12) set health states according to 4 treatment phases (i.e., induction, consolidation, intensification, and maintenance), while Lin et al. (13) added another health state "continuation" between intensification and maintenance as per a different treatment protocol. For each treatment phase, an accurate estimate of health utility is required to calculate quality-adjusted life years (QALYs) that facilitate comparison of the cost-effectiveness of health interventions from diverse areas.

Despite the significant value of health utility for a HEE model based on pALL treatment phases, very few studies measured health utilities of children with pALL at different treatment phases and all reported significant differences among treatment phases. In fact, only 4 relevant studies were identified from an updated search built on two recent reviews [one in 2017 by Fardell et al. (11) and the other in 2021 by Chen et al. (10)]. All were longitudinal studies using HUI2 and/or HUI3. One of them reported an improvement in general HRQoL, measured by the HUI3, in the first 12 months of treatment (14). The other three studies were from the same research team by Furlong et al. and showed a common trend that utility scores generally increased from induction to the post-treatment phase (15–17), despite differences in their treatment protocols and study periods. Since they are known to be influenced by varying socioeconomic and cultural backgrounds, it is not ideal to adopt health utility scores from overseas populations (18). Besides, considering potential differences in clinical practices of treating pALL, it is not clear whether the trend would remain for children with pALL in China. To the best of our knowledge, no study has yet reported health utilities of children with pALL at different treatment phases, or examined the association between pALL health utility values and treatment phases in China.

Methodologically, out of 9 multi-attribute utility instruments available for use in children and adolescents worldwide, CHU9D is the only utility instrument that was developed from its inception with young people (19). It contains 9 questions with child-friendly phrasing and each asks how respondents feel today in a distinct aspect (20). The most important reason for choosing it for this study is that only CHU9D contains a China-specific value set which is used as a scoring algorithm for converting CHU9D responses to utility scores (21). The Chinese version of CHU9D (i.e.,

CHU9D-CHN) was developed in line with ISPOR Task Force for translation and cultural adaptation (22) and found to be a satisfactory, reliable, and valid instrument to measure and value HRQoL for healthy children and adolescents in China (23, 24).

In view of the above, our study aims to measure health utility values of pALL patients in China with CHU9D-CHN, to evaluate utility weights of different treatment phases, to explore influencing factors of health utility, and to identify which aspect(s) of HRQoL were most impaired. The finding may add valuable information to the scarce evidence base and inform QALYs estimation in model-based health economic evaluations for allocation and prioritization of scarce healthcare resources for China. It may also help design targeted interventions to improve HRQoL of children with pALL.

Materials and methods

Subjects and study design

In this cross-sectional study, convenience sampling was used. Consecutive patient-caregiver dyads seen at the Shanghai Children's Medical Center (SCMC) Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China between April and November 2021 were recruited. To maximize our sample size, we also included dyads from an early pilot run in November 2019. The inclusion criteria of the patients were: Chinese residents who had a confirmed diagnosis of ALL were undergoing formal clinical treatment at the center, and were aged 4 to 18 years on the day of survey. All the eligible patients were undergoing treatments according to the Chinese Children's Cancer Group study ALL-2015 (CCCGALL-2015) protocol (25, 26). CCCGALL-2015 was a minimal residual disease (MRD)-directed, risk-stratified treatment protocol for newly diagnosed pALL, which has been widely adopted by 20 major hospitals/medical centers in 10 provinces, three central government direct-controlled municipalities and Hong Kong since 1 January 2015 (25, 26). The catchment areas of these centers contain approximately 65% of the population of China. CCCGALL-2015 comprises three treatment phases—induction (week 1–7), consolidation (week 8–15) and maintenance (week 16–125). In the early part (from week 16 to week 35) of the maintenance phase, patients undergo different continuation therapy according to their risk status, while from week 36–125 patients receive pulse therapy with dexamethasone and vincristine during maintenance therapy [details about the regimens were available in a published clinical paper (26)].

Accompanying primary informal caregivers of the eligible patients were surveyed to assess health utility values of the patients. Primary informal caregivers were defined as the family member who are most involved in providing care or ensuring provision of care to the patient. Participants who could not understand or read Chinese were excluded. Written informed consent was obtained from all participants. This study was reviewed and approved by the hospital ethical Review Board. Written informed consent was obtained from the parents, guardians, or patients, as appropriate.

Data collection

Socio-demographic and clinical data

Data were mainly retrieved from a prospective clinical database, which comprised data of initial and subsequent follow-up outpatient visits and hospitalization episodes of all pALL patients managed at the SCMC since inception. Informed consent was obtained from patients' legal guardians for their data to be included in the database as per institutional ethics board requirements. The socio-demographic data retrieved included patients' age, gender, urban residents (yes vs. no), living region (from vs. outside Shanghai), and annual disposable household income per capita level. Clinical data retrieved included co-morbid diseases, disease severity (intermediate/high risk vs. low), presence of serious adverse events in the preceding 4 weeks (yes vs. no), use of steroids today or in the preceding week (yes vs. no), and disease duration since diagnosis. Disease severity was represented by patient risk status which was determined by the leukemia molecular subtype and minimal residual disease [criteria for the classification of risk groups are described in the CCCGALL-2015 protocol (25)].

The income level was obtained from a pre-defined survey questionnaire to collect socio-demographic information of primary caregivers due to the lack of complete information in the database. The questionnaire consists of 10 items asking primary caregivers' age, relationship to child, number of children, etc. It was pretested with CHU9D-CHN in the pilot run of 10 patient-parent dyads to ensure all the questions were easy to understand. Information other than annual disposable household income per capita level is reserved for a separate analysis and thus was not reported in this study. The questionnaire and CHU9D-CHN were self-administered *via* paper and pencil by each consenting respondent in the outpatient clinic or wards of the SCMC hematological oncology department, with onsite assistance provided from a trained research nurse whenever needed. Two researchers conducted independently data entry of paper-based questionnaires and data extraction from the clinical database. Any discrepancy was resolved through either discussion with the respondent or verification against information documented in medical records.

Utility data

CHU9D is suitable for use in children and adolescents aged 7–17 years (27). The CHU9D-CHN consists of 9 items covering the following 9 dimensions–worry, sad, pain, tired, annoyed, school work/homework, sleep, daily routine, able to join in

activities. Each item contains five response levels presenting increasing degrees of severity within each dimension. The recall period is "today". The Chinese value set of CHU9D was derived from 923 students aged 9–17 years using the Best-Worst Scaling (BWS) and time trade-off (TTO) methods (18). It is known that self-reported utility scores are always preferred in theory, however, in our study, about 56% of children with pALL were < 8 years old and didn't have adequate reading and comprehension abilities to complete the survey (28). Therefore, this study presented the CHU9D-CHN responses that were assessed by accompanying primary informal caregivers, i.e., proxy-rated utility scores.

Statistical analysis

Descriptive statistics of patient characteristics and utility scores were presented and stratified by treatment phases (i.e., induction, consolidation, and maintenance). Continuous variables were presented as mean and standard deviation (SD), while categorical variables were shown as frequency and percentage. Differences among treatment phases were assessed using Fisher exact test for categorical variables and ANOVA or Kruskal–Wallis test for normally and non-normally distributed continuous variables, respectively. Next, bivariate analysis was conducted to identify potentially influencing factors of the utility scores, and the variables were selected for testing according to expected clinical relevance and findings from the previous studies.

Multivariable regression models were built to identify independent factors (such as age, gender, treatment phase etc.) that were associated with utility scores. For scores with highly skewed distribution, we used robust standard error estimates. To explore the independent effects of treatment phase after adjusting for different sets of confounding factors, stepwise multivariate linear regression models were first constructed whereby the independent variables were entered in the following steps: Step 1: treatment phases were entered in Model 1; Step 2: adjusted factors including age, gender, disease severity, and income level were added to Model 2, Model 3, Model 4, and Model 5, respectively. To check whether the ordinary least square (OLS) model was appropriate, the linearity of residuals, homogeneity of variance, collinearity, high leverage points, and normality of residuals were examined by plots using *performance* (V0.9.0) package in *R* (V4.1.0).

For model comparison, other types of models were also considered: general linear models (GLM) with various combinations of link functions and family distributions, Tobit model, and two-part model. The latter two models were usually used to analyze utility data with ceiling effects. Model diagnostics check was run to ensure model assumptions were satisfied for each type of model. The OLS model was chosen as our final model as it has the lowest Akaike Information Criteria (AIC).

The distribution of responses to each CHU9D item was presented as frequency and percentage. We combined the 5-level response to form a binary variable of high severity level (the two most severe degrees) vs. low severity level (the other 3 degrees). For dimensions with more than 30% of respondents reporting high severity level, Fisher exact test and Student's *t*-test were used to investigate the association of responses to each dimension with categorical participant characteristics and continuous ones respectively. Multivariable logistic regression models were also built to explore independent factors (such as age, gender, treatment phase etc.) that were associated with responses to the CHU9D dimensions which were most impaired.

All tests of significance were two-sided, and $P < 0.05$ was considered to indicate statistical significance. Analyses were performed by *R* version 4.1.0.

Results

Characteristics of participants

Primary caregivers of 247 eligible patients completed the survey with a response rate of 95.4%. At the day of survey, 21% of the patients were at the induction phase, 11% at the consolidation, and 68% at the maintenance. Table 1 presents patient demographic, socioeconomic, and clinical characteristics stratified by treatment phases. The mean (SD) age of patients was 8.11 (±3.47), 56.3% were male and 50.2% are urban residents. Most patients (198 [80.2%]) were not from Shanghai and reported annual disposable household income level above 50,000 CNY per capita (172 [69.6%]).

One hundred fifty participants (60.7%) were at intermediate/high risk (I/HR) according to the treatment protocol. The vast majority of patients (98.4%) had no serious adverse events in the past four weeks, and 64.8% were on steroids in the preceding week. The mean (SD) disease duration since diagnosis was 10.9 (±8.75) months. There are significant differences among three treatment phases in disease severity ($P = 0.043$), on steroids ($P < 0.001$), and disease duration ($P < 0.001$).

Only 11.3% of respondents reported full health (utility score is 1). The score distribution was shown in Supplementary Figure S1. The mean (SD) utility scores were 0.79 (±0.17), 0.73 (±0.19), 0.74 (±0.18), and 0.82 (±0.16) for all the patients, and those at the three treatment phases (induction, consolidation, and maintenance) respectively. There were significant differences in utility scores among the treatment phases ($P < 0.001$).

Influencing factors of utility scores

Table 2 shows the associations between patient characteristics and utility scores. Age (Pearson's $r = 0.134$, $P = 0.036$) and disease duration (Pearson's $r = 0.212$, $P < 0.001$)

TABLE 1 Demographic and socioeconomic characteristics of patients by treatment phases.

Characteristics	Total (N =247)	Treatment phases			P-value[a]
		Induction (N = 53)	Consolidation (N = 26)	Maintenance (N = 168)	
Age (years)					0.124
Mean (SD)	8.11 (±3.47)	7.33 (±3.00)	7.80 (±3.74)	8.41 (±3.54)	
Gender, *n* (%)					0.200
Male	169 (68.4)	32 (60.4)	16 (61.5)	121 (72.0)	
Female	78 (31.6)	21 (39.6)	10 (38.5)	47 (28.0)	
Urban residents, *n* (%)					0.695
No	116 (48.3)	24 (48.0)	10 (40.0)	82 (49.7)	
Yes	124 (51.7)	26 (52.0)	15 (60.0)	83 (50.3)	
Region, *n* (%)					0.474
From shanghai	49 (19.8)	13 (24.5)	6 (23.1)	30 (17.9)	
Not from Shanghai	198 (80.2)	40 (75.5)	20 (76.9)	138 (82.1)	
Annual disposable household income per capita (CNY), *n* (%)					0.905
Below 50K	72 (29.5)	14 (26.9)	8 (30.8)	50 (30.1)	
Above 50K	172 (70.5)	38 (73.1)	18 (69.2)	116 (69.9)	
Disease severity, *n* (%)					0.044*
LR	97 (39.3)	27 (50.9)	13 (50.0)	57 (33.9)	
I/HR	150 (60.7)	26 (49.1)	13 (50.0)	111 (66.1)	
Serious adverse events[b] *n* (%)					0.383
No	243 (98.4)	52 (98.1)	25 (96.2)	166 (98.8)	
Yes	4 (1.6)	1 (1.9)	1 (3.8)	2 (1.2)	
On steroids[c] *n* (%)					<0.001**
No	87 (35.2)	19 (35.8)	26 (100)	42 (25.0)	
Yes	160 (64.8)	34 (64.2)	0 (0)	126 (75.0)	
Disease duration (month)					<0.001**
Mean (SD)	10.9 (±8.75)	1.04 (±0.820)	3.50 (±0.957)	15.2 (±7.41)	
CHU9D parent-rated utility score					<0.001**
Mean (SD)	0.79 (±0.17)	0.73 (±0.19)	0.74 (±0.18)	0.82 (±0.16)	

SD, standard deviation; CNY, Chinese Yuan; LR, low risk; I/HR, intermediate /high risk.
*P < 0.05 and **P < 0.001, respectively.
[a]Fisher exact test was performed to examine association between two categorical variables; ANOVA was used to test differences in continuous variables among 3 treatment phases.
[b]Serious adverse events (SAE): having SAE in the preceding 4 weeks.
[c]On steroids: use of steroids today or in the preceding week.

were significantly associated with utility scores, while the rest were not.

Table 3 displays the stepwise multivariable linear regression models of utility scores. Treatment phase remained significant throughout Model 1 to 5. Utility scores during induction (*Beta* = −0.086, P = 0.005) and consolidation (*Beta* = −0.074, P = 0.043) were constantly lower than those during maintenance even after adjusting for age, gender, disease severity, and income level in Model 5.

Responses to CHU9D-CHN by dimensions

Figure 1 summarizes the utility scores of patients in 9 dimensions. Less than 3% reported the highest severity level on each of the following 7 dimensions: "worry" (2.46%), "sad" (2.04%), "pain" (2.02%), "tired" (2.83%), "annoyed" (2.43%), "sleep" (1.62%), and "daily routine" (1.22%). More than 30% reported higher severity levels (i.e., Level 4 and 5) on "school

TABLE 2 Bivariate analysis of associations between patient characteristics and utility scores.

Patient characteristics	$N = 247$	Utility scores		P-value[a]
		Mean	SD	
Age	247	r = 0.134		0.036*
Disease duration	247	r = 0.212		<0.001**
Gender				0.119
Male	169	0.803	0.162	
Female	78	0.763	0.196	
Disease severity				0.201
LR	97	0.773	0.169	
I/HR	150	0.802	0.177	
Serious adverse events				0.934
No	243	0.790	0.175	
Yes	4	0.797	0.109	
On steroids				0.543
No	87	0.781	0.158	
Yes	160	0.795	0.183	
Region				0.813
From shanghai	49	0.785	0.149	
Not from Shanghai	198	0.792	0.180	
Urban residents				0.319
No	116	0.803	0.180	
Yes	124	0.780	0.168	
Annual disposable household income per capita(CNY)				0.285
Below 50 K	72	0.808	0.169	
Above 50 K	172	0.782	0.176	

SD, standard deviation; CNY, Chinese Yuan; LR, low risk; I/HR, intermediate /high risk.
*P < 0.05 and **P <0.001, respectively.
[a]t-test was performed to examine association between utility scores and each categorical variable; Pearson's r was computed to test association between utility scores and each continuous variable.

work/homework" (34.43%), and "able to join in activities" (35.63%).

Table 4 presents the association between patient characteristics and responses (high vs. low severity level) to "school work/homework" and "able to join in activities" dimensions in bivariate analysis. Lower age and induction phase were found to be significantly associated with high severity reported on the "school work/homework" dimension.

Table 5 indicates that induction phase (vs. maintenance, OR = 3.46, $P < 0.001$) and lower school age (OR = 2.38, $P = 0.004$) were independently associated with the high severity level reported on the "school work/homework" dimension, after adjusting for age, gender, disease severity, and income level. Only the induction phase (vs. maintenance, OR = 2.24, $P = 0.016$) was independently associated with the high severity level reported on the "able to join in activities" dimension.

Discussion

This is the first study that measured the health utility of children with acute lymphoblastic leukemia in China, and evaluated utility weights of different treatment phases. The mean health utility score was 0.79 (SD = 0.17) for all children with pALL in our sample, approaching the lower end of the range (0.78–0.92) from previous studies that reported self-rated CHU9D scores of health Chinese children (23, 24, 29–32). The mean health utility scores were 0.73 (SD = 0.19) for those in induction, and 0.74 (SD = 0.18) for consolidation. Both were

TABLE 3 Multiple linear regression analyses of factors associated with CHU9D-CHN utility scores.

Variables	Model 1			Model 2			Model 3			Model 4			Model 5		
	Beta	SE	*P*-value	Beta	SE	*P*–value	Beta	SE	*P*-value	Beta	SE	*P*-value	Beta	SE	*P*-value
Intercept	0.819	0.013	<0.001**	0.775	0.031	<0.001**	0.782	0.031	<0.001**	0.781	0.032	<0.001**	0.800	0.038	<0.001**
Treatment phases															
Induction	−0.094	0.029	0.001*	−0.088	0.030	0.004*	−0.084	0.030	0.005*	−0.084	0.031	0.007**	−0.086	0.031	0.005*
Consolidation	−0.081	0.036	0.028*	−0.077	0.036	0.033*	−0.074	0.036	0.041*	−0.073	0.036	0.044**	−0.074	0.036	0.043*
Maintenance	Ref			Ref			Ref			Ref			Ref		
Age (year)				0.005	0.003	0.097	0.005	0.003	0.086	0.005	0.003	0.124	0.005	0.003	0.153
Gender															
Male							Ref			Ref			Ref		
Female							−0.032	0.025	0.205	−0.032	0.025	0.203	−0.029	0.025	0.245
Disease severity															
LR										Ref			Ref		
I/HR										0.005	0.024	0.848	0.005	0.025	0.843
Annual disposable household income per capita (CNY)															
Below 50K													Ref		
Above 50K													−0.025	0.034	0.300

SE, standard error; CNY, Chinese Yuan; LR, low risk; I/HR, intermediate /high risk.
*P < 0.05 and **P < 0.001, respectively.

10.3389/fpubh.2022.1069336

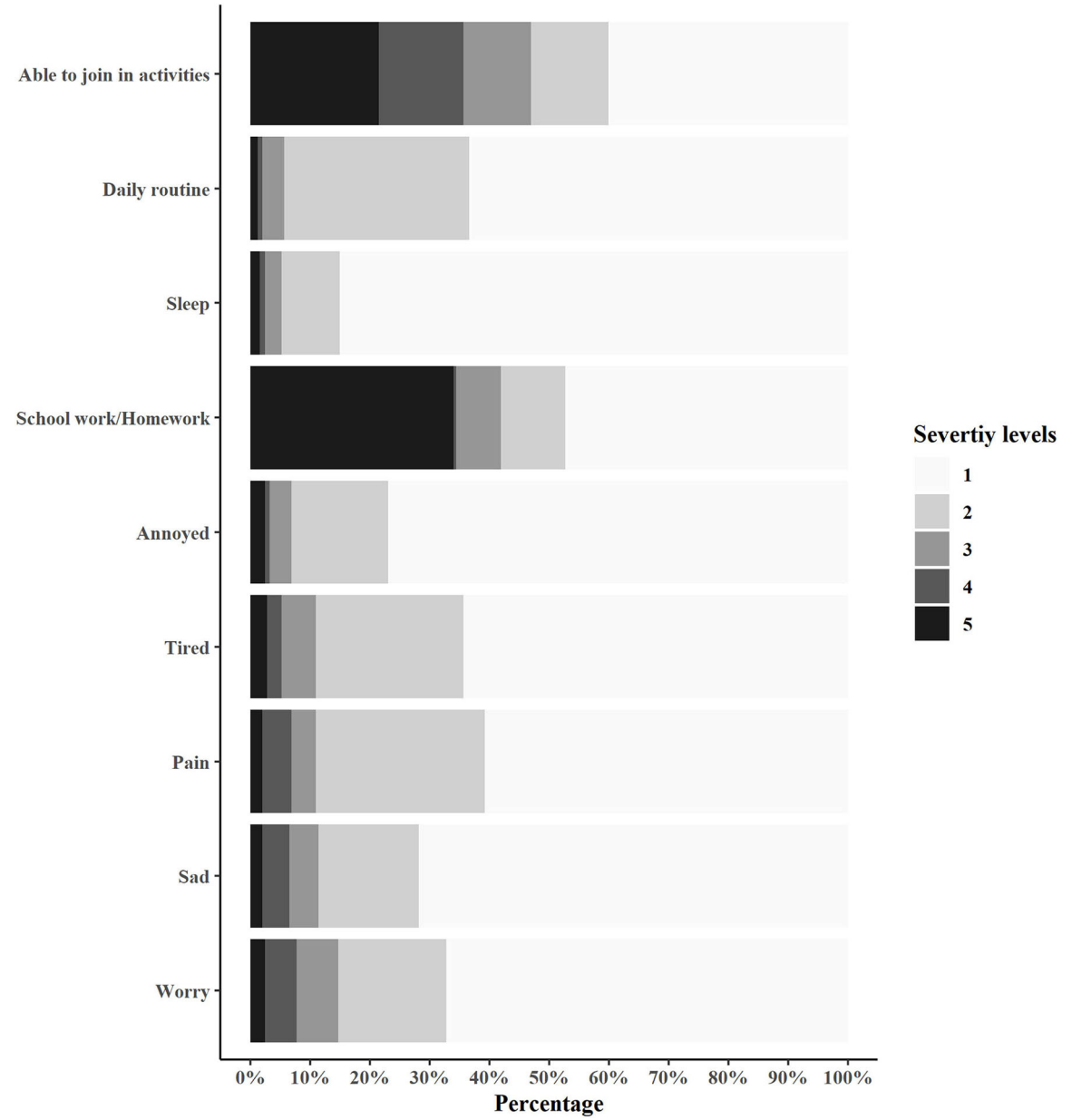


FIGURE 1
Distribution of responses to each item of CHU9D–CHN.

significantly lower than that in maintenance (mean = 0.82, SD = 0.16).

Impact of treatment phase on CHU9D utility scores

Treatment phase was found to be an independent influencing factor of health utility after adjusting for socioeconomic factors. Patients during induction seemed to have the lowest health utility. This finding was not surprising, because patients tend to experience the most pain and discomfort from adverse effects and/or complications during the induction phase whereby the chemotherapy was most intensive. Meanwhile, their primary caregivers (usually parents) were most overwhelmed by what patients suffered from the chemotherapy during this very initial stage of treatment. When it came to the second phase—consolidation, both patients and their caregivers were likely to be less overwhelmed and slowly getting used to the cancer treatment routines,

TABLE 4 Bivariate analysis of associations between patient characteristics and responses to "school work/homework" and "able to join in activities" dimensions.

Variables	School work/homework			Able to join in activities		
	Low ($N = 158$)	High ($N = 83$)	*P*-value[a]	Low ($N = 159$)	High ($N = 88$)	*P*-value[a]
Age (year)			0.002*			0.060
Mean (SD)	8.67 (±3.49)	7.24 (±3.30)		7.80 (±3.37)	8.67 (±3.59)	
School age			0.001*			0.337
Age below 6	45 (28.5)	42 (50.6)		63 (39.6)	29 (33.0)	
Age 6 or above	113 (71.5)	41 (49.4)		96 (60.4)	59 (67.0)	
Gender, *n* (%)			0.772			0.392
Male	107 (67.7)	58 (69.9)		112 (70.4)	57 (64.8)	
Female	51 (32.3)	25 (30.1)		47 (29.6)	31 (35.2)	
Disease severity, *n* (%)			0.330			0.686
LR	57 (36.1)	36 (43.4)		64 (40.3)	33 (37.5)	
I/HR	101 (63.9)	47 (56.6)		95 (59.7)	55 (62.5)	
Annual disposable household income per capita (CNY), *n* (%)			0.372			0.772
Below 50K	49 (31.4)	21 (25.3)		45 (28.8)	27 (30.7)	
Above 50K	107 (68.6)	62 (74.7)		111 (71.2)	61 (69.3)	
Treatment phases, *n* (%)			<0.001**			0.041*
Induction	23 (14.6)	28 (33.7)		28 (17.6)	25 (28.4)	
Consolidation	14 (8.9)	12 (14.5)		14 (8.8)	12 (13.6)	
Maintenance	121 (76.6)	43 (51.8)		117 (73.6)	51 (58.0)	

SD, standard deviation; CNY, Chinese Yuan; LR, low risk; I/HR, intermediate /high risk.

*P < 0.05 and **P < 0.001, respectively.

[a]Fisher exact test was performed to examine association between two categorical variables; t-test was used to examine differences in continuous variables between the two severity groups.

contributing to slightly higher utility values. In the last phase—maintenance, as the treatment intensity went lower and patient condition became more stable, their health state would be largely improved, leading to the highest health utility values.

This trend was generally consistent with the previous studies (15–17), where health utility values increased as treatment began from the induction phase, through CNS and intensification, to the continuation phase as defined by different treatment protocols. To enable a fair comparison, their CNS, intensification and continuation phases were considered equivalent to our consolidation, early (week 16–35) and late (week 36–125) maintenance phases respectively, after a careful examination of relevant treatment protocols by experienced clinicians. More specifically, our mean utility score of the induction phase (0.73) was well within the range reported by the 3 previous studies (0.66–0.81). The same finding was also shown when comparing those for the maintenance phase (i.e., 0.82 vs. 0.78–0.91).

However, for the consolidation phase, our estimate (0.74) was slightly lower than those presented by Furlong and colleagues (0.75–0.90) (15–17). A plausible explanation could be that we chose to use CHU9D, a very different utility instrument from HUI2/3 adopted by all the previous 3 studies. HUI2 and HUI3 comprise 7 and 8 dimensions respectively, covering sensation, self-care, fertility, vision, hearing, speech, ambulation/mobility, dexterity, emotion, cognition, and pain in total. Clearly, both HUI2/3 and CHU9D measure physical, mental and cognitive functions, but only CHU9D contains social function assessed by the dimension "able to join in activities." It was quite common for pALL patients who were undergoing active treatment to feel unable to join in routine activities with their peers due to frequent hospitalizations, hospital visits and/or weak physical condition. Therefore, without measuring such an important aspect of impact on HRQoL, HUI2/3 would probably over-estimate utility values of some health states at baseline and consequently under-estimate the true effect of an intervention that aims to enhance HRQoL. These

TABLE 5 Multivariable logistic regression analyses of factors associated with high severity levels reported in two CHU9D-CHN dimensions.

Variables	School work/homework		Able to join in activities	
	Odds ratio (95%CI)	*P*-value	Odds ratio (95%CI)	*P*-value
Treatment phases				
Induction	3.46 (1.75–6.92)	<0.001**	2.24 (1.16–4.34)	0.016*
Consolidation	2.19 (0.90–5.27)	0.081	2.10 (0.88–4.97)	0.090
Maintenance	Ref		Ref	
School year				
Age below 6	2.38 (1.32–4.17)	0.004*	0.66 (0.37–1.18)	0.162
Age 6 or above	Ref		Ref	
Gender				
Male	Ref		Ref	
Female	0.77 (0.41–1.42)	0.414	1.18 (0.66–2.08)	0.579
Disease severity				
LR	Ref		Ref	
I/HR	1.01 (0.56–1.83)	0.981	1.16 (0.66–2.06)	0.612
Annual disposable household income per capita (CNY)				
Below 50K	Ref		Ref	
Above 50K	1.34 (0.72–2.58)	0.363	0.91 (0.51–1.65)	0.755

CNY, Chinese Yuan; LR, low risk; I/HR, intermediate /high risk.
*P < 0.05 and **P < 0.001, respectively.

results suggested that the use of HUI2/3, as opposed to CHU9D, in cost-utility analysis may lead to less favorable estimates of incremental effects and consequently less attractive incremental cost-effectiveness ratios (ICERs) with lower chance of adopting more expensive but also more effective treatment alternatives. Nonetheless, due to the absence of head-to-head comparison between HUI2/3 and CHU9D, it clearly calls for primary studies to provide empirical evidence to elucidate this issue. More importantly, it is suggested that the choice of a preference-based HRQoL instruments for economic evaluations should be carefully justified, as it may have an important impact on the decision making based on the results of such evaluations.

Lack of association between disease severity and CHU9D utility scores

Unexpectedly, disease severity was found not associated with health utility regardless of adjustment for other influencing factors. Unlike other types of childhood cancer (10, 33), disease severity of pALL was not so much about symptoms or clinical presentation, but rather determined by molecular measurements in order to guide treatment. Nevertheless, patients with intermediate or high-risk ALL received more intensive chemotherapy than those with low risk mainly during induction and consolidation phases, hence they were more likely to have adverse events and consequently lower health utility. However, our data showed otherwise. This is not entirely consistent with previous studies either. In fact, only one of the previous pALL studies investigated the association between disease severity and health utilities. Rae and colleagues (16) found that it was only during intensification that the high-risk group had significantly lower utility scores for both HUI2 and HUI3, but there were no significant differences between groups during induction, CNS therapy, or continuation, nor at 2 years after treatment. They believed that the plausible reason could be during this stage of therapy, only high-risk patients were given doxorubicin and 3-fold higher doses of steroids that are commonly associated with nausea/vomiting and mucositis. In contrast, a possible explanation of our finding could be that treatment at any phases was well tolerated in our study sample as only very few patients ($n = 4$, 1.6%) reported serious adverse effects, thus the perceived health status of patients might be quite similar, regardless of their disease severity.

Most impaired CHU9D dimensions

Moreover, among the 9 dimensions of CHU9D-CHN, our patients were found to have more difficulties in "school work/homework," and "able to join in activities" dimensions, especially in the induction phase. Due to more frequent hospitalizations and intensive treatment, patients in the induction phase were expected to have less social contact, combined with physical and psychological discomfort at the beginning of treatment, consequently accomplish less in routine learning activities and participate less in social activities (34). Making it even worse, ever since the COVID-19 pandemic outbreak, due to ward visiting restrictions, ward volunteers were no longer allowed to access and thus had to cease our routine "sunshine care" program, which used to provide a variety of engaging activities specially designed for sick children such as playing board games, reading or general conversation for companionship and support. Without access to "sunshine care," our patients had fewer opportunities to take part in any activities during hospitalization, making the difference between induction and maintenance phases more obvious. Therefore, we call for healthcare providers, social workers, charitable organizations and other stakeholders to collaborate for developing a virtual version of "sunshine care" to provide live, online patient support, unbound by restrictions of time or location. In the future, it is hoped that "sunshine care" and similar in-person support programs be implemented for all hospitalized children with cancer. More research is also encouraged to explore the effect of other innovative interventions such as tailored physical activities (35–37) and art therapy (38) in enhancing HRQoL of pALL patients.

Another interesting finding is that our data showed preschool children were more likely to experience difficulties in the dimension of "school work/homework," than school-age children. While all the previous utility-based studies (14–17) found no association between age and HRQoL, more equivocal evidence emerged regarding child age from non-utility studies (11). In general, most of them reported either no association or an association between increasing age of a child and worse overall HRQL and/or specific HRQL domains including school and cognition subscales (11). Our seemingly counter-intuitive result in the "school work/homework" dimension might be due to the lack of clarity in the description of this particular CHU9D-CHN item. The item is phrased as "school work/homework (e.g., reading a book, writing, undertaking assignments)" by literal translation. In China, preschool children do not usually have formal "school work/homework" especially in the form of reading or writing assignments. So it is very likely that some Chinese parents might perceive this item to be irrelevant on the first sight, and thus had mistaken the last response level "can't do" for "not applicable" without a second thought. While this speculation is now being investigated in our follow-up study, researchers are strongly encouraged to pay extra attention to this particular item in a pilot testing of CHU9D-CHN to avoid potential misconceptions. Another suggestion would be adding more examples to this item, such as "reading picture books, drawing or singing songs/chants," to help parents of young preschool children relate to their routine learning activities.

Study limitations

Last but not least, our results need to be interpreted in light of several study limitations. First, a major limitation is that our patients were recruited exclusively from a designated national children's medical center—a well-known center of excellence for pediatric diseases, which may limit the study's generalizability to all pALL patients in China as it is more vulnerable to selection bias than population-based studies. The center is located in Shanghai, yet serves most provinces and cities in China and patients outside the city account for more than 70%. The selection bias favoring more severely impaired patients and/or patients from more affluent families cannot be excluded. Indeed, our patients may be from a higher socioeconomic class as suggested by their income level and willingness to seek the best specialist healthcare. Nevertheless, our sample consists of about 80% of patients outside Shanghai, 50% from rural areas, still representing a good catchment. Second, we cannot exclude residual confounding effects by unmeasured variables (e.g., family functioning level, undergoing lumbar punctures on the day of survey) in the multivariate analyses. However, we did examine treatment phases which provide comprehensive information about the treatment intensity. Third, the cross-sectional design of our study does not provide any information on changes in HRQoL as the disease progresses. Forth, despite that CHU9D is so far the only utility instrument for children with a China-specific value set, it should be noted that the value set was derived from Chinese students aged 8–17 and thus may not represent preference of the general population, or specifically, the taxpayers in China. Future studies are encouraged to investigate this speculation by developing general population-based value sets for CHU9D or other child-friendly instruments and then conducting head-to-head comparative analyses. Lastly, compared with self-rated utility scores, caregivers may overestimate or underestimate how children perceived their own HRQoL as suggested by some studies (39–41). Nonetheless, caregiver-rated utility scores were probably the best available data for our sample as about 56% were < 8 years old and didn't have adequate reading and comprehension abilities to complete the survey. More work is warranted in the future to collect and then compare utility scores from both eligible pALL children and their parents for clarifying any bias that caregiver ratings may introduce.

Conclusion

In conclusion, our study measured for the first time health utility of Chinese children with pALL. The utility scores may enable comparisons using QALYs across different clinical populations to inform decision makers for establishing healthcare priorities. We examined health utility weights of different treatment phases, providing important inputs into economic models that are utilized commonly for health resource allocations. Our results also add valuable information to the scarce evidence on factors that influence caregiver ratings of patient preferences in the area of pALL. Moreover, "school work/homework" and "able to join in activities" were identified as the most impaired aspects of HRQoL. These can have meaningful implications in developing appropriately tailored interventions to improve patient support and ultimately HRQoL for children with pALL in China.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of Shanghai Children's Medical Center. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

WW and JC contributed to conception and design of the study. MJ and XZ led the data collection. YD and WW performed the statistical analysis and wrote the first draft of the manuscript. JC and XZ contributed to interpretation of data. All authors contributed to critical revision of the manuscript for important intellectual content, read, and approved the final version.

Acknowledgments

We thank all the patients and their families, as well as the colleagues at Department of Hematology and Oncology, Shanghai Children's Medical Center for their support to this study.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2022.1069336/full#supplementary-material

References

1. Ward E, DeSantis C, Robbins A, Kohler B, Jemal A. Childhood and adolescent cancer statistics, 2014. *Cancer J Clin*. (2014) 64:83–103. doi: 10.3322/caac.21219

2. Teachey DT, Pui CH. Comparative features and outcomes between paediatric T-cell and B-cell acute lymphoblastic leukaemia. *Lancet Oncol*. (2019) 20:e142–e154. doi: 10.1016/S1470-2045(19)30031-2

3. Cooper SL, Brown PA. Treatment of pediatric acute lymphoblastic leukemia. *Pediatr Clin Am*. (2015) 62:61–73. doi: 10.1016/j.pcl.2014.09.006

4. The Subspecialty Groups of Hematology, the Society of Pediatrics, Chinese Medical Association. Recommendations for the treatment of childhood acute lymphoblastic leukemia (fourth revision). *Chin J Pediatr*. (2014) 52:641–4. doi: 10.3760/cma.j.issn.0578-1310.2014.09.001

5. PDQ® Pediatric Treatment Editorial Board. *PDQ Childhood Acute Lymphoblastic Leukemia Treatment*. Bethesda, MD: National Cancer Institute (2022). Available online at: https://www.cancer.gov/types/leukemia/patient/child-all-treatment-pdq (accessed November 17, 2022).

6. Waldeck AR, Botteman MF, White RE, van Hout BA. The importance of economic perspective and quantitative approaches in oncology value frameworks of drug selection and shared decision making. *J Manag Care Spec Ph*. (2017) 23:S6–S12. doi: 10.18553/jmcp.2017.23.6-a.s6

7. Schnipper LE, Davidson NE, Wollins DS, Blayney DW, Dicker AP, Ganz PA, et al. Updating the american society of clinical oncology value framework: Revisions and reflections in response to comments

received. *J Clin Oncol.* (2016) 34:2925–34. doi: 10.1200/JCO.2016.68.2518

8. National Centre for Comprehensive Evaluation of Drugs and Health Technologies, National Center for Drugs and Contraceptives Management, National Health Commission of the People's Republic of China. *Guidelines for the Management of Comprehensive Clinical Evaluation of Drugs*. (2020). Available online at: http://www.nhc.gov.cn/yaozs/s7656/202011/d11ddc32fae84121a0dfca36b015a31d/files/863665c55fd74f29a656990d1a4ea5ce.pdf (accessed November 17, 2022).

9. Liu G, Hu S, Wu JH, Wu J, Dong Z, Li H. *China Guideline for Pharmacoeconomic Evaluations? 2020 Chinese and English Version?* China Market Press (2020).

10. Chen P, Melissa HM, Li M, Huang I. Health utilities in pediatric cancer patients and survivors: a systematic review and meta-analysis for clinical implementation. *Qual Life Res.* (2021) 31:343–74. doi: 10.1007/s11136-021-02931-0

11. Fardell JE, Vetsch J, Trahair T, Mateos MK, Grootenhuis MA, Touyz LM, et al. Health-related quality of life of children on treatment for acute lymphoblastic leukemia: A systematic review. *Pediatr Blood Cancer.* (2017) 64:e26489. doi: 10.1002/pbc.26489

12. Health Quality Ontario. Minimal Residual Disease Evaluation in Childhood Acute Lymphoblastic Leukemia: An Economic Analysis. *Ont Health Technol Assess Ser*. (2016) 16:1–83.

13. Lin JK, Muffly LS, Spinner MA, Barnes JI, Owens DK, Goldhaber-Fiebert JD. Cost effectiveness of chimeric antigen receptor T-cell therapy in multiply relapsed or refractory adult large B-cell lymphoma. *J Clin Oncol.* (2019) 37:2105–19. doi: 10.1200/JCO.18.02079

14. Schilstra CE, McCleary K, Fardell JE, Donoghoe MW, McCormack E, Kotecha RS, et al. Prospective longitudinal evaluation of treatment-related toxicity and health-related quality of life during the first year of treatment for pediatric acute lymphoblastic leukemia. *BMC Cancer.* (2022) 22:985. doi: 10.1186/s12885-022-10072-x

15. Rae C, Furlong W, Jankovic M, Moghrabi A, Naqvi A, Sala A, et al. Economic evaluation of treatment for acute lymphoblastic leukaemia in childhood. *Eur J Cancer Care.* (2014) 23:779–85. doi: 10.1111/ecc.12173

16. Rae C, Furlong W, Feeny D, Couchman R, Silverman L, Sallan S, et al. Health status and health-related quality of life measurement in pediatric cancer clinical trials: an examination of the DFCI 00-01 acute lymphoblastic leukemia protocol. *J Pediat Hematol Onc.* (2018) 40:580–7. doi: 10.1097/MPH.0000000000001235

17. Furlong W, Rae C, Feeny D, Gelber RD, Laverdiere C, Michon B, et al. Health-related quality of life among children with acute lymphoblastic leukemia. *Pediatr Blood Cancer*. (2012) 59:717–24. doi: 10.1002/pbc.24096

18. Chen G, Xu F, Huynh E, Wang Z, Stevens K, Ratcliffe J. Scoring the child health utility 9D instrument: estimation of a Chinese child and adolescent-specific tariff. *Qual Life Res.* (2019) 28:163–76. doi: 10.1007/s11136-018-2032-z

19. Chen G, Ratcliffe J. A review of the development and application of generic multi-attribute utility instruments for paediatric populations. *Pharmacoeconomics.* (2015) 33:1013–28. doi: 10.1007/s40273-015-0286-7

20. Stevens K. Developing a descriptive system for a new preference-based measure of health-related quality of life for children. *Qual Life Res.* (2009) 18:1105–13. doi: 10.1007/s11136-009-9524-9

21. He S, Wo T, Chen L, Xi X. A review of the generic multi-attribute utility instrumentsin the economic evaluation of the pediatric medication. *Chin. Health Econom*. (2020) 39:80–3.

22. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of good practice for the translation and cultural adaptation process for Patient-Reported Outcomes (PRO) measures: report of the ISPOR task force for translation and cultural adaptation. *Value Health.* (2005) 8:94–104. doi: 10.1111/j.1524-4733.2005.04054.x

23. Zanganeh M, Adab P, Li B, Frew E. An assessment of the construct validity of the Child Health Utility 9D-CHN instrument in school-aged children: evidence from a Chinese trial. *Health Qual Life Out.* (2021) 19:205. doi: 10.1186/s12955-021-01840-7

24. Yang P, Chen G, Wang P, Zhang K, Deng F, Yang H, et al. Psychometric evaluation of the Chinese version of the Child Health Utility 9D (CHU9D-CHN): a school-based study in China. *Qual Life Res.* (2018) 27:1921–31. doi: 10.1007/s11136-018-1864-x

25. Shen S, Chen X, Cai J, Yu J, Gao J, Hu S, et al. Effect of dasatinib vs imatinib in the treatment of pediatric Philadelphia chromosome-positive acute lymphoblastic leukemia: a randomized clinical trial. *JAMA Oncol.* (2020) 6:358–66. doi: 10.1001/jamaoncol.2019.5868

26. Cai J, Yu J, Zhu X, Hu S, Zhu Y, Jiang H, et al. Treatment abandonment in childhood acute lymphoblastic leukaemia in China: a retrospective cohort study of the Chinese Children's Cancer Group. *Arch Dis Child.* (2019) 104:522–9. doi: 10.1136/archdischild-2018-316181

27. Sheffield TUO. *CHU9D - Measuring Health and Calculating QALYs for Children and Adolescents*. (2022). Available online at: https://licensing.sheffield.ac.uk/product/CHU-9D

28. Fayed N, Avery L, Davis AM, Streiner DL, Ferro M, Rosenbaum P, et al. Parent proxy discrepancy groups of quality of life in childhood epilepsy. *Value Health.* (2019) 22:822–8. doi: 10.1016/j.jval.2019.01.019

29. Zanganeh M, Adab P, Li B, Pallan M, Liu WJ, Rong L, et al. Relationship between weight status and health-related quality of life in school-age children in China. *J Health Econ Outcomes Res.* (2022) 9:75–81. doi: 10.36469/jheor.2022.32414

30. Qin Z, Wang N, Ware RS, Sha Y, Xu F. Lifestyle-related behaviors and health-related quality of life among children and adolescents in China. *Health Qual Life Out*. (2021) 19:1–8. doi: 10.1186/s12955-020-01657-w

31. Qiao H, Wang X, Qin Z, Wang N, Zhang N, Xu F. The relationship between health literacy and health-related quality of life among school-aged children in regional China. *Health Qual Life Out.* (2021) 19:262. doi: 10.1186/s12955-021-01895-6

32. Qi S, Qin Z, Wang N, Tse LA, Qiao H, Xu F. Association of academic performance, general health with health-related quality of life in primary and high school students in China. *Health Qual Life Out.* (2020) 18:339. doi: 10.1186/s12955-020-01590-y

33. Crothers A, Haeusler GM, Slavin MA, Babl FE, Mechinaud F, Phillips R, et al. Examining health-related quality of life in pediatric cancer patients with febrile neutropenia: factors predicting poor recovery in children and their parents. *Eclin. Med.* (2021) 40:101095. doi: 10.1016/j.eclinm.2021.101095

34. Kiernan G, Meyler E, Guerin S. Psychosocial issues and care in pediatric oncology. *Cancer Nurs.* (2010) 33:E12–20. doi: 10.1097/NCC.0b013e3181d5c476

35. Saultier P, Vallet C, Sotteau F, Hamidou Z, Gentet J, Barlogis V, et al. A randomized trial of physical activity in children and adolescents with cancer. *Cancers*. (2021) 13:121. doi: 10.3390/cancers13010121

36. Beulertz J, Prokop A, Rustler V, Bloch W, Felsch M, Baumann FT. Effects of a 6-month, group-based, therapeutic exercise program for childhood cancer outpatients on motor performance, level of activity, and quality of life. *Pediatr Blood Cancer.* (2016) 63:127–32. doi: 10.1002/pbc.25640

37. Morales JS, Santana-Sosa E, Santos-Lozano A, Baño-Rodrigo A, Valenzuela PL, Rincón-Castanedo C, et al. Inhospital exercise benefits in childhood cancer: A prospective cohort study. *Scand J Med Sci Spor.* (2019) 30:126–34. doi: 10.1111/sms.13545

38. Aguilar BA. The efficacy of art therapy in pediatric oncology patients: an integrative literature review. *J Pediatr Nurs.* (2017) 36:173–8. doi: 10.1016/j.pedn.2017.06.015

39. Jiang M, Ma Y, Li M, Meng R, Ma A, Chen P. A comparison of self-reported and proxy-reported health utilities in children: a systematic review and meta-analysis. *Health Qual Life Out*. (2021) 19:45. doi: 10.1186/s12955-021-01677-0

40. Khanna D, Khadka J, Mpundu-Kaambwa C, Lay K, Russo R, Ratcliffe J, et al. Are we agreed? self- versus proxy-reporting of paediatric health-related quality of life (hrqol) using generic preference-based measures: a systematic review and meta-analysis. *Pharmacoeconomics.* (2022) 40:1043–67. doi: 10.1007/s40273-022-01177-z

41. Khadka J, Kwon J, Petrou S, Lancsar E, Ratcliffe J. Mind the (inter-rater) gap. An investigation of self-reported versus proxy-reported assessments in the derivation of childhood utility values for economic evaluation: a systematic review *Soc Sci Med.* (2019) 240:112543. doi: 10.1016/j.socscimed.2019.112543

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 11 January 2023
DOI 10.3389/fpubh.2022.1050835

OPEN ACCESS

EDITED BY
Jagdish Chandra,
ESIC Model Hopsital and PGIMSR,
Basaidarapur, India

REVIEWED BY
Sezaneh Haghpanah,
Shiraz University of Medical Sciences, Iran
Arlette Setiawan,
Padjadjaran University, Indonesia

*CORRESPONDENCE
Hui-Jun Zhou
✉ lockgo2luck@126.com
Pei Wang
✉ wang_p@fudan.edu.cn

†These authors have contributed equally to this work

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

RECEIVED 22 September 2022
ACCEPTED 22 December 2022
PUBLISHED 11 January 2023

CITATION
Sun Y, Zhou H-J, Shen A, Wu B, Wang W, Luo N and Wang P (2023) A cross-sectional study evaluating health-related quality of life of Chinese pediatric patients with hematological malignancies using EQ-5D-Y. *Front. Public Health* 10:1050835. doi: 10.3389/fpubh.2022.1050835

A cross-sectional study evaluating health-related quality of life of Chinese pediatric patients with hematological malignancies using EQ-5D-Y

Yue Sun[1,2†], Hui-Jun Zhou[3*†], Anle Shen[4], Bin Wu[5], Wei Wang[6], Nan Luo[7] and Pei Wang[2,6*]

[1]Division of Medical Affairs, Ruijin Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China, [2]Key Lab of Health Technology Assessment, National Health Commission of the People's Republic of China (Fudan University), Shanghai, China, [3]Business School, University of Shanghai for Science and Technology, Shanghai, China, [4]Department of Pharmacy, Shanghai Children's Medical Centre, School of Medicine, Shanghai Jiaotong University, Shanghai, China, [5]Medical Decision and Economic Group, Department of Pharmacy, Renji Hospital, School of Medicine, Shanghai Jiaotong University, Shanghai, China, [6]School of Public Health, Fudan University, Shanghai, China, [7]Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore

Background: The study aimed to assess health-related quality of life (HRQoL) and to estimate the health utility of pediatric patients with hematological malignancies (HMs) in China.

Method: A cross-sectional study recruited a series of pediatric inpatients diagnosed with HM from November 2018 to May 2019 in the Shanghai Children's Medical Center. Subjects were interviewed to collect sociodemographic information about themselves and their guardians. The EQ-5D-Y was completed by each patient to rate their own HRQoL, which later derived the health utility. The health status was also assessed by clinicians following the Eastern Cooperative Oncology Group (ECOG) system. Upon the descriptive analysis and univariate analysis, multivariate generalized linear models were built to explore the associations of risk factors with HRQoL measures of utility, Visual Analog Scale (VAS) score, and the five EQ-5D-Y domains.

Results: The 96 subjects had a mean age of 10.5 years and included 62 (64.4%) boys. There were 46 (47.9%) and 25 (26.0%) children diagnosed with acute lymphoblastic leukemia and non-Hodgkin's lymphoma, respectively. The means (SD) of utility and EQ-VAS scores were 0.88 (0.10) and 85.8 (15.1), respectively. Twenty-six (27.1%) patients were graded poor health by the ECOG standard (score 2/3). Both univariate and multivariate analyses found strong correlations between ECOG and HRQoL. After adjusting for covariates, poor ECOG score was significantly associated with an impaired utility and VAS of −0.103 and −8.65, respectively. With regard to individual HRQoL domains, worse ECOG was more likely to report health problems with an increased risk of 2.94 to 12.50; residence, income, guardians' education, and disease duration were also found to be significantly related to either the utility or certain health domains.

Conclusion: The HRQoL of Chinese pediatric patients with HM is considered relatively poor and of great concern to healthcare. With the strong correlations between EQ-5D-Y-related HRQoL measures and the traditional clinical index ECOG, the EQ-5D-Y is able to provide valuable evidence for clinical decision-making at the individual level. At the same time, its health utility can inform resource allocation at a macro level.

KEYWORDS

health-related quality of life, health utility assessment, pediatric patient, hematological malignancies, EQ-5D-Y

Introduction

Health-related quality of life (HRQoL) of pediatric patients is an important topic in resource allocation and healthcare service optimization in pediatric cancer care. Health policies informed by utility-based economic evaluations require reliable estimates of utility (1). To yield valid HRQoL for pediatric patients, an age-appropriate instrument was required given that HRQoL measurement in children is different from adults. The EQ-5D-Y (youth) is a generic HRQoL instrument adapted from the adult EQ-5D questionnaire, specifically to measure the health status of children between 8 and 18 years (2). As one of the EQ-5D questionnaire series, it is able to derive health utility from the five dimensions describing health status. Its psychometric properties have been validated in local pediatric populations (3), laying the foundation to be applied in clinical practice.

Hematological malignancies (HMs) are common tumors in the child populations and made the second highest mortality among pediatric tumors in China (4). The prevalence of HMs was 4.2/100,000 in children younger than 14 years old, and ∼15,000 new cases were diagnosed every year in those below 18 years (5). The treatment for HM was costly with annual treatment costs between 100 and 300 K Chinese Yuan (CNY) for one patient, thus imposing a heavy financial burden on the family and society as well (6).

The health of patients with HM was debilitated not only by the HM-specific symptoms such as anemia, headache, fatigue, and pain but also by the adverse effects related to radiotherapy and chemotherapy. Patients had to be hospitalized frequently in the course of treatment and recovery (7, 8). Studies have reported worse health status of adult patients with HM compared to the general population (9). The situation can only be worse in the younger patients who are usually unable to handle the health problems on their own. After all, children are in development when their cognitive, physiological, and psychological functions are not mature. Both overseas and local studies in China have reported decreased HRQoL in pediatric patients with HM (10–12). Their physical activities, school performance, and HRQoL were impaired and associated with younger age, inattentive parenting patterns, lower guardians' educational level, and poor economic status of the family (8–10, 12).

With the advancement of HM care, children with HM now can survive to adulthood and most of them expect a normal life span. Therefore, the HRQoL of such patients turned up to be a practical issue. Unfortunately, the evidence of HRQoL in this population is scarce so far in China. Health utility was hardly reported. The cost-utility analyses hence lack the primary input of utility to support evidence-based policymaking in the care of children with HMs.

Under this circumstance, we hypothesized that Chinese pediatric patients with HM have impaired HRQoL and lower utility compared to their peers. The utility data about this population would accelerate the economic studies to better the oncological care of HM. Therefore, we conducted this study: (1) to assess the HRQoL in children/adolescents who were diagnosed with HM in China; (2) to estimate their health utility; and (3) to determine the significant factors for the HRQoL.

Methods

Study sample

The study population was defined as Chinese pediatric patients diagnosed with any type of HM in mainland China. Due to the specificity of the disease, we chose Shanghai Children's Medical Center, a tertiary hospital specializing in pediatric care, as the study site. A series of patients were approached once they were admitted to the hospital and settled down in the ward. The inclusion criteria were: (1) a confirmed diagnosis of HM; (2) aged 8–18 years; (3) no concurrent debilitating conditions; (4) no major invasive procedures in the past 3 months; and (5) able to understand questions. Considering the availability of eligible patients, constraints of research resources and the number of independent variables to be explored in the multivariate analysis, n = 100 was determined as the targeted sample size. During the study period from November 2018 to May 2019, the 96 eligible patients who were approached all consented to participate in

the study. The informed consent form was signed by parents or next-of-kin before the patient interview and data collection.

The study adhered to the tenets of the Declaration of Helsinki involving human participants and ethics approval was obtained from the Institutional Review Board of the School of Public Health, Fudan University.

Data collection

Data were collected through face-to-face interviews conducted by the trained interviewers with patients, their guardians, and doctors-in-charge. During the interview, information about patients' demographics, socioeconomic status, and guardians was gathered. Data about guardians including age, education level, occupation, and monthly income of the family were collected. At the same time, patients rated their own health status by completing the EQ-5D-Y questionnaire. Doctors-in-charge evaluated each patient in terms of the Eastern Cooperative Oncology Group (ECOG) performance score, mental consciousness, reactions, complexion, and petechia.

ECOG performance score

The Eastern Cooperative Oncology Group (ECOG) performance score, which was given by the doctor-in-charge, ranks a cancer patient's ability to perform daily activities in five grades (13): 0–fully active and no performance restrictions; (1) strenuous physical activity restricted, fully ambulatory and able to carry out light work; (2) capable of self-caring but unable to work, over 50% of daytime to get up and have activities; (3) capable of only limited self-caring, confined to bed or chair over 50% of daytime; (4) completely disabled and cannot self-care, totally confined to bed or chair; (5) death. An ECOG score of 0 or 1 indicates good health while higher scores (2–4) indicate poor health (14).

Health-related quality of life measurement

With the guidance of the trained interviewers, the patients completed the EQ-5D-Y rating of their health on the day of the survey. EQ-5D-Y comprises two sections, a descriptive system and a visual analog scale (EQ-VAS). The former asks a subject to evaluate his/her HRQoL on five dimensions, that is, "walking about," "looking after myself," "doing usual activities," "having pain or discomfort," and "feeling worried, sad, or unhappy" with each dimension categorized into three severity levels (no problems, some problems, and a lot of problems). The EQ-VAS ranging from 0 (the worst imaginable health) to 100 (the best imaginable health) inquires about the overall health rated by a patient itself. Recent HRQoL studies on Chinese pediatric patients with HM reported good or very good test–retest reliability for EQ-5D-Y as the Gwet agreement coefficients varied from 0.628 to 0.901 for the individual dimensions and intraclass correlation coefficients were 0.833 for the EQ-VAS (15). Convergent validity and known-groups validity were tested moderately in healthy Chinese children aged 8–18 years (16).

Health utility would be derived by incorporating patients' responses to the five dimensions following specific value sets for utility calculation. In this study, we were obliged to adopt the Japanese value set as the Chinese one was under development (17). As per the Japanese value set, the worst HRQoL corresponds to a utility of 0.067, while full health takes a utility of 1.0.

Statistical analysis

Descriptive statistics summarized the demographic, socioeconomic, and clinical characteristics of the sample. The internal consistency of EQ-5D-Y in measuring HRQoL was checked with Cronbach's Alpha. Then, the HRQoL was illustrated by the proportions of "reporting problems" on the five individual dimensions. The overall HRQoL was estimated as the means (standard deviation, SD) of utility and EQ-VAS scores. The general health grades in the ECOG system were also reported as proportions. The chi-square test, *t*-test, and analysis of variance were used to compare health outcomes among different levels of each variable. Correlation analyses explored the relationship between continuous variables and utility and VAS.

Multivariate generalized linear models (GLMs) were built separately on the utility and EQ-VAS score to explore their associations with potential risk factors, that is, patients' age, gender, education level (middle school vs. primary school), residence (rural area vs. urban), disease duration, diagnosis, and ECOG score (poor vs. good), and their guardians' education level (primary school or below/middle school vs. college or above), and monthly income (<CNY 5,000 vs. ≥5,000) (18, 19). For the categorical variable put in the model, each category must have 20 observations at a minimum. To assess the relationship between individual health dimensions and the factors, five GLMs were established with a binary dependent variable indicating the existence of any problems in each of the five dimensions.

The analyses were performed using SPSS (version 19) at a significance level of 0.05.

Results

As summarized in Table 1, the patient's mean age (SD) was 10.5 (2.2) years, with the boys accounting for 64.6%. The

TABLE 1 Demographic, socioeconomic, and clinical characteristics of patients with hematological malignancies.

Patients (%, *N*)		Guardians (%, *N*)	
Age (year)*	10.5 (2.2)	Age (year)*	40.1 (9.3)
Gender		Marriage	
Boy	64.6% (62)	Unmarried	3.1% (3)
Girl	35.4% (34)	Married	95.8% (92)
Education		Religion	
Primary school	79.2% (76)	Yes	16.7% (16)
Middle school	20.8% (20)	No	78.1% (75)
Residence		Relationship	
Urban	53.1% (51)	Father	20.8% (20)
Rural area	44.8% (43)	Mother	67.7% (65)
Diagnosis		Other	11.5% (11)
Acute lymphoblastic leukemia	47.9% (46)	Occupation	
Non-Hodgkin's lymphoma	26.0% (25)	Employee	21.9% (21)
Acute myeloid leukemia	10.4% (10)	Civil servant/public institution	12.5% (12)
Rhabdomyosarcoma	4.2% (4)	Housewife	19.8% (19)
Osteosarcoma	2.1% (2)	Retired	3.1% (3)
Other hematological malignancies	9.4% (9)	Peasant	18.8% (18)
Disease duration (months)*	8.3 (16.5)	Temporary worker	3.1% (3)
ECOG		Unemployed	15.6% (15)
0	16.7% (16)	Education	
1	56.3% (54)	Primary school or below	18.8% (18)
2	20.8% (20)	Middle school	45.8% (44)
3	6.3% (6)	College or above	33.3% (32)
Clinical characteristics		Monthly income (CNY)	
Mental consciousness (Good)	100.0% (96)	<5,000	40.6% (39)
Reactions (Good)	100.0% (96)	5,000–10,000	16.7% (16)
Complexion (Good)	91.7% (88)	10,001–30,000	12.5% (12)
Petechia (No)	86.5% (83)	>30,000	6.3% (6)

*Mean (SD).

majority (79.2%) of the sample were in primary education. The most common diagnosis was acute lymphoblastic leukemia (47.9%), followed by non-Hodgkin's lymphoma (26.0%). The mean (SD) disease duration was 8.3 (16.5) months with a wide range from 3 days to 7.6 years. There were 16 (16.7%) and 54 (56.3%) patients graded at ECOG score of 0 or 1 respectively, indicating good physical function.

Cronbach's alpha of EQ-5D-Y was 0.713, meaning the internal consistency of EQ-5D-Y in measuring this sample was acceptable. The utility score had a mean (SD) of 0.88 (0.10) and a range from 0.508 to 1. The self-rated global health by VAS score ranged from 45 to 100 around the mean (SD) of 85.8 (15.1). The HRQoL profile composed of the individual health domains is shown in Figure 1. In general, about 51.0–65.6% of patients had no problems on the five dimensions, and patients who had some problems accounted for 33.3–42.7% ("mobility": 33.3%, "having pain or discomfort": 42.7%). The most affected domain was "looking after myself" seeing that seven (7.29%) subjects reported the worst state.

As shown in Table 2, VAS and utility were different in categories defined by individual patient characteristics. Patients in the guard of parents finishing middle school or above rated higher VAS than those cared for by parents with an education level of primary school or below. Patients graded as having

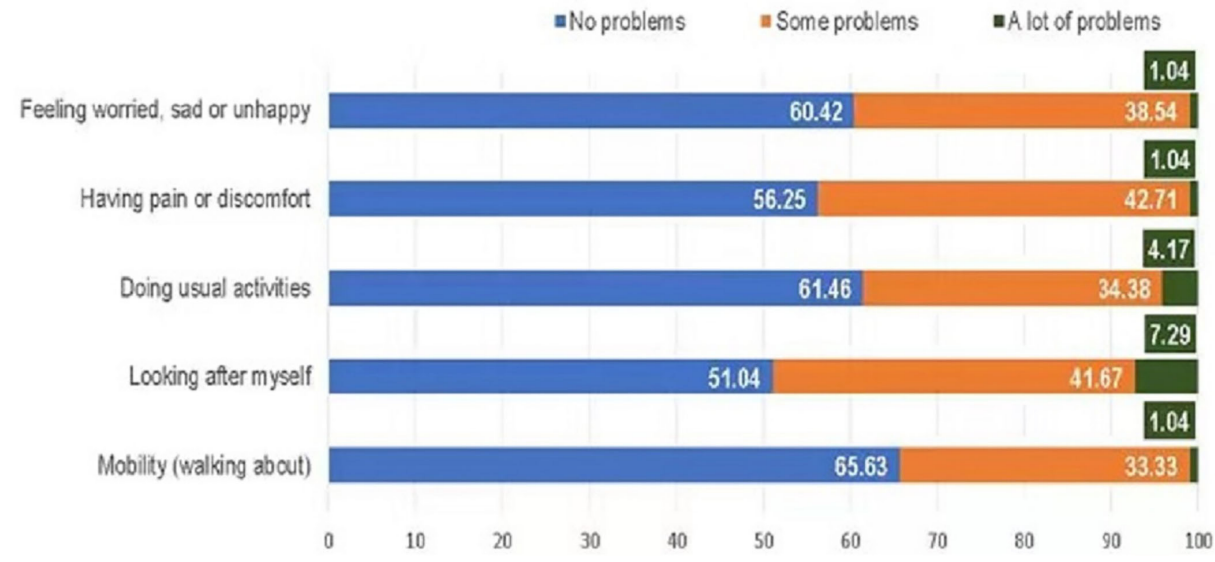


FIGURE 1
Health-related quality of life profile in EQ-5D-Y dimensions. Numbers are percentages.

a good performance by ECOG (0/1) had mean (SD) utility and VAS of 0.91 (0.09) and 88.9 (10.7), respectively, which was significantly higher than patients graded ECOG 2/3 with the mean (SD) utility and VAS of 0.81 (0.09) and 77.3 (21.2), respectively. ECOG was linked to the HRQoL profile (Figure 2). The patients graded ECOG 2/3 reported significantly more problems in four of the five EQ-5D-Y dimensions than those graded ECOG 0/1, except for the dimension "feeling worried, sad, or unhappy." The most affected domains in this group were "looking after myself" and "doing usual activities" both captured 15.38% of the sample reporting "a lot of problems."

According to the GLM examining the associations of patients' characteristics with HRQoL indices (Table 3), ECOG emerged as the strongest factor to predict the concurrent utility and VAS. After adjusting for other factors, good performance by ECOG contributed to 0.103 utility and 8.65 VAS compared to poor ECOG performance. In addition, disease duration showed a significant positive association with utility with an efficiency of 0.001. Patients diagnosed with acute lymphoblastic leukemia or non-Hodgkin's lymphoma tended to achieve higher utility and VAS than those diagnosed with other types of HM; however, the differences did not demonstrate statistical significance.

The odds ratio (OR) quantifying the associations of risk factors with health problems on individual EQ-5D-Y dimensions is summarized in Table 4. After controlling for the covariates, significant associations were observed between ECOG and health dimensions each, showing that poor ECOG indicated bigger probabilities of having HRQoL problems ("working about" [OR = 5.88; 95% CI: 2.08, 20.03], "looking after myself" [OR = 9.09; 95% CI: 2.56, 50.04], "doing usual activities" [OR = 3.70; 95%CI: 1.35, 11.01], "having pain or discomfort" [OR = 12.50; 95% CI: 3.85, 47.62], and "feeling worried, sad, or unhappy" [OR = 2.94; 95% CI: 1.11, 9.09]). The patients living on a monthly income below CNY 5,000 were less likely to report problems on "having pain or discomfort" (OR = 0.22; 95%CI: 0.05, 0.82) compared to those who lived on CNY 5,000 or more. Significant associations were observed between guardians' educational level with "looking after myself" and "feeling worried, sad, or unhappy" dimensions. Compared with the patients raised by parents with college or higher education, patients whose guardians' educational level was middle school tended to report fewer problems (OR = 0.28; 95% CI: 0.08, 0.92) on "feeling worried, sad, or unhappy," while those raised by parents with the educational level of primary school or below reported more problems on "looking after myself" with an increased likelihood (OR = 10.05; 95% CI: 1.35, 92.73). In addition, patients in rural areas significantly reported fewer problems in "looking after myself" dimension compared to those living in the urban area (OR = 0.07; 95% CI: 0.01, 0.28).

Discussion

Health-related quality of life of pediatric cancer patients is of great importance in optimizing oncological care for children. With the prolonged survival time and the consequently growing number of childhood cancer survivors, evaluating the HRQoL of such populations became both a medical and public health necessity. This study selected a sample of children with HM and evaluated their HRQoL with an age-appropriate generic

TABLE 2 Univariate analyses showing VAS and utility distribution for different risk factors.

	N	Utility		VAS	
		R	p	r	p
Age	96	−0.021	0.842	−0.072	0.488
Disease duration (months)	96	0.136	0.187	0.03	0.774
		Mean (SD)	p	Mean (SD)	p
Gender					
Female	34	0.885 (0.106)		84.21 (19.15)	
Male	62	0.877 (0.097)	0.816	86.65 (12.52)	0.453
Education					
Primary school	76	0.879 (0.102)		85.68 (15.73)	
Middle school	20	0.885 (0.093)	0.816	86.15 (13.06)	0.903
Residence					
Rural	43	0.892 (0.101)		84.44 (18.43)	
Urban	53	0.87 (0.1)	0.287	86.87 (11.92)	0.438
Guardians' education					
Primary school or below	20	0.855 (0.109)		76.2 (23.56)	
Middle school	44	0.897 (0.1)		89.36 (11.04)	
College or above	32	0.872 (0.094)	0.258	86.84 (10.74)	0.004
Monthly income (CNY)					
<5,000	47	0.887 (0.102)		88.85 (10.3)	
≥5,000	49	0.874 (0.099)	0.534	82.84 (18.29)	0.049
Diagnosis					
Acute lymphoblastic leukemia	46	0.877 (0.101)		87.13 (12.62)	
Non-Hodgkin's lymphoma	25	0.902 (0.08)		89.32 (7.97)	
Others	25	0.863 (0.116)	0.389	79.76 (22.17)	0.057
ECOG					
Good (0/1)	70	0.908 (0.089)		88.94 (10.74)	
Poor (2, 3)	26	0.806 (0.091)	<0.001	77.27 (21.21)	0.012

instrument EQ-5D-Y. We found that the health utility of HM children was 0.88 on average and the self-rated global health VAS score had a mean of 85.8. The clinical index ECOG had strong correlations with the overall HRQoL indices and all the five EQ-5D-Y domains. Family income, disease duration, guardian's education, and residence were also significant factors exerting a differential effect on utility, VAS, "looking after myself," and specific domains.

It is well-known that patients with cancer had worse health status than their healthy peers. In our study, the global health rated as EQ-5D-Y VAS was 85.8 on average. Compared to the healthy school-going children in China who achieved a mean VAS of 87.6 (16), the difference seems to be negligible. However, the HRQoL profile of our sample did not appear as satisfactory as those of healthy populations. There were 33.3–41.7 and 1.0–7.3% of our sample reporting "some problems" and "a lot of problems," respectively on "mobility," "looking after myself," and "doing usual activities" dimensions. These numbers were much higher than the proportions (2.4–7.8 and 0.1–1.9%, respectively) reported in the general Chinese population (20). The findings indicate that HMs may have negative impacts on the physical functions of patients. Similar results have been reported in another study in mainland China showing that children with HMs have a lower level of physical activity (21). The dimension "looking after myself" presented the highest prevalence of problems among the five health dimensions. This can be explained by the fact that children/adolescents were not mature enough to take care of themselves, let alone

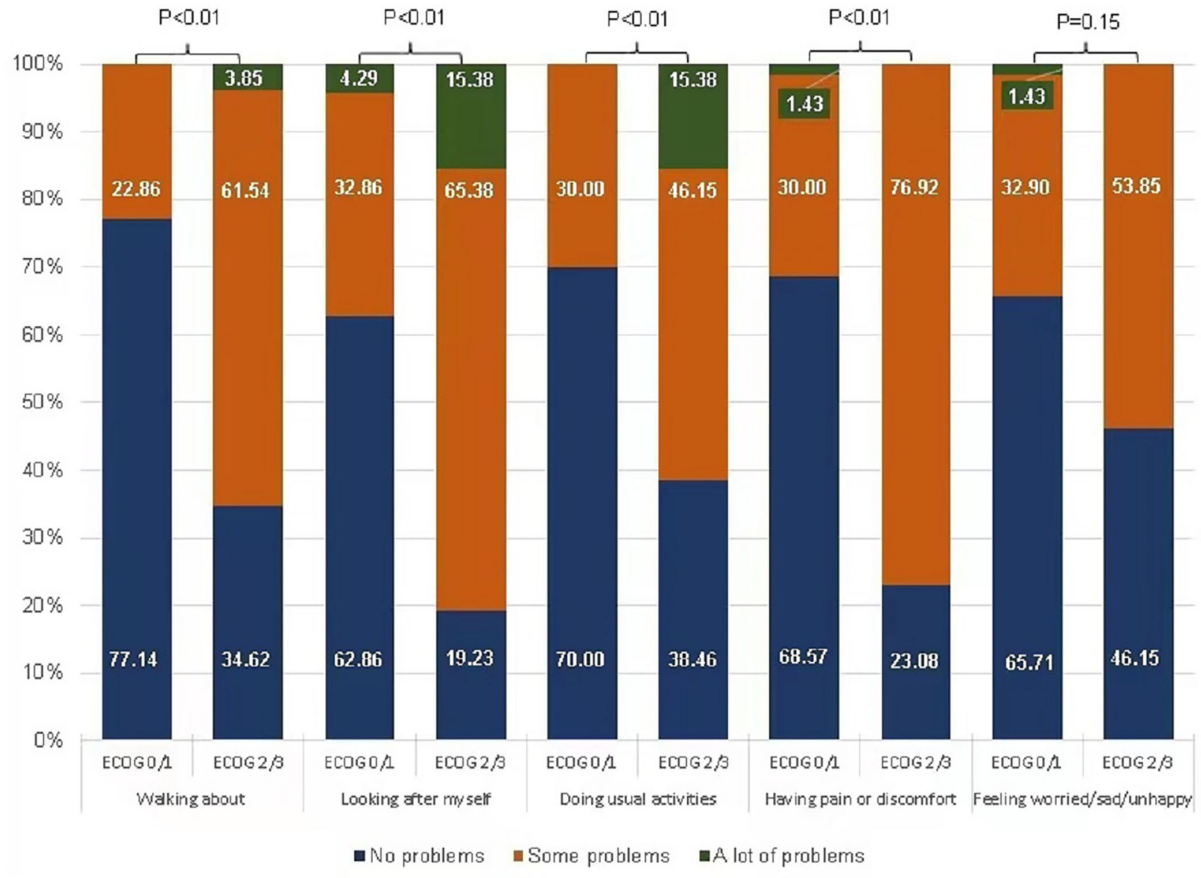


FIGURE 2
Health-related quality of life profile in EQ-5D-Y dimensions by ECOG status. Numbers are percentages.

those inflicted with a malignancy. The two psychological-related dimensions, "having pain or discomfort" and "feeling worried, sad or unhappy" also captured 38.5 and 42.7% of patients reporting problems, which were higher than the proportions (17.6 and 20.5%) reported in the general students (20). It has been acknowledged that treatments such as radiotherapy and chemotherapy could lead to fatigue and uncomfortableness (21), thus the patients tend to be stressed, depressed, and anxious. The relationship between the patients with HM and mental health was also identified previously (22).

The proxy rating has an important role in assessing the HRQoL of pediatric patients. ECOG grades the physical function of a patient with cancer by the doctor-in-charge. Our study showed that ECOG is the strongest factor predicting the concurrent utility, VAS, and all five EQ-5D-Y dimensions. Patients who were graded poor health in ECOG performance scores had lower utility and EQ-VAS scores. A similar relationship has been observed in Chinese adult patients with HM (23). As for the individual domains, more problems were reported if the patients were graded as having poor ECOG scores with an effect size (OR) ranging from 2.94 on "Feeling worried, sad or unhappy" to 12.50 "have pain or discomfort."

Standing on the view of HRQoL measurement, health status evaluated with the ECOG system and HRQoL instrument are concordant for our sample of pediatric patients with HM. Both ECOG and HRQoL are independent health evaluations performed by clinicians and patients themselves. It has been suggested that the health status rated by the pediatric patients alone is not comprehensive (24) where bigger differences existed between proxy-reported and self-reported health outcomes in measuring pediatric patients than adults (25). Therefore, the agreement between the two health indicators can serve as a validation of our results. Considering ECOG is a well-established clinical approach assessing the health outcome of patients with cancer (26), ECOG and HRQoL measurements are complementary to each other to better understand patients' health and clinical needs.

The educational level of guardians was also found to have significant associations with dimensions of "looking after

TABLE 3 Multivariate analyses evaluating associations of factors with utility and VAS.

	EQ-5D-Y utility score		EQ-VAS score	
	β (95% CI)	*P*	β (95% CI)	*P*
Age (years)	−0.002 (−0.014, 0.01)	0.755	−0.67 (−2.08, 0.74)	0.350
Gender (boy vs. girl)	−0.009 (−0.046, 0.027)	0.613	−1.59 (−6.34, 3.16)	0.512
Education (primary school vs. middle school)	−0.029 (−0.087, 0.029)	0.327	−3.85 (−10.99, 3.29)	0.290
Residence (rural vs. urban)	0.026 (−0.013, 0.066)	0.194	2.10 (−3.88, 8.09)	0.491
Monthly income (CNY)				
<5,000 vs. ≥5,000	0.03 (−0.003, 0.064)	0.075	5.07 (−0.17, 10.3)	0.058
Guardians' education				
Primary school or below vs. college or above	−0.016 (−0.08, 0.048)	0.625	−3.48 (−12.73, 5.77)	0.461
Middle school vs. college or above	0.026 (−0.01, 0.061)	0.156	4.05 (−1.66, 9.75)	0.165
Diagnosis				
Acute lymphoblastic leukemia vs. others	0.026 (−0.022, 0.074)	0.284	4.84 (−0.62, 10.29)	0.082
Non-Hodgkin's lymphoma vs. others	0.030 (−0.021, 0.081)	0.249	4.71 (−1.43, 10.84)	0.133
Disease duration (months)	0.001 (0, 0.002)	0.014	0.060 (−0.07, 0.19)	0.344
ECOG (poor vs. good)	−0.103 (−0.14, −0.066)	<0.001	−8.65 (−13.66, −3.64)	0.001

myself" and "feeling worried, sad or unhappy." The findings were consistent with a study conducted among children with lymphoblastic leukemia in mainland China (22). Additionally, patients in rural areas had fewer problems with "looking after myself" and "Doing usual activities" dimensions compared to those living in urban areas, this might be due to the fact that patients residing in urban areas would rely more on caregivers and be less independent than children in the rural places.

Nowadays, pediatric patients with HM can survive to adulthood with a generally normal life expectancy. The care after the acute treatment phase bears great significance for this patient population to live a life as good as their peers, which is achievable as they are in the development and maturity processes. Targeted recovery care, home care, and patient and guardian's education should be promoted to improve the HRQoL of pediatric patients considering their particular physical, psychological, and clinical conditions. The effect of such interventions has been reported in interventional studies (27–30).

EQ-5D-Y is new to Chinese clinicians yet its wide use in clinical practice is expected in the near future, seeing that validation studies are either ongoing or published in Chinese children (3, 16). Pioneering studies such as the present one can shed a light on its performance in the target population. The ceiling effect of EQ-5D-Y was not prominent in this study, as only 21.88 and 18.75% of the sample reported full health in utility and VAS. A higher ceiling effect of 44.1% was observed in measuring healthy children (16). The above phenomenon is in accordance with clinical reality. It further shows that HRQoL assessed by EQ-5D-Y in pediatric patients with HM is unlikely to be biased upward by the ceiling effect. The correlation of VAS and utility in our study was moderate at a coefficient of 0.51, whereas in the adult population, these two EQ-5D measures are always strongly correlated. This may reflect an issue of separation between VAS and utility in measuring HRQoL of pediatric patients. The VAS was given by the children themselves to show their personal perception of health, whereas utility was calculated against the value set usually developed on the preference of adults, that is, taxpayers according to the principle (31). Controversy has not been solved about whose preference should be used to establish the value set for the health status of children (24). The current study was obliged to apply the Japanese value set reflecting the preference of adults aged 20–79 years (17). However, a child tends to be more optimistic about their status than adult proxies (32). The separation between VAS and utility may limit the usefulness of potential cost-utility analysis on pediatric patients. The optimal services recommended by a cost-utility analysis may not serve the target population as well as it is projected. Additionally, our findings direct to the need of conducting research on China-specific value sets.

Currently, studies investigating the health utility of pediatric hematological malignancies are rare in China. Consequently, economic evaluations based on utility will lack an important input of health outcomes to inform policymaking in pediatric cancer. There is a real need for health

TABLE 4 Multivariate analyses evaluating associations of factors with each of the five health domains.

	Mobility		Looking after myself		Doing usual activities		Having pain or discomfort		Feeling worried, sad or unhappy	
	OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*
Age (years)	0.97 (0.69, 1.37)	0.867	0.92 (0.63, 1.32)	0.631	0.87 (0.61, 1.22)	0.383	1.02 (0.71, 1.45)	0.919	1.07 (0.78, 1.46)	0.686
Gender (girl vs. boy)	1.79 (0.62, 5.31)	0.246	0.66 (0.23, 1.92)	0.462	1.64 (0.6, 4.63)	0.317	2.19 (0.76, 6.72)	0.155	0.85 (0.31, 2.29)	0.743
Education (middle school vs. primary school)	1.72 (0.32, 9.82)	0.564	1.08 (0.17, 7.09)	0.933	0.60 (0.11, 3.13)	0.509	1.35 (0.23, 8.36)	0.720	1.91 (0.4, 9.73)	0.413
Residence (rural area vs. urban)	0.63 (0.18, 2.13)	0.449	0.07 (0.01, 0.28)	<0.001	0.27 (0.07, 0.87)	0.033	0.70 (0.19, 2.48)	0.567	1.67 (0.53, 5.54)	0.397
Monthly income (CNY)										
<5,000 vs. ≥ 5,000	0.86 (0.25, 2.89)	0.797	0.33 (0.07, 1.25)	0.080	1.33 (0.42, 4.21)	0.600	0.22 (0.05, 0.82)	0.024	0.58 (0.18, 1.79)	0.333
Guardians' education										
Primary school or below vs. college or above	2.35 (0.4, 14.93)	0.400	10.05 (1.35, 92.73)	0.013	3.09 (0.54, 19.23)	0.266	1.12 (0.16, 8.07)	0.908	0.50 (0.09, 2.55)	0.398
Middle school vs. college or above	1.21 (0.35, 4.2)	0.756	1.82 (0.51, 6.93)	0.341	1.19 (0.38, 3.84)	0.771	0.29 (0.07, 1.07)	0.096	0.28 (0.08, 0.92)	0.029
Diagnosis										
Acute lymphoblastic leukemia vs. others	0.87 (0.26, 2.9)	0.817	0.70 (0.19, 2.41)	0.567	0.83 (0.26, 2.62)	0.763	1.16 (0.33, 4.15)	0.831	0.58 (0.19, 1.78)	0.366
Non-Hodgkin's lymphoma vs. others	0.59 (0.15, 2.29)	0.445	0.53 (0.13, 2.13)	0.407	0.67 (0.18, 2.41)	0.564	1.39 (0.34, 5.9)	0.670	0.99 (0.29, 3.43)	0.990
Disease duration (months)	1 (0.95, 1.03)	0.695	0.97 (0.92, 1.01)	0.105	0.99 (0.95, 1.02)	0.537	0.95 (0.89, 1)	0.192	0.99 (0.96, 1.02)	0.717
ECOG (poor vs. good)	5.88 (2.08, 20.03)	0.003	9.09 (2.56, 50.04)	<0.001	3.70 (1.35, 11.10)	0.018	12.50 (3.85, 47.62)	<0.001	2.94 (1.11, 9.09)	0.043

10.3389/fpubh.2022.1050835

utility studies given the huge health burden caused by the HM. The present study may have made some contributions to that.

Our study is with several limitations. The sample comprised 96 inpatients only, which might be too small and too selective to represent the entire population of interest. The utility is likely to be underestimated as outpatients are generally better off than inpatients. The Japanese value set may not be fully transferrable to Chinese patients, which may introduce certain imprecisions in the utility estimates. Due to the limitations of data accessibility, patients' treatment information is lacking. The impact of medical interventions on the HRQoL cannot be explored in the study. Finally, the cross-sectional nature has precluded an effective inference on the casual effect relationship.

Conclusion

Chinese pediatric patients with HM live in relatively poor health status and their HRQoL is of great concern to healthcare. The strong associations of health utility, VAS, and individual health dimensions with the traditional clinical index ECOG show that the evidence of ECOG and EQ-5D-Y are complementary to each other for producing a holistic evaluation of a patient's health. Therefore, the EQ-5D-Y makes a valuable tool for clinical decision-making and further generates utility to inform resource allocation in the oncological care of pediatric HM.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by Institutional Review Board of School of Public Health, Fudan University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

YS and H-JZ conceptualized the study, analyzed the data, and drafted the manuscript. AS interviewed the patients and managed the database. BW and WW handled the issues of project operation, data quality, and interpreted the results. NL designed the study and optimized the questionnaire. PW conceptualized the study, oversighted the project, and wrote several sections. All authors revised the manuscript critically and approved the manuscript for publication.

Funding

The study has been mainly sponsored by the National Natural Science Foundation of China (Project No. 72274037) awarded to PW. It is also co-supported by The Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning awarded to H-JZ by Shanghai Municipal Education Commission (Project No. 10-20-303-601).

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. Ravens-Sieberer U, Erhart M, Wille N, Wetzel R, Nickel J, Bullinger M. Generic health-related quality-of-life assessment in children and adolescents: methodological considerations. *Pharmacoeconomics.* (2006) 24:1199–220. doi: 10.2165/00019053-200624120-00005

2. Wille N, Badia X, Bonsel G, Burstrom K, Cavrini G, Devlin N, et al. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. *Qual Life Res.* (2010) 19:875–86. doi: 10.1007/s11136-010-9648-y

3. Shen A, Zhang S, Luo N, Wu B. Feasibility and validity of EQ-5D-Y used in evaluating health related quality of life in children with hematology malignancies. *China Pharmacist.* (2020) 23:665–70.

4. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. *CA Cancer J Clin.* (2016) 66:115–32. doi: 10.3322/caac.21338

5. Sun YQ, Huang XJ. The challenges and strategies for treatment of hematological malignancies in China. *Zhonghua Nei Ke Za Zhi.* (2021) 60:857–9.

6. Xue M. Analysis of inpatient payment per visit of Leukemia in China from 2010 to 2017. *Chin J Health Stat.* (2019) 36:196–9.

7. Geue K, Götze H, Friedrich M, Leuteritz K, Mehnert-Theuerkauf A, Sender A, et al. Perceived social support and associations with health-related quality of life in young vs. older adult patients with haematological malignancies. *Health Qual Life Outcomes.* (2019) 17:145. doi: 10.1186/s12955-019-1202-1

8. Allart-Vorelli P, Porro B, Baguet F, Michel A, Cousson-Gélie F. Haematological cancer and quality of life: a systematic literature review. *Blood Cancer J.* (2015) 5:e305. doi: 10.1038/bcj.2015.29

9. Lam KK Li WH, Chiu SY, Chan GC. The impact of cancer and its treatment on physical activity levels and quality of life. *Eur J Oncol Nurs.* (2016) 21:83–9. doi: 10.1016/j.ejon.2016.01.007

10. Jankowska-Polańska B, Sliwiński M, Swiatoniowska N, Butrym A, Mazur G. Quality of life in children with acute lymphoblastic leukaemia. *Scand J Caring Sci.* (2020) 34:380–89. doi: 10.1111/scs.12739

11. Kobayashi K, Nakagami-Yamaguchi E, Hayakawa A, Adachi S, Hara J, Tokimasa S, et al. Health-related quality of life in Japanese children with acute lymphoblastic leukemia during and after chemotherapy. *Pediatr Int.* (2017) 59:145–53. doi: 10.1111/ped.13092

12. Furlong W, Rae C, Feeny D, Gelber RD, Laverdiere C, Michon B, et al. Health-related quality of life among children with acute lymphoblastic leukemia. *Pediatr Blood Cancer.* (2012) 59:717–24. doi: 10.1002/pbc.24096

13. 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 11). *Eur J Cancer.* (2009) 45:228–47. doi: 10.1016/j.ejca.2008.10.026

14. Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, et al. Toxicity and response criteria of the Eastern cooperative oncology group. *Am J Clin Oncol.* (1982) 5:649–55.

15. Zhou W, Shen A, Yang Z, Wang P, Wu B, Herdman M, et al. Patient-caregiver agreement and test–retest reliability of the EQ-5D-Y-3L and EQ-5D-Y-5L in paediatric patients with haematological malignancies. *Eur J Health Econ.* (2021) 22:1103–13. doi: 10.1007/s10198-021-01309-w

16. Wang P, Sun Y, Yang ZH, Zhang RY, Wu B, Luo N. Testing measurement properties of two EQ-5D youth versions and KIDSCREEN-10 in China. *Eur J Health Econ.* (2021) 22:1083–93. doi: 10.1007/s10198-021-01307-y

17. Shiroiwa T, Ikeda S, Noto S, Fukuda T, Stolk E. Valuation survey of EQ-5D-Y based on the international common protocol: development of a value set in Japan. *Med Decis Making.* (2021) 2021:272989X211001859. doi: 10.1177/0272989X211001859

18. Lien N, Friestad C, Klepp KI. Adolescents' proxy reports of parents' socioeconomic status: how valid are they? *J Epidemiol Community Health.* (2001) 55:731–7. doi: 10.1136/jech5510731

19. Åström M, Persson C, Lindén-Boström M, Rolfson O, Burström K. Population health status based on the EQ-5D-Y-3L among adolescents in Sweden. *Qual Life Res.* (2018) 27:2859–71. doi: 10.1007/s11136-018-1985-2

20. Pan CW, Zhong H, Li J, Suo C, Wang P. Measuring health-related quality of life in elementary and secondary school students using the Chinese version of the EQ-5D-Y in rural China. *BMC Public Health.* (2020) 20:982. doi: 10.1186/s12889-020-09116-3

21. Xia W, Li HC, Lam KW, Chung OK, Song P, Chiu SY, et al. The impact of hematologic cancer and its treatment on physical activity level and quality of life among children in mainland china: a descriptive study. *Cancer Nurs.* (2019) 42:492–500. doi: 10.1097/NCC0000000000000661

22. Papathanasiou IV, Kelepouris K, Valari C, Papagiannis D, Tzavella F, Kourkouta L, et al. Depression, anxiety and stress among patients with hematological malignancies and the association with quality of life: a cross-sectional study. *Med Pharm Rep.* (2020) 93:62–8. doi: 10.15386/mpr-1502

23. Zeng X, Sui M, Liu R, Qian X, Li W, Zheng E, et al. Assessment of the health utility of patients with leukemia in China. *Health Qual Life Outcomes.* (2021) 19:65. doi: 10.1186/s12955-021-01711-1

24. Kreimeier S, Greiner W. EQ-5D-Y as a health-related quality of life instrument for children and adolescents: the instrument's characteristics, development, current use, and challenges of developing its value set. *Value Health.* (2019) 22:31–7. doi: 10.1016/j.jval.2018.11.001

25. Khanna D, Khadka J, Mpundu-Kaambwa C, Lay K, Russo R, Ratcliffe J, et al. Are we agreed? Self- vs. proxy-reporting of paediatric health-related quality of life (HRQoL) using generic preference-based measures: a systematic review and meta-analysis. *Pharmacoeconomics.* (2022) 40:1–25. doi: 10.1007/s40273-022-01177-z

26. Henderson TO, Parsons SK, Wroblewski KE, Chen L, Hong FX, Smith SM, et al. Outcomes in adolescents and young adults with Hodgkin lymphoma treated on US cooperative group protocols: an adult intergroup (E2496) and Children's Oncology Group (COG AHOD0031) comparative analysis. *Cancer Am Cancer Soc.* (2018) 124:136–44. doi: 10.1002/cncr.30979

27. Cheng KKF, Tan LML. A pilot study of the effect of a home-based multimodal symptom-management program in children and adolescents undergoing chemotherapy. *Cancer Rep Hoboken NJ.* (2021) 4:e1336. doi: 10.1002/cnr2.1336

28. LeBlanc TW, Roeland EJ, El-Jawahri A. Early palliative care for patients with hematologic malignancies: is it really so difficult to achieve? *Curr Hematol Malig Rep.* (2017) 12:300–8. doi: 10.1007/s11899-017-0392-z

29. Hoenigl M, Duettmann W, Raggam RB, Huber-Krassnitzer B, Theiler G, Seeber K, et al. Impact of structured personal on-site patient education on low posaconazole plasma concentrations in patients with haematological malignancies. *Int J Antimicrob Agents.* (2014) 44:140–4. doi: 10.1016/j.ijantimicag.2014.03.013

30. Gatta B, LeBlanc TW. Palliative care in hematologic malignancies: a multidisciplinary approach. *Expert Rev Hematol.* (2020) 13:223–31. doi: 10.1080/17474086.2020.1728248

31. Torrance GW. Measurement of health state utilities for economic appraisal. *J Health Econ.* (1986) 5:1–30.

32. Russell KMW, Hudson M, Long A, Phipps S. Assessment of health-related quality of life in children with cancer: consistency and agreement between parent and child reports. *Cancer Am Cancer Soc.* (2006) 106:2267–74. doi: 10.1002/cncr.21871

frontiers | Frontiers in Pediatrics

TYPE Original Research
PUBLISHED 01 March 2023
DOI 10.3389/fped.2023.1037238

OPEN ACCESS

EDITED BY
Jing Wu,
Tianjin University, China

REVIEWED BY
Cemil Çolak,
İnönü University, Türkiye
Takeshi Yokoyama,
Kyushu University, Japan

*CORRESPONDENCE
Yosef Belay Bizuneh
✉ phanuelyosef@gmail.com

SPECIALTY SECTION
This article was submitted to Children and Health, a section of the journal Frontiers in Pediatrics

RECEIVED 05 September 2022
ACCEPTED 13 February 2023
PUBLISHED 01 March 2023

CITATION
Molla MT, Bizuneh YB, Nigatu YA and Melesse DY (2023) High incidence rate of postoperative sore throat in intubated children at Northwest Amhara Comprehensive Specialized Hospitals, Ethiopia. A multicenter study.
Front. Pediatr. 11:1037238.
doi: 10.3389/fped.2023.1037238

High incidence rate of postoperative sore throat in intubated children at Northwest Amhara Comprehensive Specialized Hospitals, Ethiopia. A multicenter study

Misganaw Terefe Molla[1], Yosef Belay Bizuneh[2]*, Yonas Addisu Nigatu[2] and Debas Yaregal Melesse[2]

[1]Department of Anesthesia, College of Medicine and Health Sciences, Bahirdar University, Bahirdar, Ethiopia, [2]Department of Anesthesia, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia

Background: Postoperative sore throat is the most frequent complication in pediatric patients after general anesthesia. This study aimed to investigate the incidence of postoperative sore throat in patients undergoing general anesthesia with tracheal intubation or laryngeal mask airway.
Methods: A hospital-based multicenter prospective observational cohort study was conducted. Proportional allocation was done with a total of 424 patients from March 1 to June 30, 2022. The information was entered into the Epi-Data software version 4.6 and analyzed with Stata 14. Socio–demographic, surgical, and anesthetic-related characteristics were analyzed using descriptive statistics. A *p*-value of less than 0.2 was the cutpoint of bivariate logistic regression analysis, and *p*-values of less than 0.05 were regarded as statistically significant in multivariate logistic regression to determine the presence and strength of association between independent variables and postoperative sore throat.
Results: A total of 411 patients were included in this study, with a response rate of 96.9%. The overall proportion of patients who developed postoperative sore throat was 45% (95% CI: 40.18–49.84). Patients who had anesthesia for more than two hours (AOR = 8.23: 95% CI = 4.08–16.5), those who were intubated by undergraduate anesthesia students (AOR = 2.67: 95% CI = 1.53–4.67), and those who had been intubated using tracheal tube (AOR = 2.38: 95% CI = 1.15–4.92) were significantly associated with the level of postoperative sore throat.
Conclusions and recommendations: We concluded that intubated children with ETT have a high incidence of post-operative sore throat. Tracheal tube usage, intubation by undergraduate students, and more than two hours of anesthesia duration were associated factors. The incidence of sore throat can be decreased with the use of a laryngeal mask airway, intubation by a senior anesthetist, and shortening of anesthesia time.

KEYWORDS

children, anesthesia complications, intubation, postoperative, sore throat

Abbreviations
AOR, Adjusted Odds Ratio; CI, Confidence Interval; COR, Crude Odds Ratio; ETT, Endotracheal Tube; LMA, Laryngeal Mask Airway; NPO, Nil Per Os; POST, Postoperative Sore Throat.

1. Introduction

General anesthesia is the most commonly used type of anesthesia for a surgical procedure involving an ETT or LMA for the maintenance of airway patency, resulting in postoperative complications (1–3).

A postoperative sore throat is a pain or discomfort in the patient's throat after receiving general anesthesia under tracheal tube or laryngeal mask airway (4–6). In the pediatric population, the incidence of postoperative sore throat following general anesthesia is not well understood and is thought to occur in 5%–41% of children with tracheal intubation (3, 7–11) and post-operative sore throat is also reported to be lower at 4%–13% in LMA use (12–16). In adult patients, up to 12%–60% get postoperative sore throat, which is a common and undesirable problem (6, 17, 18).

The multiple factors that contribute to post-operative sore throat include the choice of airway device, high tracheal tube cuff pressure, surgical manipulation of the airway, varying circuit humidification, airway suctioning, dehydration, and the type and length of anesthesia and surgery (7, 8, 19–21).

The exact time at which POST can occur is not clearly understood in the pediatric age group, even if some studies state that it occurs as early as 15 min after immediate postoperative time and reduces its incidence from 12 to 24 h (22, 23). However, the duration may continue after discharge, affecting patients' oral intake due to postoperative throat pain (8, 24).

Untreated POST will result in a longer discharge time, a negative experience with anesthesia, a longer postoperative recovery time, severe patient discomfort, distress, anxiety, dissatisfaction, increased hoarseness, and a reduction in patient life activity (3, 25–29).

The preventive and management techniques for POST in pediatric patients are essential for the early reduction of severe complications, including reduction of stimulation, use of the small-sized tube, reducing the number of attempts, use of a non–steroidal anti–inflammatory drug, lidocaine, nebulized magnesium and preemptive ketamine and steroids (30–34).

Postoperative pediatric sore throat is not extensively studied in sub-Saharan countries, even if the problem is significant. Therefore, we aimed to assess the incidence and associated factors of postoperative sore throat among pediatric patients in the Northwest Amhara Referral Hospitals, Ethiopia.

2. Methods

2.1. Study design, setting and period

A multicenter prospective observational cohort study was conducted. The study was conducted at Northwest Amhara comprehensive specialized hospitals, Ethiopia, from March 1 to June 30, 2022. The study took place at the University of Gondar Comprehensive Specialized Hospital, which has seven general surgery rooms; two obstetrics; two gynecology; and one ophthalmology operation room. Tibebe Ghion Specialized Hospital has seven general surgery rooms; two obstetrics and gynecology; two orthopedic operation rooms; and one ophthalmology operation room. Felege Hiwot Comprehensive Specialized Hospital has six general surgery rooms; two orthopedic; two obstetrics and gynecology rooms; and one ophthalmology operation room. The Debre Tabor Comprehensive Specialized Hospital has three general surgery rooms; two orthopedics; two obstetrics and gynecology operation rooms; and one ophthalmology operation room. The Debre Markos Comprehensive Specialized Hospital has three general surgery rooms; one orthopedic; two obstetrics and gynecology operation rooms; and one ophthalmology operation room. There is no distinct pediatric operating room in the above referral institutions.

Anaesthesia

Anaesthetic management was standardized. Standard patient monitoring, including an electrocardiogram, end-tidal carbon dioxide, pulse oximetry, and non-invasive arterial blood pressure, was performed. Because the study was cross-sectional and not controlled, anesthesia induction techniques were dependent on patient condition and anesthetist experience level. The cuff was inflated manually with air to a clinical endpoint of loss of an audible leak. The intra-cuff pressure was not measured because a hand-held manometer is not available in our setup.

The age, sex, weight, and physical status of the patients were recorded on a standardized form. Routine informed consent for surgery and anesthesia was obtained before the patient's arrival in the operating room. The type and duration of surgery, the intraoperative airway device used [Endotracheal tube (ETT), Laryngeal Mask Airway (LMA)], the grade of intubation, and the patient's position during surgery were noted. After the end of surgery, the airway devices were removed when patients were able to respond to commands. The patients were taken to the post-anesthesia care unit, from which they were shifted to the ward after adequate recovery.

2.2. Study population, inclusion and exclusion criteria and study variables

All pediatric surgical patients admitted at Comprehensive Specialized Hospitals of Northwest Amhara, Ethiopia, from March 1 to June 30, 2022, were considered the source population, and all pediatric surgical patients operated under general anesthesia with ETT or LMA whose age was 6–18 years were considered the study population.

All pediatric surgical patients, both elective and emergency procedures, aged 6–18 years who were operated on under general anesthesia with ETT or LMA within the study period were included in the study, whereas surgical procedures in oral, nasal,

and throat surgery; recent respiratory infection; recent preoperative throat pain; postoperative need for a mechanical ventilator; and intensive care were excluded from the study.

The incidence of postoperative sore throat was the primary outcome variable, and the level of severity and associated factors of postoperative sore throat were the secondary outcome variables.

2.3. Sample size determination

The sample size was determined by applying a *p*-value of 50% to a single proportion formula. The sample size was calculated with a 95% confidence interval (CI) and a 5% margin of error. Because no other study with the same sociodemographic and study population characteristics has been carried out. So

$$\frac{n = z^2(pq)}{d^2}$$

$$\frac{n = 1.96^{2*}0.5^{*}0.5}{0.05^2}$$

Then $n = 385$. Where (p) represent the estimated proportion, (z) level of confidence, (d) margin of error, and $n =$ desired sample size. Using 385 plus a 10% non-response rate of 39, the total sample size was calculated to be $n = 424$.

2.4. Sampling procedure

A multi-center consecutive sampling method was used until the target sample was fully obtained. Additionally, we distributed the sample within each institution using the proportionate allocation approach. We determined the final sample in each hospital by multiplying the total number of calculated sample sizes by the number of patients operated on in the last four months in each hospital, divided by the total number of cases operated in five hospitals. The final number of samples used during the analysis was 411, and because of incomplete data, 13 patients were removed.

2.5. The assessment methods and questionnaire

Self-reporting of pain is the ideal method for children who are able to do so. The pain assessment process may also include parents or guardians because their pain ratings are closely related to those of their children. The direct self-report technique can be used to evaluate postoperative patient complaints to help identify their magnitude and contributing factors. A child's self-reported throat pain was graded on a four-point severity scale, and the existence or absence of a sore throat was evaluated by a Yes/No response in an observational cohort study. The pain scale was greater than or equal to 1, and if there was no postoperative throat pain, it was rated as No on a four-point categorical scale with 0 denoting no throat pain, 1 denoting mild throat pain (complain when asked), 2 denoting moderate throat pain (complain of throat pain by itself), and 3 denoting severe throat pain (including change in voice or hoarseness).

2.6. Operational definitions

Junior anesthetist: BSC holder *and* senior anesthesist: MSc holder
Pediatric patients: Age range from 6 years to 18 years (35–37)

To determine whether POST is present or absent, the category pain scale is used. The absence of a sore throat was reported as no, whereas mild, moderate, and severe sore throats were rated as yes (35, 36, 38).

2.7. Data collection procedures

The questionnaire and checklist were adapted through the review of different previous similar studies (9, 23, 39–41). The questionnaire was initially prepared in English, then translated to Amharic, and then translated back into English three times to check for consistency. A structured questionnaire was used to collect the data through a face-to-face interview and a checklist for reviewing client charts. The Amharic version of the questionnaire was used for data collection.

The data was gathered by five trained BSC anesthetists, and one senior anesthetist was working as a supervisor. An interview and a review of the records are both steps in the data collection process. The data collection method used for POST was the direct self-report method. It was collected in the post-anesthesia care unit and patient wards. The presence or absence of a sore throat was noted at intervals of two hours, six hours, and twenty-four hours. Additionally, the degree of pain was assessed and classified as mild, moderate, and severe. Patients with both elective and emergency procedures were included. A record review that included information on the patient, anesthesia, and surgical procedures was also conducted in addition to the data collector interview. The questionnaires and checklists were filled out completely and collected daily after checking the completeness and consistency of the data.

2.8. Data quality assurance

A pretest at UOGCSH was conducted to assess the questionnaire's clarity, language accent, and grammatical corrections. 22 patients were included in this pretest. To maintain the quality of the data, the data collectors and supervisors were trained for 3 days on the objective of the study, the content of the questionnaire, how to fill out the questionnaire, respondent rights, informed consent, and interview technique, as well as how to maintain the confidentiality and privacy of the study subjects.

The principal investigators and supervisors gave feedback and corrections daily to the data collectors. The data were cleaned, coded, and imported into EpiData version 4.6 before being exported to Stata version 14.

2.9. Data processing and analysis

Each completed questionnaire was coded on a pre-arranged coding sheet by the principal investigator to minimize errors. Data was entered into a computer using EpiData version 4.6 and exported to Stata version 14 for additional cleaning and analysis. Sorting and cross-tabulation were used to clean the data. Multicollinearity was examined using the variance inflation factor (VIF), which had a value of 1.17. Using the Shapiro-Wilk test, normality was examined for continuous variables. A Hosmer and Lemeshow test was conducted to evaluate the goodness of fit of the binary logistic regression model, and a p-value of 0.2 was obtained. The findings were presented according to the variables through text, tables, and graphs.

Initially, bivariable logistic regression analysis was performed between the dependent variable and each of the independent variables. Then all variables with a p-value <0.2 from the bivariable logistic regression analysis were fitted into the multivariable logistic regression model to control for possible confounders. An adjusted odds ratio (AOR) with a 95% confidence interval (CI) was used to measure the strength and significance of the association. A p-value <0.05 also indicated the presence of a statistically significant association between postoperative sore throat and independent variables.

3. Results

3.1. Socio-demographic characteristics of the respondents

A total of 411 pediatric patients participated in this study, with an overall response rate of 96.9%. In this study, the median age was 10 years [interquartile range (IQR): 7–13]. 52.6% of the total participants were male. 81.8% of the pediatric patients had an ASA I. Preoperative NPO status for the patient was 56% for less than or equal to 8 h. 68.9% of patients had scheduled operations on an elective basis. Abdominal and urological operations accounted for the majority of the surgical patients (54.7%) (Table 1).

At two hours after surgery, the postoperative surgical pain level assessment revealed that 26.3% of patients reported no pain, 43.8% reported mild pain, 23.1% reported moderate pain, and 6.8% reported severe pain.

At six hours, 15.3% reported no pain, 37.0% reported moderate discomfort, 44.7% mild pain, and 3.0% reported moderate to severe pain.

At 24 hours, 62.8% of patients reported no pain, while 33.9%, 4.1%, and 0.2%, respectively, reported mild, moderate, and severe pain.

TABLE 1 Patients sociodemographic and surgery-related characteristics of pediatric patients at Northwest Amhara Referral Hospitals, from March 1 to June 30, 2022 ($N = 411$).

Variable	Postoperative sore throat, *n* (%)		Total ($n = 411$)
	Yes	No	
Sex			
Male	97 (44.9%)	119 (55.1%)	216 (52.6%)
Female	88 (45.1)	107 (54.9%)	195 (47.4%)
Postoperative NPO			
≤8 h	117 (50.9%)	117 (49.1%)	230 (55.9%)
>8 h	72 (39.8%)	109 (60.2%)	181 (44.1%)*
PONV			
No PONV	90 (41.5%)	127 (58.5%)	217 (52.8%)
Nausea	16 (41.0%)	23 (59.0%)	39 (9.5%)
Vomiting	45 (52.9%)	40 (47.1%)	85 (20.7%)
Both PONV	34 (48.6%)	36 (51.4%)	70 (17.0%)
NGT			
Yes	23 (60.5%)	15 (39.5%)	38 (9.2%)*
No	162 (43.4%)	211 (56.6%)	373 (90.8%)

NPO, Nil per Os; *n* (%), Frequency and Percentage; NGT, Naso Gastric Tube; PONV, Postoperative Nausea and Vomiting.
*Met χ^2 assumption with p-value <0.05.

TABLE 2 Anesthesia-related variables of pediatric patients at northwest amhara referral hospitals, from march 1 to June 30, 2022 ($N = 411$).

Variables	Postoperative sore throat, *n* (%)		Total ($N = 411$)
	Yes	No	
Education level of anesthetist			
Master	42 (31.3%)	92 (68.7%)	134 (32.6%)
BSc	42 (36.8%)	72 (63.2%)	114 (27.7%)
Undergraduate student	101 (62.0%)	62 (38.0%)	163 (39.7%)*
Tube type			
ETT	172 (49.3%)	177 (50.7%)	349 (84.9%)*
LMA	13 (21%)	49 (79%)	62 (15.1%)
NMB			
Yes	160 (46.4%)	185 (53.6%)	345 (83.9%)
No	25 (37.9%)	41 (62.1%)	66 (16.1%)

NMB, Neuromuscular Blocker; *N* (%), Frequency and Percentage; ETT, tracheal intubation Tube; LMA, Laryngeal Mask Airway; BSc, Bachler of Science.
*Met χ^2 assumption with p-value <0.05.

3.2. Anesthesia-related characteristics of pediatric patients

Among 411 patients, 84.9% of pediatric patients underwent ETT intubation, while all remaining patients underwent LMA. 94.2% of patients had intubation performed with a grade I laryngoscopy view. All patients under ETT intubation had cuffed tubes. During the perioperative period, airways are not used in 88.1% of patients. The median tube size was 5 centimeters [interquartile range (IQR) 4.5–5.5] (Table 2).

Regarding the induction of anesthesia, 44.5% of patients were induced with ketamine, 21.7% of patients were induced with

propofol, 4.1% of patients were induced with thiopental, and 29.7% of patients were induced with both ketamine and propofol. In terms of analgesia, 91.7% of patients received an opioid, 1.7% received diclofenac, 6.1% received paracetamol, and 0.5% received ketamine analgesia. Steroid (dexamethasone) was administered to 98% of the patients.

3.3. Incidence of postoperative sore throat, in each hospital, and the occurrence in each time of observation among pediatric patients

The overall proportion of patients who developed postoperative sore throat was 45% (95% CI: 40.18–49.84). The incidence of POST was included in a patient who underwent general anesthesia with ETT or LMA up to 24 h (Table 3).

POST occurred in 17.8% of patients at two hours, with mild, moderate, and severe cases accounting for 79.5%, 12.3%, and 8.2% of the cases, respectively. The incidence of POST at 6 h was 22.4%, and mild, moderate, and severe cases made up 58.7%, 30.4%, and 10.9%, respectively. At 24 h, the incidence was 4.8%, with mild and moderate cases accounting for 80.0% and 20.0%, respectively (Figure 1).

TABLE 3 The distribution of patients in each hospital in frequency and percentage at Northwest Amhara Referral Hospitals, from March 1 to June 30, 2022($N = 411$).

Hospital	Postoperative sore throat, *N* (%)		Total ($N = 411$)
	Yes	No	
UOGCSH	39 (45.9%)	46 (54.1%)	85 (20.7%)
TGCSH	45 (45.5%)	54 (54.5%)	99 (24.1%)
FHCSH	31 (34.8%)	58 (65.2%)	89 (21.7%)
DTCSH	38 (55.9%)	30 (44.1%)	68 (16.5%)
DMCSH	32 (45.7%)	38 (54.3%)	70 (17.0%)

N (%), Frequency and Percentage.
DMCSH, Debre Markos Comprehensive Specialized Hospital; DTCSH, Debre Tabor Comprehensive Specialized Hospital; FHCSH, Felege Hiwot Comprehensive Specialized Hospital; TGCSH, Tibebe Ghion Comprehensive Specialized Hospital; UOGCSH, University of Gondar Comprehensive Specialized Hospital.

3.4. Factors associated with postoperative sore throat among pediatric surgical patients

3.4.1. Bi-variable analysis

In the bi-variate logistic regression analysis, variables with a *p*-value of <0.2 were considered to be factors associated with postoperative sore throat among pediatric surgical patients. These variables included the educational level of the anesthetists who intubated the patients, the tube type (ETT vs. LMA), the presence of NGT, the length of anesthesia, and the postoperative NPO time longer than 8 h (Table 4).

3.4.2. Multi-variable analysis

In multivariable logistic analysis, a *p*-value of <0.05 was considered statistically significant. The tracheal intubation, the length of anesthesia, and the education level of the anesthetists were factors associated with postoperative sore throat.

Anesthesia lasting more than two hours was 8.23 times more likely to occur POST compared with less than two hours, with 82.5% (AOR = 8.23; CI = 4.08–16.5). Using ETT was 2.38 times more likely to occur POST in comparison to using LMA, with an occurrence of 49.3% (AOR = 2.38; CI = 1.15–4.92). Intubations performed by undergraduate students were 2.67 times more likely to result in POST (62%) than those by junior and senior anesthetists (AOR: 2.67, CI = 1.53–4.67) (Table 4).

4. Discussion

Children who undergo ETT or LMA under general anesthesia frequently experience a postoperative sore throat. To the best of our knowledge, no published study has assessed the incidence of POST

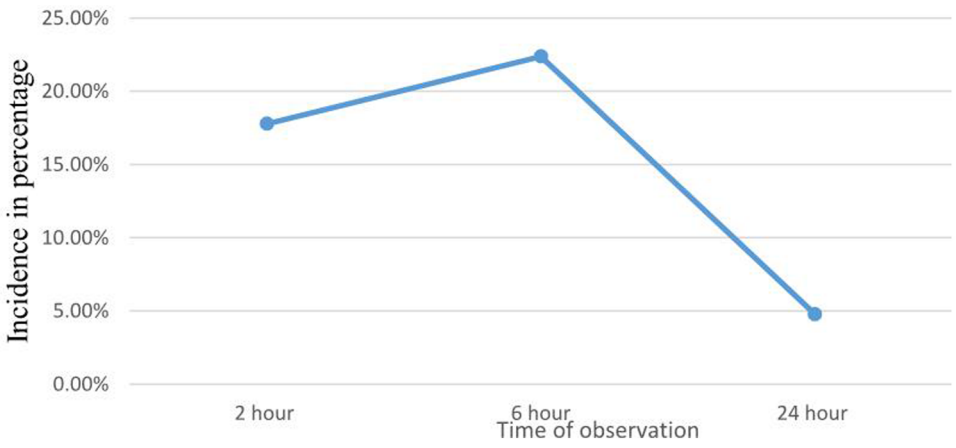


FIGURE 1
Incidence of POST in each time of observation among pediatric patients at Northwest Amhara Referral Hospitals, from March 1 to June 30, 2022 ($N = 411$).

10.3389/fped.2023.1037238

TABLE 4 Bivariate and multivariate logistic regression for POST among pediatric patients at northwest amhara referral hospitals, from march 1 to June 30, 2022($N = 411$).

Variable	Postoperative sore throat, *n* (%)		COR (95% CI)	AOR (95% CI)	*p*-value
	Yes	No			
Education					
Master	42 (31.3%)	92 (68.7%)	1	1	
BSc	42 (36.8%)	72(63.2%)	1.28(0.75–2.16)	1.06(0.58–1.94)	0.85**
Undergraduate	101(62%)	62 (38%)	3.57 (2.2–5.78)	2.67 (1.53–4.67)	0.001*
Students					
Tube					
LMA	13 (21%)	49 (79%)	1	1	
ETT	172 (49.3%)	177 (50.7%)	3.66 (1.92–6.99)	2.38 (1.15–4.92)	0.019*
Duration of Anesthesia					
<1 h	25 (29.7%)	59 (70.3%)	1	1	
1–2 h	56 (27.8%)	145 (72.2%)	0.91 (0.52–1.59)	0.78 (0.43–1.41)	0.41**
≥2 h	104 (82.5%)	22 (17.5%)	11.2 (5.79–21.5)	8.23 (4.08–16.5)	<0.001*
NGT					
No	162 (43.4%)	211 (56.6%)	1	1	
Yes	23 (60.5%)	15 (39.5%)	2.27 (1.24–4.17)	0.99 (0.46–2.11)	0.98**
Postoperative NPO					
≤8 h	53 (37.3%)	89 (62.7%)	1	1	
>8 h	132 (49.1%)	137 (50.9%)	1.62 (1.07–2.45)	0.94 (0.56–1.56)	0.79**

n (%)= Frequency and Percentage.
COR, Crude Odds Ratio; AOR, Adjusted Odds Ratio; CI, Confidence Interval; ETT, Endotracheal Tube; LMA, laryngeal Mask Airway; NGT, Nasogastric Tube; NPO, Nil Per Os; BSc, Bachler of Science.
*Statistical Significant.
**Statistical Insignificant 1 = Reference Variable.

and its risk factors in pediatric patients in Ethiopia. If prevention and treatment strategies are not implemented during the perioperative phase, postoperative sore throat increases postoperative patient dissatisfaction, anxiety and decreasing the postoperative quality of recovery (42). The main focus of this study was to assess the incidence and associated factors of POST in pediatric surgical patients after general anesthesia ETT vs. LMA.

In this study, the incidence of POST among children after surgery was 45%, with a 95% confidence interval (CI) of (40.1–49.84). This finding is in line with studies conducted in the UK and Korea that showed 41% and 49.5%, respectively (35, 43). The possible explanation might be due to a similar data collection technique. In this study, the incidence of POST among pediatric patients was higher as compared to studies conducted in the UK and Australia, with an incidence of 36.5% and 22.6%, respectively (8, 9). The possible reasons for this difference might be sample size and the study population's socio–demographic characteristics. The incidence of POST was assessed in this study at 2, 6, and 24 h, with a higher incidence at 6 h than at 2 and 24 h. The majority of the patients experience mild POST, followed by moderate and severe. This finding is supported by a study conducted in Pakistan (36), the UK (8), and Korea (35).

POST in pediatric patients is affected by different variables. Among the factors that displayed a statistically significant link in the multivariate logistic regression were patients intubated by undergraduate students, duration of anesthesia greater than 2 h, and use of ETT for intubation (**Table 4**).

In this study, anesthesia lasting longer than two hours was a statistically significant factor in the occurrence of POST. Anesthesia lasting more than two hours was 8.23 times more likely to result in POST than less than two hours, with an incidence rate of 82.5% (AOR = 8.23; CI = 4.08–16.5). This finding was supported by a study conducted in the UK that showed a longer duration of anesthesia can cause POST with an occurrence of 55.2% (8). The possible explanation for this finding, in general, is that the longer the tracheal tube was in place, the longer the anesthesia lasted, and this caused irritation, inflammation, and ischemia to the airway structure and postoperative dehydration (22). Therefore, keep in mind that shortening the process can help to reduce complications.

In this study, there was a statistically significant association between the use of ETT and POST in pediatric patients when compared to the use of LMA. Using ETT was 2.38 times more likely to occur POST in comparison to using LMA, with an occurrence of 49.3% (AOR = 2.38; CI = 1.15–4.92). This finding was supported by a study conducted in the UK (8), and Australia (9), with 54.7%, and (19%) respectively. In contrast to ETT, POST was less likely to develop when LMA was used, even though it continues to have a role in the development of POST. In our study, 84.9% of patients underwent general anesthesia for surgery utilizing an ETT rather than using an LMA (15.1%). The

idea that ETT use on a patient's airway or trachea could cause structural injury, airway mucosal inflammation, and congestion in addition to activating pain receptors in the trachea, which may result in POST, accounts for this finding (44, 45). As a result, even though ETT is typically employed, smooth intubation and extubation may reduce the risk of POST.

Intubation by undergraduate anesthesia students was significantly related to POST and was 2.67 times more likely to result in POST (62%) than those by junior and senior anesthetists (AOR 2.67, CI 1.53–4.67). This finding, which is supported by a study conducted in Switzerland, showed that complications were more likely to emerge when intubation was carried out by lower training individuals, with a 1.7-fold increased risk of postoperative complications (46). Multiple stimulations and attempts during tube insertion and extubation time could be the causes. Therefore, by being aware of the triggering circumstances, the likelihood of complications during anesthesia practice may be decreased.

The use of NGT and prolonged postoperative NPO duration were not significantly associated during multivariate logistic regression analysis in this study, with an AOR of 0.99 (CI = 0.46–2.11) and an AOR of 0.94 (CI = 0.56–1.56), respectively. An investigation carried out in the UK supported this prolonged postoperative NPO period and was not associated with POST in pediatric patients (8). Furthermore, the adult study discovered that the use of NGTs is a significant risk factor for the development of POST in the United Kingdom (47) and Ethiopia (48). However, research on pediatric patients found no conclusive correlation.

Due to the lack of a cuff manometer, we were unable to determine whether or not there was a significant correlation between cuff pressure and POST in this study. However, a previous study on cuff pressure revealed that cuff pressure greater than 20 cmH2O was a predictor of the occurrence of POST (9, 49). This can be brought on by increased airway pressure, which might result in ischemia, pressure sores, and the onset of throat pain (50). Cuffed ETT was not statistically linked in our study. However, a study using un–cuffed ETT demonstrates that it is a risk factor for the emergence of POST (9). The justification offered was that using an un–cuffed one causes an increase in the rate of reintubation, repeated stimulation, and mucosal injury (51). In our study, there was no correlation between POST and patients who underwent orotracheal intubation and we did not compare with nasotracheal intubation, because there were no patients who underwent nasotracheal intubation. However, a Canadian study found that nasotracheal intubation is a risk factor for the development of POST (22).

5. Strength and limitations of the study

5.1. Strength

According to our research, this is the first study of its kind to be conducted in Ethiopia. It also lays the groundwork for more research. Sufficient sample sizes were used in the investigation, which comprised multiple study centers.

5.2. Limitations

All pediatric age groups were not addressed in the study. Response bias may emerge as a result of the use of self-report data gathering techniques. The suctioning and extubation techniques as well as number of attempt to intubate were not evaluated. Due to the lack of a cuff manometer, we were unable to determine ETT cuff pressure.

6. Conclusions

Post-operative sore throat was common among pediatric patients in the study areas within the first 24 h following anesthesia. Significant risk factors for postoperative sore throat in pediatric patients after surgery and anesthesia were longer than two hours of anesthesia, ETT intubation, and intubation with undergraduate anesthesia students.

7. Recommendations

To the anesthetists: We recommend anesthetists use LMA when appropriate, reducing anesthesia time wasted during surgery in one or more ways. Remove the tube early, and put more of a priority on training for atraumatic intubation attempts. High-risk patients are better intubated by senior anesthetists. Based on other study findings, it is better to use airway blunt medication to reduce postoperative sore throat.

To the researchers: We encourage researchers to carry out randomized control trials based on ETT vs. LMA and awake vs. deep extubation. The effect of tracheal intubation cuff pressure on postoperative sore throat by the use of a cuff manometer.

Data availability statement

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

Ethics statement

An ethical clearance was obtained from the school of medicine, College of Medicine and Health Sciences on behalf of the University of Gondar ethical review board with the number SOM/14/04/2022. "All methods were performed in accordance with the relevant guidelines and regulations". A written support

letter was obtained from Debre Markos Comprehensive Specialized Hospital, Debre Tabor Comprehensive Specialized Hospital, Felege Hiwot Comprehensive Specialized Hospital, and Tibebe Ghion Comprehensive Specialized Hospital. After the purpose and objective of the study had been informed, written participant assent and parental consent were obtained from each study participant. Participants were also informed that their participation was voluntarily and that they could stop or leave the participation at any time if they were not comfortable. The data collection tools were anonymous and kept participants' privacy during the interview by interviewing them alone to maintain the confidentiality of any information provided by participants.

Author contributions

Conceptualization: MTM, YBB, YAN, and DYM. Data curation: MTM, YBB, YAN, and DYM. Formal analysis: MTM, YBB, YAN, and DYM. Investigation: MTM. Methodology: MTM, YBB, YAN, and DYM. Project administration: MTM. Resources: MTM. Software: MTM, YBB, YAN, and DYM. Supervision: MTM, YBB, YAN, and DYM. Validation: MTM, YBB, YAN, and DYM. Visualization: MTM. Writing – original draft: MTM and YBB. Writing – review & editing: YAN and DYM. All authors contributed to the article and approved the submitted version.

Acknowledgments

We would like to forward our deepest appreciation and thanks to the University of Gondar, College of medicine and health sciences, and school of medicine for providing Ethical clearance. We would also like to extend our thanks to Comprehensive specialized Hospitals administrators for the provision of support letter and staff cooperation during the data collection period. We are deeply grateful to study participants, data collectors, and supervisors who participated in this 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

References

1. Brown EN, Purdon PL, Van Dort CJ. General anesthesia and altered states of arousal: a systems neuroscience analysis. *Annu Rev Neurosci*. (2011) 34:601. doi: 10.1146/annurev-neuro-060909-153200

2. Jaensson M, Lassinantti Olowsson L, Nilsson U. *Does the size of the endotracheal tube affect the risk for sore throat following surgery in woman?: a randomized controlled study. 9th World congress for nurse anesthetists*, Haag, Nederländerna, 4–8 juni 2010 (2010).

3. Higgins P, Chung F, Mezei G. Postoperative sore throat after ambulatory surgery. *Br J Anaesth*. (2002) 88(4):582–4. doi: 10.1093/bja/88.4.582

4. McHardy F, Chung F. Postoperative sore throat: cause, prevention and treatment. *Anaesthesia*. (1999) 54(5):444–53. doi: 10.1046/j.1365-2044.1999.00780.x

5. Scuderi PE. Postoperative sore throat: more answers than questions. *Anesth Analg*. (2010) 111(4):831–2. doi: 10.1213/ANE.0b013e3181ee85c7

6. Macario A, et al. Which clinical anesthesia outcomes are important to avoid? The perspective of patients. *Anesth Analg*. (1999) 89(3):652. doi: 10.1213/00000539-199909000-00022

7. Cozine K, Stone J. Determinants of postoperative sore throat. In: *Anesthesiology*. Philadelphia, PA, United States: Lippincott-Raven Publ 227 East Washington SQ (1993).

8. Miskovic A, et al. A prospective observational cohort study on the incidence of postoperative sore throat in the pediatric population. *Pediatr Anesth*. (2019) 29 (12):1179–85. doi: 10.1111/pan.13757

9. Calder A, et al. Predictors of postoperative sore throat in intubated children. *Pediatr Anesth*. (2012) 22(3):239–43. doi: 10.1111/j.1460-9592.2011.03727.x

10. Splinter W, et al. Postoperative sore throat in children and the laryngeal mask airway. *Can J Anaesth*. (1994) 41(11):1081–3. doi: 10.1007/BF03015658

11. Miyerbekov Y, Mutagirov V, Batyrhanov M. Can we reduce airway management complications in children? Laryngeal mask vs. Endotracheal tube: 19AP5–2. *Eur J Anaesthesiol*. (2010) 27(47):256. doi: 10.1097/00003643-201006121-00826

12. Abdi W, et al. Sparing the larynx during gynecological laparoscopy: a randomized trial comparing the LMA supreme™ and the ETT. *Acta Anaesthesiol Scand*. (2010) 54(2):141–6. doi: 10.1111/j.1399-6576.2009.02095.x

13. McGinn G, Haynes S, Morton N. An evaluation of the laryngeal mask airway during routine paediatric anaesthesia. *Pediatr Anesth*. (1993) 3(1):23–8. doi: 10.1111/j.1460-9592.1993.tb00029.x

14. Füllekrug B, et al. The laryngeal mask airway: anesthetic gas leakage and fiberoptic control of positioning. *J Clin Anesth*. (1993) 5(5):357–63. doi: 10.1016/0952-8180(93)90097-X

15. Janssens M, Marechal J. The laryngeal mask–Liège experience. *Acta Anaesthesiol Belg*. (1991) 42(4):199–206.

16. Maltby J, Loken R, Watson N. The laryngeal mask airway: clinical appraisal in 250 patients. *Surv Anesthesiol*. (1991) 35(1):12. doi: 10.1097/00132586-199102000-00013

17. Biro P, Seifert B, Pasch T. Complaints of sore throat after tracheal intubation: a prospective evaluation. *Eur J Anaesthesiol*. (2005) 22(4):307–11. doi: 10.1017/S0265021505000529

18. Mandøe H, et al. Sore throat after endotracheal intubation. *Anesth Analg*. (1992) 74(6):897–900.

19. Loeser EA, et al. The influence of endotracheal tube cuff design and cuff lubrication on postoperative sore throat. *Surv Anesthesiol*. (1983) 27(6):341. doi: 10.1097/00132586-198312000-00012

20. Bennett M, Isert P, Cumming R. Postoperative sore throat and hoarseness following tracheal intubation using air or saline to inflate the cuff—a randomized controlled trial. *Anaesth Intensive Care*. (2000) 28(4):408–13. doi: 10.1177/0310057X0002800409

21. Chandler M. Tracheal intubation and sore throat: a mechanical explanation: apparatus. *Anaesthesia*. (2002) 57(2):155–61. doi: 10.1046/j.1365-2044.2002.02329.x

22. El-Boghdadly K, Bailey CR, Wiles MD. Postoperative sore throat: a systematic review. *Anaesthesia*. (2016) 71(6):706–17. doi: 10.1111/anae.13438

23. Lee JY, et al. Incidence and risk factors of postoperative sore throat after endotracheal intubation in Korean patients. *J Int Med Res*. (2017) 45(2):744–52. doi: 10.1177/0300060516687227

24. Calder A, et al. Predictors of postoperative sore throat in intubated children-reply. *Pediatr Anesth*. (2012) 22:596–7. doi: 10.1111/j.1460-9592.2012.03871.x

25. Lehmann M, et al. Postoperative patient complaints: a prospective interview study of 12,276 patients. *J Clin Anesth*. (2010) 22(1):13–21. doi: 10.1016/j.jclinane.2009.02.015

26. Kim MS, et al. A randomized comparison of the i-gel™ with the self-pressurized air-Q™ intubating laryngeal airway in children. *Pediatr Anesth*. (2015) 25(4):405–12. doi: 10.1111/pan.12609

27. Hameed M, Samad K, Ullah H. Comparison of two supraglottic airway devices on postoperative sore throat in children: a prospective randomized controlled trial. *Rev Bras Anestesiol*. (2020) 70:240–7.

28. Tanaka Y, et al. Lidocaine for preventing postoperative sore throat. *Cochrane Database Syst Rev*. (2015) 7.

29. Lee JH, et al. Effects of topical dexamethasone in postoperative sore throat. *Korean J Anesthesiol*. (2017) 70(1):58–63. doi: 10.4097/kjae.2017.70.1.58

30. Al-Qahtani AS, Messahel FM. Quality improvement in anesthetic practice–incidence of sore throat after using small tracheal tube. *Middle East J Anaesthesiol*. (2005) 18(1):179–83.

31. Ahuja V, Mitra S, Sarna R. Nebulized ketamine decreases incidence and severity of post-operative sore throat. *Indian J Anaesth*. (2015) 59(1):37–42. doi: 10.4103/0019-5049.149448

32. Lee J, et al. Combined intraoperative paracetamol and preoperative dexamethasone reduces postoperative sore throat: a prospective randomized study. *J Anesth*. (2017) 31(6):869–77. doi: 10.1007/s00540-017-2411-6

33. Zhao X, Cao X, Li Q. Dexamethasone for the prevention of postoperative sore throat: a systematic review and meta-analysis. *J Clin Anesth*. (2015) 27(1):45–50. doi: 10.1016/j.jclinane.2014.06.014

34. Gupta S, et al. Nebulized magnesium for prevention of postoperative sore throat. *Br J Anaesth*. (2012) 108(1):168–9. doi: 10.1093/bja/aer437

35. Yhim HB, Yoon SH. Effects of benzydamine hydrochloride on postoperative sore throat after extubation in children: a randomized controlled trial. *BMC Anesthesiology*, (2020) 20(1):77.

36. Hameed M, Samad K, Ullah H. [Comparison of two supraglottic airway devices on postoperative sore throat in children: a prospective randomized controlled trial]. *Braz J Anesthesiol*. (2020) 70(3):240–7. doi: 10.1016/j.bjan.2020.03.006

37. Butler EK, et al. Epidemiology of pediatric surgical needs in low-income countries. *PLoS One*. (2017) 12(3):e0170968. doi: 10.1371/journal.pone.0170968

38. Canbay O, et al. Ketamine gargle for attenuating postoperative sore throat. *Br J Anaesth*. (2008) 100(4):490–3. doi: 10.1093/bja/aen023

39. Hunduma JC, Merga HT. Predictors of postoperative sore throat among surgical patients at Ethiopian teaching hospitals. *J Clin Med Res*. (2016) 8(1):1–11. doi: 10.5897/JCMR2016.0280

40. Chang J-E, et al. Effect of endotracheal tube cuff shape on postoperative sore throat after endotracheal intubation. *Anesth Analg*. (2017) 125(4):1240–5. doi: 10.1213/ANE.0000000000001933

41. Wong JGL, et al. Impact of laryngeal mask airway cuff pressures on the incidence of sore throat in children. *Pediatr Anesth*. (2009) 19(5):464–9. doi: 10.1111/j.1460-9592.2009.02968.x

42. Yu J, Ren L, Min S. Nebulized pharmacological agents for preventing postoperative sore throat: a systematic review and network meta-analysis. *PLOS ONE* (2020) 15(8):e0237174.

43. Perrott C, et al. Perioperative experiences of anesthesia reported by children and parents. *Pediatric Anesthesia* (2018) 28(2):149–56.

44. Navarro RM, Baughman VL. Lidocaine in the endotracheal tube cuff reduces postoperative sore throat. *J Clin Anesth*. (1997) 9(5):394–7. doi: 10.1016/S0952-8180(97)00068-8

45. Mitobe Y, et al. A literature review of factors related to postoperative sore throat. *J Clin Med Res*. (2022) 14(2):88–94. doi: 10.14740/jocmr4665

46. Mamie C, et al. Incidence and risk factors of perioperative respiratory adverse events in children undergoing elective surgery. *Paediatr Anaesth*. (2004) 14(3):218–24. doi: 10.1111/j.1460-9592.2004.01169.x

47. Kloub R. Sore throat following tracheal intubation. *Middle East J Anaesthesiol*. (2001) 16(1):29–40.

48. Gemechu BM, Gebremedhn EG, Melkie TB. Risk factors for postoperative throat pain after general anaesthesia with endotracheal intubation at the University of Gondar Teaching Hospital, Northwest Ethiopia, 2014. *Pan Afr Med J*. (2017) 27:127. doi: 10.11604/pamj.2017.27.127.10566

49. Felten ML, et al. Endotracheal tube cuff pressure is unpredictable in children. *Anesth Analg*. (2003) 97(6):1612–6. doi: 10.1213/01.ANE.0000087882.04234.11

50. Hoffman RJ, et al. Linear correlation of endotracheal tube cuff pressure and volume. *West J Emerg Med*. (2009) 10(3):137.

51. Khine HH, et al. Comparison of cuffed and uncuffed endotracheal tubes in young children during general anesthesia. *Anesthesiology*. (1997) 86(3):627–31; discussion 27A. doi: 10.1097/00000542-199703000-00015

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 09 March 2023
DOI 10.3389/fpubh.2023.1137933

OPEN ACCESS

EDITED BY
Jing Wu,
Tianjin University, China

REVIEWED BY
Arlette Setiawan,
Padjadjaran University, Indonesia
Linhua Liu,
Guangdong Medical University, China

*CORRESPONDENCE
Qiaoyuan Yang
✉ qiaoyuan_yang@gzhmu.edu.cn

†These authors have contributed equally to this work

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

RECEIVED 05 January 2023
ACCEPTED 20 February 2023
PUBLISHED 09 March 2023

CITATION
Zheng Y, Chen S, Chen Y, Li J, Xu B, Shi T and Yang Q (2023) Association between $PM_{2.5}$-bound metals and pediatric respiratory health in Guangzhou: An ecological study investigating source, health risk, and effect. *Front. Public Health* 11:1137933. doi: 10.3389/fpubh.2023.1137933

Association between $PM_{2.5}$-bound metals and pediatric respiratory health in Guangzhou: An ecological study investigating source, health risk, and effect

Yi Zheng[1†], Sili Chen[1†], Yuyang Chen[2], Jingye Li[1], Binhe Xu[3], Tongxing Shi[4,5] and Qiaoyuan Yang[1,5*]

[1]Department of Preventive Medicine, School of Public Health, Guangzhou Medical University, Guangzhou, China, [2]Department of Anesthesiology, School of Anesthesiology, Southern Medical University, Guangzhou, China, [3]Department of Clinical Medicine, Basic Medicine College, Zunyi Medical University, Zunyi, China, [4]Department of Environmental Health, Guangzhou Center for Disease Control and Prevention, Guangzhou, China, [5]Department of Environmental Health, Institute of Public Health, Guangzhou Medical University, Guangzhou, China

Background: The adverse effects of 2.5-μm particulate matter ($PM_{2.5}$) exposure on public health have become an increasing concern worldwide. However, epidemiological findings on the effects of $PM_{2.5}$-bound metals on children's respiratory health are limited and inconsistent because $PM_{2.5}$ is a complicated mixture.

Objectives: Given the vulnerability of children's respiratory system, aim to pediatric respiratory health, this study evaluated the potential sources, health risks, and acute health effects of ambient $PM_{2.5}$-bound metals among children in Guangzhou, China from January 2017 to December 2019.

Methods: Potential sources of $PM_{2.5}$-bound metals were detected using positive matrix factorization (PMF). A health risk assessment was conducted to investigate the inhalation risk of $PM_{2.5}$-bound metals in children. The associations between $PM_{2.5}$-bound metals and pediatric respiratory outpatient visits were examined with a quasi-Poisson generalized additive model (GAM).

Results: During 2017–2019, the daily mean concentrations of $PM_{2.5}$ was 53.39 $\mu g/m^3$, and the daily mean concentrations of $PM_{2.5}$-bound metals range 0.03 ng/m^3 [thorium (Th) and beryllium (Be)] from to 396.40 ng/m^3 [iron (Fe)]. $PM_{2.5}$-bound metals were mainly contributed by motor vehicles and street dust. $PM_{2.5}$-bound arsenic (As), cadmium (Cd), cobalt (Co), chromium (Cr)(VI), nickel (Ni), and lead (Pb) were found to pose a carcinogenic risk (CR). A quasi-Poisson GAM was constructed that showed there were significant associations between $PM_{2.5}$ concentrations and pediatric outpatient visits for respiratory diseases. $PM_{2.5}$ was significantly associated with pediatric outpatient visits for respiratory diseases. Moreover, with a 10 $\mu g/m^3$ increase in Ni, Cr(VI), Ni, and As concentrations, the corresponding pediatric outpatient visits for respiratory diseases increased by 2.89% (95% *CI*: 2.28–3.50%), acute upper respiratory infections (AURIs) increased by 2.74% (2.13–3.35%), influenza and pneumonia (FLU&PN) increased by 23.36% (20.09–26.72%), and acute lower respiratory infections (ALRIs) increased by 16.86% (15.16–18.60%), respectively.

Conclusions: Our findings showed that $PM_{2.5}$ and $PM_{2.5}$-bound As, Cd, Co, Cr(VI), Ni, and Pb had adverse effects on pediatric respiratory health during the study period. New strategies are required to decrease the production of $PM_{2.5}$ and $PM_{2.5}$-bound metals by motor vehicles and to reduce levels of street dust to reduce children's exposure to these pollutants and thereby increase child health.

KEYWORDS

$PM_{2.5}$-bound metals, source identification, children, health risk, time-series study

Introduction

The International Agency for Research on Cancer (IARC) defines outdoor particulate matter (PM) as a Group 1 carcinogen in humans (1). Exposure to fine particulate matter ($PM_{2.5}$; particulate matter with an aerodynamic diameter of $\leq$2.5 μm) has significantly increased the burden of disease globally, mainly by increasing the risk of diabetes, cardiovascular diseases, poor respiratory health, and lung cancer (2, 3). In China, $PM_{2.5}$ exposure is one of the top four risk factors for death from stroke, ischemic heart disease, chronic obstructive pulmonary disease (COPD), and lung cancer (2).

The small size and large specific surface area of $PM_{2.5}$ means that it can easily bind toxic compounds, including metals. Among these, $PM_{2.5}$-bound metals are key contributors to $PM_{2.5}$ toxicity and can cause severe adverse health effects. $PM_{2.5}$-bound metals can enter the body by ingestion, dermal contact, or by inhalation, which allows them to affect cardiopulmonary function and distribute to organs through the blood circulation (4). Several epidemiological studies have found that exposure to $PM_{2.5}$-bound nickel (Ni), vanadium (V), lead (Pb), and zinc (Zn) can increase cardiopulmonary diseases, and that exposure to $PM_{2.5}$-bound arsenic (As), cobalt (Co), Ni, manganese (Mn), and chromium (Cr) can trigger the production of reactive oxygen species and thereby cause respiratory inflammation (5–9). However, owing to on the inconsistency of $PM_{2.5}$-bound metals, there is a limited epidemiological understanding of their effects on respiratory health (10–13). Therefore, it is critical to investigate the health effects of $PM_{2.5}$-bound metals further.

A better understanding of the sources of $PM_{2.5}$-bound metals and the identification of those metals that are most harmful to children's health would assist decision-makers to develop air pollution regulations and provide directions for basic research. Positive matrix factorization (PMF) is an analytical method that has been widely employed to determine the contributions of various sources of $PM_{2.5}$ to air pollution and is recommended by the United States Environmental Protection Agency (USEPA) (14, 15). Moreover, health risk assessments are recommended by the USEPA and the National Health Commission of the People's Republic of China (16, 17) for identifying heavy metals that pose carcinogenic and non-carcinogenic risks to humans via inhalation (8). A generalized additive model (GAM) with a quasi-Poisson link can be used for time series analysis of the short-term association between ambient air pollution and outpatient visits, as it enables control for long-term trends, day of the week (DOW), temperature, and humidity (18). Thus, we used this model to determine the most significant associations between $PM_{2.5}$-bound metals and respiratory diseases in children.

Guangzhou is the third-largest city in China and is one of the top four first-tier cities in the country. Due to rapid social and economic development, air pollution condition in Guangzhou has become more serious. Present study on the sources of $PM_{2.5}$-bound metals in Guangzhou mainly focused on total PM, and the results show that the main sources are traffic emissions, soil dust, and biomass burning (19). However, since the distribution spectrum of metal is specific to different sources, it is of public health importance to study the sources of $PM_{2.5}$-bound metals and to discover the effects on the population for better regulation of different sources in Guangzhou. In addition, many studies have demonstrated that children are more vulnerable than adults to the adverse effects of air pollution for many reasons. These include the fact that children breathe twice as quickly as adults, children's lungs are still growing, and children spend more time outside than adults (20, 21). Hence, comprehensive studies are needed to determine the characteristics of air pollution and the relationship between metal-containing pollution and respiratory diseases in children.

The major objectives of this study were to (1) use PMF to determine the sources of $PM_{2.5}$ and $PM_{2.5}$-bound metals air pollution and estimate their contribution to overall $PM_{2.5}$ and $PM_{2.5}$-bound metals air pollution (as a ratio); (2) identify the non-carcinogenic and carcinogenic risks (CR) posed by $PM_{2.5}$-bound metals to children; and (3) use a GAM to investigate the short-term relationships between air concentrations of $PM_{2.5}$-bound metals and daily pediatric outpatient visits for respiratory diseases.

Materials and methods

Study location

Guangzhou (22°26′-23°56′ N, 112°57′-114°3′ E) is located in the south of China and is the primary economic and cultural city in the Pearl River Delta. It covers an area of 7,434 km^2 and had a residential population of 8,970,000 in 2017. Due to its location, Guangzhou is an influential port and transportation hub for Guangdong province. Guangzhou has a humid, warm, tropical/subtropical climate, so its weather is hot and humid in summer and mild and dry in winter. The annual average temperature is 22°C, and the annual average humidity is 80%. Yuexiu district is one of the "four old districts" in Guangzhou and from 2017 to 2019 was ranked 10th out of the 11 districts of Guangzhou in terms of air pollution, and Liwan district ranked 11th. Yuexiu district also has the largest population of all districts in Guangzhou, with children younger than 18 years accounting for 14.41% of its population.

Hospital outpatients

Data from January 1, 2017 to December 31, 2019 were obtained from Guangzhou Yuexiu District Children's Hospital, which is one of the two specialized children's hospitals in Guangzhou. During the study period, individuals who visited the respiratory pediatric department of this hospital and were diagnosed with respiratory disease [International Statistical Classification of Diseases and Related Health Problems 10th revision (ICD-10) codes: J00–J99 and R04–R9.3] were identified. Specifically, patients who over 18 years of age would be excluded, and the respiratory diseases would be further divided into three kind of specific diseases: acute upper respiratory infections (AURIs, ICD-10 codes J00–J06), influenza and pneumonia (FLU&PN, J09–J18), and acute lower respiratory infections (ALRIs, J20–J22).

$PM_{2.5}$-bound metals and meteorological data

We collected 24-h mean concentrations of $PM_{2.5}$ and 26 metals from the air monitoring station of the Guangdong Environment Monitoring Center in Yuexiu district for 265 days. The 26 metals were silver (Ag), aluminum (Al), As, barium (Ba), beryllium (Be), bismuth (Bi), Cd, Co, chromium (Cr), copper (Cu), iron (Fe), mercury (Hg), lithium (Li), Ni, Mn, molybdenum (Mo), Pb, antimony (Sb), selenium (Se), tin (Sn), strontium (Sr), thorium (Th), thallium (Tl), uranium (U), V, and Zn. The samples were obtained approximately in the middle of every month, for an average of approximately 10 days, from 2017 to 2019. The air pollution monitor was located approximately 10–20 m above ground level. It was regularly maintained in accordance with the standard operating procedures described in the China National Quality Control Assurance Plan. We also obtained the daily 24-h mean $PM_{2.5}$ mass from 2017 to 2019 (674 days).

The daily meteorological data of temperature (°C) and relative humidity (%) in the study period were obtained from the Guangdong Meteorological Service.

Statistical analysis

PMF

We used EPA PMF 5.0 to quantify the source contributions of $PM_{2.5}$-bound metals. The number of sources was calculated based on Cattell's scree test. The mathematical expression of PMF can be written as follows (Equation 1):

$$X_{ij} = \sum_{k=1}^{p} G_{ik}F_{kj} + E_{ij} \quad (1)$$

where X_{ij} is the concentration of metal j in sample i; p is the number of pollution sources; G_{ik} is the factor contribution of source p to sample i; F_{kj} is the factor concentration of pollutant j from source p; and E_{ij} is the residual.

G_{ik} and F_{kj} can be calculated by minimizing the objective function Q (Equation 2), as follows:

$$Q = \sum_{i=1}^{n}\sum_{j=1}^{m}\left(\frac{X_{ij} - \sum_{k=1}^{p} G_{ik}F_{kj}}{U_{ij}}\right) = \sum_{i=1}^{n}\sum_{j=1}^{m}\left(\frac{E_{ij}}{U_{ij}}\right)^2 \quad (2)$$

where U_{ij} is the uncertainty in the concentration of each $PM_{2.5}$-bound metal component.

Finally, because we did not know the detection limit for the measurement method, we calculated the uncertainty as follows (Equation 3) (22):

$$U_{ij} = \sqrt{(c_j \times SD)^2 + (0.05 \times X_{ij})^2} \quad (3)$$

where c_j is the relative uncertainty for high values of the measured parameter, and SD is the standard deviation of X_{ij}.

Health risk assessment

To assess the health risk of inhaled $PM_{2.5}$-bound metals in children, we adopted the model recommended by the standards of both the USEPA and the National Health Commission of the People's Republic of China. According to previous research (8, 23), Al, As, Ba, Cd, Co, Cr(VI), Mn, Ni, Se, and V are considered to be non-carcinogenic, while As, Cd, Co, Cr(VI), Ni, and Pb are considered to be both carcinogenic and non-carcinogenic. Daily respiratory tract exposure concentrations were estimated by Equation (4), as follows:

$$EC = (C \times ET \times EF \times ED \times ASF) \div AT \quad (4)$$

where EC is the average daily concentration of metal exposure through inhalation ($\mu g/m^3$); C is the concentration of metals in $PM_{2.5}$ ($\mu g/m^3$); ET is the exposure time (24 h/d for children), EF is the exposure time (365 d/year for children), ED is the exposure duration (18 years for children); ASF is the age sensitivity factor (3 for children); and AT is the average contact time (18 $years \times 365\ d/year \times 24\ h/d$ for non-carcinogenic risk and 70 $years \times 365\ d/year \times 24\ h/d$ for carcinogenic risk).

The carcinogenic and non-carcinogenic risks were further estimated based on EC. The non-carcinogenic risk level was calculated by Equation (5), as follows:

$$HQ = EC \div (RfC \times 1000) \quad (5)$$

where HQ is the hazard quotient, which represents the non-carcinogenic risk for $PM_{2.5}$-bound metals, and the RfC is the maximum daily reference concentration (Table 1), above which there will be a non-carcinogenic risk over a child's lifetime (mg/m^3). $HQ < 1$ indicates that there will be no adverse health effects, while $HQ > 1$ indicates that non-carcinogenic effects are possible.

Carcinogenic risk (CR), which represents the risk of developing cancer due to exposure to $PM_{2.5}$-bound metals, was calculated by Equation (6), as follows:

$$CR = IUR \times EC \quad (6)$$

where IUR is the inhalation unit risk ($m^3/\mu g$). $CR < 1 \times 10^{-6}$ indicates a negligible carcinogenic risk in children; $CR = 1 \times 10^{-6}$-1×10^{-4} indicates a possible carcinogenic risk in children; and $CR > 10^{-4}$ indicates a high carcinogenic risk in children and therefore attention is needed.

GAM

We used two-stage Poisson regression in a GAM to examine the associations between pediatric outpatient visits for respiratory diseases and daily $PM_{2.5}$ and $PM_{2.5}$-bound metal concentrations, respectively. This model complies with the recommendations of the National Health Commission of the People's Republic of China (17). Due to missing data—the lack of daily concentration data for $PM_{2.5}$-bound metals—we followed previous studies by computing the ratio of the monthly mean concentration of $PM_{2.5}$-bound metals to the daily $PM_{2.5}$ concentration (13, 24). We controlled for seasonality and long-term trends in the model by using penalized smoothing splines with six degrees of freedom (df) per year. We also included an indicator variable for the DOW and the daily mean temperature and relative humidity, each with three df. Consistent with a previous study (25), we calculated the effect of $PM_{2.5}$ exposure with a lag of 0–7 days and a moving average of lag days

10.3389/fpubh.2023.1137933

TABLE 1 Summary of non-carcinogenic and carcinogenic risks from exposure to $PM_{2.5}$-bound metals via inhalation in children.

Measure	Al[a]	As[a]	Ba[b]	Cd[a]	Co[a]	Cr(VI)[a]	Mn[a]	Ni[a]	Pb[a]	Se[a]	V[c]
IUR[c]	–	4.30×10^{-3}	–	1.80×10^{-3}	9.00×10^{-3}	8.40×10^{-2}	–	2.40×10^{-4}	1.20×10^{-5}	–	–
RfC[c]	5.00×10^{-3}	1.50×10^{-3}	5.00×10^{-4}	1.00×10^{-5}	6.00×10^{-6}	1.00×10^{-4}	5.00×10^{-5}	1.40×10^{-5}	–	2.00×10^{-2}	1.00×10^{-4}
HQ[e]	6.85×10^{-2}	1.21×10^{-2}	1.06×10^{-1}	3.66×10^{-1}	1.30×10^{-1}	1.61×10^{-2}	1.51*	5.29×10^{-1}	–	3.81×10^{-4}	1.37×10^{-1}
CR[e]	–	7.80×10^{-5}*	–	6.59×10^{-6}*	7.02×10^{-6}*	1.35×10^{-4}**	–	1.78×10^{-6}*	1.16×10^{-6}*	–	–

[a]Data from Integrated Risk Information System (IRIS), Health Effects Assessment Summary Tables (HEAST), Agency for Toxic Substances and Disease Registry (ATSDR).

[b]"–" indicates that no recommended parameters were identified or the parameter did not have units.

[c]*IUR* units are $(\mu g/m^3)^{-1}$, *RfC* units are $(mg/m^3)^{-1}$.

[d]"*" indicates $HQ > 1$ or *CR* between 1×10^{-6} and 1×10^{-4}, "**" indicates $CR > 1 \times 10^{-4}$.

[e]*IUR* represents the inhalation unit risk, *RfC* represents the maximum daily reference concentration *HQ* represents the hazard quotient, *CR* represents carcinogenic risk.

01–07, in addition to the effect of $PM_{2.5}$-bound metal exposure with a lag of 0 days.

In the first stage, we fitted a time series analysis of daily pediatric outpatient visits for respiratory disease and daily $PM_{2.5}$ concentration (Equation 7). In the second stage, we used the $PM_{2.5}$ concentration, the ratio between the monthly mean concentration of the $PM_{2.5}$-bound metal and the daily $PM_{2.5}$ concentration, and the number of daily pediatric outpatient visits for respiratory diseases in Equation (8). These two equations are presented below:

$$\log E(Y_t) = \beta_0 Z_t + s(time, df) + s(X_t, df) + DOW + \alpha$$
$$\log E(Y_t) = \beta_1 Z_t + s(time, df) + s(X_t, df) + DOW + \alpha + \beta_2\ monthly\ metal\ concentration/PM_{2.5} + \quad (7)$$

$$\beta_3\ monthly\ metal\ concentration/PM_{2.5} \times PM_{2.5} \quad (8)$$

where $E(Y_t)$ is the expected number of daily pediatric outpatient visits for respiratory diseases on day t; β_0–β_3 are the regression coefficients, where β_0–β_2 represent the main effects of $PM_{2.5}$ and the ratio of the monthly mean concentration of $PM_{2.5}$-bound metals to the daily $PM_{2.5}$ concentration and β_3 is the interaction term; Z_t is the daily $PM_{2.5}$ concentration; s is the penalized smoothing spline function; df are the degrees of freedom; X_t is the meteorological factor, which includes daily mean temperature and relative humidity, DOW is the day of week, which was used to adjust for the DOW effect; and a is the intercept. We reported the results as the percent change in excess risk (ER) with a 95% confidence interval (CI) for every 10 $\mu g/m^3$ increase in $PM_{2.5}$ concentration or for every interquartile range (IQR) increase in the ratio of the monthly mean $PM_{2.5}$-bound metal concentration to daily $PM_{2.5}$ concentration.

All analyses were conducted using the *mgcv* package in R (version 4.0.2) software. $P < 0.05$ was considered statistically significant.

Results

Table 2 shows the means and standard deviations of $PM_{2.5}$ concentrations, $PM_{2.5}$-bound metal concentrations, and pediatric outpatient visits for respiratory disease from 2017 to 2019. During the study period, there were 1,707,346 pediatric outpatient visits for respiratory diseases, ranging from 176 to 3,046 visits per day. During the 265-day sampling period, the daily mean concentrations of $PM_{2.5}$, Cd, As, Cr, and Hg were 53.39 $\mu g/m^3$, 1.22, 6.05, 3.22, and 0.06 ng/m^3, respectively. Based on the secondary and primary concentration limits in GB 3085-2012 and stipulated by the World Health Organization, 19.60%, 69.00%, and 89.10% of the days exceeded the $PM_{2.5}$ concentration limit (75.00, 35.00, and 25.00 $\mu g/m^3$), and 1.10%, 35.80%, and 100% of the days exceeded the Cd, As, and Cr concentration limits, respectively. Cr(VI) accounted for 16.66% of the total Cr concentration (26).

As shown in Figure 1, the $PM_{2.5}$ concentrations were higher in winter and lower in summer. The time series analysis required $PM_{2.5}$ concentration information on consecutive dates; however,

TABLE 2 Daily means and standard deviations of $PM_{2.5}$ concentration, $PM_{2.5}$-bound metal concentrations, and pediatric respiratory outpatient visits from 2017 to 2019.

Item	Units	$\bar{x} \pm s$	Min	P25	P50	P75	Max
$PM_{2.5}$	($\mu g/m^3$)	53.39 ± 30.10	6.30	32.45	46.00	65.75	199.00
Ag	(ng/m^3)	0.22 ± 0.17	0.03	0.10	0.18	0.28	1.45
Al	(ng/m^3)	114.13 ± 70.66	19.98	66.60	93.70	142.00	495.00
As	(ng/m^3)	6.05 ± 5.35	1.01	2.68	4.56	7.40	42.70
Ba	(ng/m^3)	17.66 ± 16.63	4.69	10.00	13.50	19.40	171.19
Be	(ng/m^3)	0.03 ± 0	0.03	0.03	0.03	0.03	0.03
Bi	(ng/m^3)	1.64 ± 1.56	0.17	0.62	1.35	2.11	13.20
Cd	(ng/m^3)	1.22 ± 1.07	0.21	0.52	0.95	1.58	8.30
Co	(ng/m^3)	0.26 ± 0.14	0.06	0.16	0.23	0.34	0.81
Cr	(ng/m^3)	3.22 ± 4.83	0.35	1.51	2.47	3.63	70.70
Cu	(ng/m^3)	21.02 ± 13.04	4.89	11.40	17.00	26.30	85.70
Fe	(ng/m^3)	396.40 ± 318.27	130.69	243.00	319.00	461.00	4,420.00
Hg	(ng/m^3)	0.06 ± 0.09	0.00	0.02	0.05	0.08	0.90
Li	(ng/m^3)	0.65 ± 0.48	0.03	0.33	0.54	0.82	3.36
Mn	(ng/m^3)	25.24 ± 14.21	7.17	14.70	21.30	30.70	91.40
Mo	(ng/m^3)	0.76 ± 0.43	0.11	0.46	0.65	0.97	2.65
Ni	(ng/m^3)	2.47 ± 2.11	0.06	1.00	1.91	3.22	14.60
Pb	(ng/m^3)	32.21 ± 25.77	5.19	12.40	25.00	41.60	155.00
Sb	(ng/m^3)	3.49 ± 2.45	0.81	1.70	2.88	4.31	16.25
Se	(ng/m^3)	2.54 ± 1.59	0.43	1.38	2.10	3.27	9.25
Sn	(ng/m^3)	5.21 ± 3.95	1.19	2.80	4.13	6.51	31.40
Sr	(ng/m^3)	2.57 ± 3.93	0.33	1.31	1.84	2.77	42.76
Th	(ng/m^3)	0.03 ± 0.00	0.03	0.03	0.03	0.03	0.03
Tl	(ng/m^3)	0.29 ± 0.24	0.04	0.13	0.23	0.36	1.37
U	(ng/m^3)	0.04 ± 0.03	0.02	0.02	0.04	0.05	0.14
V	(ng/m^3)	4.55 ± 5.05	0.12	1.00	3.14	5.67	34.00
Zn	(ng/m^3)	139.15 ± 90.75	36.00	81.80	108.00	173.00	736.00
Respiratory diseases outpatient visits	*n*	1,571 ± 455	176	1,284	1,546	1,805	3,046
AURIs	*n*	860 ± 262	102	665	850	1,048	1,559
FLU&PN	*n*	67 ± 52	2	30	51	94	404
ALRIs	*n*	297 ± 128	37	185	296	401	649

$\bar{x}$, mean; s, standard deviation; P25, lower quartile; P50, median; P75, upper quartile; Max, maximum value; Min, minimum value; AURIs, acute upper respiratory infections; FLU&PN, influenza and pneumonia; ALRIs, acute lower respiratory infections.

there were incomplete data for the 24-h mean $PM_{2.5}$ concentrations for the 26 metals (data from only 265 days were available). However, the trends in the $PM_{2.5}$ concentrations on these 265 days and the trends in the $PM_{2.5}$ concentrations over the entire study period were similar. Therefore, we used the continuously monitored $PM_{2.5}$ data to fit the incomplete $PM_{2.5}$ concentration data by correlation and linear regression, and the results showed a strong correlation ($r = 0.74$, $P < 0.01$) and were significant ($\beta = 0.97$, $P < 0.05$).

Source classification

We used PMF to analyze the sources of $PM_{2.5}$-bound metals. First, Cattle's scree test was performed to determine the number of principal components. Next, based on the source spectrum database in China (15, 27, 28), EPA PMF 5.0 was employed to quantify $PM_{2.5}$-bound metal sources. Based on the Kaiser criterion in the Cattle's scree test, we retained four principal factors as pollution sources. Figure 2 presents the PMF results

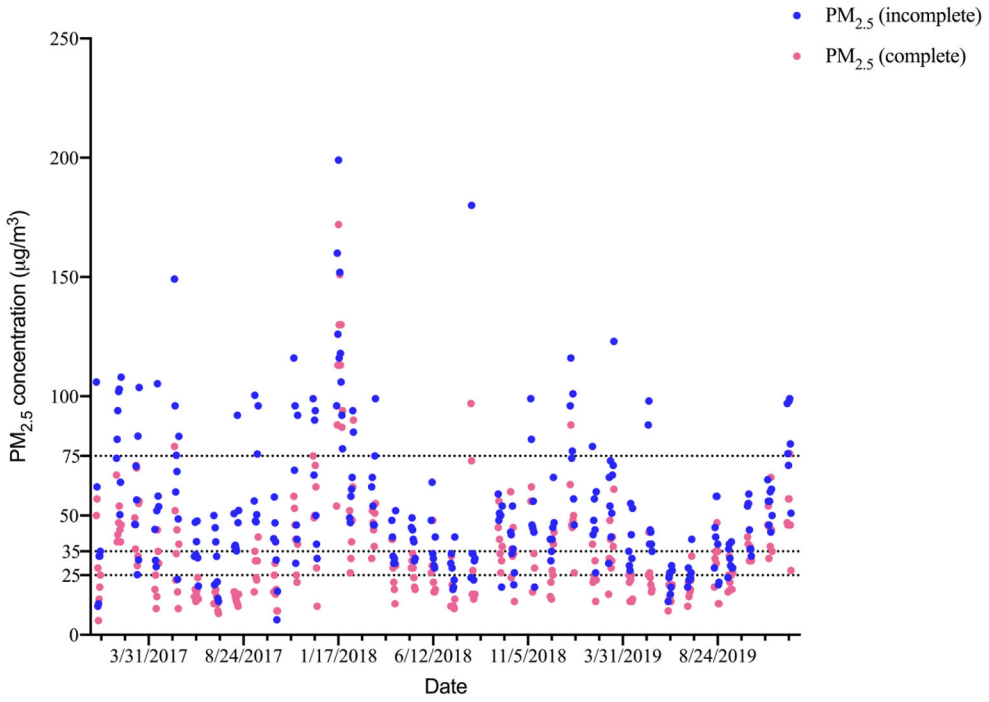


FIGURE 1
Trends in the complete and incomplete $PM_{2.5}$ concentration ($\mu g/m^3$) on the sampling dates in Yuexiu district and a comparison with World Health Organization and GB 3085-2012 standards. The three lines represent $PM_{2.5}$ concentration limits in GB 3085-2012 (75.00 and 35.00 $\mu g/m^3$) and World Health Organization (25.00 $\mu g/m^3$).

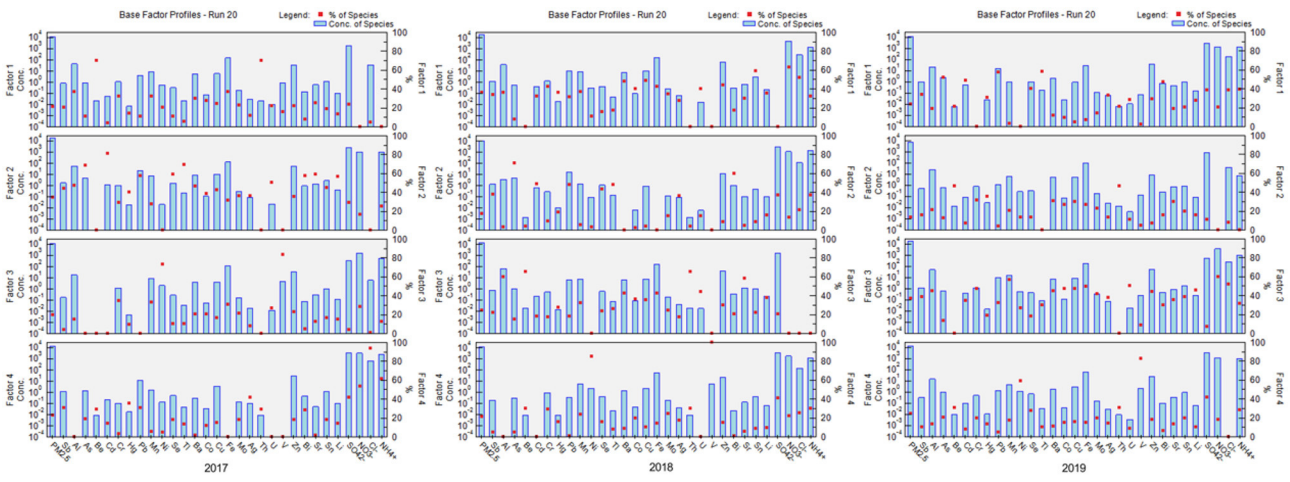


FIGURE 2
Positive matrix factorization (PMF) analysis of the distribution of sources of $PM_{2.5}$-bound metal from 2017 to 2019. Red dot means percentage ratio of each kind of species, blue bar means concentration of each species.

from 2017 to 2019. These results reveal that the six potential sources chosen by the model were street dust, motor vehicles, ships and heavy oil, secondary nitrates, boiler combustion, and aerospace manufacturing.

Among these six potential sources, street dust was characterized by high concentrations of Al, Cr, and Fe; motor vehicles were characterized by high concentrations of Pb and Zn; ships and heavy oil were characterized by high concentrations of Ni and V; secondary nitrates were characterized by high concentrations of sulfate (SO_4^{2-}), nitrate (NO_3^-), and ammonium (NH_4^+); boiler combustion was characterized by high concentrations of As; and aerospace manufacturing was characterized by high concentrations of Be.

The source contributions to $PM_{2.5}$-bound metals from 2017 to 2019 are shown in Figure 3. The highest contributions in the 3 years

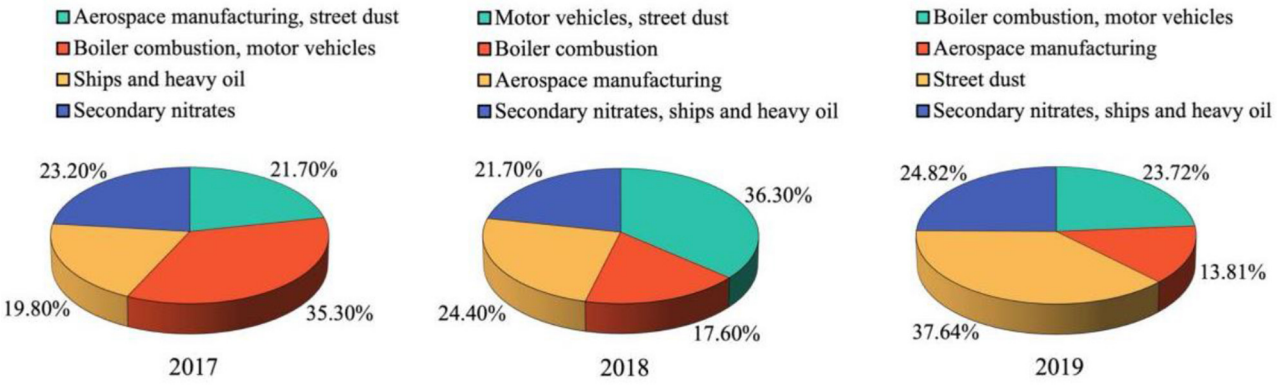


FIGURE 3
Source contributions to $PM_{2.5}$-bound metals from 2017 to 2019.

were from motor vehicles and boiler combustion (2017), motor vehicles and street dust (2018), and street dust (2019).

Human health risk assessment of exposure to $PM_{2.5}$-bound metals

Table 1 shows the exposure parameters and non-carcinogenic risk assessment results for exposure to $PM_{2.5}$-bound Al, As, Ba, Cd, Co, Cr(VI), Mn, Ni, Se, and V, and the CRs for exposure to $PM_{2.5}$-bound As, Cd, Co, Cr(VI), Ni, and Pb. Only Mn posed a non-carcinogenic health risk, and thus the other $PM_{2.5}$-bound metals were not included in the time series analysis. The mean CRs of As, Cd, Co, Cr(VI), Ni, and Pb exceeded the acceptable risk level (1×10^{-6}); therefore, the relatively high exposures of children to these $PM_{2.5}$-bound metals are of concern.

Associations between concentrations of $PM_{2.5}$ and its metal constituents and pediatric respiratory outpatient visits

Figure 4 shows the $PM_{2.5}$-associated *ER* for total pediatric outpatient visits for diseases and pediatric outpatient visits for three respiratory diseases. The lag structure of changes in pediatric outpatient visits for respiratory diseases was associated with a 10 $\mu g/m^3$ increase in $PM_{2.5}$ daily concentration. $PM_{2.5}$ concentration significantly affected the number of daily pediatric outpatient visits due to AURIs, FLU&PN, and ALRIs.

Regarding the lag effect, lag3 was specifically associated with a 1.49% (95% *CI*: 1.39–1.59%) increase in daily pediatric outpatient visits for respiratory diseases per 10 $\mu g/m^3$ increase in $PM_{2.5}$ concentrations. Regarding pediatric outpatient visits for the three causes of respiratory diseases, the highest lag effects were a 3.04% (95% *CI*: 2.59–3.49%) increase in FLU&PN at lag0, a 2.18% (95% *CI*: 2.04–2.31%) increase in AURIs at lag0, and a 1.36% (95% *CI*: 1.13–1.59%) increase in ALRIs at lag5.

In terms of the cumulative lag effect, the strongest effect was at lag07 for all but FLU&PN—the FLU&PN daily outpatient visits reached a peak at lag05. In terms of total pediatric outpatient visits for respiratory diseases, each 10 $\mu g/m^3$ increase in the concentration of $PM_{2.5}$ was associated with a 3.18% (95% *CI*: 3.01–3.35%) increase in visits. FLU&PN visits were most affected, with an increase of 4.84% (95% *CI*: 4.20–5.49%), which was consistent with the results of the lagged effects model.

Relationships between $PM_{2.5}$-bound metal concentrations and respiratory diseases in children

We included As, Cd, Co, Cr(VI), Mn, Ni, and Pb in the GAM because these metals were found in the risk assessment to have carcinogenic and non-carcinogenic risks in children. Next, as mentioned, the ratio of the monthly mean concentration of $PM_{2.5}$-bound metals to the daily $PM_{2.5}$ concentration was calculated to compensate for the absence of daily metal information. The *ER* for respiratory disease outpatient visits is given per *IQR* increase in the monthly mean ratio of metal concentration to $PM_{2.5}$ concentration.

Figure 5 shows the changes in total pediatric outpatient visits for respiratory diseases and pediatric outpatient visits for the three respiratory diseases associated with an *IQR* increase in the monthly mean metal-to-$PM_{2.5}$ concentration ratio. The increase in monthly mean Ni concentration/$PM_{2.5}$ concentration had the greatest effect, increasing total pediatric outpatient visits for respiratory diseases by 2.89% (95% *CI*: 2.28–3.50%) and for daily pediatric outpatient visits for FLU&PN by 23.36% (95% *CI*: 20.09–26.72%). An *IQR* increase in the monthly mean Cr(VI) and As concentrations/$PM_{2.5}$ concentration had the greatest effects on daily pediatric outpatient visits for AURIs and ALRIs, respectively.

The exposure–response relationship curves in Figure 6 indicate that the daily concentrations of $PM_{2.5}$ and $PM_{2.5}$-bound metals were associated with total daily pediatric outpatient visits for respiratory diseases at lag0 (Figure 6). The relationships between exposure to $PM_{2.5}$, As, Cd, Co, Cr(VI), Mn, and total daily pediatric outpatient visits for respiratory diseases show an "N" shape, suggesting that lower concentrations of $PM_{2.5}$ and these six metals increased the total daily pediatric outpatient visits for

10.3389/fpubh.2023.1137933

Respiratory diseases	ER	95% CI
lag0	1.49	1.39–1.59
lag1	1.25	1.14–1.35
lag2	1.27	1.17–1.37
lag3	1.49	1.40–1.59
lag4	1.24	1.14–1.34
lag5	1.35	1.26–1.45
lag6	1.00	0.90–1.10
lag7	0.72	0.62–0.82
lag01	1.78	1.66–1.90
lag02	2.11	1.98–2.24
lag03	2.54	2.40–2.68
lag04	2.79	2.64–2.94
lag05	3.04	2.89–3.20
lag06	3.14	2.98–3.30
lag07	3.18	3.01–3.35
Acute upper respiratory infections		
lag0	2.18	2.04–2.31
lag1	1.86	1.72–2.00
lag2	1.84	1.70–1.97
lag3	1.94	1.80–2.07
lag4	1.61	1.48–1.74
lag5	1.55	1.42–1.68
lag6	1.18	1.05–1.31
lag7	0.93	0.80–1.06
lag01	2.62	2.47–2.78
lag02	3.09	2.92–3.26
lag03	3.60	3.41–3.78
lag04	3.91	3.71–4.10
lag05	4.16	3.95–4.36
lag06	4.27	4.05–4.48
lag07	4.34	4.11–4.56
Influenza and pneumonia		
lag0	3.04	2.59–3.49
lag1	2.75	2.29–3.21
lag2	2.55	2.10–3.00
lag3	2.36	1.92–2.81
lag4	2.01	1.57–2.44
lag5	1.74	1.30–2.18
lag6	0.97	0.53–1.41
lag7	0.62	0.18–1.06
lag01	3.60	3.09–4.11
lag02	4.09	3.54–4.65
lag03	4.49	3.91–5.08
lag04	4.72	4.10–5.34
lag05	4.84	4.20–5.49
lag06	4.75	4.08–5.42
lag07	4.63	3.93–5.33
Acute lower respiratory infections		
lag0	0.51	0.27–0.75
lag1	0.54	0.29–0.78
lag2	0.85	0.62–1.09
lag3	1.10	0.87–1.34
lag4	0.81	0.57–1.04
lag5	1.36	1.13–1.59
lag6	1.14	0.91–1.38
lag7	0.97	0.74–1.20
lag01	0.68	0.41–0.96
lag02	1.01	0.70–1.31
lag03	1.37	1.04–1.70
lag04	1.52	1.18–1.87
lag05	1.83	1.47–2.20
lag06	1.99	1.62–2.37
lag07	2.10	1.71–2.50

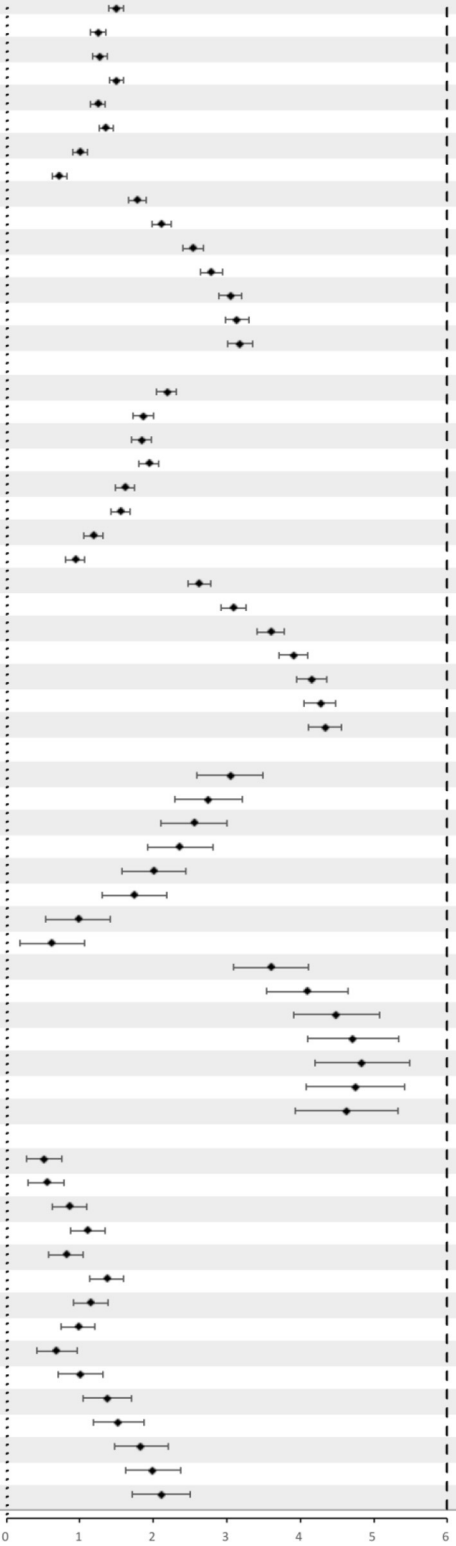


FIGURE 4
Excess risk and 95% *CI*s for the percentage increase in outpatient visits for respiratory diseases per 10 $\mu g/m^3$ increase in the concentration of $PM_{2.5}$. *CI*, confidence interval, the black dot indicates the mean excess risk, the black vertical line corresponds to 95%*CI*, and the vertical dotted line at 0 represents the reference value.

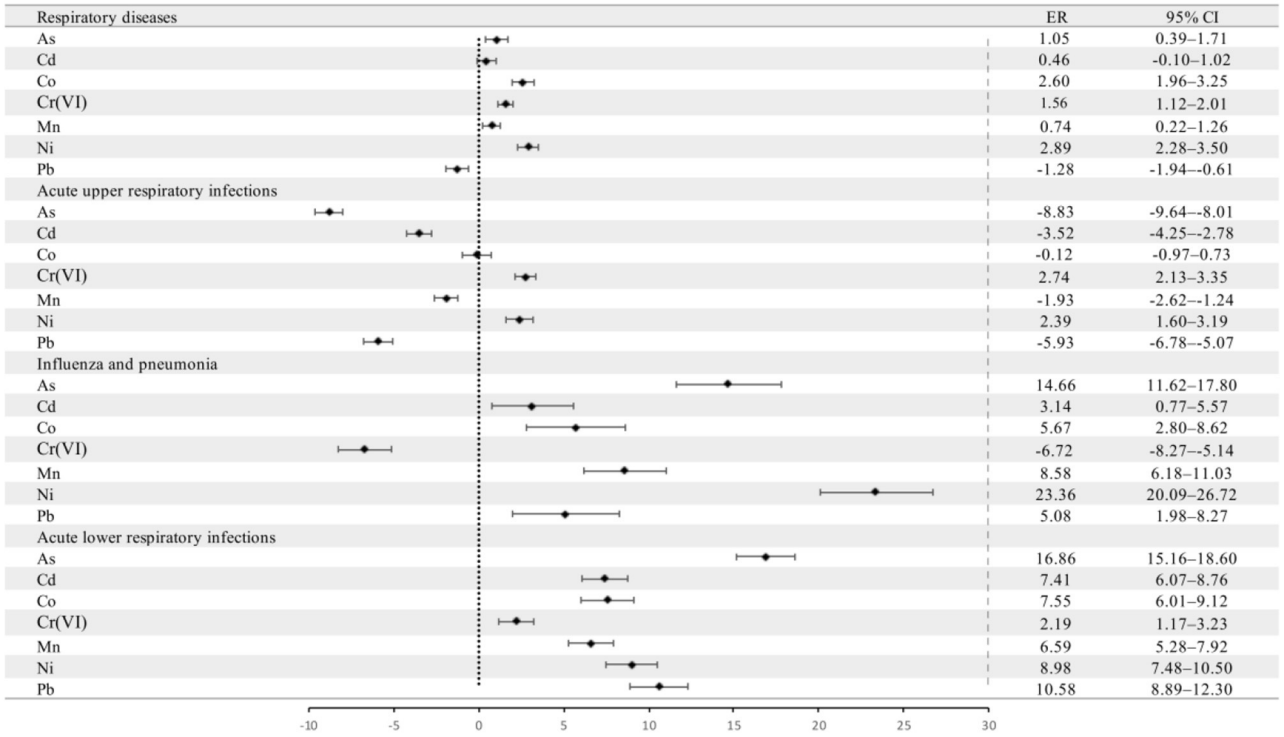


FIGURE 5
Percentage change (*ER* and 95% *CI*) in pediatric outpatient visits for various respiratory diseases per *IQR* increase in the concentrations of As, Cd, Co, Cr(VI), Mn, Ni, and Pb in $PM_{2.5}$ at time lag0. *CI*, confidence interval, black dot indicates the mean excess risk, the black vertical line corresponds to 95%*CI*, and the vertical dotted line at 0 represents the reference value.

respiratory diseases. The exposure–response curve for Ni shows a "∼" trend, indicating that as Ni exposure increased, the total daily pediatric outpatient visits for respiratory diseases initially decreased, then rose, and then finally decreased.

Discussion

This study was conducted in three stages. In the first stage, we investigated the pollution characteristics of $PM_{2.5}$-bound metals through PMF source analysis. In the second stage, we identified metals that pose a potential carcinogenic risk and non-carcinogenic risk for children. In the third stage, we used GAM to analyze the associations between ambient $PM_{2.5}$ and $PM_{2.5}$-bound metals and the risk of daily pediatric outpatient visits for respiratory diseases in Yuexiu district, Guangzhou, China, from 2017 to 2019. During the study period, the daily mean $PM_{2.5}$ concentration was 53.39 $\mu g/m^3$ in Yuexiu district; by comparison, a similar study found a concentration of 50.30 $\mu g/m^3$ in Panyu district, Guangzhou (29). Yuexiu district has poor air quality and severe ambient pollution. The $PM_{2.5}$ concentration is typically higher in spring and winter and lower in summer and autumn; these changes might be related to the decrease in temperature and relative humidity in spring and winter, in addition to firecracker burning that takes place during the Spring Festival in China (30).

The PMF source analysis results showed that the primary sources of $PM_{2.5}$-bound metals were motor vehicles and present in street dust; secondary sources were boiler combustion and ships and heavy oil. These results are consistent with previous studies that found that $PM_{2.5}$-bound metals originated mostly from vehicle exhaust and boiler combustion in Zhuhai (28) and Guangzhou (27). According to the source spectrum database in China (15, 27, 28), Al, Cr, and Fe are mainly found in street dust, Pb and Zn are mainly emitted by motor vehicles, Ni and V are mainly produced by ship fuel and heavy oil combustion, and As is mainly produced by boiler combustion.

The health risk assessment outcomes showed that $PM_{2.5}$-bound As, Cd, Co, Cr(VI), Mn, Ni, and Pb posed a risk to children. Only Mn had a non-carcinogenic risk, a finding that is consistent with a similar study (9). In the carcinogenic risk assessment, As, Cd, Co, Cr(VI), Ni, and Pb were found to be potentially carcinogenic in children, with a *CR* order of Cr(VI) > As > Co > Cd > Ni > Pb. This is consistent with the results of a previous study in Yuexiu district (As > Cd) (9). At present, the IARC classifies As, Cd, and Cr(VI) as "I" (Group 1 carcinogens) and Ni as "IIB" (Group 2B carcinogens); hence, more attention should be given to the health effects of As, Cd, Co, Cr(VI), Mn, Ni, and Pb exposure.

We found significant associations between $PM_{2.5}$ concentrations and pediatric outpatient visits for various respiratory diseases. We also found that the effect of $PM_{2.5}$ concentrations on pediatric outpatient visits for various respiratory diseases was cumulative. Regarding the lag effect, we observed that the $PM_{2.5}$ concentration had a significant positive effect on total pediatric outpatient visits for respiratory diseases at lag3 (1.49%;

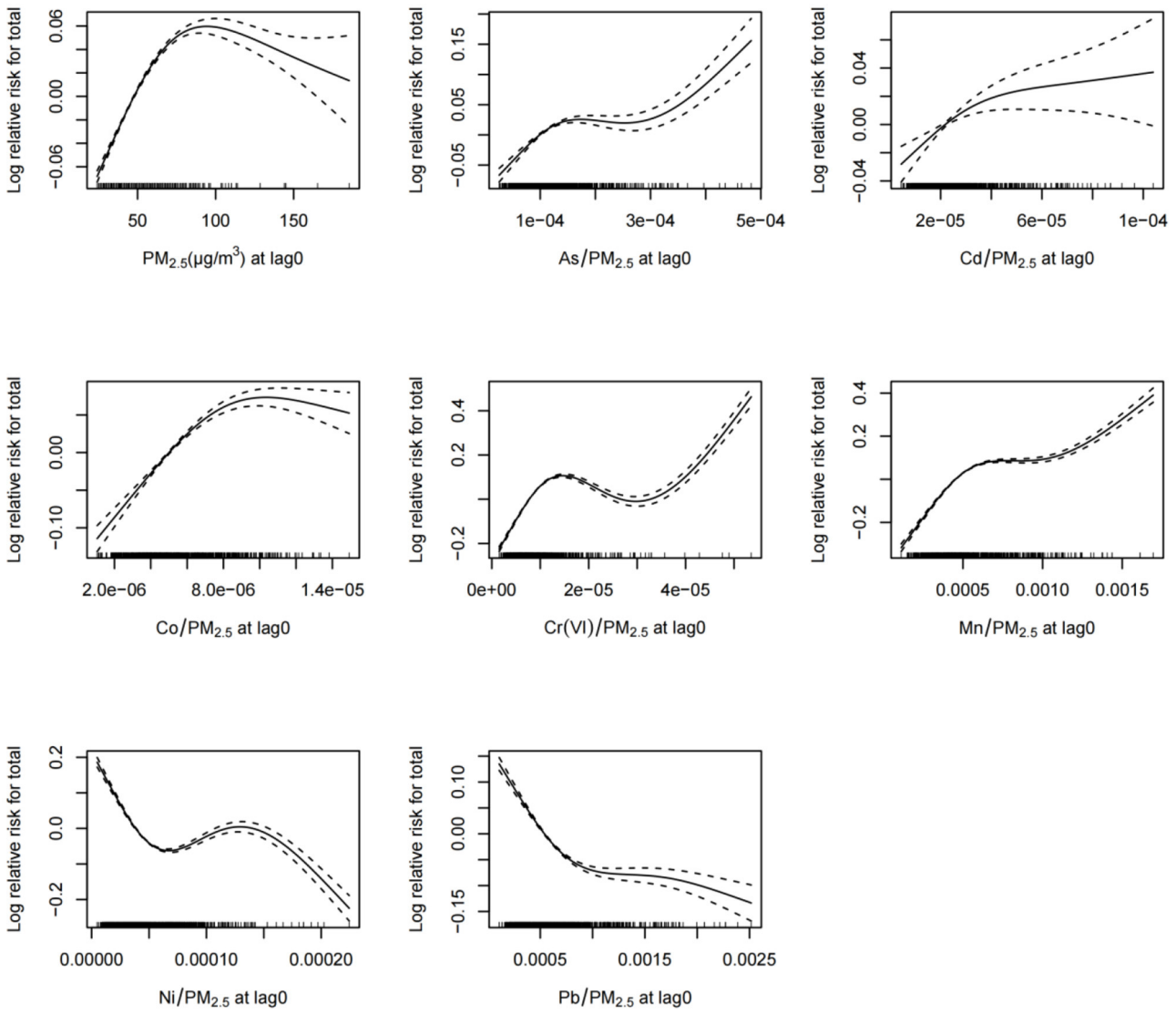


FIGURE 6
Exposure–response relationships between the concentrations of ambient $PM_{2.5}$ and $PM_{2.5}$-bound metals and total pediatric outpatient visits for respiratory diseases at lag0. *CI*, confidence interval, dotted line represents confidence interval; solid line represents log relative risk.

95% *CI*: 1.39–1.59%). This result is similar to that in a previous study in the same city. However, it is a stronger association than has been found in a different city (25, 31). Regarding the cumulative lag effect, the effect of $PM_{2.5}$ concentration on pediatric outpatient visits for various respiratory diseases increased with the number of cumulative lag days. The most significant effect was for lag07, where each 10 $\mu g/m^3$ increase in $PM_{2.5}$ concentration was associated with a 3.18% (95% *CI*: 3.01–3.35%) increase in total daily pediatric outpatient visits for respiratory diseases. This finding is similar to those of previous studies (25, 31).

Regarding pediatric respiratory outpatient visits for AURIs, FLU&PN, and ALRIs, the *ER* was most significant at lag0 for AURIs and FLU&PN. The *ER* increase slowed as cumulative lag days increased. A few studies have examined the relationship between outpatient visits for different respiratory diseases and $PM_{2.5}$, and most have also only found relationships between $PM_{2.5}$ concentrations and AURIs and FLU&PN (25, 32, 33). Therefore, further exploration of the relevant mechanisms and epidemiological evidence is necessary. Our analysis revealed that As, Cd, Co, Cr(VI), Mn, Ni, and Pb concentrations were associated with pediatric outpatient visits for respiratory diseases. In a similar analysis, Valdés et al. (13) found that Ni and Cr concentrations were significantly associated with higher respiratory mortality. In line with our findings, a study in Canada (6) reported associations between respiratory hospitalizations and Sulfur (S) and Cu, and Mn exposure, and a study in Xian (11) found a significant association between Ni exposure and respiratory mortality.

The mechanisms by which inhaled As, Cd, Co, Cr(VI), Mn, Ni, and Pb adversely affect children's respiratory health have been widely studied (34–39). As is known to cause lung cancer by

disrupting immune function and inhibiting immune function-related genes, and as exposure is also associated with non-malignant respiratory diseases (e.g., acute respiratory infections and symptoms of decreased lung function) (36). Low levels of Cd exposure can lead to lung function alterations that can cause pulmonary fibrosis, emphysema, and lung tumors, and Cd toxicity increases oxidative stress that depletes protein-bound glutathione and sulfhydryl groups. Although Cd can enter the body through inhalation, the current study shows that high Cd exposure mainly arises through food intake and smoking. This suggests that particular attention should be given to preventing children being exposed to secondhand smoke (35). Chronic inhalation of Co compounds can lead to respiratory tumors, and Co exposure often leads to interstitial lung disease (34). Inhalation is the primary route of exposure to Cr and can lead to lung cancer, chromosomal damage, asthma, cough, acute bronchitis, and pulmonary edema (35). Exposure to low concentrations of Mn can result in adverse health effects on respiratory organs; Mn is also an immunotoxin and is cytotoxic to lung macrophages, thus making children susceptible to small airway bronchial injury (37). Ni exposure can increase respiratory disease morbidity and mortality, and it promotes the expression of inflammatory factors, contributing to airway inflammation (39). Pb exposure downregulates interleukin-13 expression and can increase the risk of asthma-related immunomodulatory abnormalities in preschool children (38).

Children are a population group that is sensitive to air pollution because they are in a period of growth and development. In addition, children have a relatively rapid respiratory rate and are more vulnerable to respiratory diseases such as asthma, cough, and respiratory inflammation. Studies have shown that respiratory diseases have become the most frequent cause of childhood outpatient visits for illness, and these diseases are a primary medical problem adversely affecting the health and growth of children (40).

Thus, the government should strengthen the supervision and control of three sources of pollution—waste incineration, metal smelting, and traffic—in addition to improving purification measures and rationalizing the separation of traffic flow to relieve traffic pressure. In addition, parents can lower the risk of respiratory diseases in their children by reducing their children's exposure to $PM_{2.5}$, As, Cd, Co, Cr(VI), Mn, Ni, and Pb. This can be achieved by parents not bringing their children to certain areas and not exposing them to secondhand smoke. Finally, parents should also ensure that their children have adequate physical exercise to strengthen their immunity.

This study has several strengths. First, to the extent of our knowledge, it is the first study to use a comprehensive and systematic approach to quantify the associations between exposure to $PM_{2.5}$-bound metals and pediatric outpatient visits for respiratory disease. Second, it examined the total pediatric outpatient visits for respiratory diseases, AURIs, FLU&PN, and ALRIs, while previous studies have investigated only some of these conditions. Third, to obtain reliable results, it used a quasi-Poisson GAM with adjustments for lag effects, DOW, temperature, and humidity.

The limitations of our study should also be noted. First, we included clinical data from only one hospital, which cannot accurately represent all of the pediatric outpatient visits in Yuexiu district. Second, the air pollution exposure concentration we used was obtained from regional air pollution monitoring data, and thus there may have been some errors in the exposure measurement. Third, a certain amount of error was introduced by compensating for the lack of daily data on metal concentrations by calculating the monthly mean metal concentration to daily $PM_{2.5}$ concentration ratio. Few studies have been conducted in this way; therefore, precise comparisons with previous studies were not possible. Nevertheless, relative changes were examined.

Conclusion

Our findings suggest that $PM_{2.5}$ and $PM_{2.5}$-bound As, Cd, Co, Cr(VI), Ni, and Pb pose potential health risks to children in Guangzhou, China, and can contribute to respiratory disease in this population. Government departments can lower children's risk of respiratory diseases by introducing measures that decrease the production of these pollutants by motor vehicles and reduce the amount of street dust.

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

QY designed the study and supervised the research, including funding, text review, and overall quality assurance and control. YZ and SC were helped with the formulation of research methods, software analysis and interpretation of the results, and wrote the original draft of the text. YC and JL were helped with the investigation and review of the data and assisted in the preparation of the original draft of the text. BX and TS were assisted in the implementation of research, data management and investigation, and supervision. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Natural Science Foundation of China (Grant Number: 81773385 to QY), Natural Science Foundation of Guangdong Province (Grant Numbers: 2019A1515011298 and 2022A1515010727 to QY), and Key scientific research projects of Guangdong Provincial Department of Education (Grant Number: 2022ZDZX2047 to QY).

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. Cancer IAfRo. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans.* Lyon: International Agency for Research on Cancer (2013, 2016). Available online at: https://monographs.iarc.who.int/list-of-classifications

2. Bu X, Xie Z, Liu J, Wei L, Wang X, Chen M, et al. Global PM2.5-attributable health burden from 1990 to 2017: estimates from the Global Burden of disease study 2017. *Environ Res.* (2021) 197:111123. doi: 10.1016/j.envres.2021.111123

3. Zhang Q, Meng X, Shi S, Kan L, Chen R, Kan H. Overview of particulate air pollution and human health in China: evidence, challenges, and opportunities. *Innovation (Camb).* (2022) 3:100312. doi: 10.1016/j.xinn.2022.100312

4. Zhang L, Jin X, Johnson AC, Giesy JP. Hazard posed by metals and As in PM25 in air of five megacities in the Beijing-Tianjin-Hebei region of China during APEC. *Environ Sci Pollut Res Int.* (2016) 23:17603–12. doi: 10.1007/s11356-016-6863-2

5. Zhao Y, Wang S, Lang L, Huang C, Ma W, Lin H. Ambient fine and coarse particulate matter pollution and respiratory morbidity in Dongguan, China. *Environ Pollut.* (2017) 222:126–31. doi: 10.1016/j.envpol.2016.12.070

6. Korsiak J, Lavigne E, You H, Pollitt K, Kulka R, Hatzopoulou M, et al. Air pollution and pediatric respiratory hospitalizations: effect modification by particle constituents and oxidative potential. *Am J Respir Crit Care Med.* (2022) 206:1370–8. doi: 10.1164/rccm.202205-0896OC

7. Chang Q, Liu S, Chen Z, Zu B, Zhang H. Association between air pollutants and outpatient and emergency hospital visits for childhood asthma in Shenyang city of China. *Int J Biometeorol.* (2020) 64:1539–48. doi: 10.1007/s00484-020-01934-9

8. Wang YS, Chang LC, Chang FJ. Explore regional PM25 features and compositions causing health effects in Taiwan. *Environ Manage.* (2021) 67:176–91. doi: 10.1007/s00267-020-01391-5

9. Chen R, Jia B, Tian Y, Feng Y. Source-specific health risk assessment of PM(25)-bound heavy metals based on high time-resolved measurement in a Chinese megacity: insights into seasonal and diurnal variations. *Ecotoxicol Environ Saf.* (2021) 216:112167. doi: 10.1016/j.ecoenv.2021.112167

10. Achilleos S, Kioumourtzoglou MA, Wu CD, Schwartz JD, Koutrakis P, Papatheodorou SI. Acute effects of fine particulate matter constituents on mortality: a systematic review and meta-regression analysis. *Environ Int.* (2017) 109:89–100. doi: 10.1016/j.envint.2017.09.010

11. Cao J, Xu H, Xu Q, Chen B, Kan H. Fine particulate matter constituents and cardiopulmonary mortality in a heavily polluted Chinese city. *Environ Health Perspect.* (2012) 120:373–8. doi: 10.1289/ehp.1103671

12. Hwang SH, Lee JY Yi SM, Kim H. Associations of particulate matter and its components with emergency room visits for cardiovascular and respiratory diseases. *PLoS ONE.* (2017) 12:e0183224. doi: 10.1371/journal.pone.0183224

13. Valdés A, Zanobetti A, Halonen JI, Cifuentes L, Morata D, Schwartz J. Elemental concentrations of ambient particles and cause specific mortality in Santiago, Chile: a time series study. *Environ Health.* (2012) 11:82. doi: 10.1186/1476-069X-11-82

14. US EPA. *Positive Matrix Factorization Model for Environmental Data Analyses.* (2022). Available online at: https://www.epa.gov/air-research/positive-matrix-factorization-model-environmental-data-analyses

15. Ho WY, Tseng KH, Liou ML, Chan CC, Wang CH. Application of positive matrix factorization in the identification of the sources of $PM_{2.5}$ in Taipei city. *Int J Environ Res Public Health.* (2018). 15:1305. doi: 10.3390/ijerph15071305

16. US EPA. *Risk Assessment.* (2022). Available online at: https://www.epa.gov/risk

17. China NHCotPsRo. *Technical Specifications for Health Risk Assessment of Ambient Air Pollution. WS/T 666-2019.* Beijing: Standards Press of China (2019).

18. Kan H, London SJ, Chen G, Zhang Y, Song G, Zhao N, et al. Differentiating the effects of fine and coarse particles on daily mortality in Shanghai, China. *Environ Int.* (2007) 33:376–84. doi: 10.1016/j.envint.2006.12.001

19. Yang M, Zeng HX, Wang XF, Hakkarainen H, Leskinen A, Komppula M, et al. Sources, chemical components, and toxicological responses of size segregated urban air PM samples in high air pollution season in Guangzhou, China. *Sci Total Environ.* (2023) 865:161092. doi: 10.1016/j.scitotenv.2022.161092

20. Friedrich MJ. UNICEF reports on the impact of air pollution on children. *JAMA.* (2017) 317:246. doi: 10.1001/jama.2016.19034

21. Saadeh R, Klaunig J. Child's development and respiratory system toxicity. *J Environ Anal Toxicol.* (2014) 4:5. doi: 10.4172/2161-0525.1000233

22. Brown SG, Eberly S, Paatero P, Norris GA. Methods for estimating uncertainty in PMF solutions: examples with ambient air and water quality data and guidance on reporting PMF results. *Sci Total Environ.* (2015). 518–519:626–35. doi: 10.1016/j.scitotenv.2015.01.022

23. Xiao SH, Cai MJ, Li X, Huang Z, Wang J, Zhu QD, et al. [Characterization and health risk assessment of heavy metals in $PM_{2.5}$ in Xiamen port]. *Huan Jing Ke Xue.* (2022) 43:3404–15 (Chinese). doi: 10.13227/j.hjkx.202110163

24. Mostofsky E, Schwartz J, Coull BA, Koutrakis P, Mittleman MA. Modeling the association between particle constituents of air pollution and health outcomes. *Am J Epidemiol.* (2012) 176:317–26. doi: 10.1093/aje/kws018

25. Li Y, Li C, Liu J, Meng C, Xu C, Liu Z, et al. An association between PM(25) and pediatric respiratory outpatient visits in four Chinese cities. *Chemosphere.* (2021) 280:130843. doi: 10.1016/j.chemosphere.2021.130843

26. Brown RJC, Van Aswegen S, Webb WR, Goddard SL. UK concentrations of chromium and chromium (VI), measured as water soluble chromium, in PM_{10}. *Atmos Environ.* (2014) 99:385–91. doi: 10.1016/j.atmosenv.2014.10.008

27. Wang X, Xinhui BI, Sheng G, Jiamo FU. Chemical composition and sources of PM_{10} and $PM_{2.5}$ aerosols in Guangzhou, China. *Environ Monit Assess.* (2006). 119:425–30. doi: 10.1007/s10661-005-9034-3

28. Liang Z, Zhao X, Chen J, Gao L, Zhu A, Wang Z, et al. Seasonal characteristics of chemical compositions and sources identification of PM2.5 in Zhuhai, China. *Environ Geochem Health.* (2018) 41:715–28. doi: 10.1007/s10653-018-0164-2

29. Min L, Yan-hong G, Ling-chuan G, Xiao-wen L, Jing Y. Pollution characteristics of atmospheric PM(2.5) metal components in two urban areas of Guangzhou. *J Environ Occup.* (2016). 33. doi: 10.16241/j.cnki.1001-5914.2016.05.011

30. Zhao JP, Xu Y, Zhang FW, Chen JS. Atmospheric pollution characteristic during fireworks burning time in spring festival in Quanzhou suburb. *Huan Jing Ke Xue.* (2011) 32:1224–30. doi: 10.13227/j.hjkx.2011.05.020

31. Liu J, Li Y, Li J, Liu Y, Tao N, Song W, et al. Association between ambient PM(25) and children's hospital admissions for respiratory diseases in Jinan, China. *Environ Sci Pollut Res Int.* (2019) 26:24112–20. doi: 10.1007/s11356-019-05644-7

32. Xiao D, Guo W, Xu D, Chen J, Liang Z, Zhang X. Three exposure metrics for fine particulate matter associated with outpatient visits for acute lower respiratory infection among children in Guangzhou, China. *Front Public Health.* (2022) 10:876496. doi: 10.3389/fpubh.2022.876496

33. Liang Z, Meng Q, Yang Q, Chen N, You C. Size-specific particulate matter associated with acute lower respiratory infection outpatient visits in children: a counterfactual analysis in Guangzhou, China. *Front Public Health.* (2021) 9:789542. doi: 10.3389/fpubh.2021.789542

34. Adams TN, Butt YM, Batra K, Glazer CS. Cobalt related interstitial lung disease. *Respir Med.* (2017) 129:91–7. doi: 10.1016/j.rmed.2017.06.008

35. Osman MA, Yang F, Massey IY. Exposure routes and health effects of heavy metals on children. *BioMetals.* (2019) 32:563–73. doi: 10.1007/s10534-019-00193-5

36. Sanchez TR, Perzanowski M, Graziano JH. Inorganic arsenic and respiratory health, from early life exposure to sex-specific effects: a systematic review. *Environ Res.* (2016) 147:537–55. doi: 10.1016/j.envres.2016.02.009

37. McRae N, Gennings C, Rivera Rivera N, Tamayo-Ortiz M, Pantic I, Amarasiriwardena C, et al. Association between prenatal metal exposure and adverse respiratory symptoms in childhood. *Environ Res.* (2022) 205:112448. doi: 10.1016/j.envres.2021.112448

38. Zeng Z, Xu X, Zhu Y, Wang Q, Zhang Y, Huo X. Pb and Cd exposure linked with Il-10 and Il-13 gene polymorphisms in asthma risk relevant immunomodulation in children. *Chemosphere.* (2022) 294:133656. doi: 10.1016/j.chemosphere.2022.133656

39. Gray DL, Wallace LA, Brinkman MC, Buehler SS, La Londe C. Respiratory and cardiovascular effects of metals in ambient particulate matter: a critical review. *Rev Environ Contam Toxicol.* (2015) 234:135–203. doi: 10.1007/978-3-319-10638-0_3

40. Saravanos GL, Ramos I, Britton PN, Wood NJ. Respiratory syncytial virus subtype circulation and associated disease severity at an Australian paediatric referral hospital, 2014–2018. *J Paediatr Child Health.* (2021) 57:1190–5. doi: 10.1111/jpc.15419

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 09 March 2023
DOI 10.3389/fpubh.2023.1046729

OPEN ACCESS

EDITED BY
Nan Luo,
National University of Singapore, Singapore

REVIEWED BY
Anam Mehmood,
Harbin Medical University, China
Ayan Paul,
Northeastern University, United States
Pei Wang,
Fudan University, China

*CORRESPONDENCE
Min Ouyang
✉ 21726798@QQ.com

†These authors share first authorship

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

RECEIVED 17 September 2022
ACCEPTED 15 February 2023
PUBLISHED 09 March 2023

CITATION
Yao Z, Pang L, Xie J, Shi S and Ouyang M (2023) The relationship between social anxiety and self-injury of junior high school students: Mediation by intolerance of uncertainty and moderation by self-esteem.
Front. Public Health 11:1046729.
doi: 10.3389/fpubh.2023.1046729

The relationship between social anxiety and self-injury of junior high school students: Mediation by intolerance of uncertainty and moderation by self-esteem

Zhendong Yao[1†], Lu Pang[2], Jin Xie[3], Seqin Shi[4] and Min Ouyang[5,6*†]

[1]Normal College, Hunan University of Arts and Science, Changde, China, [2]School of Preschool Education, Hunan College for Preschool Education, Changde, China, [3]Mental Health Service Center, Huanghuai University, Zhumadian, China, [4]No. 1 Middle School of Loudi, Loudi, China, [5]Students Affairs Department, Xiangnan University, Chenzhou, China, [6]College of Educational Science and Law, Xiangnan University, Chenzhou, China

Objective: The problem of adolescents' self-injury has gradually attracted social attention, however, a lack of research exists on the internal mechanism between social anxiety and self-injury. This study explored the relationship between social anxiety and self-injury in Chinese junior high school students.

Method: An adolescent self-injury questionnaire, social anxiety scale, intolerance of uncertainty questionnaire and self-injury questionnaire were used to survey 614 junior high school students.

Results: The results showed that: (1) social anxiety had a significant positive predictive effect on self-injury; (2) intolerance of uncertainty had a significant mediating effect between social anxiety and self-injury; and (3) self-esteem had a significant moderating effect on the mediating effect of intolerance of uncertainty.

Conclusion: The study suggested that social anxiety in junior high school students has an impact on self-injury through mediation of intolerance of uncertainty and modulation of self-esteem.

KEYWORDS

junior high school student, self-injury, social anxiety, intolerance of uncertainty, self-esteem

Introduction

Adolescence, often referred to as the "storm period," includes the junior high school stage, an important period for individual physical and mental development. The psychological characteristics of junior high school students have two aspects: autonomy and dependence. They not only want to be independent, but also want to be helped by others (1). The way teenagers express their emotions changes transitions from externalism to concealment, and sometimes they hide their true emotions without expressing themselves (2). For this reason, many junior high school students often lose their psychological balance, causing various problems arise, such as self-injury, to arise.

The relationship between social anxiety and self-injury

Non-suicidal self-injury refers to the behavior of intentionally and repeatedly hurting one's body without suicidal intention, which is not accepted by society and will not lead to death (3). In China, teenagers comprise a high-risk group for self-injury (4). A significant public health problem, the incidence of self-injury among adolescents is as high as 30% (5). Although it has a certain universality, it is often a hidden problem (6). These venting behaviors do not always occur in public. Some behaviors vent in unreasonable and potentially negative ways when the individual is alone, prompting negative coping styles that may increase the risk of non-suicidal self-injury (7). What is the cause of self-injury in individuals? According to the system model, self-injury is a symptom manifestation of environmental dysfunction. Environmental influences may unintentionally enhance self-injury because self-injurious can transfer dysfunction. Additionally, individuals can use self-injurious behavior to obtain external attention (8). The core feature of self-injury is an emotional management disorder (9). In summary, social anxiety may be the cause of self-injury.

Social anxiety refers to the psychological phenomenon of nervousness, fear, and embarrassment in the process of normal interpersonal interaction with others (10). Individual social anxiety begins in adolescence (11), because compared with the previous primary school period, the pupils' compliance with parental arrangements and patience with teachers' instruction begin to gradually shift to the pursuit of "self." Studies have shown that self-awareness has some effect on social anxiety (12). Family and school are the main environments where junior high school students communicate placing the focus on the family members, classmates, and teachers. The junior high school period is a time when students' self-consciousness takes a great leap (13). Studies have proven the existence of a relationship between anxiety and self-injury (14). Further, research has indicated that specific extrinsic actions are necessary for people with intrinsic anxiety, because, they can express their feelings through extrinsic self-injury as an outlet (15). Social anxiety is a type of anxiety that corresponds with the above research results. It has been suggested that individuals with a higher degree of social anxiety can interpret facial expressions as negative (16). Moreover, the emotional instability of people with anxiety disorders may lead to non-suicidal self-injury (17). Although previous studies have demonstrated the influence of social anxiety on self-injury, few studies have explored the underlying mechanism. To prevent self-injury behavior and provide effective intervention measures for the self-injury of junior high school students, we examined a moderated mediation model to uncover the possible mechanisms underlying this relationship.

The mediating role of intolerance of uncertainty

Prior research indicates that individuals with social anxiety have negative processing bias toward positive stimulation (18), and tend to often be emotionally unstable, impatient and impulsive (19). Such individuals experience more discomfort, which leads to poor or even little to no tolerance of uncertain events. Intolerance of uncertainty refers to the aversive response of an individual due to their inability to perceive significant or sufficient information, sustained by uncertainty perception (20). Intolerance of uncertainty is especially closely related to social anxiety (21). The inability to tolerate uncertainty may be an important component of anxiety (22) and may be one of the key features of anxiety disorder referrals (23). Therefore, it is logical reason that social anxiety can be positively related to intolerance of uncertainty.

Moreover, intolerance of uncertainty not only affects one's physical and mental health, but also one's decision-making and problem solving ability. Individuals with high intolerance of uncertainty are more biased toward the occurrence of potentially adverse events (24). In daily diagnoses, intolerance of uncertainty can be a powerful indicator in examining the rehabilitation of anxiety disorders (25). Because of the close relationship between negative emotions and self-injury, negative emotions are often used as the primary indicator to assess self-injury functioning (26). Junior high school students are faced with increasing learning tasks and complicated interpersonal communication. During this period, when social anxiety is common, students often do not know how to relieve their inner discomfort, prompting them to inflict non-suicidal self-injury to rescue themselves. Based on the above, this paper put forward the assumption:

> **Hypothesis 1.** Social anxiety is positively related to intolerance of uncertainty, which in turn would be positively related to self-injury.

The moderating role of self-esteem

Although social anxiety may increase the risk of self-injury through the mediating role of intolerance of uncertainty, not all individuals with high social anxiety experience higher levels of intolerance of uncertainty and self-injury. Therefore, it is necessary to explore the potential moderating variables between social anxiety and self-injury. The present study tests the hypothesis that the links between social anxiety and self-injury are moderated by self-esteem.

According to the hypothesis regarding the relief of anxiety, self-esteem is a natural buffer to cope with anxiety. Social maladjustment causes teenagers to lack effective countermeasures when facing uncertain external events (27), making young people unable to tolerate increased levels of uncertainty. It is believed that the flexible space provided by the self-regulation mechanism of self-esteem can help people alleviate social anxiety (28). Self-esteem is an important variable among the indicators of psychosocial adjustment of individuals. It refers to a person's cognition, attitude and view of one's own personal ability or personal value (29), and plays an important role in monitoring individual interpersonal relationships (30). Low self-esteem may be a psychological variable related to self-injury (31). The higher the level of self-esteem, the lower the risk of anxiety and the less possibility of anxiety related behaviors (32). For instance, left-behind children with higher levels of self-esteem experience a weaker negative impact on self-injury by stressful life events than left-behind children with lower levels of self-esteem (33). According to the above literature, we put forward the following assumptions:

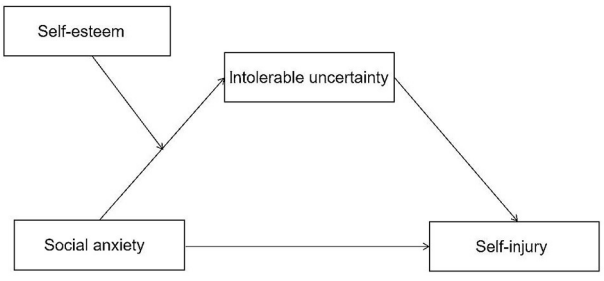


FIGURE 1
The proposed moderated mediation model.

Hypothesis 2. Self-esteem moderates the relationships between social anxiety and self-injury.

The present study

The objectives of this study were as follows: (a) to examine the mediating role of intolerance of uncertainty between social anxiety and self-injury, and (b) to examine whether the relational pathway between social anxiety and self-injury is moderated by self-esteem. This study conducted a moderated mediation analysis to test these hypotheses. Therefore, Figure 1 presents a moderated mediation model.

Methods

Participants

Through convenient sampling, four junior high schools of Hunan Province, China were investigated. A total of 700 questionnaires were distributed and 614 valid questionnaires were returned, with an effective rate of 87.71%. Upon analysis of the sample population, 263 (42.80%) were males and 351 (57.20%) were females. The mean age of the participants was 13.91 years (SD = 1.01), with a range of 13–15 years. Regarding grade level, 205 (33.40%) were in the first grade, 223 (36.30%) in the second grade, and 223 (36.30%) in the third grade.

Measures

Social anxiety scale

The Social Anxiety Scale (SAS) is a subscale derived from the self-awareness scale compiled by Fenigstein et al. (34). The Social Anxiety Scale is one of the subscales. The Chinese version of the Social Anxiety Scale was selected from the Mental Health Rating Scale Manual edited by Wang et al. (35), which has a total of six question items, such as "I often speak and feel anxious in front of the population." In this study, a 5-point Likert scale was used (1 = very inconsistent, 2 = somewhat inconsistent, 3 = unclear, 4 = somewhat consistent, and 5 = extremely consistent), indicating the degree of social anxiety. In this study, Cronbach's α of the Social Anxiety Scale was 0.74.

Adolescent self-injury questionnaire

This paper selects from Feng (36), which was based upon the questionnaire developed by Zheng and the self-injurious behavior scale developed by Graze, a revised localized and standardized scale, which has good psychometric indicators. The questionnaire consists of 18 items, such as "deliberately grasping oneself violently and achieving the degree of scarring or bleeding," and 1 open-ended item (In this study, the number of people who filled in this question was less, so only 18 closed-ended items were counted). Among them, the options for the number of self-injury occurrences are 0, 1, 2 – 4, and more than 5; the options for the degree of self-injury to the body are: none, mild, moderate, severe, and very severe. In this study, the Cronbach's alpha of the self-injury questionnaire was 0.95.

Intolerance of uncertainty scale

This item was selected from the Chinese version of the Intolerable Uncertainty Scale (simplified version) revised by Lijuan et al. (37). This questionnaire consists of 12 question items, such as "When I am not sure, I cannot do well," which are divided into three dimensions, anticipatory behavior, inhibitory behavior, and anticipatory emotion. Responses were given on a 5 point Likert scale: 1 = completely inconsistent, 2 = somewhat inconsistent, 3 = basically in conformity, 4 = very in conformity and 5 = completely in conformity. In this study, the Cronbach's alpha of the self-injury questionnaire was 0.80.

Self-esteem scale

The Self-Esteem Scale (SES) was developed by Rosenberg, and is used to measure teenagers' overall feelings about self-worth and self-acceptance (38). The scale includes 10 items in the scale, such as "generally speaking, I am satisfied with myself." A 4 point Likert scale was used: very good agreement = 4, conformity = 3, non-conformity = 2, and extreme non-conformity = 1. In this study, the Cronbach's α of self-injury questionnaire was 0.84.

Data analysis

In this study, IBM SPSS Statistics 26.0 was used to analyze the collected data, Hayes' PROCESS macro program (2012) was used to analyze the mediating effect, and bootstrap method was used to test the significance of regression coefficient (39). The sample distribution was reconstructed by random sampling with replacement. In this study, a total of 5,000 samples were constructed, each with a sample size of 614. From this, the standard errors and confidence intervals of the parameter estimation were calculated. If the confidence interval does not include zero, it means the results are significant.

Results

Preliminary statistics

There may be common method bias in the data of the self-assessment questionnaire. In this study, the Harman's single-factor

test was used to test the effect of procedural control (40). The four variables of social anxiety, intolerable of uncertainty, self-esteem, and self-injury were combined and texted using IBM SPSS 26.0. The results indicated 11 factors with eigenvalues >1.0, and the first common factor is 14.39% of the variance, which was far less than the critical standard of 40%. Therefore, common method bias was unlikely in this study.

The results of the correlation analysis of social anxiety, intolerance of uncertainty, self-esteem and self-injury are presented in Table 1. There were significant positive correlations between all variables, which were consistent with theory and previous research. Social anxiety was positively correlated with intolerance of uncertainty ($r = 0.50, p < 0.01$) and self-injury ($r = 0.18, p < 0.01$). Additionally, intolerance of uncertainty was positively correlated with self-injury ($r = 0.18$, $p < 0.01$). Moreover, self-esteem was related inversely related to the social anxiety ($r = -0.30, p < 0.01$) and intolerance of uncertainty ($r = -0.30, p < 0.01$) and self-injury ($r = -0.15, p < 0.01$).

Testing for mediation effect

Considering the advantages of applying the Bootstrap method in testing the mediating effect (41), the PROCESS PROCEDURE for SPSS developed by Hayes was used to test the mediating effect (model 4). The results showed that the mediating effect value is $F_{(1,614)} = 50.57$, $P < 0.001$, $R^2 = 0.25$, the mediating effect value is 0.061, and the confidence interval of 95% is [0.020, 0.11], indicating that the mediating effect is significant, accounting for 57.49% of the total effect and 61.00% of the direct effect. Therefore, Hypothesis 1 was supported.

TABLE 1 Descriptive statistics and correlations between study variables.

Variables	*M*	*SD*	1	2	3	4
1. Social anxiety	18.50	4.57	1			
2. Intolerance of uncertainty	32.79	7.31	0.50**	1		
3. Self-esteem	28.00	4.87	−0.30**	−0.30**	1	
4.Self-injury	6.58	14.31	0.18**	0.18**	−0.15**	1

**$P < 0.01$.

Test of the moderated mediation model

According to the research theory of Muller et al. (42) and others, this study used three regression equation tests to check the mediation models with regulation one by one. Equation 1 examined the moderating effect of self-esteem on social anxiety and self-injury of junior high school students. Equation 2 examined the moderating effect of self-esteem on the relationship between social anxiety and intolerance of uncertainty. Equation 3 examined the moderating effect of self-esteem on the relationship between intolerance of uncertainty and self-injury, and the moderating effect of social anxiety and self-injury participation. Among the three regression equations, any significant moderating effect can show that self-esteem plays a moderating role in the relationship between social anxiety and self-injury of junior high school students.

In this study, independent variables, mediating variables, and moderating variables were all standardized, and gender, grade, and age were controlled by means of hierarchical regression (see results in Table 2). In Equation 1, self-esteem was significant for the relationship between social anxiety and self-injury in junior high school students ($\beta = -0.082$, $P = 0.023$). In Equation 2 self-esteem was significant for the relationship between social anxiety and intolerance of uncertainty in junior high school students ($\beta = -0.63$, $P = 0.047$). In Equation 3, self-esteem modulated the effects of social anxiety and self-injury in junior high school students and the participation effects of social anxiety and self-injury in junior high school students ($\beta = -0.048$, $P = 0.16$). In summary, self-esteem has a moderating effect on the social anxiety of intolerance of uncertainty. Therefore, Hypothesis 2 was supported.

In order to analyze the moderating effect of self-esteem on the mediating effect of intolerance of uncertainty more clearly, a simple slope test was conducted on the interaction, as demonstrated in

TABLE 2 The moderating effect of self-esteem qon the mediation of intolerance of uncertainty.

	Model 1 (Self-injury)		Model 2 (Intolerance of uncertainty)		Model 3 (Self-injury)	
	β	*t*	*β*	*t*	*β*	*t*
Gender	0.023	0.29	0.038	0.53	0.024	0.30
Age	0.14	2.26*	0.096	1.74	0.13	2.02*
Grade	−0.30	−3.86**	−0.080	−1.16	−0.29	−3.73**
Social anxiety	0.14	3.41**	0.45	12.43**	0.090	1.95
Self-esteem	−0.088	−2.10	−0.16	−4.25**	−0.070	−1.63
Social anxiety × Self-esteem	−0.082	−2.28*	−0.063	−2.00*	−0.047	−1.40
Intolerance of uncertainty					0.10	2.17*
Intolerance of uncertainty × Self-esteem					−0.048	−1.41

*$P < 0.05$ and **$P < 0.01$.

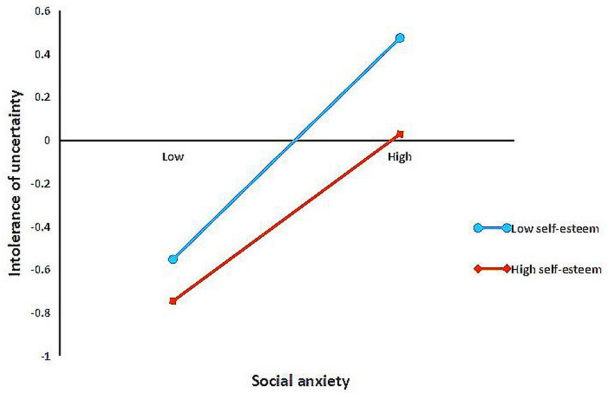


FIGURE 2
Association between social anxiety and self-injury at higher and lower levels of self-esteem.

Figure 2. The specific method was as follows: based upon their score of self-esteem, the subjects are divided into high (Z $\geq$ 1 SD and low (Z $\leq$ - 1 SD) groups, to investigate the intolerance of uncertainty on different levels of self-esteem among junior middle school students, and to draw a simple effect analysis diagram. The results demonstrate that when the self-esteem level of junior high school students is at a low level, social anxiety has a significant negative predictive effect on intolerance of uncertainty ($\beta = 0.52$, $p < 0.00$). When the self-esteem level of junior high school students is high, the influence of social anxiety on intolerable of uncertainty is reduced ($\beta = 0.39$, $p < 0.00$).

Discussion

Social anxiety and self-injury of junior high school students

The phenomenon of self-injury is very common among adolescents, while it may also be in an ascending state (43). According to Mathew et al. (44) research, self-injured individuals scored higher on anxiety assessment. According to Benjet et al. (45), the occurrence of anxiety is associated with that of self-injury. That, anxiety may become a negative emotion that causes self-injury (46). What is the reason for the social anxiety of junior high school students? Prior research has indicated that social anxiety derives from self-awareness. If individuals are unaware, it is difficult for them to generate anxiety. Feelings of anxiety may lead to a decrease in the individual's motivation in social competition and the emergence of avoidance strategies (47).

Why do anxious individuals tend to self-injure? According to Hasking et al. (48) research, the emotional instability of patients with anxiety disorders may lead to self-injury and the client's impulsivity is very strong, leading to non-suicidal self-injury. For junior high school students, the emergence of social anxiety is related to their developmental stage—e.g., the rapid development of the adolescent physiology of boys and girls leads to dramatic changes in their respective psychology, adding complex factors to their processing of interpersonal relationships. In addition, due to the increase of academic pressure among junior high school students, they must experience the process of navigating self-study, personal life, and getting along with others. Coordinating the relationships with multiple people can be a challenge, thus, dysregulated symptoms can emerge among various social contacts.

In this study, junior high school students' social anxiety has a significant positive predictive effect on self-injury. The results of this study are consistent with the above theory. It is evident that individuals prone to self-injury exhibit higher scores on anxiety items in self-assessment; and their anxiety experience is clear. Thus, junior high school students are more prone to self-injury due to the painful experience of social anxiety. Just as previous studies have indicated, interpersonal relationships are an important driving force for the development of adolescent self-injury (49). Regarding individuals within in a certain social structure who have self-injured, anxiety will cause them to vent through self-injury (50).

The mediating role of intolerance of uncertainty

Upon further examination, this study found that the mediating effect of intolerance of uncertainty between social anxiety and self-injury of junior high school students is significant. Per prior research, social anxiety is related to intolerance of uncertainty (51). Moreover, there is a significant positive correlation between social anxiety and intolerance of uncertainty (52). The results of this study are consistent with those of previous studies.

Intolerance of uncertainty reflects an individual's perception of uncertainty, and, is not related to the real situation or the threat (53). Therefore, intolerance of uncertainty reflects the individual's inner psychological security. Individuals with high intolerance of uncertainty tend to cognitively process uncertain or vague situations in a negative way, so they are more likely to produce negative emotions (54). These difficulties in regulating emotions may lead them to hurt themselves (55). Previous studies on adolescent self-injury also corresponded to the above research. Adolescents experencing non-suicidal self-injury have difficulty with emotional regulation (56). Based on the above logic, the treatment of anxiety and intolerance of uncertainty may help to resolve issues with social problems (57), such as self-injury.

The moderating role of self-esteem

The results showed that intolerance of uncertainty was a mediating variable between social anxiety and self-injury in junior high school students, and that self-esteem was a moderating variable between social anxiety and intolerance of uncertainty in junior high school students. The results of this study are consistent with the prior theoretical logic. Per prior research, improving the self-esteem of Chinese emerging adults (58) can help reduce their social anxiety (59). Why does self-esteem help reduce anxiety? According to the theory of fear management,

the flexible space provided by the self-regulation mechanism of self-esteem can help individuals effectively relieve anxiety (60). Therefore, when an individual's self-esteem is enhanced, anxiety is alleviated, and more effective actions are under taken to help the individual to better return to a balanced life. In other words, self-esteem may be one of the protective factors of junior high school students. These findings are consistent with the study's hypothesis 2.

Further analysis of the findings indicate that individuals with higher self-esteem can reduce the negative influence of anxiety on their psychology and behavior (32). Conversely, individuals with low self-esteem demonstrate higher intolerance of uncertainty (61). Consequently, this study confirms that self-esteem is an important protective factor that can reduce the negative impact of social anxiety and ameliorate the intolerance of uncertainty.

Limitations and implications

There are several limitations in this study. First, this study used self-reported methods to collect data. Although, there is no serious collinearity in this study, future studies should collect data in a variety of ways, such as collecting other comments, and through observation and interview methods. Second, a cross-sectional research design was used to detect the relationship between social anxiety and self-injury in junior high school students. However, cross-sectional studies do not easily reveal causal relationships, thus, future studies could use experimental or longitudinal methods to examine them in depth. Third, the sample selected for the study was only a subset of junior high school students in Hunan Province, China, which lacked representativeness. The study needs to further recruit a different representative sample. This study only explored the moderating role of self-esteem between social anxiety and intolerance of uncertainty, and did not explore the moderating role of self-esteem between social anxiety and self-injury. Future research could simultaneously examine the moderating role of the direct and indirect between social anxiety and self-injury.

Although there are limitations in the research, the results of the study have implications for theory and practice. First, this study expanded upon previous research by confirming the mediating role of intolerance of uncertainty and the moderating role of self-esteem. This could contribute to a better understanding of how and when social anxiety can be related to self-injury. Second, our study demonstrated that self-esteem played a protective role to reduce negative influence. This means that it is meaningful to improve junior high school students' self-esteem, thereby reducing the incidence rate of self-injury. Third, we should intentionally guide junior high school students to develop positive self-esteem, so as to enhance their ability to resist negative stimuli. School education can be enlightened as follows: in the prevention of self-injury in junior high school students, it can improve the self-esteem level of individuals by ameliorating their social anxiety and intolerance of uncertainty, and through mental health education activities, such as a second classroom and other forms, to achieve the expected results (62).

Conclusion

This study revealed that self-injury served as a potential mechanism by which serious social anxiety were associated with self-injury in a sample of Chinese junior high school students. Moreover, the association between social anxiety and intolerance of uncertainty had a more significant impact for junior high school students with a lower level of self-esteem compared with those with a higher level of self-esteem.

Data availability statement

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

Author contributions

ZY and MO designed the study protocol, provided financial support, as well as guided the first draft of the manuscript, provided guidance on the overall design of the study, and the revision of the manuscript. JX and LP performed the statistical analysis, drafted the manuscript, completed the literature review, and participated in the study design and interpretation analysis. SS guided the statistical analysis, interpretation of the results, edited the final manuscript, and provided guides in the process of revision. All authors contributed, read, and approved the submitted version.

Funding

This study were supported by grants from Scientific Research Project of Xiangnan University, "A Study on Core Self-evaluation and Mental Health of College Students" (2015XC10), and China's National Research and Innovative Learning Experimental Program for College Students, "Research on the Construction and Application of Peer Psychological Assistance Platform for College Students in the Era of Internet +" (201810545003).

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. Lu-Jia GE, Fei LI. The ideas and implementation of junior middle school students'mental health education in the positive psychology perspective. *J Northeast Normal Univ.* (2016) 3:244–8.

2. Li-Yan MA. Life events, instant messaging, and loneliness of early adolescents. *Psychol Develop Educ.* (2008) 4:106–12.

3. Nock MK, Favazza AR. Nonsuicidal self-injury: definition and classification. In: Nock MK, editor. *Understanding Nonsuicidal Self-Injury Origins Assessment and Treatment*. Washington, DC: American Psychological Association (2009). p. 9–18.

4. Wang YP, Zhu XZ, Jin-Yao YI, Yao SQ. Self-harm. and suicide behaviors in middle school students. *Chin J Clin Psychol.* (2012) 20:666–7. doi: 10.16128/j.cnki.1005-3611.2012.05.037

5. Xin XH, Yao SQ. Direct self-injurious behavior in adolescents:prevalence and its association with life events. *Chin J Clin Psychol.* (2016) 24:124–8.

6. BjeŽančević M, Groznica HrŽić IG, Dodig-Curković K. Self-injury in adolescents: a five-year study of characteristics and trends. *Psychiatr Danub.* (2019) 31:413–20. doi: 10.24869/psyd.2019.413

7. Wan Y, Chen R, Wang S, Clifford A, Zhang S, Orton S, et al. Associations of coping styles with nonsuicidal self-injury in adolescents: do they vary with gender and adverse childhood experiences? *Child Abus Neglect.* (2020) 104:104470. doi: 10.1016/j.chiabu.2020.104470

8. Suyemoto KL, Macdonald ML. Self-cutting in female adolescents. *Psychother Theor Res Pract Train.* (1995) 32:162–71. doi: 10.1037/0033-3204.32.1.162

9. Ying M, Jiang G, Yu L, Lu T. Patterns of reinforcement sensitivity for non-suicidal self-injury in college students. *Acta Psychol Sinica.* (2016) 48:258. doi: 10.3724/SP.J.1041.2016.00258

10. Zhong J, Wang A, Qian M, Zhang L, Gao J, Yang J, et al. Shame, personality, and social anxiety symptoms in Chinese and American nonclinical samples: a cross-cultural study. *Depress Anxiety.* (2008) 25:449–60. doi: 10.1002/da.20358

11. Jin X, Jing X. The relationship between cognitive emotional regulation strategies, self-esteem, personality trait and social anxiety. *Chin J Health Psychol.* (2013) 21:609–12. doi: 10.13342/j.cnki.cjhp.2013.04.020

12. Tomita T, Mizuko M, Kanemitsu Y. A consideration of social anxiety through a cognitive construct model: the relationship between self-consciousness, self-evaluation, social anxiety. *Kawasaki Med Welfare J.* (1999) 9:49–54.

13. Bao X, Zhang W, Chengfu YU, Zhu J, Bao Z, Yanping J, et al. Perceived school climate and Internet gaming disorder among junior middle school students: the mediating role of academic self-efficacy and the moderating role of parental academic involvement. *Psychol Develop Educ.* (2016) 32:358–68. doi: 10.16187/j.cnki.issn1001-4918.2016.03.13

14. Jza B. Anxiety symptoms mediates the influence of cybervictimization on adolescent non-suicidal self-injury: the moderating effect of self-control. *J Affect Disord.* (2021) 285:144–51. doi: 10.1016/j.jad.2021.01.004

15. Sulak B. *Non-suicidal self-injury, anxiety, and self-esteem among undergraduate college students (Doctoral dissertation)*. Ohio University, Athens, OH, United States. (2015). Available online at: http://rave.ohiolink.edu/etdc/view?acc_num=ohiou1427382976

16. Ikeda S. Social anxiety enhances sensitivity to negative transition and eye region of facial expression. *Pers Individ Dif.* (2020) 163. 110096. doi: 10.1016/j.paid.2020.110096

17. Peters EM, Bowen R, Balbuena L. Mood instability contributes to impulsivity, non-suicidal self-injury, and binge eating/purging in people with anxiety disorders. *Psychol Psychother.* (2019) 92:422–38. doi: 10.1111/papt.12192

18. Vassilopoulos SP, Banerjee R. Social interaction anxiety and the discounting of positive interpersonal events. *Behav Cogn Psychother.* (2010) 38:597–609. doi: 10.1017/S1352465810000433

19. Wu WL, Liu XH. Relationship between adult attachment styles of social anxiety disorder and its personality characters. *Chin J Clin Rehab.* (2006) 14:18–22.

20. Castilho P, Martins MJ, Pinto AM, Viegas R, Carvalho S, Madeira N. Understanding the effect of attachment styles in paranoid ideation: the mediator role of experiential avoidance. *J Context Behav Sci.* (2017) 6:42–6. doi: 10.1016/j.jcbs.2016.11.007

21. Boelen PA, Vrinssen I, van Tulder F. Intolerance of uncertainty in adolescents correlations with worry, social anxiety, and depression. *J Nerv Mental Dis.* (2010) 198:194–200. doi: 10.1097/NMD.0b013e3181d143de

22. Carleton RN, Sharpe D, Asmundson GJG. Anxiety sensitivity and intolerance of uncertainty: requisites of the fundamental fears? *Behav Res Ther.* (2007) 45:2307–16. doi: 10.1016/j.brat.2007.04.006

23. Carleton RN, Mulvogue MK, Thibodeau MA, McCabe RE, Antony MM, Asmundson GJG. Increasingly certain about uncertainty: intolerance of uncertainty across anxiety and depression. *J Anxiety Disord.* (2012) 26:468–79. doi: 10.1016/j.janxdis.2012.01.011

24. Zhang YX, Yang ZH, Wen-Qian HE, Zhang LX, Wei XU. Neuroticism and social anxiety: effects of intolerance of uncertainty and meta-worry. *Chin Ment Health J.* (2015) 29:145–9.

25. Shapiro MO, Gros DF, McCabe RE. Intolerance of uncertainty and social anxiety while utilizing a hybrid approach to symptom assessment. *Int J Cogn Ther.* (2020) 13:189–202. doi: 10.1007/s41811-020-00068-5

26. Walsh B. Clinical assessment of self-injury: a practical guide. *J Clin Psychol.* (2007) 63:1057–68. doi: 10.1002/jclp.20413

27. Barber BK. Parental psychological control: revisiting a neglected construct. *Child Dev.* (1996) 67:3296–319. doi: 10.2307/1131780

28. Ding Z, Wang X, Liu Q. The relationship between college students' self-esteem and cyber aggressive behavior: the role of social anxiety and dual self-consciousness. *Psychol Develop Educ.* (2018) 34:171–80. doi: 10.16187/j.cnki.issn1001-4918.2018.02.06

29. Rosenberg M. Society and the adolescent self-image. *Social Forces.* (1965) 3:280. doi: 10.1515/9781400876136

30. Leary MR, Tambor ES, Terdal SK, Downs DL. Self-esteem as an interpersonal monitor: the sociometer hypothesis. *J Pers Soc Psychol.* (1995) 68:518–30. doi: 10.1037/0022-3514.68.3.518

31. Forrester RL, Slater H, Jomar K, Mitzman S, Taylor PJ. Self-esteem and non-suicidal self-injury in adulthood: a systematic review. *J Affect Disord.* (2017) 221:172–83. doi: 10.1016/j.jad.2017.06.027

32. Rushton JP. *Review of E. E. Maccoby: Social Development: Psychological Growth and the Parent-Child Relationship*. New York, NY: Holt; Rinehart & Winston (1980).

33. Wang S, Xu H, Zhang S, Wan Y, Tao F. Mediating effects of self-esteem in the relationship between childhood maltreatment and non-suicidal self-injury among adolescents: the roles of sex and only-child status. *Soc Sci Med.* (2020) 249:112847. doi: 10.1016/j.socscimed.2020.112847

34. Fenigstein A, Scheier MF, Buss AH. Public and private self-consciousness: assessment and theory. *J Consult Clin Psychol.* (1975) 43:522–7. doi: 10.1037/h0076760

35. Wang XD, Wang XL, Ma H. *Handbook of Mental Health Assessment Scale (Revised Version)*. Beijing: China Mental Health Journal Press.

36. Feng Y. *The Relation of Adolecents' Self-Harm Behaviors Individual Emotion Characteristics and Family Environment Factors*. Central China Normal University (2008).

37. Li-Juan WU, Wang JN, Qi XD. Validity and reliability of the intolerance of uncertainty scale-12 in middle school students. *Chin Ment Health J.* (2016) 30:700–5.

38. Tian L. Shortcoming and merits of Chinese version of Rosenberg 1965 self-esteem scale. *Psychol Explorat.* (2006) 2:88–91.

39. Wen ZL, Ye BJ. Analyses of mediating effects: the development of methods and models. *Adv Psychol Sci.* (2014) 22:731–45. doi: 10.3724/SP.J.1042.2014.00731

40. Podsakoff PM, MacKenzie SB, Lee JY, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. *J Appl Psychol.* (2003) 88:879–903. doi: 10.1037/0021-9010.88.5.879

41. Preacher KJ, Hayes AF. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav Res Methods.* (2008) 40:879–91. doi: 10.3758/BRM.40.3.879

42. Muller D, Judd CM, Yzerbyt VY. When moderation is mediated and mediation is moderated. *J Pers Soc Psychol.* (2005) 89:852–63. doi: 10.1037/0022-3514.89.6.852

43. Plener PL. Tailoring treatments for adolescents with nonsuicidal self-injury. *Eur Child Adolescent Psychiat.* (2020) 29:893–5. doi: 10.1007/s00787-020-01523-6

44. Mathew AS, Davine TP, Snorrason I, Houghton DC, Woods DW, Lee HJ. Body-focused repetitive behaviors and non-suicidal self-injury: a comparison of clinical characteristics and symptom features. *J Psychiatr Res.* (2020) 124:115–22. doi: 10.1016/j.jpsychires.2020.02.020

45. Benjet C, González-Herrera I, Castro-Silva E, Méndez E, Borges G, Casanova L, et al. Non-suicidal self-injury in Mexican young adults: prevalence, associations with suicidal behavior and psychiatric disorders, and DSM-5 proposed diagnostic criteria. *J Affect Disord.* (2017) 215:1–8. doi: 10.1016/j.jad.2017.03.025

46. Giner-Bartolome C, Mallorquí-Bagué N, Tolosa-Sola I, Steward T, Jimenez-Murcia S, Granero R, et al. Non-suicidal self-injury in eating disordered patients: associations with heart rate variability and state-trait anxiety. *Front Psychol.* (2017) 8:1163. doi: 10.3389/fpsyg.2017.01163

47. Gu R, Shi Y, Yang J, Shi J, Cai H. The influence of anxiety on social decision behavior. *Adv Psychol Sci.* (2015) 23:547. doi: 10.3724/SP.J.1042.2015.00547

48. Hasking PA, Di Simplicio M, McEvoy PM, Rees CS. Emotional cascade theory and non-suicidal self-injury: the importance of imagery and positive affect. *Cogn Emot.* (2018) 32:941–52. doi: 10.1080/02699931.2017.1368456

49. Huang Y, Zhao Q, Li C. How interpersonal factors impact the co-development of depression and non-suicidal self-injury in Chinese early

adolescents. *Acta Psychologica Sinica*. (2021) 53:515–26. doi: 10.3724/SP.J.1041.2021.00515

50. Brossard B, Steggals P. The sociological implications of taking self-injury as a practice: an author meets critic interview. *Soc Theory Health*. (2020) 18:211–23. doi: 10.1057/s41285-020-00131-3

51. Counsell A, Furtado M, Iorio C, Anand L, Canzonieri A, Fine A, et al. Intolerance of uncertainty, social anxiety, and generalized anxiety: differences by diagnosis and symptoms. *Psychiatry Res*. (2017) 252:63–9. doi: 10.1016/j.psychres.2017.02.046

52. Zhiyong LI. The relationship between college students'self-esteem and social anxiety: the mediatingeffect of intolerance of uncertainty. *Chin J Special Educ*. (2013) 5:72–6.

53. Tanovic E, Gee DG, Joormann J. Intolerance of uncertainty: neural and psychophysiological correlates of the perception of uncertainty as threatening. *Clin Psychol Rev*. (2018) 60:87–99. doi: 10.1016/j.cpr.2018.01.001

54. Zhang GH, Dai BB. A summary of research on intolerance of uncertainty. *J Capital Normal Univ*. (2012) 2:124–30.

55. Wong SL, Chung MC. Child abuse and non-suicidal self-injury among Chinese university students: the role of emotion dysregulation and attachment style. *Curr Psychol*. (2021). doi: 10.1007/s12144-021-01832-2

56. In-Albon T, Tschan T, Schwarz D, Schmid M. Emotion regulation in adolescents with nonsuicidal self-injury. *Prax Kinderpsychol Kinderpsychiatr*. (2015) 64:386–403. doi: 10.13109/prkk.2015.64.5.386

57. Sternheim L, Danner U, Elburg AV, Harrison A. Do anxiety, depression, and intolerance of uncertainty contribute to social problem solving in adult women with anorexia nervosa? *Brain Behav*. (2020) 10:1588. doi: 10.1002/brb3.1588

58. Cheng G, Zhang D, Ding F. Self-esteem and fear of negative evaluation as mediators between family socioeconomic status and social anxiety in Chinese emerging adults. *Int J Soc Psychiatry*. (2015) 61:569–76. doi: 10.1177/0020764014565405

59. Chen C, Qin J. Emotional abuse and adolescents' social anxiety: the roles of self-esteem and loneliness. *J Fam Violence*. (2020) 35:497–507. doi: 10.1007/s10896-019-00099-3

60. Zhang Y, Zuo B. Self-esteem terror management theory. *Adv Psychol Sci*. (2006) 14:273.

61. Ming-Zheng LI, Tao L. Relationships between self-esteem, intolerance of uncertainty, career decision-making difficulties and job anxieties. *Chin J Clin Psychol*. (2012) 20:564–6.

62. Zhang Y, Li S, Yu G. The relationship between self-esteem and social anxiety: a meta-analysis with Chinese students. *Adv Psychol Sci*. (2019) 27:1005. doi: 10.3724/SP.J.1042.2019.01005

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 17 April 2023
DOI 10.3389/fpubh.2023.1152151

OPEN ACCESS

EDITED BY
Nan Luo,
National University of Singapore, Singapore

REVIEWED BY
Arlette Setiawan,
Padjadjaran University, Indonesia
Luenda Charles,
Centers for Disease Control and Prevention (CDC), United States

*CORRESPONDENCE
Chen-Wei Pan
✉ pcwonly@gmail.com
Hua Zhong
✉ zhoculist@163.com
Chun-Hua Zhao
✉ chzhao@njmu.edu.cn

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

RECEIVED 27 January 2023
ACCEPTED 28 March 2023
PUBLISHED 17 April 2023

CITATION
Li D-L, Nie X-Y, Li J, Tao Y-J, Zhao C-H, Zhong H and Pan C-W (2023) Factors associated with sleep disorders among adolescent students in rural areas of China. *Front. Public Health* 11:1152151. doi: 10.3389/fpubh.2023.1152151

Factors associated with sleep disorders among adolescent students in rural areas of China

Dan-Lin Li[1], Xin-Yi Nie[1], Jun Li[2], Yi-Jin Tao[3], Chun-Hua Zhao[4]*, Hua Zhong[3]* and Chen-Wei Pan[1]*

[1]School of Public Health, Medical College of Soochow University, Suzhou, China, [2]Department of Ophthalmology, The Second People's Hospital of Yunnan Province, Kunming, China, [3]Department of Ophthalmology, The First Affiliated Hospital of Kunming Medical University, Kunming, China, [4]Department of General Medicine, Big Data Center, The Affiliated Suzhou Hospital of Nanjing Medical University, Suzhou Municipal Hospital, Nanjing Medical University, Suzhou, China

Background: This study aimed to determine sleep patterns and the prevalence and association factors of sleep disorders in a regionally representative sample in Mo Jiang, China.

Methods: A total of 2,346 (participation rate 93.5%) Grade 7 students (aged 13–14years) from 10 middle schools, including 1,213 (51.7%) boys and 1,133 (48.3%) girls, participated in the study. All the participants were invited to complete questionnaires that acquired information on sleep patterns, academic performance, academic stress, and sociodemographic factors. Sleep disorders were assessed using the Chinese version of the Children's Sleep Habits Questionnaire. Logistic regression models were used to investigate factors associated with sleep disorders.

Results: The prevalence of sleep disorders among rural adolescents was 76.4%, which is higher than that among urban adolescents. Compared with previous findings in urban areas, our results indicate that sleep loss is much more severe in rural adolescents. Sleep disorders were positively associated with factors, such as watching TV [odds ratio (OR)=1.22, p=0.001], academic performance (OR=1.80, p<0.001), and academic stress (OR=1.38, p=0.04). In addition, girls were more likely to suffer from sleep disorders than boys (OR=1.36, p=0.01).

Conclusion: Insufficient sleep and sleep disorders have become common health problems in rural Chinese adolescents.

KEYWORDS

sleep pattern, sleep disorder, adolescent, rural China, student

Introduction

Good quality and adequate sleep plays an important role in the healthy development and optimal daytime function of adolescents (1, 2). However, sleep disorders frequently occur in adolescents. The prevalence of sleep disorders in adolescents varies from 20 to 60% in Western populations (3, 4) and is even reported to be higher in Asians (5, 6), especially in Chinese adolescents (7). Interethnic discrepancy in sleep disorders between Westerners and Asians are not completely understood. Moreover, sleep disorders in adolescents may be associated with a

broad range of factors, such as sociodemographic, sociocultural, environmental, familial, and individual factors (8).

In China, a series of studies conducted in urban areas demonstrate that sleep disorders are a widespread problem in adolescents with an estimated persistence rate of 15–70% (9–11). Li et al. conducted a multicenter study on 19,299 adolescents from eight cities in China: Shanghai, Canton, Wuhan, Chengdu, Xi'an, Urumqi, Hohhot, and Harbin found that more than two-thirds of adolescents aged over 11 years have sleep disorders (12). In a study in Beijing (13), the prevalence of sleep loss (total sleep time (TST) less than 9 h/day) among 9,198 children aged 3–14 years was 11.8%. Zhou et al. reported that 34.3% of children and adolescents aged 6–14 in Shanghai have poor sleep habits (11). Of the 1,365 adolescents between the ages of 12 and 18 years old in Shandong (14), 16.9% reported symptoms of insomnia, and the prevalence of difficulty initiating sleep, difficulty maintaining sleep and early morning awakening was 10.8, 6.3, and 2.1%, respectively. In another group of 1,056 adolescents from Shandong (15), 26.2% were unsatisfied with their sleep. About 18.8% reported that their sleep quality was poor, 16.1% had insomnia and 17.9% had daytime sleepiness. In coastal cities such as Guangzhou (16), a sizeable percentage of adolescents also struggle to sleep. In a study of 912 Shenzhen (8) adolescents aged 6–14, 69.3% had sleep disorders, including bedtime resistance (22.9%) and daytime sleepiness (20.0%). In addition, 23.8% reported suffering from repeated sleep loss (TST less than 9 h/day). The students in grade 7 are under high schooling pressure and their sleep status is the most concerned problem of public health practitioners in China However, data on Chinese school students in this age, especially in rural communities, are still lacking.

China is the world's most populous country, and populations residing in rural areas account for more than half of the country's population.[1] Furthermore, with the rapid development of the rural economy in China, tremendous changes in lifestyles have occurred among young generations in these areas. They are faced with a new range of time use choices and prolonged commuting times, which may encroach on sleep time, interfere with sleep patterns, and result in sleep disorders. Therefore, understanding sleep patterns and sleep disorders may have important implications from a public health perspective. This study aims to examine sleep patterns, sleep disorders and associated factors among a school-based sample of 7th graders aged 13–14 in rural China. We hypothesize that sleep disorders among rural stumaildents is common and might be associated with lifestyles and school factors.

Methods

Study participants

The present findings were based on data collected from the Mojiang Myopia Progression Study on students of Grade 7 from 10 secondary schools, which was conducted in 2016 (17). Mojiang is a small rural county located in Southwestern China with a population of 0.36 million and an area of 5,312 km^2. It was chosen as the study site due to its relatively stable demographic structure and similar socioeconomic status to the average of rural China. An explanatory letter about the nature of the study was sent out, and a consent form for taking full measurements was obtained from at least one parent or legal guardian of each participant. Telephone calls or home visits were made when parents could not be contacted. Ultimately, a total of 2,346 (participation rate 93.5%) Grade 7 students including 1,213 (51.7%) boys and 1,133 (48.3%) girls participated in the study. More details of the project were described in previous publications (17, 18).

Our study adhered to the tenets of the Declaration of Helsinki for research involving human subjects and was approved by the Institutional Review Board of Kunming Medical University.

Chinese version of the children's sleep habits questionnaire

The Chinese version of the Children's Sleep Habits Questionnaire (CSHQ) has been used in extensive research, both in children and adolescents (19–21). The questionnaire about adolescents' sleep habits in a typical recent week consists of 33 items grouped into eight subscales: bedtime resistance, sleep onset delay, sleep duration, sleep anxiety, night waking, parasomnias, sleep-disordered breathing, and daytime sleepiness. Each item was scored on a three-point Likert scale: 1 ("rarely, 0–1 time in a week"), 2 ("sometimes, 2–4 times in a week"), and 3 ("usually, 5–7 times in a week"). The adequate internal consistency of the overall questionnaire ranged from 0.68 to 0.78, and the test–retest reliability ranged from 0.62 to 0.79. Sleep disorders were diagnosed by a total CSHQ score of 41 points or more (19). In addition, previous studies suggested a subscale score > standard deviation (SD) above the mean score of a community sample to identify specific sleep disorders of clinical significance (19).

Sleep-related variables and demographics

A predesigned questionnaire was used to collect information regarding demographic characteristics that were potentially associated with sleep disorders, such as academic performance, time spent on TV, time spent surfing the internet, time spent outdoors, family structure, and parental education level. With the help of parent (s), the adolescents completed the questionnaire at home for 2 days. After completion of the questionnaires, the research assistants checked the questionnaires to ensure that all questions were properly answered.

Academic performance was determined *via* self-report measures. The adolescents responded to the question "how is your academic performance?" on a three-point Likert scale from 1 ("excellent"), 2 ("moderate"), and 3 ("underachievement"). Adolescents filled out their usual bedtime and wake time on weekdays and weekends, and the research assistants calculated the total sleep duration.

Statistical analysis

Sleep patterns, including morning wake time, bedtime, and total sleep time (TST) on weekends and weekdays, were described with means and SD. We calculated the prevalence rate of sleep disorders (total score of 41 points or more) and the number and ratio of

1 http://www.stats.gov.cn/ztjc/zdtjgz/zgrkpc/dlcrkpc/

individuals who exceeded cut-off points. Logistic regression models were used to explore the associations of sleep disorders with potential factors. First, we examined variables of interest such as lifestyle-related factors and school factors in univariate models. Second, we included sex and factors with a *p* value of less than 0.10 in univariate analyses in the multivariate analysis models. Odds ratios (ORs) and 95% confidence intervals (CIs) were calculated to quantify the associations. Statistical tests were set with a significance level of 0.05. Statistical Package for Social Science version 20.0 was used for all the analyses.

Results

Descriptions of demographic characteristics, media (computer and TV), family structure, and academic performance are shown in Table 1. The mean age of the adolescents was 13.6 (SD: 0.5) years old, and among the total number of adolescents, girls accounted for 48.3% (1,133). Most of their parents graduated from junior high school or below, with only 17.0% (399) of their fathers and 12.4% (270) of their mothers who graduated from college or had a higher degree.

TABLE 1 Distributions of children by their characteristics.

	N=2,346
Age in years, mean (SD)	13.6 (0.5)
Girls, no. (%)	1,133 (48.3)
Compute Use (yes), no. (%)	323 (13.8)
Time on computer per day, hours (SD)	0.9(0.9)
Time on watching TV per day, hours (SD)	1.4(0.9)
Time outdoors per day, hours (SD)	1.3(1.0)
Father's Education, no. (%)	
Informal education	84 (3.6)
Primary school	760 (32.4)
Junior high school	1,029 (43.9)
High school	292 (12.4)
College school or above	107 (4.6)
Mother's education, no. (%)	
Informal education	253 (10.8)
Primary school	982 (41.9)
Junior high school	717 (30.6)
High school	222 (9.5)
College school or above	68 (2.9)
Family structure, no. (%)	
Nuclear family	743 (31.7)
Extended family	1,504 (64.1)
Single family or others	99 (4.2)
Academic performance, no. (%)	
Excellence	416 (17.7)
Moderate	1,195 (50.9)
Underachievement	667 (28.4)
Academic stress (yes), no. (%)	2090 (89.1)

SD, standard deviation.

Regarding school factors, 28.4% of adolescents underestimated their academic achievement and 89.1% of the students thought they experienced academic pressure.

Their sleep patterns during weekdays and weekends and the distribution of TST are shown in Table 2 and Figure 1, respectively. In our study, the adolescents woke up much earlier on weekdays than on weekends, but their bedtime was similar, accounting for significantly longer weekend TST than weekday TST. Figure 1 shows that 27.8% of the adolescents could sleep 9 h on weekends, while during weekdays 46.5% of the adolescents could sleep 7 h, and 10% of the adolescents could only sleep 6 h or less. Compared to urban adolescents, rural adolescents go to bed later, which results in an average loss of one-hour sleeping time.

The total CSHQ and subscale scores are listed in Tables 2, 3, respectively. Approximately 76.4% of adolescents with a CSHQ total score above 41 suffer from sleep disorders, and sleep duration (50.3%) is the most prevalent sleep disorder among the eight specific sleep disorders. Based on the CSHQ score, 72.6% of adolescents reported that they "fall asleep in their own bed" 5–7 times a week, which belongs to the bedtime resistance subscale. Except for sleep onset delay, other specific sleep disorders were reported more frequently by the urban adolescents, as shown in Table 3.

Logistic regression analyses were used to determine the factors associated with sleep disorders, and the results are shown in Table 4. Univariate analyses revealed that sleep disorders were significantly associated with gender, time spent on watching TV, family structure, academic performance, and academic stress (all $p < 0.05$). In the multivariate analyses, time spent watching TV (OR: 1.22; 95% CI: 1.08–1.37) and academic stress (OR: 1.38; 95% CI: 1.03–1.91) were positively associated with sleep disorders. Girls were 1.36 (95% CI: 1.31, 2.49: 1.10, 1.68) times more likely to have sleep disorders than boys. Furthermore, sleep disorders are positively associated with adolescents' academic performance in school. After controlling for gender, academic stress, time spent watching TV, and family structure, adolescents with average academic performance were 1.32 (95% CI: 1.06, 1.67) times and underachieving adolescents were 1.80 (95% CI: 1.31, 2.49) times more likely to have sleep disorders compared with academically excellent adolescents.

Discussion

The prevalence rates reported in the present study are much higher than those reported from cities in China. In the present sample, the prevalence of sleep disorders was 76.4%. A total of 94.9% of rural adolescents had insufficient sleep (TST less than 9 h/day) on weeknights, whereas the percentage decreased to 49.1% on weekends. Moreover, a number of positive factors of sleep disorders, such as over 2 h of watching TV, poor academic performance, and academic stress, occur among rural adolescents. For rural adolescents, watching TV less than 1 h on school days was associated with lower odds of sleep disorders.

Compared with previous findings in urban areas (8, 16), our results indicate that sleep loss is much more severe in rural adolescents. In our study, the prevalence of sleep loss [total sleep time (TST) less than 9 h/day] was 94.9%. About 56.5% of the adolescents reported that they sleep less than 8 h/day, and 10.0% only sleep 6 h/day or less. The discrepancy in sleep patterns in rural adolescents may be attributed to

TABLE 2 CSHQ score and sleep pattern.

	Rural			Urban[a]
	Mean (SD)	Cut-off	Score>cut-off (%)	
Sleep patterns (Weekday)				
Morning wake-up time, h	6.5 (0.5)			6.0 (0.5)
Bedtime, h	22.4 (1.1)			21.4 (0.8)
Total sleep time, h	8.0 (1.2)	8.0	56.5	9.0 (0.8)
Sleep patterns (Weekend)				
Morning wake-up time, h	7.8 (1.4)			8.2 (1.0)
Bedtime, h	22.5 (0.9)			22.2 (0.9)
Total sleep time, h	9.3 (1.5)	8.0	24.5	9.8 (1.1)
CSHQ subscales				
Bedtime resistance	9.0 (1.4)	10.8	9.7	33.1
Sleep onset delay	1.9 (0.8)	2.3	28.6	25.4
Sleep duration	5.4 (1.5)	5.3	50.3	70.1
Sleep anxiety	4.8 (1.3)	7.8	5.9	41.4
Nighttime awakening	3.5 (0.9)	5.3	3.7	27.4
Parasomnias	8.3 (1.7)	10.6	10.4	52.4
Sleep disordered breathing	3.4 (0.8)	4.5	9.2	11.9
Daytime sleepiness	12.5 (3.6)	15.2	23.5	71.2
Total CSHQ score	48.9 (8.0)	41.0	76.4	

CSHQ, children's sleep habits questionnaire.
SD, standard deviation.[a]The study sample was recruited from eight urban cities in China, including Urumqi, Chengdu, Xi'an, Hohhot, Wuhan, Canton, Shanghai, and Harbin, aged 11 years (12).

their bedtime being much later for night lessons, which is a typical practice in rural Chinese schools (adolescents have to study in school during weeknights). The percentage of sleep loss decreases to 49.1% on weekends, which supports this notion. Given that the rural Chinese school system is different from that of urban schools, research on sleep schedules between rural and urban areas separately provides great insights into adolescents' sleep patterns. Sleep loss in rural adolescents requires far more attention from the government and education system.

In the present study, the prevalence of sleep disorders was 78.8% in rural Chinese adolescents. Sleep duration (50.3%) was the most prevalent among the eight specific sleep disorders. These findings may partly explain why the CSHQ subscales include items that are related to strict school schedules (e.g., "going to bed at the same time," "falls asleep in own bed," and "sleeping the right amount"). Furthermore, the prevalence of sleep onset delay and daytime sleepiness in this study is considerably higher (e.g., 28.6% for sleep onset delay and 23.5% for daytime sleepiness) compared with a previous study conducted in eight Chinese cities on adolescents aged over 11 (12). Other reasons may also explain these findings. First, sleep is strongly shaped and interpreted by cultural values, ideological beliefs and parenting values. In China, especially in rural areas, adolescents are highly influenced by their parents and teachers. The notion of 'knowledge changes destiny' is widespread in rural China, and rural adolescents face great academic pressure to enter a highly regarded school. Second, the cut-offs used in this study were calculated from American children aged 4–12 (19). As the current sample is composed of Chinese adolescents and their ages extend beyond 12 years old, the prevalence determined based on these cut-offs may cause bias. Third, cosleeping (room sharing) in school is a popular practice in rural China. Positive correlations between sleep quality and cosleeping among school-aged students have been identified in previous studies (22–24).

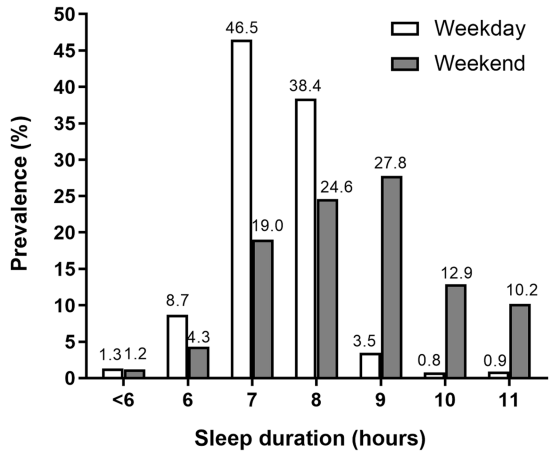


FIGURE 1
Distribution of weekday and weekend total sleep time (TST) in 2,346 Chinese adolescents aged 13–14years.

Our study supports previous findings on urban adolescents showing that watching TV is associated with increased odds of sleep disorders in adolescents (25–27). These studies report that adolescents who watch TV more than 2 h a day experience more daytime sleepiness and problematic sleep. Another study also shows that each hour spent watching TV reduces sleep duration by 7 min (28).

TABLE 3 The prevalence of specific sleep disorders.

CSHQ subscales	Usually, No. (%)	Sometimes, No. (%)	Rarely, No. (%)
Bedtime resistance			
Going to bed at same times	726 (30.9)	977 (41.6)	643 (27.4)
Falling asleep in own bed	1,703 (72.6)	254 (10.8)	389 (16.6)
Falling asleep in other's bed	53 (2.3)	166 (7.1)	2,127 (90.7)
Needing parent in room to sleep	17 (0.7)	112 (4.8)	2,217 (94.5)
Struggling at bedtime	17 (0.7)	79 (3.4)	2,250 (95.9)
Afraid of sleeping alone	95 (4.0)	248 (10.6)	2,003 (85.4)
Sleep onset delay			
Falling asleep in 20 min	671 (28.6)	824 (35.1)	851 (36.3)
Sleep duration			
Sleeping too little	362 (15.4)	759 (32.4)	1,225 (52.2)
Sleeping the right amount	877 (37.4)	726 (30.9)	743 (31.7)
Sleeping same amount each day	485 (20.7)	779 (33.2)	1,082 (46.1)
Sleep Anxiety			
Needing parent in room to sleep	17 (0.7)	112 (4.8)	2,217 (94.5)
Afraid of sleeping in the dark	131 (5.6)	351 (15.0)	1,864 (79.5)
Afraid of sleeping alone	95 (4.0)	248 (10.6)	2,003 (85.4)
Trouble sleeping away	164 (7.0)	457 (19.5)	1,725 (73.5)
Night awakening			
Moving to other's bed in night	16 (0.7)	73 (3.1)	2,257 (96.2)
Awakening once during night	126 (5.4)	596 (25.4)	1,624 (69.2)
Awakening more than once	52 (2.2)	225 (9.6)	2,069 (88.2)
Parasomnias			
Wetting the bed at night	8 (0.3)	67 (2.9)	2,271 (96.8)
Talking during sleep	54 (2.3)	347 (14.8)	1,945 (82.9)
Restless and moves a lot	170 (7.2)	434 (18.5)	1,742 (74.3)
Sleepwalking	23 (1.0)	65 (2.8)	2,258 (96.2)
Grinding teeth during sleep	44 (1.9)	136 (5.8)	2,166 (92.3)
Awakening screaming, sweating	36 (1.5)	297 (12.7)	2,013 (85.8)
Alarmed by nightmare	148 (6.3)	647 (27.6)	1,551 (66.1)
Sleep disordered breathing			
Snoring loudly	25 (1.1)	191 (8.1)	2,130 (90.8)
Stopping breathing	28 (1.2)	150 (6.4)	2,168 (92.4)
Snorting and gasping	55 (2.3)	288 (12.3)	2,003 (85.4)
Daytime sleepiness			
Awakening by himself	715 (30.5)	523 (22.5)	1,102 (47.0)
Awakening up in negative mood	241 (10.3)	471 (20.1)	1,634 (69.7)
Awakening by other's	318 (13.6)	480 (20.5)	1,548 (66.0)
Hard time getting out of bed	448 (19.1)	542 (23.1)	1,356 (57.8)
Takes long time to be alert	220 (9.4)	444 (18.9)	1,682 (71.7)
Seems tired during the day	206 (8.8)	659 (28.1)	1,481 (63.1)
Asleep while watching TV	734 (31.3)	559 (23.8)	1,053 (44.9)
Asleep while riding in car	335 (14.3)	528 (22.5)	1,483 (63.2)

CSHQ, children's sleep habits questionnaire.

TABLE 4 Logistic regression model of factors associated with sleep disorders (N=2,346).

Variables	Univariate logistic regression		Multiple logistic regression	
	OR (95%CI)	*p*	OR (95%CI)	*p*
Sex (Girls vs. Boys)	1.27 (1.04, 1.55)	0.02	1.36(1.10, 1.68)	0.01
Ethnicity (Han vs. Monitory)	1.09 (0.84, 1.42)	0.25	-	
Time spent on TV, per hour increase	1.22 (1.08, 1.36)	0.001	1.22(1.08, 1.37)	0.001
Time spent outdoors, per hour increase	0.95 (0.84, 1.06)	0.34	-	
Computer use (yes)	1.05 (0.62, 1.15)	0.28	-	
Father's education (category)				
Informal education	**Reference**			
Primary school	1.28 (0.70, 2.33)	0.43		
Junior high school	1.35 (0.75, 2.45)	0.32		
High school	1.21 (0.64, 2.31)	0.55		
College school or above	1.27 (0.61, 2.67)	0.53		
Mother's education (category)				
Informal education	**Reference**			
Primary school	1.07 (0.76, 1.51)	0.71		
Junior high school	1.00 (0.70, 1.43)	0.99		
High school	1.05 (0.67, 1.65)	0.82		
College school or above	1.54 (0.83, 2.83)	0.17		
Family structure				
Nuclear family	**Reference**		**Reference**	
Single family or others	1.32 (1.07, 1.63)	0.01	1.23(0.65, 1.53)	0.06
Extended family	1.14 (0.69, 1.88)	0.42	1.34(0.78, 2.28)	0.28
Academic performance				
Excellence	**Reference**		**Reference**	
Moderate	1.32 (1.06, 1.66)	0.02	1.32 (1.06, 1.67)	0.02
Underachievement	1.70 (1.24, 2.33)	0.001	1.80 (1.31, 2.49)	< 0.001
Academic stress (yes)	1.37 (1.01, 1.85)	0.04	1.38 (1.03, 1.91)	0.04

OR, odds ratio; CI, confidence interval.

Although most studies cannot identify a causal relationship between watching TV and sleep disorders, sleeping patterns of urban adolescents can be influenced by watching TV. In our study, decreasing TV watching by 1 h on weekdays alleviated sleep disorders in rural adolescents. One hypothesis is that appropriate time spent watching TV may relieve rural adolescents from intensive learning tasks. Academic load and stress are risk factors for sleep disorders in adolescents (29). Further research should explore in more detail how much time spent watching TV is associated with sleep disorders and how much time spent watching TV may be beneficial for adolescents. Furthermore, no significant differences in terms of computer use with sleep disorders were reported. A possible explanation is that almost all families have TV, but computers were rarely available (13.8%) in this community.

In our study, academic underachievement was a common and serious problem that affected 28.4% of rural adolescents. Consistent with previous studies (30, 31), our study found a positive link between sleep and academic performance. In addition, 89.1% of rural adolescents suffer from tremendous academic stress. We assume that underachieving adolescents may suffer more from academic stress and tend to spend more time studying at the expense of their sleep. We further found that academic stress is positively related to sleep disorders, which supports our hypothesis. In addition, girls are more likely to suffer from sleep disorders, which is in line with a previous study in Anhui (32) and Taiwan (33) in China. Girls experiencing the menstrual period undergo fluctuating levels of estrogen, progesterone, melatonin, and cortisol, leaving them susceptible to negative effects, which may lead to sleep disorders (34). Parental educational level and family structure are not related to sleep disorders. This finding is inconsistent with the study by BaHammam et al. (35). They found that mothers' educational level is related to adolescents' bedtime. Another study from Maha et al. (36) reported that adolescents with a high CSHQ score had fathers with low educational levels. The effects of parental educational level and family structure on sleep disorders in adolescents warrant further investigation.

Our study provides detailed information on rural adolescents' sleep, and we hope it can aid policymakers and support targeted interventions. At the same time, we hope it can assist professional health organizations in identifying problematic sleep and helping families to promote their children's sleep quality.

Several limitations should also be acknowledged. First, we cannot identify a causal relationship between sleep disorders and potential factors through the cross-sectional design. Second, all measures relied on parent-reported questionnaires rather than objective assessments, which may have led to report bias. Third, although the Chinese version of the CSHQ has adequate internal consistency and test–retest reliability, the exact relationship of the parent-reported sleep disorder variables and results from objective assessment is still uncertain. Last, factors, such as seasonal effects and sleep environments, which may also affect sleep patterns and sleep disorders, should be considered in future studies.

Conclusion

To summarize, the sample population used in the present study had an extremely higher prevalence rate of sleep disorders and sleep loss than urban adolescents. Governments and health policy makers should be aware of this issue and adopt appropriate strategies to improve rural adolescents' sleep quality.

Data availability statement

The datasets presented in this article are not readily available because privacy policy. Requests to access the datasets should be directed to pcwonly@gmail.com.

Ethics statement

The studies involving human participants were reviewed and approved by the study adhered to the tenets of the Declaration of Helsinki for research involving human subjects and was approved by the Institutional Review Board of Kunming Medical University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

C-WP, HZ, and C-HZ concepted the study. D-LL conducted the statistical analyses and wrote the original manuscript. JL, C-WP, and HZ participated the investigation of the study. X-YN, JL, Y-JT, HZ, C-HZ, and C-WP edited the manuscript. JL got the funding acquisition. C-WP and HZ supervised the study. All authors contributed to the article and approved the submitted version.

Funding

The research was funded by the National Natural Science Foundation of China (grant no. 81560169).

Acknowledgments

We gratefully acknowledge all the participants and data acquisition staff for their on-site cooperation during the data acquisition process.

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. Astill RG, Van der Heijden KB, Van Ijzendoorn MH, Van Someren EJ. Sleep, cognition, and behavioral problems in school-age children: a century of research meta-analyzed. *Psychol Bull*. (2012) 138:1109–38. doi: 10.1037/a0028204

2. Kirshenbaum JS, Coury SM, Colich NL, Manber R, Gotlib IH. Objective and subjective sleep health in adolescence: associations with puberty and affect. *J Sleep Res*. (2022):e13805. doi: 10.1111/jsr.138058

3. Knutson KL. The association between pubertal status and sleep duration and quality among a nationally representative sample of US. *Adolesc Am J Hum Biol*. (2005) 17:418–24. doi: 10.1002/ajhb.20405

4. van Litsenburg RR, Waumans RC, van den Berg G, Gemke RJ. Sleep habits and sleep disturbances in Dutch children: a population-based study. *Eur J Pediatr*. (2010) 169:1009–15. doi: 10.1007/s00431-010-1169-8

5. Otsuka Y, Kaneita Y, Spira AP, Mojtabai R, Itani O, Jike M, et al. Trends in sleep problems and patterns among Japanese adolescents: 2004 to 2017. *Lancet Reg Health West Pac*. (2021) 9:100107. doi: 10.1016/j.lanwpc.2021.100107

6. Matsuoka M, Nagamitsu S, Iwasaki M, Iemura A, Yamashita Y, Maeda M, et al. High incidence of sleep problems in children with developmental disorders: results of a questionnaire survey in a Japanese elementary school. *Brain Dev*. (2014) 36:35–44. doi: 10.1016/j.braindev.2012.12.004

7. Liang M, Guo L, Huo J, Zhou G. Prevalence of sleep disturbances in Chinese adolescents: a systematic review and meta-analysis. *PLoS One*. (2021) 16:e247333. doi: 10.1371/journal.pone.0247333

8. Wang G, Xu G, Liu Z, Lu N, Ma R, Zhang E. Sleep patterns and sleep disturbances among Chinese school-aged children: prevalence and associated factors. *Sleep Med*. (2013) 14:45–52. doi: 10.1016/j.sleep.2012.09.022

9. Wang M, Zhong JM, Hu RY, Gong WW, Yu M. Sleep duration and behavioral correlates in middle and high school students: a cross-sectional study in Zhejiang province. *Chin Sleep Med*. (2021) 79:55–61. doi: 10.1016/j.sleep.2021.01.008

10. Jiang X, Hardy LL, Baur LA, Ding D, Wang L, Shi H. Sleep duration, schedule and quality among urban Chinese children and adolescents: associations with routine after-school activities. *PLoS One*. (2015) 10:e115326. doi: 10.1371/journal.pone.0115326

11. Zhou HQ, Shi WB, Wang XF, Yao M, Cheng GY, Chen PY, et al. An epidemiological study of sleep quality in adolescents in South China: a school-based study. *Child Care Health Dev*. (2012) 38:581–7. doi: 10.1111/j.1365-2214.2011.01300.x

12. Li S, Jin X, Wu S, Jiang F, Yan C, Shen X. The impact of media use on sleep patterns and sleep disorders among school-aged children in China. *Sleep*. (2007) 30:361–7. doi: 10.1093/sleep/30.3.361

13. Li XD, Tai J, Xu ZF, Wang GX, Wu YX, Du JN, et al. Sleep duration and factors related to sleep loss in 3-14-year-old children in Beijing: a cross-sectional survey. *Chin Med J*. (2018) 131:1799–807. doi: 10.4103/0366-6999.237403

14. Liu X, Uchiyama M, Okawa M, Kurita H. Prevalence and correlates of self-reported sleep problems among Chinese adolescents. *Sleep*. (2000) 23:27–34.

15. Yang QZ, Bu YQ, Dong SY, Fan SS, Wang LX. A comparison of sleeping problems in school-age children between rural and urban communities in China. *J Paediatr Child Health*. (2009) 45:414–8. doi: 10.1111/j.1440-1754.2009.01530.x

16. Wang J, Adab P, Liu W, Chen Y, Li B, Lin R, et al. Prevalence of adiposity and its association with sleep duration, quality, and timing among 9-12-year-old children in Guangzhou. *Chin J Epidemiol*. (2017) 27:531–7. doi: 10.1016/j.je.2016.11.003

17. Pan CW, Wu RK, Wang P, Li J, Zhong H. Reduced vision, refractive errors and health-related quality of life among adolescents in rural China. *Clin Exp Optom*. (2018) 101:758–63. doi: 10.1111/cxo.12680

18. Qian DJ, Zhong H, Nie Q, Li J, Yuan Y, Pan CW. Spectacles need and ownership among multiethnic students in rural China. *Public Health*. (2018) 157:86–93. doi: 10.1016/j.puhe.2017.11.015

19. Owens JA, Spirito A, Mcguinn M. The Children's sleep habits questionnaire (CSHQ): psychometric properties of a survey instrument for school-aged children. *Sleep*. (2000) 23:1043–51.

20. Churchill SS, Kieckhefer GM, Bjornson KF, Herting JR. Relationship between sleep disturbance and functional outcomes in daily life habits of children with down syndrome. *Sleep*. (2015) 38:61–71. doi: 10.5665/sleep.4326

21. Lucas I, Mulraney M, Sciberras E. Sleep problems and daytime sleepiness in children with ADHD: associations with social, emotional, and behavioral functioning at school, a cross-sectional study. *Behav Sleep Med*. (2019) 17:411–22. doi: 10.1080/15402002.2017.1376207

22. Chandrasegaran B, Devikittu AS. Pattern and profile of co-sleeping in school-aged children. *Indian Pediatr*. (2022) 59:250–1. doi: 10.1007/s13312-022-2479-4

23. Gupta R, Kandpal SD, Goel D, Mittal N, Dhyani M, Mittal M. Sleep-patterns, co-sleeping and parent's perception of sleep among school children: comparison of domicile and gender. *Sleep Sci*. (2016) 9:192–7. doi: 10.1016/j.slsci.2016.07.003

24. Kim E, Lee R, Cain KC. Cosleeping, sleep disturbances, children's behavioral problems, and parenting self-efficacy among Korean American families. *J Child Adolesc Psychiatr Nurs*. (2017) 30:112–20. doi: 10.1111/jcap.12182

25. Mak YW, Wu CS, Hui DW, Lam SP, Tse HY, Yu WY, et al. Association between screen viewing duration and sleep duration, sleep quality, and excessive daytime sleepiness among adolescents in Hong Kong. *Int J Environ Res Public Health*. (2014) 11:11201–19. doi: 10.3390/ijerph111111201

26. Hale L, Guan S. Screen time and sleep among school-aged children and adolescents: a systematic literature review. *Sleep Med Rev*. (2015) 21:50–8. doi: 10.1016/j.smrv.2014.07.007

27. Tzischinsky O, Haimov I. Comparative study shows differences in screen exposure, sleep patterns and sleep disturbances between Jewish and Muslim children in Israel. *Acta Paediatr*. (2017) 106:1642–50. doi: 10.1111/apa.13961

28. Singh M, Drake CL, Roehrs T, Hudgel DW, Roth T. The association between obesity and short sleep duration: a population-based study. *J Clin Sleep Med*. (2005) 1:357–63. doi: 10.5664/jcsm.26361

29. Sun WQ, Spruyt K, Chen WJ, Jiang YR, Schonfeld D, Adams R, et al. The relation among sleep duration, homework burden, and sleep hygiene in Chinese school-aged children. *Behav Sleep Med*. (2014) 12:398–411. doi: 10.1080/15402002.2013.825837

30. Anjarwala ZM, Ali NS, Nanji K, Beg AA, Karmani VK. Sleep and its relation with academic performance among adolescents: an analytical cross sectional study in Gulshan town. *Karachi J Pak Med Assoc*. (2020) 70:1948–51. doi: 10.5455/JPMA.25469

31. Wang G, Ren F, Liu Z, Xu G, Jiang F, Skora E, et al. Sleep patterns and academic performance during preparation for college entrance exam in Chinese adolescents. *J Sch Health*. (2016) 86:298–306. doi: 10.1111/josh.12379

32. Ouyang F, Lu BS, Wang B, Yang J, Li Z, Wang L, et al. Sleep patterns among rural Chinese twin adolescents. *Sleep Med*. (2009) 10:479–89. doi: 10.1016/j.sleep.2008.04.011

33. Lin LN, Chang LY, Hurng BS, Wu CC, Yen LL, Chang HY. Sex differences in sleep patterns and changes in 7th to 12th graders: a longitudinal follow-up study in Taiwan. *Sleep*. (2018) 41. doi: 10.1093/sleep/zsx211

34. Campbell IG, Grimm KJ, de Bie E, Feinberg I. Sex, puberty, and the timing of sleep EEG measured adolescent brain maturation. *Proc Natl Acad Sci U S A*. (2012) 109:5740–3. doi: 10.1073/pnas.1120860109

35. Bahammam A, Bin SA, Al-Faris E, Shaikh S. Sleep duration and its correlates in a sample of Saudi elementary school children. *Singap Med J*. (2006) 47:875–81.

36. Abou-Khadra MK. Sleep patterns and sleep problems among Egyptian school children living in urban, suburban, and rural areas. *Sleep Biol Rhythms*. (2009) 7:84–92. doi: 10.1111/j.1479-8425.2009.00392.x

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 20 April 2023
DOI 10.3389/fpubh.2023.1156755

OPEN ACCESS

EDITED BY
Pei Wang,
Fudan University, China

REVIEWED BY
Dirk Bruland,
Bielefeld University of Applied Sciences, Germany
Kenneth Land,
Duke University, United States

*CORRESPONDENCE
Lucas Arrais Campos
✉ lucas.arraisdecampos@tuni.fi

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

RECEIVED 01 February 2023
ACCEPTED 27 March 2023
PUBLISHED 20 April 2023

CITATION
Silva BNS, Martins BG, Campos LA, Marôco J and Campos JADB (2023) Subjective wellbeing of preschool children.
Front. Public Health 11:1156755.
doi: 10.3389/fpubh.2023.1156755

Subjective wellbeing of preschool children

Bianca Núbia Souza Silva[1], Bianca Gonzalez Martins[2], Lucas Arrais Campos[1,3,4,5]*, João Marôco[6,7] and Juliana Alvares Duarte Bonini Campos[2]

[1]Department of Morphology and Children's Clinics, São Paulo State University (UNESP), School of Dentistry, Araraquara, São Paulo, Brazil, [2]Department of Biological Sciences, School of Pharmaceutical Sciences, São Paulo State University (UNESP), Araraquara, São Paulo, Brazil, [3]Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland, [4]Department of Ear and Oral Diseases, Tampere University Hospital, Tampere, Finland, [5]Institute of Dentistry, Faculty of Health Sciences, University of Eastern Finland, Kuopio, Finland, [6]William James Center for Research (WJCR), University Institute of Psychological, Social, and Life Sciences (ISPA), Lisbon, Portugal, [7]Flu Pedagogy, Nord University, Bodø, Norway

Objective: The aim of the present study was to evaluate the psychometric properties of the *Autoquestionnaire Qualité de Vie Enfant Imagé* (AUQEI) in pre-school children and estimate the influence of demographic characteristics on their subjective wellbeing.

Methods: Construct validity was estimated using confirmatory analysis and the chi-square per degrees of freedom ratio (χ^2/df), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Root Mean Square Error of Approximation (RMSEA). Reliability was assessed by the ordinal alpha (α) and omega (ω) coefficients and the factorial invariance by the difference in CFI (ΔCFI). Mean scores for each AUQEI item and the general score were calculated.

Results: A total of 443 Preschool children enrolled in public education institutions participated. The original 4-factor AUQEI model showed collinearity between factors and a high correlation between two items. A single factor model was tested, presenting adequate fit to the data (χ^2/df = 4.47; CFI = 0.98; TLI = 0.98; RMSEA = 0.08; α = 0.98; ω = 0.93; UniCo > 0.95, EVC > 0.85, and MIREAL < 0.30) and strict model invariance (ΔCFI < 0.01). The AUQEI model proved to be valid in relation to the external variables. Most children (76.7%) had positive subjective wellbeing. Higher scores were observed for items concerning recreation, holidays, and birthdays, and lower scores for those referring to hospitalization, medication, medical consultation, and being away from the family. The relationship between the demographic characteristics of the child or his/her mother and subjective wellbeing was not significant ($p > 0.05$).

Conclusions: The assessment of subjective wellbeing with the single-factor AUQEI model provided valid, reliable, and invariant. Thus, being a relevant and interesting instrument to assess wellbeing in young children.

KEYWORDS

subjective wellbeing, child wellbeing, children, preschool, psychometrics

1. Introduction

Subjective wellbeing is the self-assessment of life according to different personal criteria, and is based on a conscious complex and multidimensional cognitive judgment (1). In the search for definitions of wellbeing in the literature it is possible to find different ways of evaluating life or emotional experiences that involve feelings, living conditions, experiences, desire satisfaction, and the balance between pain and pleasure (1, 2). According to the

literature it is possible to find two different classifications for wellbeing, which are the hedonic and eudaimonic dimensions. Eudaimonic is driven by the cultivation of wellbeing based on long-term emotional processes, being used to refer to a combination of strengths of character encompassing facets of cooperativism (positive personal relationships), self-direction (life purpose, autonomy, environmental domain and self-acceptance) and self-transcendence (personal growth and self-realization) (3, 4). Hedonic is focused on the search for pleasure and happiness, being related to the experience of satisfaction, that is, what makes life pleasurable, the presence of positive affects and the absence of negative affects (3). Despite having different theoretical traditions, the constructs are closely related and influence each other (3).

Thus, instead of focusing in the difference between hedonic and eudaimonic components, these evaluations can be either in terms of affect (positive and negative) or cognitive reflections (5). Positive and negative affects are more related to the frequency with which people experience emotions than with the intensity of those emotions (6). Positive affects relate to pleasing emotions such as enthusiasm, energy, concentration, and satisfaction (7). On the other hand, negative affects are conditions of displeasure that involve mood swings, anger, and distress. The cognitive component of wellbeing is the self-assessment of life experiences, which is developed according to a set of self-imposed standards (7).

Children's Subjective wellbeing has been widely studied the past several decades (8). Some of the studies in children were carried out from the parents' perspective, as if children are unable to properly assess and understand issues related to their own lives (9), or were based on childhood sociology to predict future outcomes (10). However, considering children as present-moment members of the community moving toward future adults, with the right to immediate wellbeing to optimally develop their skills, studies using children-based data are relevant (10).

Childhood is a time of rapid change, and it is at this time that the trajectories of health and wellbeing are established for life, and that will impact adult life (11). Being a stage for the development of psychological difficulties and mental health problems (12, 13). The main challenge for the study of wellbeing in children and adolescents is to seek more sensitive assessment methods according to the stage of development in which the children are, therefore, the identification of characteristics that lead to the beginning, course and result requires projects that involve a younger age group to understand their vulnerabilities (13). Thus, wellbeing and quality of life measuring instruments for children must account for the stage of development of each age (14). An instrument should have different formats that are appropriate for the age group to be assessed, taking into account the normal age limitations and other factors, especially for very young or sick children who may have difficulty in providing accurate information (15). In addition, the response options should be age-appropriate (16). Some instruments use images so that children can more easily identify the best answer (17, 18).

Despite the importance of the knowledge of aspects that influence children's life according to their own point of view and their perception of interpersonal relationships, few instruments are available for measuring these aspects (19). Among them, the *Autoquestionnaire Qualité de Vie Enfant Imagé* (AUQEI) was initially proposed in French by Manificat and Dazord (20) and later translated into Portuguese (17) with the objective of assessing the subjective wellbeing of children between 4 and 12 years of age, based on the premise they are capable of expressing their own feelings. The instrument has been mostly used in children with health problems such as autism (21), cystic fibrosis (22), cerebral palsy (23, 24), spinal muscular atrophy (25), orofacial clefts (26), and born prematurely (27). However, no study has used confirmatory analysis to verify the validity of the data obtained with the AUQEI. The selection of the instrument for subjective wellbeing investigation in children must be guided not only by the pediatric context and the stage of development of the study population (17) but also by adequate estimates of validity and reliability of the data collected with the instrument in the sample to guarantee the quality of the obtained data (28).

The knowledge of this construct is relevant in order to allow for the early identification of children at-risk for poor wellbeing in community or educational settings (29). Therefore, to help support the improvement of child services and development of public policies, this study aimed at evaluating the psychometric properties of the AUQEI when applied to pre-school children and estimate the influence of demographic characteristics on their subjective wellbeing.

2. Methods

2.1. Ethical aspects

The main investigator of the study obtained agreement of the Children's Education and Recreation Centers (CER) to conduct the study and scheduled the application of the questionnaire to the participating mothers and the interviews with the children. The STROBE tool (Strengthening the Reporting of Observational studies in Epidemiology) was used to assist the study design and data reporting (30).

A signed informed consent form was obtained by all participants. This study was approved by the Research Ethics Committee. To be included in the study, children who agreed to participate had to provide a Consent Form signed by their parents. The study followed the ethical guidelines of the National Health Council Resolution 466/2012.

2.2. Study design and sampling

This was an observational, cross-sectional study. Preschool children (4–6 years old) enrolled in municipal public education institutions in Araraquara -São Paulo- Brazil were invited to participate in the study. The sample recruitment was done by the three-stage probabilistic method. In the first stage, the clusters were defined (considering the educational institution), in the second, the sample was stratified according to the number of preschoolers enrolled in the participating centers, and in the third stage, simple probabilistic sampling was performed.

The calculation of the minimum sample size was performed using $\alpha = 5\%$, $\beta = 20\%$, $\varepsilon = 10\%$, $N = 2.272$ (total number of

preschool children enrolled in CER), and the prevalence of positive subjective wellbeing of 50%, because a reference value for this parameter was not found in the literature. Thus, the minimum sample size estimated was 329 children, and with a 15% addition to account for a potential 15% loss rate, the final sample size was 388. This sample size was also sufficient to meet the statistical analysis requirements [58 parameters: 26 items, 26 errors, and six correlations between factors; considering the need for 5 subjects per parameter (31), $N = 290$].

2.3. Sample characterization

Demographic data (gender, age, education level, marital status, work activity of mothers and economic strata of the family members) were collected by questionnaires answered by the children's mothers. The economic stratum was estimated using the Brazilian Economic Classification Criteria (32), being the participants classified according to their economic strata [mean monthly income: low—C/D/E (U$ 175.63–735.50) and high—A/B (U$ 1,330.09–5,789.67)].

2.4. Measuring instrument

The Portuguese version of AUQEI (17) consists of 26 items distributed in four factors (Autonomy, Leisure, Function, and Family). To facilitate the understanding by, and application of, the instrument in children, the responses are represented by images of faces with different emotional states (ranging from very unhappy to very happy). The original version had a response scale with four points without a neutral point. However, in this study, we chose to use a 5-point scale with the inclusion of a neutral point. This choice was based on the fact that previous studies (33–35) indicated that the greater the number of response categories, the better the sensitivity of the items and the greater the probability to discriminate structurally different individuals. In addition, the addition of the neutral point (35) was carried out in order to make respondents more comfortable, as it is possible that for some AUQEI questions the respondent does not have an opinion or experience and, therefore, the answer neutral would be the most viable alternative (Figure 1).

2.5. Content validity

The content of the AUQEI items and the response scale were then evaluated by three experts in the areas of pediatrics, psychology and psychometry to estimate the instrument's content validity. The clarity of the items, their practical relevance, theoretical relevance and scope were evaluated, following the proposal of Hernández-Nieto (36). After establishing an absolute consensus among the experts, the AUQEI was applied to the target population in order to verify whether the sentences, instructions and the response scale were understandable for the children and what would be the best format for applying the response scale (Figure 1).

A pilot study was done with 25 children. Individual plastic cards with each of the five faces were presented to children on 3 sizes: 7, 10, and 15 cm long, all 7 cm wide. The cards were arranged sequentially, from the "very unhappy" to the "very happy". The children were asked about the preferred card size to answer each question, with 88% of children choosing the 10 cm card. During the interview, the researcher asked the children about the difficulties in understanding the content of the items and all the items were well-understood. The children answered each item without time-restriction and indicated the image (answer) that best represented their feeling. The mean duration of the interview with the children was 19.8 (SD = 1.60) minutes.

2.6. Evaluation of psychometric indicators

2.6.1. Internal data validity

Data were summarized by means, medians, and standard deviations and assessed for skewness and kurtosis. Strong deviations from normality were considered if skewness and kurtosis were above 3 and 7, respectively (37).

To evaluate the validity and reliability of the data, the total sample was randomly subdivided into two parts, "Test Sample" and "Validation Sample", and the psychometric properties of the AUQEI were evaluated for the two samples, separately.

The factorial validity was estimated using confirmatory factor analysis (CFA) with the Weighed Least Squares Mean and Variance Adjusted (WLSMV) estimation method. To assess the fit of the model to the data, the chi-square per degrees of freedom ratio (χ^2/df), the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI), and the Root Mean Square Error of Approximation (RMSEA) and local adjustment will also be considered based on the assessment of the factor loading of the items (λ) (28, 38, 39) were used. The fit was considered acceptable when $\chi^2/df \leq 5.0$, CFI and TLI ≥ 0.90, RMSEA ≤ 0.10 and factor loading (λ) ≥ 0.5 (28, 37, 38). If the model fit was not acceptable, the modification indices calculated using the Lagrange Multiplier (ML) method were observed. Besides the fitness evaluation of the tetra-factorial model (originally proposed for the AUQEI model), the Unidimensional Congruence (UniCo), Explained Common Variance (EVC) and the Mean of Item Residual Absolute Loadings (MIREAL) were assessed to verify the overall fit of the unidimensional model to the data. Values of UniCo > 0.95, EVC > 0.85, and MIREAL < 0.30 were considered to indicate an adequate fit of the unifactorial model to the data (40). The replicability of the model to future studies was estimated using the H-index (H-latent: assesses how well the factor can be identified by the continuous latent response variables that underlie the observed item scores; H-observed: assesses how well it can be identified from the observed item scores) (41). Values >0.80 indicate that the items adequately represent the factor and the structure has a high probability of being replicated (41).

Convergent validity was assessed based on the average variance extracted (AVE). AVE was estimated using the proposal by Fornell and Larcker (42) and was considered adequate if ≥ 0.50.

The MPLUS v.8.3 (Muthén and Muthén, Los Angeles, CA) and FACTOR (Lorenzo-Seva and Ferrando, Tarragona, Spain) programs were used to perform the analyses.

FIGURE 1
Images for response options of the AUQEI.

AUQEI's reliability was estimated from the ordinal alpha (α) and omega (ω) coefficients calculated with the "semTools" package (43) and "lavaan" (44) in the R program (45). The values α and ω ≥ 0.70 were indicators of acceptable internal consistency (31).

To assess whether the factors obtained were constant in independent samples, factorial invariance was carried out through multi-group analysis. The model for the "Test Sample" was compared with the model of the "Validation Sample". The CFI difference (ΔCFI) was used to compare factor loading (λ), thresholds (*t*), and variance/covariance of the residuals (Cov/Res). The CFI values of the configural models (M0), the factor loading model (M1), the threshold model (M2), and the residual model (M3) were considered (Metric invariance: M1 − M0; Scalar invariance: M2 − M1; and Strict invariance: M3 − M2) (46). Invariance between the models was confirmed when the CFI difference (ΔCFI) was <0.01 (47).

2.6.2. Validity in relation to external variables

The *Satisfaction with Life Scale* (SWLS) was used to assess the validity of the AUQEI model in relation to external variables, based on the responses of the children's mothers taking part in the study and on the children's ceo-d index. The SWLS was originally developed in English (48), characterized as a unifactorial scale consisting of five items according to a 7-points Likert scale (1: fully disagree to 7: fully agree). Both the AUQEI model and the SWLS should ideally be positively correlated, with the better the mother's wellbeing the greater the child's wellbeing (positive convergent validity). Negative convergent validity was estimated between the AUQEI model and the ceo-d index (decayed, indicated extraction or filled—ceo-d index) (49). It is anticipated that the AUQEI model and a health measure the ceo-d index present a negative and significant correlation, where the higher the ceo-d index the lower the child's wellbeing. The use of the SWLS and the ceo-d index to estimate the validity of the model was chosen out of convenience, as the present study is part of a wider project for which these variables had been collected. In the wider study, the SWLS was used to assess the subjective wellbeing of the children's mothers taking part in the present study, while the ceo-d index was used to measure the experience of caries.

2.6.3. Structural model and subjective wellbeing of preschool children

A structural model was made considering the impact of demographic variables on the AUQEI. The variables sex of the child (male or female), age of the mother, marital status (married or not married) and family economic strata [low - C/D/E (mean monthly income: U$ 175.63–735.50) and high - A/B (U$ 1,330.09–5,789.67] were included in the model.

The goodness of fit of this hypothetical model was evaluated on the polychoric correlation matrix using Weighed Least Squares Mean and Variance Adjusted (WLSMV) estimation, as implemented in the R program (45). The fitting of the structural model was evaluated using the previously cited indexes (χ^2/df, CFI, TLI, and RMSEA) (28). The trajectories (β) were estimated and evaluated with the *z*-test. A significance level of 5% (28).

In addition, the average scores for the general sample were calculated and an average subjective wellbeing ≥2.46 was considered positive subjective wellbeing, based on the suggestion of the instrument's author of using the 61.5 percentile (P61.5) as a reference for adequate subjective wellbeing. Based on this recommendation, the prevalence of positive subjective wellbeing (adequate) was estimated using a 95% confidence interval (95% CI). The authors of the scale recommend the assessment of scores using the sum. However, as psychometric instruments need to be fitted to the data and context under assessment and there was a risk of "losing items", we did not use the sum of items as score but the average, despite using the cutoff percentile suggested by the authors. The average scores were also calculated for each item of the instrument by a 95% confidence interval (95%CI). The level of significance was 5%.

3. Results

A total of 443 children participated in the study [mean age: 5.19 Standard deviation (SD = 0.64) years; 52.4% male]. The average age of the participants' mothers was 33.4 (SD = 7.01) years and most were married, had a job, and belonged to economic strata B and C. The demographic information of the sample is shown in Table 1.

The descriptive statistics of the answers given to the AUQEI items by the participants of the two samples (Test and Validation)

TABLE 1 Demographic characteristics of the study participants.

Characteristic	*n* (%)	Test (*n* = 212)	Validation (*n* = 231)
Children			
Sex			
Male	232 (52.4)	107 (50.5)	125 (54.1)
Female	211 (47.6)	105 (49.5)	106 (45.9)
Mothers			
Age (years)			
<30	163 (37.9)	85 (41.3)	78 (34.8)
≥30	267 (62.1)	121 (58.7)	146 (65.2)
Marital status			
Single	124 (28.6)	63 (30.4)	61 (26.9)
Married	276 (63.6)	129 (62.3)	147 (64.8)
Divorced	30 (6.9)	12 (5.8)	18 (7.9)
Widow	4 (0.9)	3 (1.4)	1 (4.0)
Work activity			
No	142 (32.4)	65 (31.3)	77 (33.5)
Yes	296 (67.6)	143 (68.8)	153 (66.5)
Economic strata (estimated mean family income)			
A (U$ 5,789.67)	23 (5.2)	8 (3.8)	15 (6.5)
B (U$ 1,330.09–2,575.89)	200 (45.1)	89 (42.0)	90 (39.0)
C (U$ 419.54–735.50)	205 (46.3)	110 (51.9)	116 (50.2)
D-E (U$ 175.63)	15 (3.4)	5 (1.4)	10 (4.3)

are shown in Table 2. As none of the AUQEI items presented absolute values of skewness > 3 and kurtosis > 7, no strong deviation from normality was assumed, and therefore the psychometric sensitivity of the items was considered adequate.

Despite the adequate fit of the original AUQEI model (complete - M) to the total sample, the factors Leisure and Autonomy [Variance Inflation Factor (VIF) = 166.91], Function and Autonomy (VIF = 12.75), Function and Leisure (VIF = 125.25), Family and Autonomy (VIF = 55.80), Family and Leisure (VIF = 38.21), and Family and Function (VIF = 100.25) were collinear. In addition, a high correlation was found between the errors of items 2 and 3 of the instrument, which led to the non-convergence of the covariance matrix. Thus, each item (2 and 3) was excluded individually and then both were excluded, however, this strategy did not favor the matrix's convergence. Collinearity may suggest that the AUQEI applied to this study sample has a single factor model and the adequacy verification indices of the one-dimensional model reinforce this suggestion (UniCo = 0.995; ECV = 0.946; MIREAL = 0.161), therefore, this proposal was tested. In the single factor model, was observed high correlation between items 2 and 3, we excluded item 2, based on the modification indices and the theoretical content of the item. The single-factor model showed adequate fit to the sample as well as convergent validity and adequate reliability both in the total sample (λ = 0.71–0.92; χ^2/df = 4.47; CFI = 0.98; TLI = 0.98; RMSEA = 0.08; AVE = 0.70; α = 0.98; ω = 0.93) and in the Test and Validation samples (Figure 2). This model has a high probability of being applied in other studies (H-Latent = 0.984; H-Observed = 0.946).

In the analysis performed on independent samples (Test, n = 212 × Validation, n = 231), strict invariance was observed ($\Delta CFI_{M1-M0} = 0.001$; $\Delta CFI_{M2-M1} = 0.000$; $\Delta CFI_{M3-M2} = 0.001$) indicating that the factorial structure found remains in independent samples. The structural model tested presented adequate fit to the sample (χ^2/df = 3.58; CFI = 0.98; TLI = 0.99; RMSEA = 0.07). The child's sex, the mother's age, the family's economic level, the fact that the mother is or is not married, and the exercise of work activity did not impact the children's wellbeing (Table 3).

The correlational analysis between the AUQEI and the SWLS ($r = 0.70$, $p < 0.001$) pointed to an adequate positive convergent validity of the AUQEI, while the correlation between the AUQEI and the ceo-d index ($r = -0.36$, $p < 0.001$) to attest negative convergent validity of the AUQEI model.

Most children [76.7%; (95% CI: 72.7–80.5%)] had positive subjective wellbeing. The mean scores of the items from children ($n = 443$) and the baseline (P61.5; mean score = 2.46) are shown in Figure 3.

Higher scores of subjective wellbeing were given for items concerning recreation, vacations, and birthdays, and lower scores referred to hospitalization, medication, medical consultation, and being away from the family. This corroborates the theoretical framework of the instrument, pointing toward adequate levels of subjective wellbeing. Only item 14 (hospitalization) presented a lower score than the baseline.

4. Discussion

The present study confirmed the validity and reliability of subjective wellbeing data of preschool children obtained with the AUQEI. The found wellbeing estimates highlight the feasibility and importance of assessing subjective wellbeing of young children and expand the knowledge required for developing educational programs for this population.

In recent years, the study of children subjective wellbeing has increased (17, 20, 23), however, it is still a challenge to choose the best measurement tool to assess the wellbeing, especially in young children. Although the AUQEI is an interesting instrument to assess wellbeing in children (17, 20), there were still no studies that supported the validity and reliability of the measure obtained with this instrument. Some important aspects in the validation process, such as item analysis, psychometric sensitivity, and factorial validity, were not performed, which may compromise the conclusions drawn from previous studies.

The AUQEI factorial model that fitted the data adequately was the single factor structure. To obtain the best fit, a new theoretical proposal different from the original one was used. The validation and reliability analysis of the original (four-factor) and the single-factor versions were conducted, and a high correlation among the factors of the original structure was found, which compromises the variances of the parameter estimates. When at

10.3389/fpubh.2023.1156755

TABLE 2 Descriptive statistics of the Autoquestionnaire Qualité de Vie Enfant Imagé (AUQEI) responses by the participants (test sample $N = 212$ and validation sample $N = 231$).

Item	Test sample/validation					
	Mean	Standard deviation	Skewness	Kurtosis	Minimum	Maximum
Say how you feel:						
It1. at the table, with your family	2.99/3.05	0.98/0.94	−0.58/−0.68	−0.74/−0.47	1/1	4/4
It2. at night when you lie down	2.42/2.44	1.00/1.01	−0.59/−0.20	−1.00/−0.84	0/0	4/4
It3. if you have siblings, when you play with them	3.12/3.15	0.83/0.86	−0.64/−0.78	−0.20/0.29	0/0	4/4
It4. at night, when sleeping	2.39/2.45	1.03/0.99	−0.04/−0.18	−1.00/−0.86	0/0	4/4
It5. in the classroom	2.84/2.86	0.88/0.85	−0.36/−0.20	−0.58/−0.78	1/1	4/4
It6. when you see a picture of yourself	2.97/3.00	0.80/0.80	−0.21/−0.19	−0.80/−0.95	1/1	4/4
It7. at play time, during school recess	3.51/3.48	0.56/0.58	−0.76/−0.71	−0.50/0.30	1/1	4/4
It8. when you go to a doctor's appointment	1.83/1.90	1.27/1.16	−0.04/−0.05	−1.22/−1.09	0/0	4/4
It9. when you play a sport	3.27/3.32	0.78/0.77	−0.89/−0.84	0.30/−0.10	1/1	4/4
It10. when you think of your father	2.96/3.08	1.07/0.97	−0.81/−0.72	−0.22/−0.45	0/0	4/4
It11. on your birthday	3.47/3.46	0.67/0.67	−1.26/−0.95	1.87/0.07	1/1	4/4
It12. when you do your homework	2.85/2.94	0.95/0.93	−0.47/−0.42	−0.22/−0.66	0/0	4/4
It13. when you think of your mother	3.31/3.29	0.76/0.79	−1.31/−1.06	2.82/1.41	0/0	4/4
It14. when you are admitted to the hospital	1.02/1.10	0.89/0.92	0.27/0.54	−0.90/−0.32	0/0	4/4
It15. when you play alone	2.18/3.16	1.23/1.18	−0.41/0.12	−1.35/−1.35	0/0	4/4
It16. when your dad or mom talk about you	2.72/3.72	1.20/1.14	−0.57/−0.64	−0.74/−0.45	0/0	4/4
It17. during sleepovers	2.46/3.46	1.24/1.25	−0.49/−0.52	−0.87/−0.86	0/0	4/4
It18. when someone asks you to show something you know how to do	3.07/4.03	0.86/0.87	−0.80/−0.60	0.76/0.01	0/0	4/4
It19. when friends talk about you	2.74/3.73	1.12/1.10	−0.66/−0.72	−0.53/−0.40	0/0	4/4
It20. when you take medicines	1.64/2.59	1.02/1.10	0.26/0.24	−0.37/−0.56	0/0	4/4
It21. during the holidays	3.41/4.36	0.78/0.77	−1.11/−1.10	0.36/0.60	1/1	4/4
It22. when you think of being a grown-up	3.11/4.13	0.82/0.83	−0.57/−0.58	−0.08/−0.24	0/0	4/4
It23. when you are away from your family	2.02/2.95	1.11/1.10	0.30/0.39	−0.76/−0.66	0/0	4/4
It24. when you get grades from school	3.11/4.10	0.83/0.89	−0.35/−0.56	−1.09/−0.70	1/1	4/4
It25. when you are with your grandparents	3.42/4.44	0.67/0.69	−0.80/−1.07	−0.30/0.77	1/1	4/4
It26. when you watch television	3.54/4.53	0.58/0.62	−0.83/−0.99	−0.29/−0.06	2/2	4/4

least one of the variables is redundant, an estimate of negative variance occurs (28, 37). Therefore, to avoid problems in model estimation, we tested the single-factor model, which adequately fitted the data. This proposal also provided adequate validity of the AUQEI in relation to external variables, ensuring, therefore, three measures of validity to the instrument (content, internal structure and external variables). This finding should raise the alert for future users of the instrument about the importance to obtain estimates of adequate quality for each sample and study context. The process of adapting an instrument for a population is only completed after confirmation of validity and reliability applied to different samples (28, 37). The comparison of the present findings with those of others was not possible since no previous study verified the fit of the AUQEI factor model to other samples using a confirmatory analysis.

Item 2 ("Say how you feel: at night, when you lie down") presented a high correlation with item 3 ("Say how you feel: if you have siblings, when playing with them"), and we found that item 2 had a high theoretical similarity to item 4 ("Say how you feel: at night, when sleeping") causing collinearity, which may have resulted in the high correlation between items 2 and 3. Considering that such young children, in our sample, could not differentiate the concepts in items 2 and 4, item 2 was excluded. Evidence from the literature on child development and psychology shows that children under 6 years of age have a limited distinction of language and verbalization of emotions, especially when referring to past

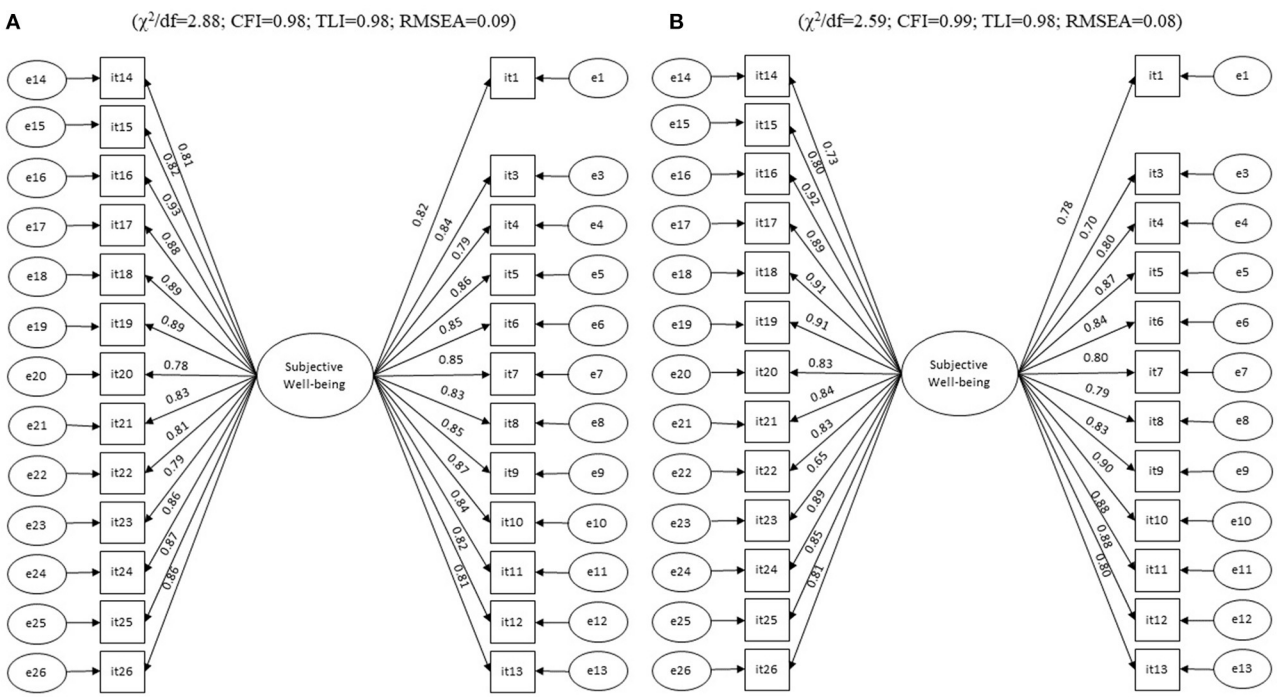


FIGURE 2
Factorial model fitted for the test sample (n = 212) **(A)** and validation sample (n = 231) **(B)** of preschool children.

events, and abstract thinking capacity develops after the age of 6 (50, 51). Future studies that include children from different stages of development are suggested to assess the best theoretical/factorial proposal for the AUQEI in each of these phases.

The strong invariance of the AUQEI observed between independent samples confirmed the stability of the single-factor model applied to young children. Thus, a valid and reliable set of data obtained with the AUQEI model is provided.

Currently, investigation of wellbeing has shifted, aiming at individuals without a specific complaint or disease, focusing on promotion of wellbeing and not just its impair (52). However, most of the studies are carried out on children with health problems (20, 23, 24). So far, only the study by Assumpção et al. (17) was performed in healthy children; however, the authors did not present evidence related to the validity of the proposed model. In addition, their model was developed for a sample of children at different stages of development (4–12 years old), which may explain the difference between the data obtained in the present study in relation to the factorial model. In the study by Assumpção et al. (17) only 49 children were aged 4–6 years and, therefore, their data may have a low contribution in the 4-factor model presented. Thus, the present study contributes to the field of study by providing a specific model for children under 6 years of age; further studies may confirm or contest our findings. In addition, future studies aiming at defining the best model for assessing subjective wellbeing using the AUQEI in children at different stages of development are recommended.

No difference in the scores of subjective wellbeing was found between mothers' age, marital status, work activity, and income (see structural model Table 3). Although income has a role in wellbeing, it is not its main indicator (19), and other conditions such as culture and social aspects should be considered in the relationship between income and wellbeing (53). The lack of an effect of the above variables in the responses shows that children do not have a clear perception of the interference of these conditions in their daily life, and provide a more self-focused assessment, based on the effects of their own world on the daily routine, without considering broader aspects.

TABLE 3 Structural model for assessing demographic contribution to children's subjective wellbeing (AUQEI).

Pathway	β	β_S	SE	*p*
Sex → AUQEI	−0.019	−0.010	0.104	0.520
Age → AUQEI	0.002	−0.013	0.008	0.803
Work activity → AUQEI	0.201	0.092	0.116	0.082
Marital status → AUQEI	−0.115	−0.054	0.111	0.300
Economic strata → AUQEI	0.113	0.056	0.108	0.297

βS, β standardized; SE, standard error.

Children presented a greater wellbeing score for items related to recreation, vacations, and birthdays, a finding that corroborates with previous studies (17, 54). Activities that break the routine and provide fun moments, such as free-time activities, are valued in childhood, being significant for wellbeing, especially among younger children (19). In contrast, the items with the lowest scores were those that refer to hospitalization and being away from the family, which is in line with the work of Assumpção et al. (17)

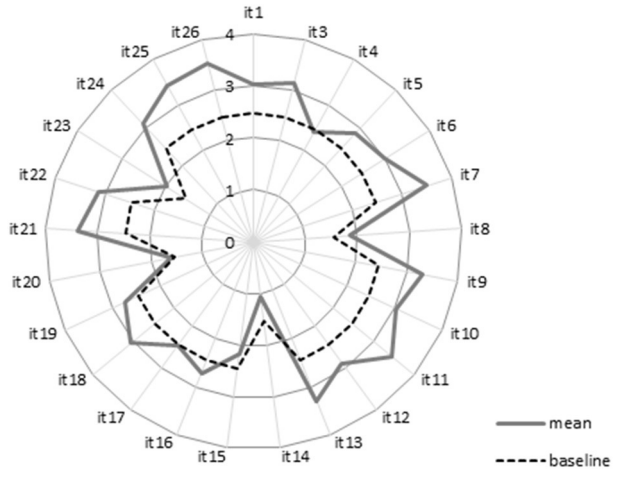


FIGURE 3
Mean item score for the AUQEI given by children.

not only in relation to the items, but also with regard to the overall positive wellbeing for most children. Despite presenting lower scores, the average subjective wellbeing was positive for most participants, which is explained by the concept of subjective wellbeing being a balance between positive and negative emotions, which not always is affected by some conditions, be they material, health-related, financial, or security-related. The influence of these factors depends on expectations and values of each person, the people around him or her, and the community in which he or she lives. Therefore, subjective wellbeing involves a global judgment of all aspects of life, and although some conditions can affect wellbeing, the emphasis is usually placed on the general judgment of a person's life (55). Despite the high prevalence rate of subjective wellbeing found in the sample, approximately one-fourth of the sample presented lower values than expected. Thus, further monitoring and investigation of the factors involved is suggested.

As limitation of the study, the data were collected in a single national context, namely, Brazil. To broaden the scope of research, it is recommended that further studies be conducted using samples from different national contexts. While the main factorial structure is expected to be consistent, cross-cultural comparisons may unveil variations in factor loadings and associations with other variables, fostering discussions on sociocultural factors' role in children's wellbeing. In addition, the cross-sectional design of the study does not allow the establishment of a cause-and-effect relationship. Despite this limitation, this study provides information about the psychometric properties of the AUQEI model for pre-school children and identifies the subjective wellbeing of preschool children, which can contribute to health professionals and researchers from different areas of knowledge in guiding and developing preventive strategies focusing not only on treating health problems, but also in promoting wellbeing.

The AUQEI model provided valid and reliable data, being thus an interesting instrument to assess the wellbeing of young children. This assessment can broaden the view of education and health professionals in order to better estimate, re-establish and monitor the wellbeing of children. The demographic characteristics of the mother or family analyzed in the present study, or even the child's sex, had no significant impact on the children's wellbeing. The study observed a high prevalence of children reporting a high subjective wellbeing.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by Research Ethics Committee of São Paulo State University (UNESP). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

BS and BM involved in conception and design, data acquisition, analysis and interpretation, and drafted and critically revised the manuscript. JM, JC, and LC involved in conception and design, data interpretation, and drafted the manuscript. All authors contributed to the article and approved the submitted version.

Funding

The work was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—grant Code 001 and the São Paulo Research Foundation (FAPESP) (grants # 2019/17200-9 and # 2019/24424-0).

Conflict of interest

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

Publisher's note

References

1. Cross MP, Hofschneider L, Grimm M, Pressman SD. Subjective well-being and physical health. In: Diener E, Oishi S, Tay L, editors. *Handbook of Well-Being*. Champaign, IL: DEF Publications (2018).

2. Diener E, Pressman SD, Hunter J, Delgadillo-Chase D. If, why, and when subjective well-being influences health, and future needed research. *Appl Psychol Health Wellbeing*. (2017) 9:133–67. doi: 10.1111/aphw.12090

3. Giuntoli L, Vidotto G. Exploring Diener's multidimensional conceptualization of well-being through network psychometrics. *Psychol Rep*. (2020). 124:896–919. doi: 10.1177/0033294120916864

4. Karunamuni N, Weerasekera R. Theoretical foundations to guide mindfulness meditation: a path to wisdom. *Curr Psychol*. (2019) 38:627–46. doi: 10.1007/s12144-017-9631-7

5. Ryff C. Eudaimonic well-being: highlights from 25 years of inquiry. In: Shigemasu K, Kuwano S, Sato T, Matsuzawa T, editors. *Diversity in Harmony - Insights From Psychology: Proceedings of the 31st International Congress of Psychology*. Chichester: John Wiley & Sons Ltd. (2018). p. 375–95.

6. Kobau R, Seligman ME, Peterson C, Diener E, Zack MM, Chapman D, et al. Mental health promotion in public health: perspectives and strategies from positive psychology. *Am J Public Health*. (2011) 101:e1–9. doi: 10.2105/AJPH.2010.300083

7. Gasparetto LG, Bandeira C, Giacomoni CH. Subjective well-being and personality traits in children: a possible relationship? *Temas Psicol*. (2017) 25:447–57. doi: 10.9788/TP2017.2-03

8. Newland LA, Giger JT, Lawler MJ, Roh S, Brockevelt BL, Schweinle A. Multilevel analysis of child and adolescent subjective well-being across 14 countries: child- and country-level predictors. *Child Dev*. (2019) 90:395–413. doi: 10.1111/cdev.13134

9. Soares AH, Martins AJ, Lopes MC, Britto JA, Oliveira CQ, Moreira MC. Quality of life of children and adolescents: a bibliographical review. *Ciência Saúde Coletiva*. (2011) 16:3197–206. doi: 10.1590/S1413-81232011000800019

10. Ben-Arieh A, Casas F, Frønes I, Korbin JE. Multifaceted concept of child well-being. *Handb Child Wellbeing*. (2014) 1:1–27. doi: 10.1007/978-90-481-9063-8_134

11. Daelmans B, Darmstadt GL, Lombardi J, Black MM, Britto PR, Lye S, et al. Early childhood development: the foundation of sustainable development. *Lancet*. (2017) 389:9–11. doi: 10.1016/S0140-6736(16)31659-2

12. Dalsgaard S, Thorsteinsson E, Trabjerg BB, Schullehner J, Plana-Ripoll O, Brikell I, et al. Incidence rates and cumulative incidences of the full spectrum of diagnosed mental disorders in childhood and adolescence. *JAMA Psychiatry*. (2020) 77:155–64. doi: 10.1001/jamapsychiatry.2019.3523

13. Ortuño-Sierra J, Aritio-Solana R, Fonseca-Pedrero E. New evidences about subjective well-being in adolescence and its links with neurocognitive performance. *Int J Environ Res Public Health*. (2020) 17:1866. doi: 10.3390/ijerph17061866

14. Matza LS, Swensen AR, Flood EM, Secnik K, Leidy NK. Assessment of health-related quality of life in children: a review of conceptual, methodological, and regulatory issues. *Value Health*. (2004) 7:79–92. doi: 10.1111/j.1524-4733.2004.71273.x

15. Reppold CT, Serafini AJ, Ramires DA, Gurgel LG. Análise dos manuais psicológicos aprovados pelo SATEPSI para avaliação de crianças e adolescentes no Brasil. *Aval Psicol*. (2017) 16:19–28. doi: 10.15689/ap.2017.1601.03

16. Buczynski AK, Castro GF, de Souza IPR. The impact of oral health on the quality of life of HIV infected children: a literature review. *Ciencia Saude Coletiva*. (2008) 13:1797. doi: 10.1590/S1413-81232008000600014

17. Assumpção JR FB, Kuczynski E, Sprovieri MH, Aranha EMG. Escala de avaliação de qualidade de vida (AUQEI - autoquestionnaire qualité de vie enfant imagé): validade e confiabilidade de uma escala para qualidade de vida em crianças de 4 a 12 anos. *Arq Neuropsiquiatr*. (2000) 58:119–27. doi: 10.1590/S0004-282X2000000100018

18. Gherunpong S, Tsakos G, Sheiham A. Developing and evaluating an oral health-related quality of life index for children; the CHILD-OIDP. *Commun Dent Health*. (2004) 21:161–9.

19. Giacomoni CH, de Souza LK, Hutz CS. Children's view on happiness. *Psicol Escolar Educ*. (2014) 18:143–50. doi: 10.1590/S1413-85572014000100015

20. Manificat S, Dazord A. Évaluation de la qualité de vie de l'enfant: validation d'un questionnaire, premiers résultats. *Neuropsychiatr l'Enfance l'Adolesc*. (1997) 45:106–14.

21. Elias A, Assumpção Jr F. Quality of life and autism. *Arq Neuropsiquiatr*. (2006) 64:295–9. doi: 10.1590/S0004-282X2006000200022

22. Pfeifer LI, Silva MA. Assessmente of quality of life in children with cystic fibrosis. *Rev Nufen*. (2009) 1:118–30.

23. Barbosa-Resende W, de Oliveira Rangel V, Frontarolli AC, Araújo RRH, da Silva CHM, Pinto R, et al. Psychometric properties of the Autoquestionnaire Qualite de Vie Enfant Image (AUQEI) applied to children with cerebral palsy. *PLoS ONE*. (2015) 10:115643. doi: 10.1371/journal.pone.0115643

24. Borges MBS, Werneck MJS, Silva ML, Gandolfi L, Pratesi R. Therapeutic effects of a horse riding simulator in children with cerebral palsy. *Arq Neuropsiquiatr*. (2011) 69:799–804. doi: 10.1590/S0004-282X2011000600014

25. de Oliveira CM, Araujo AP. Self-reported quality of life has no correlation with functional status in children and adolescents with spinal muscular atrophy. *Eur J Paediatr Neurol*. (2011) 15:36–9. doi: 10.1016/j.ejpn.2010.07.003

26. Tannure P, Soares F, Küchler E, Motta L, Costa M, Granjeiro J. Measuring the impact of quality of life of children treated for orofacial clefts: a case-control study. *J Clin Pediatr Dentist*. (2013) 37:381–4. doi: 10.17796/jcpd.37.4.f185423j04m66753

27. Martini JA, Perosa GB, Padovani FHP. Qualidade de vida de escolares nascidos prematuros, o relato do cuidador e o auto-relato infantil. *Ciência Saúde Coletiva*. (2019) 24:4699–706. doi: 10.1590/1413-812320182412.18062017

28. Marôco J. *Análise de Equações Estruturais*. 2nd ed. Lisboa: Report Number (2014).

29. WHO (2018). *Copenhagen Consensus of Mayors–Healthier and Happier Cities for All: A Transformative Approach for Safe, Inclusive, Sustainable and Resilient Societies*. Copenhagen: WHO Regional Office for Europe: World Health Organisation Copenhagen.

30. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol*. (2008) 61:344–9. doi: 10.1016/j.jclinepi.2007.11.008

31. Hair JF, Black WC, Babin BJ, Anderson RE. *Multivariate Data Analysis*. Hampshire: Cengage Learning (2019).

32. ABEP. *Critério de Classificação Econômica Brasil Brazilian Criteria 2018*. (2019). Available from: http://www.abep.org/criterio-brasil (acessed April 3, 2023).

33. Carifio J, Perla R. Resolving the 50-year debate around using and misusing Likert scales. *Med Educ*. (2008) 42:1150–2. doi: 10.1111/j.1365-2923.2008.03172.x

34. Churchill Jr GA, Peter JP. Research design effects on the reliability of rating scales: a meta-analysis. *J Market Res*. (1984) 21:360–75. doi: 10.1177/002224378402100402

35. Coelho PS, Esteves SP. The choice between a five-point and a ten-point scale in the framework of customer satisfaction measurement. *Int J Mark Res*. (2007) 49:313–39. doi: 10.1177/147078530704900305

36. Hernández-Nieto RA. *Contributions to Statistical Analysis*. Mérida: Universidad de Los Andes (2002) 193 p.

37. Kline RB. *Principles and Practice of Structural Equation Modeling*. New York, NY: The Guilford Press (2015).

38. Brown TA. *Confirmatory Factor Analysis for Applied Research*. New York, NY: The Guilford Press (2015).

39. Byrne BM. *Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming (Multivariate Applications Series)*. New York, NY: Taylor and Francis Group (2010).

40. Ferrando PJ, Lorenzo-Seva U. Assessing the quality and appropriateness of factor solutions and factor score estimates in exploratory item factor analysis. *Educ Psychol Meas*. (2018) 78:762–80. doi: 10.1177/0013164417719308

41. Hancock GR, Mueller RO. Rethinking construct reliability within latent variable systems. In: Cudeck R, Toit S, Sorbom D, editors. *Structural Equation Modeling: Present and Future—A Festschrift in Honor of Karl Joreskog*. Lincolnwood, IL: Scientific Software International (2001). p. 195–216.

42. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. *J Market Res*. (1981) 18:39–50. doi: 10.1177/002224378101800104

43. Jorgensen TD, Pornprasertmanit S, Schoemann SAM, Rosseel Y. *Sem Tools: Useful Tools for Structural Equation Modeling*. (2022). Available online at: https://cran.r-project.org/web/packages/semTools/index.html (accessed November 2, 2022).

44. Rosseel Y. Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *J Stat Softw*. (2012) 48:1–36. doi: 10.18637/jss.v048.i02

45. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. (2022). Available online at: https://www.R-project.org/ (accessed November 2, 2022).

46. Wu H, Estabrook R. Identification of confirmatory factor analysis models of different levels of invariance for ordered categorical outcomes. *Psychometrika*. (2016) 81:1014–45. doi: 10.1007/s11336-016-9506-0

47. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Eq Model.* (2002) 9:233–55. doi: 10.1207/S15328007SEM0902_5

48. Diener E, Emmons RA, Larsen RJ, Griffin S. The satisfaction with life scale. *J Pers Assess.* (1985) 49:71–5. doi: 10.1207/s15327752jpa4901_13

49. WHO. Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychol Med*. (1998) 28:551–8. doi: 10.1017/S0033291798006667

50. Hetherington EM, Parke RD, Locke VO. *Child Psychology: A Contemporary Viewpoint*. Boston, MA: McGraw-Hill (1999).

51. Rebok G, Riley A, Forrest C, Starfield B, Green B, Robertson J, et al. Elementary school-aged children's reports of their health: a cognitive interviewing study. *Qual Life Res.* (2001) 10:59–70. doi: 10.1023/A:1016693417166

52. Ben-Arieh A. The child indicators movement: past, present, and future. *Child Indic Res.* (2008) 1:3–16. doi: 10.1007/s12187-007-9003-1

53. Machado L. *Subjective Well-being: Implications for Psychiatry and Medical Psychology*. 1st ed. Rio de Janeiro: Medbook (2017).

54. Rezende BA, Lemos SMA, de Medeiros AM. Quality of life and health self-perception in children with poor school performance. *Rev Paul Pediatr*. (2017) 35:415. doi: 10.1590/1984-0462/;2017;35;4;00009

55. Diener E. New findings and future directions for subjective well-being research. *Am Psychol.* (2012) 67:590. doi: 10.1037/a0029541

frontiers | Frontiers in Pediatrics

TYPE Original Research
PUBLISHED 09 May 2023
DOI 10.3389/fped.2023.1135415

OPEN ACCESS

EDITED BY
Tim S. Nawrot,
University of Hasselt, Belgium

REVIEWED BY
Doina Anca Plesca,
Carol Davila University of Medicine and Pharmacy, Romania
Arlette Setiawan,
Padjadjaran University, Indonesia

*CORRESPONDENCE
Caroline Jung-Sievers
✉ cjungsievers@ibe.med.uni-muenchen.de

†These authors share last authorship

SPECIALTY SECTION
This article was submitted to Children and Health, a section of the journal Frontiers in Pediatrics

RECEIVED 31 December 2022
ACCEPTED 29 March 2023
PUBLISHED 09 May 2023

CITATION
Schillok H, Coenen M, Rehfuess EA, Kuhlmann PH, Matl S, Kindermann H, Maison N, Eckert J, von Both U, Behrends U, Frühwald MC, Neubert A, Woelfle J, Melter M, Liese J, Hübner J, Klein C, Kern A and Jung-Sievers C (2023) Changes in behavior and quality of life in German young children during the COVID-19 pandemic—results from the COVID kids bavaria study.
Front. Pediatr. 11:1135415.
doi: 10.3389/fped.2023.1135415

Changes in behavior and quality of life in German young children during the COVID-19 pandemic—results from the COVID kids bavaria study

Hannah Schillok[1,2], Michaela Coenen[1,2], Eva A. Rehfuess[1,2], Pia H. Kuhlmann[3], Stefan Matl[3], Hannah Kindermann[3], Nicole Maison[3,4,5], Jana Eckert[3], Ulrich von Both[3,6], Uta Behrends[7], Michael C. Frühwald[8], Antje Neubert[9], Joachim Woelfle[9], Michael Melter[10], Johannes Liese[11], Johannes Hübner[3], Christoph Klein[3], Anna Kern[3†] and Caroline Jung-Sievers[1,2*†]

[1]Institute for Medical Information Processing, Biometry, and Epidemiology (IBE), Chair of Public Health and Health Services Research, LMU Munich, Munich, Germany, [2]Pettenkofer School of Public Health, Munich, Germany, [3]Department of Pediatrics, Dr. von Hauner Children's Hospital, University Hospital LMU Munich, Munich, Germany, [4]German Center for Lung Research (DZL), Munich, Germany, [5]Institute for Asthma- and Allergy Prevention (IAP), Helmholtz Center Munich, German Research Center for Environmental Health (GmbH), Munich, Germany, [6]German Center for Infection Research (DZIF), Partner Site Munich, Munich, Germany, [7]Department of Pediatrics, Kinderklinik München Schwabing, StKM GmbH und Klinikum Rechts der Isar, Technische Universität München, Munich, Germany, [8]Paediatric and Adolescent Medicine University Hospital Augsburg, Augsburg, Germany, [9]Department of Paediatric and Adolescent Medicine, Erlangen University Hospital, Erlangen, Germany, [10]University Children's Hospital Regensburg (KUNO), Regensburg, Germany, [11]Department of Pediatrics, University Hospital Würzburg, Würzburg, Germany

Introduction: The COVID-19 pandemic with its containment measures such as closures of schools and daycare facilities led to numerous restrictions in daily life, putting developmental opportunities and health-related quality of life in children at risk. However, studies show that not every family was impacted equally by the pandemic and that this exceptional health and societal situation reinforced pre-existing health inequalities among the vulnerable. Our study aimed at analyzing changes in behavior and health-related quality of life of children attending elementary schools and daycare facilities in Bavaria, Germany in spring 2021. We also sought to identify associated factors contributing to inequalities in quality of life.
Methods: Data from a multi-center, open cohort study ("COVID Kids Bavaria") conducted in 101 childcare facilities and 69 elementary schools across all electoral districts of Bavaria were analyzed. Children attending these educational settings (aged 3-10 years) were eligible for participation in a survey on changes in behavior and health-related quality of life. The KINDL^R questionnaire (based on children's self-report and parental report) was administered about one year after the onset of the pandemic (spring 2021). Descriptive and logistic regression analyses and comparisons to pre-pandemic KiGGS (German Health Interview and Examination Survey for Children and Adolescents) data were undertaken.
Results: Among respondents, a high percentage of parents reported changes in their children's eating and sleeping behavior, sports and outdoor activities as well as altered screen time. Health-related quality of life in KINDL^R analyses compared to pre-pandemic population averages were lower in all age groups (for 3–6-year-old KINDL^R-total score: COVID Kids Bavaria MD 74.78 ± 10.57 vs KiGGS data 80.0 ± 8.1; 7-10 years-old KINDL^R-total score: COVID Kids Bavaria MD 73.88 ± 12.03 vs KiGGS data 79.30 ± 9.0). No significant differences were detected with regard to

associated factors, namely type of institution, sex of the child, migration background, household size and parental education.

Conclusion: These findings suggest a relevant impact of the COVID-19 pandemic on children's behavior and health-related quality of life one year after the onset of the pandemic. Further analyses in large-scale longitudinal studies are needed to determine the effects of specific pandemic or crisis associated factors contributing to health inequalities.

KEYWORDS

COVID-19, pandemic, mental health—related quality of life, behavior change, health-related quality of life (HRQL), health inequalities in children

1. Introduction

The global spread of the SARS-CoV-2 virus can be seen as a worldwide stress test for populations around the world, but especially for vulnerable groups. Even if children are only mildly affected by a SARS-CoV-2 infection in general, they represent a particularly vulnerable group with regards to the consequences of mitigation measures taken against the spread of the virus during different stages of the pandemic (1, 2). Being in a critical developmental phase, they were at times deprived of educational opportunities and social contacts by the closure or reduced services of schools and childcare institutions. At the same time, they were confronted with stressors such as isolation, stressed parents due to potentially insecure social and economic situations, and loss of daily structure through restricted leisure activities. In addition, they may have had less resilience capacity due to their lack of life experience compared to adults.

During the pandemic, the mental health of children has been compromised as demonstrated by many international studies and systematic reviews (3–10). A meta-analysis of studies on the prevalence of depressive and anxiety symptoms in children and adolescents during the pandemic demonstrates a two-fold increase in these symptoms (4). Certain groups, such as children with preexisting (chronic) physical or mental health conditions and lower socioeconomic family resources suffer disproportionately (11).

After a rapid increase in the first phase of the pandemic, stress and psychosocial symptoms have remained at high and relatively stable levels throughout the pandemic. Ravens-Sieberer et al. reported in the COPSY (longitudinal COVID-19 and Psychological Health) study that two-thirds of children and adolescents were highly burdened by the COVID-19 pandemic with lower health-related quality of life (QoL), more mental health problems and worsened behavior compared to pre-pandemic times. However, levels of mental health problems between the first wave in spring 2020 and the second wave in winter 2020/2021 did not differ significantly (12–16). Analyses of the German serial cross-sectional "Corona Snapshot Monitoring" (COSMO) study showed that about a third of children suffered from mental health problems according to the Strengths and Difficulties Questionnaire (SDQ) proxy-report by their parents (17). This was again higher than pre-pandemic levels while no significant differences were observed between different waves.

An additional stressing factor for children during the pandemic might stem from their parents, who themselves have been highly burdened at the same time, posing another risk factor for children's wellbeing (18). The mental health of parents and of their children are closely intertwined and the risk of violence against children increases through a burdened family environment (19). Against this backdrop, it is particularly important to monitor children's health-related QoL and psychological status and to identify how they can be strengthened and protected during the present and future crises.

Our study, a sub-study embedded in the larger COVID Kids Bavaria study, thus sought to analyze changes in behavior and health-related quality of life of children attending elementary schools and daycare facilities in Bavaria, Germany in spring 2021. More specifically, we investigated the following aspects: (a) behavior change during the pandemic compared to pre-pandemic behavior, (b) health-related QoL compared to pre-pandemic reference values among a nationally representative sample of healthy children, (c) health-related QoL comparing children's self-assessment with their parents' assessment, (d) determining factors for health inequalities in children during the pandemic, depicted as differences in QoL.

We hypothesized that children in our sample would demonstrate decreased health-related QoL as well as altered, presumably negatively, changed behavior patterns. We also hypothesized that children of different socioeconomic backgrounds would be differentially affected. The COVID Kids Bavaria main study question investigated the occurrence of asymptomatic cases of SARS-CoV-2 infections among children and staff in daycare facilities and elementary schools in Bavaria. Results of this SARS-CoV-2 infection surveillance were previously published (20).

2. Methods

COVID Kids Bavaria is a multi-center, open cohort study covering every electoral district in Bavaria, Germany. It was registered with the German Clinical Trials Register (http://www.drks.de/DRKS00022380) and approved by six local ethics committees. Methods described in this report focus on the survey component of the COVID Kids Bavaria study, whereas methods of the SARS-CoV-2 infection surveillance component of the study are reported elsewhere (20).

2.1. Population

Six study centers (Augsburg, Erlangen, Regensburg, Würzburg, Technische Universität München, Ludwig-Maximilians-Universität München) collected data in all 46 Bavarian electoral districts with selected facilities evenly distributed among them (20). Initially 149 facilities were enrolled (101 childcare facilities and 49 elementary schools) in September 2020. 147 facilities were visited by study teams. Towards the end of the study additional 20 elementary schools were recruited for administering questionnaires to further increase the study population. The regional distribution of the participating institutions is displayed in **Supplementary Figure S1**.

Participating families were eligible for enrollment if they met the following inclusion criteria: child attending the enrolled schools or daycare facilities (aged between 1 and 10 years) and written informed consent provided by parents or legal representatives. Potential participants were approached from September 2020 up to May 2021 in three consecutive phases. Eligibility to participate in the QoL KINDL^R questionnaire was granted for the age group of 3 years and older.

2.2. Data collection

At selected elementary schools and childcare facilities, an information letter about the study was distributed to all families. After providing written informed consent, parents received a personal access link to the web-based, electronic questionnaire *via* email. Online forms were pseudonymized during data entry.

Great care was taken by the consortium to ensure the comprehensibility of the questionnaire. Thus, the focus was put on the use of simple language for lay participants with the aim to generate understandable information material. Also, the questionnaire offered free field comments to point out unclear aspects of the survey. In addition, a pilot phase was conducted in July 2020 in which the questionnaire was tested and adapted accordingly. The KINDL^R represents a validated questionnaire, previously tested for comprehensibility in the respective target groups.

Extrapolating from the size and number of enrolled facilities, we expected 15 827 families with school children and 8,586 families with preschool children to be eligible. Of these, 3,166 families with children (1,891 families with school children and 1,275 families with preschool children) consented to participate in the study and to fill in an online questionnaire.

The questionnaire contained socioeconomic variables, demographics, measures taken in regard to hygiene and mitigation, behavior change, the parent or caretaker situation and external contextual factors.

The customized database for electronic data capture was programmed with the Castor software that is compliant with regulations such as ICH E6 Good Clinical Practice, GDPR, HIPAA, FDA 21 CFR Part 11, ISO 27001, and ISO 9001 [Castor Electronic Data Capture (Castor EDC, Amsterdam, the Netherlands)].

2.3. Operationalization of outcomes and variables

We assessed person-related covariates, such as the children's sex and the type of institution (elementary school, kindergarten, nursery, combined institution). Type of institution was used as an indicator for the child's age group as proxy. We defined migration background as either the child itself or at least one parent indicating a country of origin other than Germany. The family structure as well as the number of children living in the household were directly reported by the parents. Parental education was assessed based on years of education of whichever parent answered the questionnaire and then categorized into three groups (≤9 years, ≤12 years, >12 years). In addition, the children's parents were divided into system-relevant and non-system-relevant "workers", with system-relevant professions regarded as essential for the continuation of essential services, such as supermarkets, schools and healthcare settings.

Changes in behavior were assessed by parents' report on the following topics: eating behavior, sleeping behavior, sports, outdoor activities, meeting friends offline, meeting friends online, screen time (leisure), screen time (education). Parents rated the changes in behaviour (more than in pre-pandemic times, equal, less than in pre-pandemic times) for each of the topics by comparing the one-year time period after the onset of the pandemic to pre-pandemic times.

Health-related QoL was assessed using the KINDL^R questionnaire. It can be answered either by the children themselves (self-report) or by their parent or caretaker (parental assessment). Depending on the children's age and developmental stage, up to four Likert-scaled questions are asked across each of six dimensions of subjective well-being (physical well-being, emotional well-being, self-esteem, family, friends, and school). The items are answered using up to five-level response categories.

The KINDL^R was developed by Ravens-Sieberer & Bullinger and is publicly available (https://www.kindl.org/) (21, 22). It has been used in different international health contexts including by the nationally representative KiGGS/BELLA study (the mental health module of the German Health Interview and Examination Survey for Children and Adolescents (KiGGS) initiated by the Robert Koch Institute (RKI)) (21, 23–30). Three versions of the KINDL^R are available comprising 12 or 24 questions (as self-report and parental proxy-report), depending on the children's age. In this study, the Kiddy KINDL^R version for children was used for children aged 3–6 years, as well as the Kid KINDL^R version for children aged 7–10 years. Parents answered the corresponding KINDL^R caretaker version. It should be noted, however, that the child-assessed Kiddy KINDL^R is validated only for the age group 4–6 years, whereas the parent-assessed version is validated for the age group 3–6 years. While the parent-assessed (proxy) versions are identical for both KINDL^R versions and together with the self-reported Kid KINDL^R consist of 24 items that can be translated into 6 sub-scores and one total score, the Kiddy KINDL^R version differs in its structure. Due to the challenge of interviewing very young children, the self-reported version's total score has only 12 items, 2 items for each of the 6 dimensions. Therefore, it is not possible to calculate any sub-scores for this KINDL^R version. Another difference of this

version lies in the three-level response categories, which range from 1 = "never" to 3 = "quite often" in contrast to five-level response categories for children aged 7–10 (1 = "never" to 5 = "always"). A visual overview of the different KINDL^R versions as well as their respective sub-scores and dimensions used in this report can be found in the appendix (**Supplementary Figure S2**).

According to the KINDL^R manual the total score and six sub-scores are calculated independently for the KINDL^R and the Kid KINDL^R (31). The responses per scale are added up, whereby certain items need to be negatively recoded and pooled beforehand, creating a mean value substitution ("sum score"). Any scale can only be evaluated if no more than 30% of its items are missing (corresponding to max. 7 for the total score and 1 for any sub-scale, respectively). The resulting scores are then transformed onto a scale ranging from 0 to 100, representing the respective health-related QoL in any given dimension (e.g., Family, Friends, School, or Total score), with a higher value indicating a better QoL (see **Supplementary Figure S2**).

Parent-reported KINDL^R scores were additionally analysed in relation to reported changes in behavior.

2.4. Data analysis

For our analysis, we included data of questionnaires completed from April to May 2021 one year after the onset of the COVID-19 pandemic in Germany. A respective flowchart depicts the inclusion of participants (**Supplementary Figure S3**).

Descriptive statistics include the person-related key characteristics described above (age and sex of children, household size and migration background of family) as well as behavior change during the COVID-19 pandemic compared to pre-pandemic times. Missing values were kept as missing.

For the health-related QoL in children, means and standard deviation of the KINDL^R total and sub-scores were compared to pre-pandemic values using population averages from the German KiGGS/BELLA study. For this part, we distinguished between the age groups 3–6 years and 7–10 years.

Logistic regressions were undertaken to identify factors affecting children's health-related QoL during the COVID-19 pandemic. The outcome variable was defined as a lower total score on the KINDL^R questionnaire compared to the median of the respective population averages (median split). For the univariate as well as the multivariate regression, the two parent-assessed KINDL^R reports (age categories 3–6 and 7–10) were dichotomized (*via* median split) and then pooled ("health-related QoL better than median" vs. "health-related QoL worse than median"). The same pooling strategy was applied to the child-assessed KINDL^R reports of both age categories. The covariates are shown in **Table 1** and describe the socioeconomic background of the families, the daily life of the children during the pandemic, and information regarding the parents' occupation.

Children's self-assessed KINDL^R scores were contrasted with parent-assessed KINDL^R scores *via* a t-test for paired samples and Cohen's kappa to test the inter-rater reliability of the two assessments.

As an exploratory approach to observe the effect of various adaptive behavior patterns during the pandemic, we descriptively compared the KINDL^R scores (means and standard deviations) stratified by the different types of behavior change.

Univariate logistic regression analyses with pooled KINDL^R scores as outcome variables (median split) were performed and corresponding ORs (odds ratios) and 95% CIs (confidence intervals) were reported.

All analyses were undertaken in IBM SPSS Statistics 27.

TABLE 1 Characteristics of the analyzed COVID Kids Bavaria study sample one year after the onset of the pandemic in spring 2021 (N = 1,962).

	N	%
Setting		
Children in elementary schools	1,244	63.4%
Children in daycare facilities	718	36.6%
Sex		
Male	1,023	52.1%
Female	935	47.7%
Diverse	1	0.1%
Missing	3	0.2%
Migration background		
No	1,611	82.1%
Yes	340	17.3%
Missing	11	0.6%
Family structure		
Single parent	99	5.0%
Single parent and partner	38	1.9%
Both parents in household	1,807	92.1%
Changing responsibilities	13	0.7%
Other modell	5	0.3%
Number of children in household		
1	396	20.2%
2	1,044	53.2%
3	330	16.8%
>3	47	2.6%
Missing	145	7.4%
System relevant parental occupation		
No	1,172	59.7%
Yes	790	40.3%
Years of paternal education		
≤9 Years	83	0.0%
10–12 Years	132	0.1%
>12 Years	1,080	55.0%
Missing	667	34.0%
Years of maternal education		
≤9 Years	71	0.0%
10–12 Years	88	0.0%
>12 Years	1,136	57.9%
Missing	667	34.0%

3. Results

3.1. Descriptive characteristics

Comparing the expected number of participants with the actual number of participants shows a response rate of 12.9% (11.9% for school children, 14.8% for pre-schoolers) over the entire study period. We included 1962 child-parent-pairs with completed

questionnaires (Table 1). Participants and non-participants were shown to be similar with regards to demographics and educational level, as reported in the main study paper by Kern et al. (20).

The majority of children attended Bavarian elementary schools ($n = 1244$, 63.4%) while 36.6% of the children ($n = 718$) attended childcare centers. Sex distribution was balanced (boys 52.1% vs. girls 47.7%, with one child specified as diverse). 17.3% of children had at least one parent born outside of Germany and 40.3% of children had a parent engaged in a profession considered "system relevant" during the COVID-19 pandemic. A detailed overview of the distribution concerning the children's living conditions and family-related covariates can be found in Table 1.

In addition, we used variables to assess the family impact of the COVID-19 pandemic. Thus, our data showed that 4.5% of the respondents themselves and 2.3% of their partners were no longer working due to the pandemic. Regarding healthcare access during the pandemic, 5.4% of parents reported refraining from visiting a doctor with their child due to the pandemic. While 65.5% of all parents reported that all medical appointments took place, 6.2% reported that their children's medical appointments were cancelled during these times.

Concerning attitudes toward vaccination, 71.7% expressed interest in COVID-19 vaccination for their children which was not available at the time of questionnaire distribution. In total, 87.5% of included children were fully vaccinated against childhood infectious diseases according to the German Standing Committee on Vaccination (STIKO) recommendations while 3.7% were not fully vaccinated. In comparison to previously published vaccination rates, the rates in our sample are high (32).

3.2. Changes in behavior compared to pre-pandemic times and in relation to health-related quality of life

Parents were asked to judge their children's behavior patterns compared to pre-pandemic times. Regarding sport/ physical activity, the majority reported that their child did less sport. General screen time for leisure and education purposes increased, while time with friends was also reportedly to shift towards more online time. Parents also observed considerable changes concerning the eating and sleeping behavior of their children, with changes in both directions, indicating maladaptive coping strategies in each case. A detailed overview is displayed in Table 2.

Comparing the parent-assessed KINDL^R scores associated with different changes in behavior, we observed that results were similar for both age categories. Therefore, for the sake of clarity, this paragraph will always refer to the KINDL^R total score of the 7–10-year-olds. All results for all age categories are displayed in Supplementary Table A1.

Eating more and eating less than before the pandemic were both associated with worse KINDL^R total scores (eating more: MD 67.40 ± 12.76); eating less: MD 64.05 ± 13.18) compared to children whose eating behavior had not changed during the pandemic according to their parents' judgement (MD 76.04 ± 10.87). Similarly, for sleep behavior, the children who experienced changes in behavior (more or less sleep than before the pandemic) also had a clearly worse KINDL^R total score than their peers whose sleep behavior had remained the same (MD 76.56 ± 10.62), with the children who slept less than before having even worse KINDL^R scores (MD 66.91 ± 12.97) than the children who slept more than before (MD 73.34 ± 11.70). Although few in number ($n = 77$), children who did more sport during the pandemic showed a higher KINDL^R score (MD 80.93 ± 9.11) than children with unchanged sports behavior (MD 78.93 (±10.49) and especially children who played less sports during the pandemic (MD 72.13 ± 12.05). Children whose leisure time remained unchanged compared to pre-pandemic times reported better KINDL^R scores (MD 78.00 ± 10.94) than children who spent more time in front of a screen during the pandemic (MD 72.16 ± 12.07). This would also apply to children whose screen time decreased during the pandemic, but numbers were small ($n = 16$).

3.3. Health-related QoL measured with KINDL^R scores compared to pre-pandemic population averages

We compared the parent-assessed KINDL^R of the COVID Kids Bavaria cohort one year after the onset of the pandemic to pre-pandemic reference values provided by the German KiGGS/ BELLA cohort, a longitudinal study on the health of children and adolescents in Germany conducted between 2003 and 2006 (Table 3) (30, 33).

TABLE 2 Subjective changes in behavior one year after the onset of the pandemic compared to pre-pandemic times assessed by parents ($N = 1,962$).

	Eating behavior	Sleeping behavior	Sports	Outdoor activities	Meeting friends offline	Meeting friends online	Screen time (leisure time)	Screen time (education)
	n (%)	*n (%)*	*n (%)*	*n (%)*	*n (%)*	*n (%)*	*n (%)*	*n (%)*
Total *N*	***1,962***	***1,962***	***1,962***	***1,962***	***1,962***	***1,962***	***1,962***	***1,962***
More than in prepandemic times	247 (12.6%)	144 (7.3%)	77 (3.9%)	407 (20.7%)	39 (2.0%)	573 (29.2%)	1,289 (65.7%)	1,200 (61.2%)
Equal	1,580 (80.5%)	1,402 (71.5%)	658 (33.5%)	976 (49.7%)	175 (8.9%)	910 (46.4%)	633 (32.3%)	218 (11.1%)
Less than in prepandemic times	118 (6.0%)	399 (20.3%)	1,220 (62.2%)	569 (29.0%)	1,739 (88.6%)	322 (16.4%)	21 (1.1%)	16 (0.8%)
Don't know	17 (0.9%)	17 (0.9%)	7 (0.0%)	10 (0.5%)	9 (0.5%)	157 (8.0%)	19 (1.0%)	15 (0.8%)
Not applicable	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	513 (26.1%)

In general, we observed lower total scores in all KINDL^R analyses in the COVID Kids Bavaria cohort compared to pre-pandemic population averages across all age groups (for 3–6-year-old KINDL^R-total score: COVID Kids Bavaria MD 74.78 ± 10.57 vs. KiGGS data 80.0 ± 8.1; 7–10 years-old KINDL^R-total score: COVID Kids Bavaria MD 73.88 ± 12.03 vs. KiGGS data 79.30 ± 9.0).

Regarding the KINDL^R sub-scores, Emotional Well-Being, Family, Friends, and School, the KINDL^R scores for all age groups were lower one year after the onset of the pandemic compared to pre-pandemic population averages. Also, in the dimensions of Physical Well-Being (COVID Kids Bavaria MD 78.76 ± 16.26 vs. KiGGS data 80.60 ± 16.3) and Self-Esteem (COVID Kids Bavaria MD 69.69 ± 15.25 vs. KiGGS data 71.40 ± 13.2) lower KINDL^R values were found for the 7–10-year-olds. In contrast, the values for the 3–6-year-olds were slightly above pre-pandemic population averages in these dimensions (Physical Well-Being: COVID Kids Bavaria MD 82.00 ± 14.6 vs. KiGGS data 80.20 ± 15.7; Self-esteem: COVID Kids Bavaria MD 73.80 ± 13.2 vs. KiGGS 73.60 ± 13.3).

Comparing the parent-assessed KINDL^R scores to the children's self-reports, our results indicate that they do not correspond well, with the t-test also indicating a significant difference between parental and child assessment (**Supplementary Table A2, A3**). The kappa-values for interrater-reliability shown in these tables generally translate to no more than a slight strength of agreement (34).

Across all age groups, children's self-report on health-related QoL was significantly higher than the parent-assessed QoL, indicating that children consistently assessed their situation more positively than their parents, with only one exception for self-esteem, where the 7–10 year old children rated their QoL slightly worse than their parents. Our findings seem to be stable over time, as they align with the KINDL^R scores of 2068 children assessed from November 2020 – March 2021 who display similar KINDL^R outcomes compared to the time period analyzed by us (April-May 2021) (**Supplementary Table A4**).

3.4. Equity considerations in health-related QoL (univariate and multivariate analyses)

We sought to identify subgroups that may be especially vulnerable to a worsening of health-related QoL during the pandemic, thereby indicating a health disadvantage. To do so we performed binary univariate regression analyses with a pooled outcome variable of the KINDL^R score, combining all age categories into a dichotomized median split.

As shown in **Table 4**, we did not find any differences between those children with a high or low health-related QoL (median split) for the covariates type of institution, sex, migration background, family structure, household size, system relevant parental occupation and parental education, for the parent-assessed KINDL^R report for children aged 3–10 years. However, with the children-assessed KINDL^R report, our results indicate that boys were less likely to report a KINDL^R total score above the sample's median compared to girls (OR: 0.82 [0.68; 0.99]). Regarding the family concept, children living with both parents in one household were more likely to experience an above-median health-related QoL (KINDL^R total score) compared to children living with a single parent (OR: 1.62 [1.03; 2.55]).

Under the same assumptions, we ran multivariate binary regressions for both the pooled parent-assessed outcome and the child-assessed outcome (**Table 5**). For both KINDL^R outcomes, we could not detect any significant differences between the children aged 3–10 years with above-median or below-median health-related QoL considering the mostly socio-economic covariates described above.

4. Discussion

The COVID Kids Bavaria study assessed the incidence of SARS-CoV-2 and its ramifications in healthy children attending elementary schools and day care facilities in different phases of the COVID-19 pandemic. Herein, we investigated health-related QoL and changes in behavior of children with a focus on the time point one year after the onset of the COVID-19 pandemic.

Main results were that a high percentage of parents reported pronounced changes in behaviour in their children, which were found to be associated with reduced health-related QoL. Children's health-related QoL was lower one year after the onset of the pandemic compared to pre-pandemic levels while children still had a more positive QoL perception than their parents.

TABLE 3 Health-related quality of life (QoL) measured with the KINDL^R one year after the onset of the pandemic compared to reference data in healthy population (30, 33).

KINDL^R scores	Parent-assessed KINDL^R score (Kids aged 3–6)		Prepandemic reference	Parent-assessed KINDL^R score (Kids aged 7–10)		Prepandemic reference
	N	M (SD)	M (SD)	*N*	M (SD)	M (SD)
KINDL^R—Total score	633	74.78 (10.57)	80.00 (8.10)	1,034	73.88 (12.03)	79.30 (9.00)
KINDL^R—Physical		82.00 (14.60)	80.20 (15.70)		78.76 (16.26)	80.60 (16.30)
KINDL^R—Emotional		74.75 (15.62)	83.00 (11.40)		71.81 (17.31)	82.70 (12.10)
KINDL^R—Self-Esteem		73.80 (13.20)	73.60 (13.30)		69.69 (15.25)	71.40 (13.20)
KINDL^R—Family		74.89 (13.52)	80.70 (11.90)		74.86 (14.77)	79.90 (12.60)
KINDL^R—Friends		68.62 (15.86)	79.70 (12.30)		71.33 (15.86)	78.50 (13.00)
KINDL^R—School		74.69 (20.85)	83.80 (12.50)		77.00 (17.59)	83.10 (14.20)

Higher score values indicate higher QoL.

TABLE 4 Univariate logistic regression (median split of parent-assessed and children-assessed health-related quality of life) for children aged 3–10.

Likelihood of a health-related Quality of Life score above the Median (all age groups pooled together)				
	Parent-assessed KINDL^R (all age groups pooled together)		Children-assessed KINDL^R (all age groups pooled together)	
	% of participants with KINDL^R score <median	OR [95% CI]	% of participants with KINDL^R scores <median	OR [95% CI]
Type of institution				
Primary School (Ref.)	47.1%		47.9%	
Kindergarden	47.6%	1.02 [0.76; 1.37]	49.7%	0.79 [0.61; 1.03]
Nursery	50.0%	1.12 [0.53; 2.38]	60.0%	0.85 [0.59; 1.22]
Combined institution	51.0%	1.17 [0.91; 1.50]	53.8%	1.29 [0.56; 2.96]
Sex				
Female (Ref.)	50.0%		51.9%	
Male	46.0%	0.85 [0.70; 1.03]	47.0%	**0.82 [0.68; 0.99]***
Migration background				
No (Ref.)	48.2%		49.9%	
Yes	46.8%	0.95 [0.73; 1.22]	47.1%	0.90 [0.69; 1.16]
Family structure				
Single parent (Ref.)	37.8%		38.6%	
Single parent and partner	28.6%	0.66 [0.28; 1.55]	26.5%	0.57 [0.24; 1.40]
Both parents in household	48.7%	1.56 [0.99; 2.47]	50.4%	**1.62 [1.03; 2.55]***
Changing responsibilities	61.5%	2.63 [0.79; 8.77]	53.8%	1.86 [0.58; 6.03]
Other model	75.0%	4.94 [0.49; 49.56]	75.0%	4.78 [0.48; 47.97]
Number of children living in the household				
1 (Ref.)	47.2%		49.0%	
2	49.2%	1.08 [0.83; 1.40]	50.4%	1.06 [0.81; 1.38]
3	49.1%	1.08 [0.78; 1.49]	51.1%	1.09 [0.78; 1.51]
>3	60.5%	1.71 [0.89; 3.29]	57.5%	1.41 [0.72; 2.75]
System relevant parental occupation				
No (Ref.)	48.1%		50.0%	
Yes	47.7%	0.99 [0.83; 1.40]	48.4%	0.94 [0.77; 1.15]
Parental education				
≤9 years (Ref.)	42.9%		47.4%	
10–12 years	46.2%	1.15 [0.64; 2.04]	46.5%	0.97 [0.54; 1.73]
>12 years	48.5%	1.26 [0.79; 2.00]	50.0%	1.11 [0.70; 1.76]

CI: Confidence Interval.
*indicating a p-value ≤0.05; bold font indicating significance.

These findings are in line with other studies reporting a deterioration in wellbeing and mental health among the youngest during the COVID-19 pandemic nationally (15) and internationally (35–40). Prior to the pandemic, in 2017, 17.2% of children and adolescents aged 3 to 17 years in the German BELLA cohort study showed psychosocial problems measured with the SDQ (41). Similarly, a meta-analysis from 2012 reported that about one in five children were at risk of mental health problems (42) while recent publications now report one in three children to be at risk (17). With regard to KINDL^R-specific values, a German study from 2016 showed that the KINDL^R scores and therefore the health-related QoL perception of children and parents alike had increased consistently over the past 10 years before the start of the COVID-19 pandemic (43). This would point towards an even higher difference compared to pre-pandemic values.

Overall, children aged 3–10 years in our cohort had much more positive perception of their current QoL than their parents, with children generally reporting higher KINDL^R total scores and sub-scores. This, however, is in contrast to pre-pandemic studies comparing parents' and children's KINDL^R reports, where the picture is less consistent. In a German study comparing parents' and children's KINDL^R reports stemming from the KiGGS study (from 2003 to 2006), parents significantly overestimated their children's health-related QoL, both in the KINDL^R total score and in the sub-scores for the dimensions Physical Well-Being, Self-esteem, and School. At the same time, they underestimated the sub-scores Emotional Well-Being and Family (24, 44). For a direct interpretation, however, it is important to know that the children considered in the KiGGS cohort were older than in our sample, namely 11–17 years old. Similar results were also reported in a Norwegian study from 1999, in which parents generally reported significantly more positive KINDL^R scores than their 9–16-year-old children (45). Overall, our results suggest that the COVID-19 pandemic might have exacerbated pre-existing discrepancies in the perceptions of children's QoL,

TABLE 5 Multivariate logistic regression with health-related QoL outcome (median split) for children aged 3–10.

Likelihood of a health-related Quality of Life score above the Median (all age groups pooled together)		
	Parent-assessed KINDL^R	Children-assessed KINDL^R
	OR [95% CI]	OR [95% CI]
Type of institution		
Primary School (Ref.)		
Kindergarden	0.99 [0.72; 1.37]	1.05 [0.75; 1.46]
Nursery	1.13 [0.52; 2.46]	1.51 [0.66; 3.44]
Combined institution	1.11 [0.84; 1.46]	1.15 [0.87; 1.53]
Sex		
Female (Ref.)		
Male	0.86 [0.70; 1.06]	0.83 [0.67; 1.02]
Migration Background		
No (Ref.)		
Yes	0.95 [0.72; 1.25]	0.97 [0.73; 1.28]
Family structure		
Single parent (Ref.)		
Single parent and partner	0.53 [0.21; 1.35]	0.38 [0.14; 1.03]
Both parents in household	1.21 [0.72; 2.01]	1.21 [0.73; 2.01]
Changing responsibilities	2.41 [0.65; 8.91]	1.59 [0.45; 5.61]
Other model	4.16 [0.41; 42.48]	3.76 [0.37; 38.25]
Number of children living in the household		
1 (Ref.)		
2	1.06 [0.80; 1.39]	1.03 [0.78; 1.36]
3	1.07 [0.76; 1.50]	1.06 [0.75; 1.50]
>3	1.80 [0.91; 3.56]	1.34 [0.67; 2.69]
System relevant parental occupation		
No (Ref.)		
Yes	0.98 [0.78; 1.23]	0.95 [0.75; 1.20]
Parental education		
≤9 years (Ref.)		
10–12 years	1.07 [0.59; 1.96]	1.02 [0.56; 1.90]
>12 years	1.14 [0.71; 1.84]	1.07 [0.66; 1.73]
Pseudo R^2	0.014	0.017

CI: Confidence Interval.
* indicating a p-value ≤0.05; bold font indicating significance.

making it even more difficult for parents to assess their children's actual situation. Alternatively, due to the high stress levels during this situation (18), parents could have tried to compensate this by being especially concerned about their children, resulting in evaluating their QoL worse than it actually was. The QoL tends to deteriorate with the age of the children. One could also deduce that especially younger age groups in our sample perceive the threatening situation differently from their parents.

Regarding changes in behavior, nearly 20% of this selected cohort ate either more or less than before the pandemic, 30% changed their sleeping behavior, 60% of children engaged in fewer sports activities than before, and 65% of children spent more leisure time in front of a screen. These results are in line with other international studies. Thus, trajectories in psychosocial aspects have already been described in the literature, all of them reported changes in children's and adolescents' sleeping behavior, reductions in physical activity and increases in screen time (46–51). On this basis, a 2020 US study recognized the positive association between higher levels of physical activity or less screen time and better mental health outcomes during the pandemic, as measured by the SDQ (52). Furthermore, a Turkish study of children during the pandemic showed the association between increased Internet use and worsened KINDL^R scores (48).

A systematic review also stressed the association of poor mental health with sleeping problems, which could lead to a downward trajectory during the COVID-19 pandemic and could partially explain the study results (53). Screen time also tends to be related to sleeping problems and sleeping deviations, which—as shown by our data—has increased during the COVID-19 pandemic, both school-related and during leisure time. Thus, the observed changes in behavior and associated KINDL^R scores may also be interdependent to some extent.

Univariate analyses of the children self-reported KINDL^R revealed that being male was a risk factor for a below-median health-related QoL, whereas living in a household with both parents increased the likelihood of an above-median QoL. However, based on our multivariate analyses of parent-assessed and children-assessed KINDL^R outcomes, we were not able to identify any predicting factor for above- or below-median health-related QoL that could indicate a specific vulnerability for health inequalities during the pandemic.

In contrast, published evidence often reports female sex as a risk factor for poor mental health during the COVID-19 pandemic, as it is also considered a risk factor for depression in general (54–58).

4.1. Limitations

The overall low response rate limits the external validity of the study with a high likelihood for systematic bias. The results shown here should thus be interpreted with caution due to the high probability of a selection and non-responder bias in our study sample. Thus, data cannot be considered to be representative of Bavarian children as a whole. Our study cohort rather provides information about which population groups and segments are willing to participate in a questionnaire-based study in a school setting during a pandemic crisis.

Based on a non-responder analysis for COVID Kids Bavaria (20), the characteristics of the responders of this convenience sample shows that a systematic middle-class bias is likely to be present. For instance, 17.3% of all children in our cohort have a migration background, whereas the Bavaria-wide share of families with a migration background was 34.2% in 2019 and is thus strongly underrepresented in our sample and may explain why our covariate analysis for migration background did not reveal any differences for a high or low health-related QoL (median split) (59). The situation is similar for single parent status: while 15.1% of all families across Bavaria were single parents in 2019 (59), the share in our cohort is only 6.9%. The number of children per household in our sample also differs

from the Bavarian average: in 2019, 49.5% of households with children had one child, 39.1% had two children, and 11.4% had three or more children (59). Our sample overestimates the number of households with two children (53.0% instead of 39.1%) and underestimates the number of one-child households (20.0% instead of 49.5%).

Consequently, the families in our sample have less migration background, are more likely to be married and less likely to be single parents, and are more likely to have two children. Notably, these factors were also found to be generally protective against health inequalities, potentially leading to distortions in our data that may mask determinants of health disadvantage.

Based on the characteristics of our sample and the comparison with the Bavarian general population, our participants seem to come from the rather well-off middle class milieu and are therefore not representative for Bavaria as a whole. Other possible homogeneities in the cohort, such as common beliefs or motivation, may not be traceable at all. A particularly strong bias is also evident in the years of parental education: for both fathers and mothers, 34.0% did not indicate any answer. While the percentage of missing values was relatively low for most variables, the variables for paternal and maternal education both showed 34.0% of missing values, which could very likely distort the picture. Thus, our results have only a limited external validity and are therefore only transferable to Bavarian primary school and daycare facility children to some extent.

Self-assessment may represent another source of bias, potentially leading to recall bias, optimism bias, or social-desirability bias. However, as results point in a similar direction as parental external assessments and evidence from other studies, we believe that this effect should not be overestimated.

4.2. Implications for research and practice

A growing body of evidence shows that children suffer from significant psychosocial consequences during this pandemic. However, it remains a challenge to differentiate the effects of acute and chronic somatic, psychological and social sequelae of COVID-19 and its mitigation measures in individual children and pediatric populations as a whole. Therefore, further analyses in large-scale longitudinal studies and surveillance data are needed to understand how the various factors affect quality of life, changes in behavior and mental health in children. For future crises and pandemics, policymakers, researchers and clinicians should prioritize the wellbeing, psychosocial monitoring and mental health right from the beginning to identify and protect vulnerable groups, to identify targets for interventions for those at risk and to promote mental health and resilience. As it now became evident that the pandemic affects all areas of children's lives, these efforts should not only be limited to mitigation measures but also other pandemic aspects. The approach of the German Wü-KiTa-CoV-Project on the feasibility of SARS-CoV-2 testing in daycare centers gives an example of how the situation of children could not only be well monitored, but also of how to include the children's perspective through qualitative work and how to test different models for improvement through intervention studies (60–62).

In terms of policies and practice, widespread awareness, research and implementation of evidence based (mental) health protection and promotion efforts in children during crises should be undertaken. While at the same time inequalities should be balanced through focus on the disadvantaged by drawing on international and national guidance (63, 64).

Data availability statement

Inquiries regarding the raw data supporting the conclusions of this study can be directed to the corresponding authors.

Ethics statement

The studies involving human participants were reviewed and approved by Ethikkommission der Medizinischen Fakultät der LMU (Ethics committee of the Medical Faculty of the University of Munich, LMU). Processing number: 20-538. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

The COVID Kids Bavaria study team (herein represented by AK, PHK, SM, HK, NM, JE, UvB, UB, MCF, AN, JW, MM, JL, JH, and CK) were responsible for the overall COVID Kids Bavaria study design, organization and conduct. HS, MC, EAR and CJ-S were responsible for the choice of survey instruments and analysis of data presented in this study including the interpretation of findings and manuscript preparation. All authors contributed to the article and approved the submitted version.

Acknowledgments

We thank all children and parents for their participation and schools, the school authorities and the teachers for their constructive support.

Conflict of interest

The study COVID Kids Bavaria was financed by the Bavarian Ministry of Research and Art. CK reports support from the Care-for-Rare Foundation. CJS reports involvement with other COVID-19 related studies such as the COSMO study, COVerCHILD (NUM) and others. ER reports involvement with other COVID-19 related studies, notably COVerCHILD, and having co-coordinated the German evidence- and consensus-based S3 guideline on schooling during COVID-19 as well as being a

member of the WHO Technical Advisory Group on Schooling during COVID-19. HS is a doctoral candidate in the Research Training Group "PrediktOren und Klinische Ergebnisse bei depressiven ErkrAnkungen in der hausärztLichen Versorgung (POKAL, DFG-GRK 2621)" (Predictors and Clinical Outcomes of Depressive Disorders in Primary Care) POKAL, a member of the German Research Foundation (DFG). UB is employed by the Department of Pediatrics, Kinderklinik München Schwabing, StKM GmbH und Klinikum Rechts der Isar, Technische Universität München, Germany. NM is employed by the Institute for Asthma- and Allergy Prevention (IAP), Helmholtz Center Munich, German Research Center for Environmental Health (GmbH), Munich, Germany.

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/fped.2023.1135415/full#supplementary-material.

References

1. Kratzer S, Pfadenhauer LM, Biallas RL, Featherstone R, Klinger C, Movsisyan A, et al. Unintended consequences of measures implemented in the school setting to contain the COVID-19 pandemic: a scoping review. *Cochrane Database Syst Rev.* (2022) 6(6):CD015397. doi: 10.1002/14651858.CD015397

2. Krishnaratne S, Littlecott H, Sell K, Burns J, Rabe JE, Stratil JM, et al. Measures implemented in the school setting to contain the COVID-19 pandemic. *Cochrane Database Syst Rev.* (2022) 1:CD015029. doi: 10.1002/14651858.CD015029

3. Fong V, Iarocci G. Child and family outcomes following pandemics: a systematic review and recommendations on COVID-19 policies. *J Pediatr Psychol.* (2020) 45 (10):1124–43. doi: 10.1093/jpepsy/jsaa092

4. Racine N, McArthur BA, Cooke JE, Eirich R, Zhu J, Madigan S. Global prevalence of depressive and anxiety symptoms in children and adolescents during COVID-19: a meta-analysis. *JAMA Pediatr.* (2021) 175(11):1142–50. doi: 10.1001/jamapediatrics.2021.2482

5. Imran N, Aamer I, Sharif MI, Bodla ZH, Naveed S. Psychological burden of quarantine in children and adolescents: a rapid systematic review and proposed solutions. *Pak J Med Sci.* (2020) 36(5):1106–16. doi: 10.12669/pjms.36.5.3088

6. Panda P, Gupta J, Chowdhury S, Kumar R, Meena A, Madaan P, et al. Psychological and behavioral impact of lockdown and quarantine measures for COVID-19 pandemic on children, adolescents and caregivers: a systematic review and meta-analysis. *J Trop Pediatr.* (2021) 67(1)fmaa122. doi: 10.1093/tropej/fmaa122

7. Meherali S, Punjani N, Louie-Poon S, Rahim KA, Das JK, Salam RA, et al. Mental health of children and adolescents amidst COVID-19 and past pandemics: a rapid systematic review. *Int J Environ Res Public Health.* (2021) 18(7):3432. doi: 10.3390/ijerph18073432

8. Singh S, Roy D, Sinha K, Parveen S, Sharma G, Joshi G. Impact of COVID-19 and lockdown on mental health of children and adolescents: a narrative review with recommendations. *Psychiatry Res.* (2020) 293:113429. doi: 10.1016/j.psychres.2020.113429

9. Loades ME, Chatburn E, Higson-Sweeney N, Reynolds S, Shafran R, Brigden A, et al. Rapid systematic review: the impact of social isolation and loneliness on the mental health of children and adolescents in the context of COVID-19. *J Am Acad Child Adolesc Psychiatry.* (2020) 59(11):1218. doi: 10.1016/j.jaac.2020.05.009

10. Ravens-Sieberer U, Erhart M, Devine J, Gilbert M, Reiss F, Barkmann C, et al. Child and adolescent mental health during the COVID-19 pandemic: results of the three-wave longitudinal COPSY study. *J Adolesc Health.* (2022) 71(5):570–78. doi: 10.1016/j.jadohealth.2022.06.022

11. Reiss F, Meyrose AK, Otto C, Lampert T, Klasen F, Ravens-Sieberer U. Socioeconomic status, stressful life situations and mental health problems in children and adolescents: results of the German BELLA cohort-study. *PLoS One.* (2019) 14(3):e0213700. doi: 10.1371/journal.pone.0213700

12. Skoda EM, Spura A, de Bock F, Schweda A, Dörrie N, Fink M, et al. [Change in psychological burden during the COVID-19 pandemic in Germany: fears, individual behavior, and the relevance of information and trust in governmental institutions]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz.* (2021) 64 (3):322–33. doi: 10.1007/s00103-021-03278-0

13. Ravens-Sieberer U, Kaman A, Erhart M, Devine J, Schlack R, Otto C. Impact of the COVID-19 pandemic on quality of life and mental health in children and adolescents in Germany. *Eur Child Adolesc Psychiatry.* (2021) 31:879–89. doi: 10.1007/s00787-021-01726-5

14. Ravens-Sieberer U, Kaman A, Erhart M, Otto C, Devine J, Löffler C, et al. Quality of life and mental health in children and adolescents during the first year of the COVID-19 pandemic: results of a two-wave nationwide population-based study. *Eur Child Adolesc Psychiatry.* (2021) 1–14. doi: 10.1007/s00787-021-01889-1

15. Ravens-Sieberer U, Kaman A, Otto C, Adedeji A, Devine J, Erhart M, et al. Mental health and quality of life in children and adolescents during the COVID-19 pandemic—results of the copsy study. *Dtsch Arztebl Int.* (2020) 117:828–9. doi: 10.3238/arztebl.2020.0828

16. Ravens-Sieberer U, Kaman A, Otto C, Adedeji A, Napp AK, Becker M, et al. Mental health and psychological burden of children and adolescents during the first wave of the COVID-19 pandemic—results of the COPSY study. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz.* (2021) 64:1–10. doi: 10.1007/s00103-021-03291-3

17. Rathgeb C, Schillok H, Voss S, Coenen M, Schulte-Körne G, Merkel C, et al. Emotional situation of children and adolescents during the COVID-19 pandemic in Germany: results from the COVID-19 snapshot monitoring study (COSMO). *Int J Environ Res Public Health.* (2022) 19(5):2698. doi: 10.3390/ijerph19052698

18. Rabe JE, Schillok H, Merkel C, Voss S, Coenen M, de Bock F, et al. [Burden in parents of school-aged children during different phases of the COVID-19 pandemic in Germany: an analysis within the COVID-19 snapshot monitoring (COSMO) study]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz.* (2021) 64 (12):1500–11. doi: 10.1007/s00103-021-03453-3

19. Bhatia A, Fabbri C, Cerna-Turoff I, Turner E, Lokot M, Warria A, et al. Violence against children during the COVID-19 pandemic. *Bull World Health Organ.* (2021) 99 (10):730–8. doi: 10.2471/BLT.20.283051

20. Kern A, Kuhlmann PH, Matl S, Ege M, Maison N, Eckert J, et al. Surveillance of acute SARS-CoV-2 infections in elementary schools and daycare facilities in bavaria, Germany (09/2020–03/2021). *Front Pediatr.* (2022) 10:1019. doi: 10.3389/fped.2022.888498

21. Ravens-Sieberer U, Bullinger M. Assessing health-related quality of life in chronically ill children with the German KINDL: first psychometric and content analytical results. *Qual Life Res.* (1998) 7(5):399–407. doi: 10.1023/A:1008853819715

22. Ravens-Sieberer U, Bullinger M. News from the KINDL-questionnaire—a new version for adolescents. *Qual Life Res.* (1998) 7:653. doi: 10.1023/A:1008853819715

23. Bullinger M, Brütt AL, Erhart M, Ravens-Sieberer U. Psychometric properties of the KINDL-R questionnaire: results of the BELLA study. *Eur Child Adolesc Psychiatry.* (2008) 17(Suppl 1):125–32. doi: 10.1007/s00787-008-1014-z

24. Erhart M, Ellert U, Kurth BM, Ravens-Sieberer U. Measuring adolescents' HRQoL via self reports and parent proxy reports: an evaluation of the psychometric properties of both versions of the KINDL-R instrument. *Health Qual Life Outcomes.* (2009) 7(1):77. doi: 10.1186/1477-7525-7-77

25. Eser E, Yüksel H, Baydur H, Erhart M, Saatli G, Ozyurt BC, et al. [The psychometric properties of the new Turkish generic health-related quality of life questionnaire for children (kid-KINDL)]. *Turk Psikiyatri Derg.* (2008) 19 (4):409–17. PMID: 19110983.

26. Essaddam L, ben Mansour A, ben Amor A, Ravens-Sieberer U, Klein TM, ben Becher S. Validation of the Arabic and Tunisian Arabic version of the KINDL

questionnaires for children with diabetes type 1. *Libyan J Med*. (2019) 14(1):1–7. doi: 10.1080/19932820.2018.1537457

27. Fernández-López JA, Fernández Fidalgo M, Cieza A, Ravens-Sieberer U. [Measuring health-related quality of life in children and adolescents: preliminary validation and reliability of the Spanish version of the KINDL questionnaire]. *Aten Primaria*. (2004) 33(8):434–42. doi: 10.1016/S0212-6567(04)79429-9

28. Rajmil L, Serra-Sutton V, Fernandez-Lopez JA, Berra S, Aymerich M, Cieza A, et al. [The Spanish version of the German health-related quality of life questionnaire for children and adolescents: the Kindl]. *An Pediatr (Barc)*. (2004) 60(6):514–21. doi: 10.1016/S1695-4033(04)78320-4

29. Bullinger M, Schmidt S, Petersen C, Erhart M, Ravens-Sieberer U. [Methodological challenges and potentials of health-related quality of life evaluation in children with chronic health conditions under medical health care]. *Med Klin (Munich)*. (2007) 102(9):734–45. doi: 10.1007/s00063-007-1092-6

30. Ravens-Sieberer U, Ellert U, Erhart M. Gesundheitsbezogene Lebensqualität von Kindern und Jugendlichen in Deutschland. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*. (2007 May 1) 50(5–6):810–8. doi: 10.1007/s00103-007-0244-4

31. Ravens-Sieberer U, Bullinger M. Manual—kindl.org. (2000). Available from: https://www.kindl.org/english/manual/ (cited January 9, 2022).

32. Krankenkasse T, Egert-Schmidt A-M, Holtz S, Kaufmann-Kolle P, Neubert A, Spelly C, et al. Kinder und Arzneimittel – Verordnungsreport (2022).

33. Ravens-Sieberer U, Erhart M, Wille N, Bullinger M, BELLA study group. Health-related quality of life in children and adolescents in Germany: results of the BELLA study. *Eur Child Adolesc Psychiatry*. (2008) 17(Suppl 1):148–56. doi: 10.1007/s00787-008-1016-x

34. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. (1977) 33(1):159. doi: 10.2307/2529310

35. Mahase E. COVID-19: children's mental health services in England are "nowhere near sufficient," says commissioner. *Br Med J*. (2021) 372:n258. doi: 10.1136/bmj.n258

36. Cost KT, Crosbie J, Anagnostou E, Birken CS, Charach A, Monga S, et al. Mostly worse, occasionally better: impact of COVID-19 pandemic on the mental health of Canadian children and adolescents. *Eur Child Adolesc Psychiatry*. (2022) 31(4):671–84. doi: 10.1007/s00787-021-01744-3

37. Saurabh K, Ranjan S. Compliance and psychological impact of quarantine in children and adolescents due to COVID-19 pandemic. *Indian J Pediatr*. (2020) 87(7):532–6. doi: 10.1007/s12098-020-03347-3

38. Patrick SW, Henkhaus LE, Zickafoose JS, Lovell K, Halvorson A, Loch S, et al. Well-being of parents and children during the COVID-19 pandemic: a national survey. *Pediatrics*. (2020) 146(4):e2020016824. doi: 10.1542/peds.2020-016824

39. Xie X, Xue Q, Zhou Y, Zhu K, Liu Q, Zhang J, et al. Mental health status among children in home confinement during the coronavirus disease 2019 outbreak in Hubei province, China. *JAMA Pediatr*. (2020) 174(9):898–900. doi: 10.1001/jamapediatrics.2020.1619

40. Orgilés M, Morales A, Delvecchio E, Mazzeschi C, Espada JP. Immediate psychological effects of the COVID-19 quarantine in youth from Italy and Spain. *Front Psychol*. (2020) 11:2986. doi: 10.3389/fpsyg.2020.579038

41. Klasen F, Meyrose AK, Otto C, Reiss F, Ravens-Sieberer U. Leitthema Psychische Auffälligkeiten von Kindern und Jugendlichen in Deutschland Ergebnisse der BELLA-Studie. *Monatsschr Kinderheilkd*. (2017) 5(165):402–7. doi: 10.1007/s00112-017-0270-8

42. Barkmann C, Schulte-Markwort M. Prevalence of emotional and behavioural disorders in German children and adolescents: a meta-analysis. *J Epidemiol Community Health (1978)*. (2012) 66(3):194–203. doi: 10.1136/jech.2009.102467

43. Meyer M, Oberhoffer R, Hock J, Giegerich T, Müller J. Health-related quality of life in children and adolescents: current normative data, determinants and reliability on proxy-report. *J Paediatr Child Health*. (2016) 52(6):628–31. doi: 10.1111/jpc.13166

44. Ellert U, Ravens-Sieberer U, Erhart M, Kurth BM. Determinants of agreement between self-reported and parent-assessed quality of life for children in Germany-results of the German Health Interview and Examination Survey for Children and Adolescents (KiGGS). (2011). Available from: http://www.hqlo.com/content/9/1/102 (cited January 9,2022).

45. Jozefiak T, Larsson B, Wichstrøm L, Mattejat F, Ravens-Sieberer U. Quality of life as reported by school children and their parents: a cross-sectional survey. *Health Qual Life Outcomes*. (2008) 6(1):1–11. doi: 10.1186/1477-7525-6-34

46. Liu Z, Tang H, Jin Q, Wang G, Yang Z, Chen H, et al. Sleep of preschoolers during the coronavirus disease 2019 (COVID-19) outbreak. *J Sleep Res*. (2021) 30(1):e13142. doi: 10.1111/jsr.13142

47. López-Bueno R, López-Sánchez GF, Casajús JA, Calatayud J, Gil-Salmerón A, Grabovac I, et al. Health-related behaviors among school-aged children and adolescents during the spanish COVID-19 confinement. *Front Pediatr*. (2020) 8:573. doi: 10.3389/fped.2020.00573

48. Adıbelli D, Sümen A. The effect of the coronavirus (COVID-19) pandemic on health-related quality of life in children. *Child Youth Serv Rev*. (2020) 119:105595. doi: 10.1016/j.childyouth.2020.105595

49. Francisco R, Pedro M, Delvecchio E, Espada JP, Morales A, Mazzeschi C, et al. Psychological symptoms and behavioral changes in children and adolescents during the early phase of COVID-19 quarantine in three European countries. *Front Psychiatry*. (2020) 11:1329. doi: 10.3389/fpsyt.2020.570164

50. Aguilar-Farias N, Toledo-Vargas M, Miranda-Marquez S, Cortinez-O'ryan A, Cristi-Montero C, Rodriguez-Rodriguez F, et al. Sociodemographic predictors of changes in physical activity, screen time, and sleep among toddlers and preschoolers in Chile during the COVID-19 pandemic. *Int J Environ Res Public Health*. (2020) 18(1):1–13. doi: 10.3390/ijerph18010176

51. Medrano M, Cadenas-Sanchez C, Oses M, Arenaza L, Amasene M, Labayen I. Changes in lifestyle behaviours during the COVID-19 confinement in Spanish children: a longitudinal analysis from the MUGI project. *Pediatr Obes*. (2021) 16(4):e12731. doi: 10.1111/ijpo.12731

52. Tandon PS, Zhou C, Johnson AM, Gonzalez ES, Kroshus E. Association of children's physical activity and screen time with mental health during the COVID-19 pandemic. *JAMA Netw Open*. (2021) 4(10):e2127892. doi: 10.1001/jamanetworkopen.2021.27892

53. Palacio-Ortiz JD, Londoño-Herrera JP, Nanclares-Márquez A, Robledo-Rengifo P, Quintero-Cadavid CP. Psychiatric disorders in children and adolescents during the COVID-19 pandemic. *Rev Colomb Psiquiatr (Engl Ed)*. (2020) 49(4):279.

54. Panchal U, Salazar de Pablo G, Franco M, Moreno C, Parellada M, Arango C, et al. The impact of COVID-19 lockdown on child and adolescent mental health: systematic review. *Eur Child Adolesc Psychiatry*. (2021) 1:1–27. doi: : 10.1007/s00787-021-01856-w

55. Chen F, Zheng D, Liu J, Gong Y, Guan Z, Lou D. Depression and anxiety among adolescents during COVID-19: a cross-sectional study. *Brain Behav Immun*. (2020) 88:36–8. doi: 10.1016/j.bbi.2020.05.061

56. Chen S, Cheng Z, Wu J. Risk factors for adolescents' mental health during the COVID-19 pandemic: a comparison between Wuhan and other urban areas in China. *Global Health*. (2020) 16(1):1–11. doi: 10.1186/s12992-019-0531-5

57. Magson NR, Freeman JYA, Rapee RM, Richardson CE, Oar EL, Fardouly J. Risk and protective factors for prospective changes in adolescent mental health during the COVID-19 pandemic. *J Youth Adolesc*. (2021) 50(1):44–57. doi: 10.1007/s10964-020-01332-9

58. Smirni P, Lavanco G, Smirni D. Anxiety in older adolescents at the time of COVID-19. *J Clin Med*. (2020) 9(10):1–10. doi: 10.3390/jcm9103064

59. Bayerisches Landesamt für Gesundheit und Lebensmittelsicherheit (LGL). Gesundheitsatlas Bayern.Available from: https://www.lgl.bayern.de/gesundheit/gesundheitsberichterstattung/gesundheitsatlas/ia_report/atlas.html (cited March 4, 2022).

60. Gierszewski D, Kurotschka PK, Krauthausen M, Fröhlich W, Forster J, Pietsch F, et al. Parents' and childcare workers' perspectives toward SARS-CoV-2 test and surveillance protocols in Pre-school children day care centers: a qualitative study within the German Wü-KiTa-CoV project. *Front Med (Lausanne)*. (2022) 9:1251.

61. Engels G, Forster J, Streng A, Rücker V, Rudolph P, Pietsch F, et al. Acceptance of different self-sampling methods for semiweekly SARS-CoV-2 testing in asymptomatic children and childcare workers at German day care centers: a nonrandomized controlled trial. *JAMA Netw Open*. (2022) 5(9):e2231798. doi: 10.1001/jamanetworkopen.2022.31798

62. Forster J, Streng A, Rudolph P, Rücker V, Wallstabe J, Timme S, et al. Feasibility of SARS-CoV-2 surveillance testing among children and childcare workers at German day care centers: a nonrandomized controlled trial. *JAMA Netw Open*. (2022) 5(1):e2142057. doi: 10.1001/jamanetworkopen.2021.42057

63. Deutsche Gesellschaft für Epidemiologie (DGEpi), Deutsche Gesellschaft für Public Health (DGPH), Deutsche Gesellschaft für, Kinder- und Jugendmedizin (DGKJ), Deutsche Gesellschaft für pädiatrische Infektiologie (DGPI), Gesellschaft für Hygiene, et al. S3-Leitlinie – Maßnahmen zur Prävention und Kontrolle der SARS-CoV-2- Übertragung in Schulen – Lebende Leitlinie, Langfassung Version 1.1 2021. Available from: https://www.awmf.org/leitlinien/detail/ll/027-076.html (cited 2022 Jan 9).

64. RAND Mental Health Systems Initiative | RAND. Available from: https://www.rand.org/health-care/key-topics/mental-health/initiative.html (cited 2022 Jan 9).

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 31 May 2023
DOI 10.3389/fpubh.2023.1166760

OPEN ACCESS

EDITED BY
Pei Wang,
Fudan University, China

REVIEWED BY
Arlette Setiawan,
Padjadjaran University, Indonesia
Xin Li,
Nanjing Medical University, China
Shunping Li,
Shandong University, China

*CORRESPONDENCE
Xiaoyu Xi
✉ xixy@cpu.edu.cn

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

RECEIVED 15 February 2023
ACCEPTED 30 March 2023
PUBLISHED 31 May 2023

CITATION
Wang Q, Wan C, Li M, Huang Y and Xi X (2023) Mapping the Peds QL™ 4.0 onto CHU-9D: a cross-sectional study in functional dyspepsia population from China.
Front. Public Health 11:1166760.
doi: 10.3389/fpubh.2023.1166760

Mapping the Peds QL™ 4.0 onto CHU-9D: a cross-sectional study in functional dyspepsia population from China

Qiqi Wang†, Chuchuan Wan†, Maozhen Li, Yuankai Huang and Xiaoyu Xi*

The Research Center of National Drug Policy & Ecosystem, China Pharmaceutical University, Nanjing, China

Objective: The study aims to develop a mapping algorithm from the Pediatric Quality of Life Inventory™ 4. 0 (Peds QL 4.0) onto Child Health Utility 9D (CHU-9D) based on the cross-sectional data of functional dyspepsia (FD) children and adolescents in China.

Methods: A sample of 2,152 patients with FD completed both the CHU-9D and Peds QL 4.0 instruments. A total of six regression models were used to develop the mapping algorithm, including ordinary least squares regression (OLS), the generalized linear regression model (GLM), MM-estimator model (MM), Tobit regression (Tobit) and Beta regression (Beta) for direct mapping, and multinomial logistic regression (MLOGIT) for response mapping. Peds QL 4.0 total score, Peds QL 4.0 dimension scores, Peds QL 4.0 item scores, gender, and age were used as independent variables according to the Spearman correlation coefficient. The ranking of indicators, including the mean absolute error (MAE), root mean squared error (RMSE), adjusted R^2, and consistent correlation coefficient (CCC), was used to assess the predictive ability of the models.

Results: The Tobit model with selected Peds QL 4.0 item scores, gender and age as the independent variable predicted the most accurate. The best-performing models for other possible combinations of variables were also shown.

Conclusion: The mapping algorithm helps to transform Peds QL 4.0 data into health utility value. It is valuable for conducting health technology evaluations within clinical studies that have only collected Peds QL 4.0 data.

KEYWORDS

mapping, CHU-9D, Peds QL 4.0, health utility value, functional dyspepsia

Introduction

Functional dyspepsia (FD) is a clinical syndrome occurring in the gastroduodenal region. FD is a common pediatric disorder with a prevalence of ∼3% in China (1) and ∼3 to 7.6% in other countries (2, 3). Children and adolescents with FD are often associated with symptoms such as postprandial fullness, early satiation, anorexia, belching, nausea, vomiting, upper abdominal gaseous distension, pain, burning sensation, and acid regurgitation (1, 4).

Previous studies have shown that FD leads to lower health-related quality of life (HRQoL) in children and adolescents, with significant negative effects on their lives (3, 5), causing sleep disturbances, psychological distress, frequent absence from school, and less social interaction (6–8). Thus, accurate measuring and monitoring of the HRQoL of patients are valuable for understanding the harm caused by FD and managing the disease. Meanwhile, it promotes the development of health technology assessment. As a key indicator for measuring HRQoL, health utility value (HUV) can adequately reflect individual preferences and has been widely used. HUVs can be obtained by direct or indirect measurement. Indirect measures are more commonly used because of their simplicity and ease of use. Currently, the main scales that indirectly measure the quality of life and HUVs in children or adolescents include the Child Health Utility 9D (CHU-9D), EuroQoL five-dimension youth version three-level (EQ-5D-Y-3L), short-form six-dimension (SF-6D), and Health Utilities Index (HUI) (9, 10). However, Pediatric Quality of Life Inventory^{TM} 4.0 (Peds QL 4.0) is often used in clinical studies of FD (11–13). It cannot directly measure HUVs (14) as a non-preference-based scale. To make full use of existing research data and obtain HUVs of FD patients, constructing a mapping algorithm from Peds QL 4.0 to preference-based scale is necessary. The CHU-9D, a multi-attribute utility scale for children and adolescents, is used worldwide (9, 11) and has been demonstrated by previous studies to have good psychometric properties in the Chinese population (15). Unlike other preference-based scales that apply to children and adolescents, CHU-9D was originally developed for young people (16). In addition, both the Peds QL 4.0 and CHU-9D had items to investigate the physical and psychosocial health of children and adolescents. For instance, the CHU-9D sad and sleep dimensions were captured in Peds QL 4.0 emotional functional dimension, and the pain and tired dimensions were captured in physical functioning dimension. The conceptual overlap makes it meaningful to develop a mapping algorithm between Peds QL 4.0 and CHU-9D.

Up to now, mapping has been regarded as the "second-best solution" for measuring HUVs. ISPOR has issued relevant guidance to guide researchers to use this method (17, 18). A large number of studies have shown that the mapping algorithm between Peds QL 4.0 and preference-based scale can be developed to obtain the HUV (19–24). However, no such study has been conducted for developing a mapping algorithm between Peds QL 4.0 and CHU-9D in Chinese FD children and adolescents. Thus, this study was to develop an algorithm based on Chinese FD patients. In addition, the mapping algorithm enables researchers to calculate HUVs using available clinical data, which facilitates the development of health technology evaluations.

Methods

From March to May 2020, FD outpatients were recruited from several hospitals in Zhejiang province through convenient sampling. Meanwhile, we collected data related to the HRQoL of participants by CHU-9D and Peds QL 4.0. All participants voluntarily participated and signed informed consent. This study was approved by the Ethics Committee of China Pharmaceutical University.

Sample

Inclusion criteria were as follows: (1) informed and voluntary; (2) age 6 to 17 years [the CHU-9D is suitable for children and adolescents aged 7–17 years. Studies have shown that it is also suitable for children aged 6–7 years (25)]; and (3) diagnosed with FD according to the Rome IV criteria (26). The exclusion criteria were as follows: (1) non-Chinese; (2) mental patients, unconsciousness, unable to describe their own situation.

Data collection

Trained investigators went to hospitals to present the study to FD patients and their guardians and asked them about their willingness to participate. For patients who want to take part, investigators will provide them with an informed consent form and a questionnaire. After signing the forms in a quiet environment, they completed questionnaires in full view of investigators. The questionnaire includes two HRQoL tools. In addition, to avoid ranking bias, participants decided the order of filling by flipping a coin. Considering that younger patients may have difficulty in understanding the scale, the children younger than 7 years old completed the questionnaires with their guardians, while other participants completed on their own. For the basic information part of the questionnaire, the guardian should assist the participants to complete it. After the questionnaire was completed, the investigator would check the questionnaire and upload the data after the respondents confirmed the questions. After that, the auditor would review it again. For the questionnaire with obvious problems, the auditor would return the questionnaire.

Questionnaire

The literature and experts' opinions were drawn upon (11, 13, 14, 27, 28). The questionnaire was divided into two parts as follows: basic information and health status. According to the results of the pilot survey in hospitals, we modified the questionnaire and formed the final version. Its rationality, readability, and comprehensibility were affirmed by experts and supported by the results of pilot research.

There were two parts to the questionnaires. Part 1 collected sociodemographic information, including age, gender, parents' education level, and family income. Part 2 collected some health status indicators reported by patients through CHU-9D and Peds QL 4.0.

Child health utility 9D

Child Health Utility 9D (CHU-9D) was a universal scale developed for children and adolescents by Professor Katherine

Stevens from the University of Sheffield. It was used to obtain children's and adolescents' HUVs of subjects and had been widely used at present. The translation process of the Chinese version of the CHU-9D questionnaire was designed based on the recommendations of the ISPOR Task Force (29). In 2013, a pilot study was conducted in Jiangsu province, China. Its results supported the feasibility and construct validity of the Chinese version of CHU-9D for measuring and valuing the HRQoL of Chinese young people (15). The CHU-9D consisted of nine dimensions that are "worried," "sad," "pain," "tired," "annoyed," "schoolwork/homework," "sleep," "daily routine," and "activities," each dimension had five levels, a total of $5^9 = 1{,}953{,}125$ possible health states were defined. The CHU-9D was originally developed for children aged 7–11 years (16), but subsequent studies have shown that it can also be used for children aged 6 years and adolescents aged 11–17 years (25, 30). In this study, the CHU-9D utility scores were calculated using the China value set (27).

Peds QL 4.0

Pediatric Quality of Life Inventory™ 4.0 (Peds QL 4.0) was developed under the lead of Professor Varni et al. (14) and was officially released in 1999. Peds QL 4.0 was introduced in China in 2004. Yi-Yun et al. developed the Chinese version of Peds QL 4.0 based on the standard procedure of cross-cultural adaptation (translation–back translation–cultural adaptation–pre-test) to (31–33). This version was proved that it was applicable to Chinese children (14, 33–35). The self-report versions of the Peds QL 4.0 for young children (aged 5–7 years), children (aged 8–12 years), and teenagers (aged 13–18 years) and parent-report versions of the Peds QL 4.0 for young children (aged 5–7 years) were used in the study. Peds QL 4.0 contains four dimensions as follows: physical functioning (PF), emotional functioning (EF), social functioning (SF), and school functioning (ScF). There were 23 items, and the recall period was 1 month. PF contained eight items, while EF, SF, and ScF each contained five items. Items were scored on a five-point Likert scale as follows: 0 indicates "never a problem," 1 indicates "almost never a problem," 2 indicates "sometimes a problem," 3 indicates "often a problem," and 4 indicates "almost always a problem." Items were, then, transformed into a score ranging from 0 to 100 (where 0 = 100, 1 = 75, 2 = 50, 3 = 25, and 4 = 0). The dimension score was the average score of the items contained in the dimension, and the total score was the average score of all items responded.

Data analysis

Descriptive statistics

Descriptive statistics [mean and standard deviation (SD) for continuous variables and frequency and percentage for categorical variables] were used for the sample characteristics. The distributions of the CHU-9D utility score and Peds QL 4.0 score were shown through Shapiro–Wilk test and figures.

Correlation test

Mapping of the scale requires some conceptual overlap between the initial scale and the target scale (36, 37). Spearman's rank correlations (ρ) were used in this study to test the conceptual overlap strength between CHU-9D and Peds QL 4.0. Conceptual overlap was characterized by the content similarity between HRQoL result measurements. If two scales lack conceptual overlap, the mapping relationship would not be established. In addition, we tested the correlation among the variables included in the study to ensure the low collinearity among the variables included in the mapping model. The strength of correlation could be divided into four levels (very weak = 0–0.19; weak = 0.20–0.39; moderate = 0.40–0.59; strong = 0.60–0.79; and very strong = 0.80–1.00) (38, 39).

Mapping model

Mapping consists of two broad approaches, such as direct mapping and response mapping. We used six regression models for developing a simpler and more accurate mapping algorithm, based on guidelines and previous research (17), including ordinary least squares regression (OLS), general linear regression model (GLM), MM-estimator model (MM), Tobit regression model (Tobit), Beta regression model (Beta) for direct mapping, and multinomial logistic regression (MLOGIT) for response mapping.

Ordinary least squares regression (OLS) uses linear functions to construct the relationship between independent variables and dependent variables. Due to its simplicity, OLS is widely used in direct mapping studies (38, 40). However, OLS performs poorly in predicting poor or full health, and the predicted values may be outside of the reasonable range (41–43). GLM (link "logit"), a flexible form of OLS, allows the outcome variables to have non-normal error distributions (44, 45). MM can better identify outliers, thus MM is less affected by outliers and has less deviation from the fitted residuals (23, 46). When using the traditional linear regression model, the predicted value is often out of the range of the dependent variable. In view of the fact that the CHU-9D scores are deleted at the upper and lower limits, OLS and GLM tend to produce systematic bias. We also used the Tobit model and Beta model. As a censored model, the Tobit model can be used to predict the continuous but limited or truncated dependent variables, but it is more sensitive to heteroscedasticity and non-normal distribution (47, 48). The Beta model solves this problem by assuming that the value of the dependent variable is between 0 and 1, and the model is also suitable for cases with heteroscedasticity or non-normal distribution of the data (49, 50). In response mapping, we used the MLOGIT model to obtain the probability of a specific level of CHU-9D in each dimension and then calculated the utility value of CHU-9D using the expected utility value method (50).

Variables

We chose the CHU-9D total score and scores of each dimension of CHU-9D as the dependent variable for direct and response mapping, and Peds QL 4.0 total score, Peds QL 4.0 dimension scores, and Peds QL 4.0 item scores were used as independent variables for regression to generate mapping algorithm. In addition, to ensure the accuracy of the mapping algorithm, age and gender

TABLE 1 Combinations of variables.

Combination	Independent variables
Combination 1	Peds QL 4.0 total score
Combination 2	Peds QL 4.0 total score, age, gender
Combination 3	Peds QL 4.0 dimension scores
Combination 4	Peds QL 4.0 dimension scores, age, gender
Combination 5	Peds QL 4.0 item scores[a]
Combination 6	Peds QL 4.0 item scores[a], age, gender

[a]Peds QL 4.0 items were selected according to statistical significance using stepwise regression, and the statistical significance level was 0.05. Variables eventually included in the study: PF2, Hard to run; PF5, Hard to take a bath or shower; PF8, Low energy; EF1, Feel afraid or scared; EF4, Trouble sleeping; EF5, Worry about what will happen; SF1, Trouble getting along with others; SF3, Teased; SF4, Not able to do things that others can do; ScF1, Hard to pay attention in class; ScF2, Forget things; ScF3, Trouble keeping up with schoolwork; ScF4, Miss school because of not feeling well.

TABLE 2 Participant characteristics.

Characteristics ($N = 2,152$)	Mean ± SD/N (%)	Median	Min	Max
Age (years)	7.23 ± 1.47	6	6	17
Gender				
Boys	1,177 (54.69%)			
Girls	975 (45.31%)			
CHU-9D	0.88 ± 0.10	0.88	0.62	1.00
Peds QL 4.0 total score	61.52 ± 5.63	60.87	53.12	100.00
Peds QL 4.0 Physical Functioning	67.02 ± 5.49	62.50	50.00	100.00
Peds QL 4.0 Emotional Functioning	57.45 ± 5.87	55.00	40.00	100.00
Peds QL 4.0 Social Functioning	57.21 ± 10.09	55.00	40.00	100.00
Peds QL 4.0 School Functioning	62.26 ± 6.46	60.00	40.00	100.00

SD, standard deviation.

were included as independent variables according to the correlation between variables. Finally, six combination models of independent variables were developed, as presented in Table 1.

Validation and comparison of mapping algorithms

The 10-fold cross-validation method was used to predict model performance. In this method, the original sample was divided into 10 subsamples of roughly equal size. In total. One of the 10 subsamples was taken as the validation sample, and the remaining subsamples were taken as the training samples for regression (23, 45).

The mean of the HUVs predicted (mean P), root mean square error (RMSE), mean absolute error (MAE), adjusted R^2 (adj R^2), and concordance correlation coefficient (CCC) was recorded and averaged for each combination. The mapping algorithm with the best comprehensive ranking was the optimal mapping algorithm in different combinations of variables (18, 45).

All statistical analyses were performed by stata15, programs R and Microsoft® Excel 2016.

Result

Participant characteristics

A total of 2,152 eligible FD patients were enrolled in this study (Table 2), of whom 1,155 patients had their guardians complete the questionnaire on their behalf. Their mean age (SD) was 7.23 (1.47) years, the mean utility score of CHU-9D (SD) was 0.88 (0.10), and the mean score of Peds QL 4.0 (SD) was 61.52 (5.63). Participants' CHU-9D utility score and Peds QL 4.0 total score were skewed (Figure 1).

Correlation test results

Correlations between CHU-9D utility score, Peds QL 4.0 total score, Peds QL 4.0 dimension scores, Peds QL 4.0 item scores, age, and gender are presented in Appendix 1. There was a strong positive correlation between the CHU-9D utility score and Peds QL 4.0 total score ($\rho = 0.6836$, $p < 0.001$), as well as the Peds QL 4.0 dimension scores and the CHU-9D utility score, among which the SF dimension had the highest correlation ($\rho = 0.6399$, $p < 0.001$) and the PF dimension had the lowest correlation ($\rho = 0.3146$, $p < 0.001$).

Performance of mapping algorithms

The results of mapping algorithm performance based on six regression techniques are presented in Table 3.

For model combinations based on the full estimation sample, the mean P ranged from 0.8767 (Tobit of Combination 1) to 0.8862 (Beta of Combination 1), with the OLS of Combination 3 having the closest predicted score (0.8784) to the mean score. In Combination 1, Beta had the lowest MAE (0.0576), while MLOGIT performed better in RMSE (0.0716) and adj R^2 (0.4498). In Combination 2, Beta performed better than other models. In Combinations 3 to 6, MLOGIT had the lowest comprehensive ranking and better performance than other models. The Beta had the best CCC in all combinations. Among all combinations, MLOGIT in Combination 6 had the best MAE (0.0501), RMSE (0.0630), and adj R^2 (0.5745), and the Beta model had the best CCC (0.7319). In summary, the performance based on Peds QL 4.0 item scores (Combinations 5 and 6) was better than other combinations. MLOGIT had better performance in the validation index, with the best MAE, RMSE, and adj R^2, followed by the Beta model and Tobit model, and the difference between MLOGIT and Beta was very small.

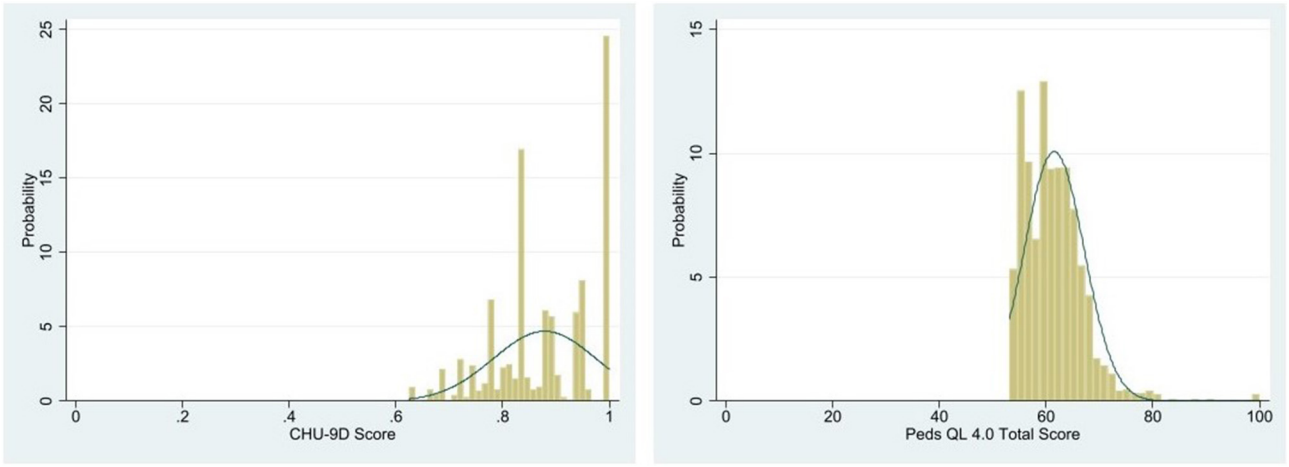


FIGURE 1
Distributions of CHU-9D utility score and Peds QL 4.0 total score.

Validation

The results showed that MLOGIT had the best performance. However, the absence of some dimension levels may lead to abnormal or biased fitting results of MLOGIT. Moreover, there was a small difference between MLOGIT and the second-best model in each index. Thus, we preliminarily concluded that choosing the second-best model would be more helpful to obtain accurate results. MLOGIT was not validated in this study.

Table 4 summarizes the validation results of the model through the 10-fold cross-validation method. In all combinations, Beta and Tobit performed better on the validation index. In Combinations 1 to 2, Beta had the lowest ranking and better performance than other models. In Combination 3, MAE and CCC of Beta were superior to Tobit, while RMSE and adj R^2 of Tobit were superior. The comprehensive ranking of the two models was consistent. Considering that RMSE was more sensitive to potential outliers, more weight could be given to MAE in this case (51). Therefore, we concluded that Beta was better in this combination. In Combinations 4 to 6, the Tobit had the lowest comprehensive ranking and better performance than other models. Of all the combinations, Tobit had the best MAE (0.0559), RMSE (0.0685), and adj R^2 (0.4973) in Combination 5, and Beta had the best CCC (0.7278) in Combination 6.

Best-performing mapping algorithm

According to the comprehensive ranking of the four indexes, the Tobit with Peds QL 4.0 item scores, gender, and age as independent variables (Combination 6) was the best model to predict the CHU-9D utility score. However, the Peds QL 4.0 item scores were difficult to obtain in reality. Hence, we provided the parameters of the optimal mapping algorithm for the CHU-9D utility score of different combinations (Table 5) and the consistency between the predicted CHU-9D utility score and the observed CHU-9D utility score (Figure 2). The Pearson correlation coefficients were 0.6836, 0.6788, 0.6845, 0.7092, 0.7455, and 0.7476 for Combinations 1 to 6, respectively. They indicated a high correlation between the observed CHU-9D utility scores and predicted CHU-9D utility scores. On the basis of these results, we suggest that researchers choose different variable combinations and corresponding mapping algorithms depending on the available data.

Mapping algorithm

Based on the conclusions from the section best-performing mapping algorithm, the mapping algorithm formula can be shown as follows:

Combination 1–3: Beta Model.

$$Beta\ Prediction = Intercept + \beta 1 * X1 + \beta 2 * X2 \ldots + \beta i * Xi$$
$$CHU - 9D\ Prediction = exp(Beta\ Prediction)/(1 + exp(Beta\ Prediction))$$

Combination 4–6: Tobit Model.

$$y = Intercept + \beta 1 * X1 + \beta 2 * X2 \ldots + \beta i * Xi$$
$$CHU - 9D\ Prediction = \begin{cases} 1 & if\ y > 1 \\ y & otherwise \end{cases}$$

[X_i is the independent variable, such as total Score, gender, and age. β_i is the coefficient (parameters are presented in Table 4)].

Discussion

Up to now, three types of research have constructed mapping algorithms from Peds QL 4.0 to CHU-9D based on children and adolescents (19–22), but none of them are based on the Chinese population. Considering the differences and uniqueness of populations in different countries, this study is the first to construct a mapping algorithm from Peds QL 4.0 to CHU-9D based on Chinese children and adolescent FD patients. This mapping algorithm is able to get HUVs by converting the non-preference-based Peds QL 4.0 into preference-based CHU-9D scores. It can

TABLE 3 Goodness-of-fit results from the full estimation sample ($N = 2,152$).

Model	Mean P (SD)	Min P	Max P	MAE	RMSE	Adjr^2	CCC
Combination 1							
OLS	0.8787 (0.0551)	0.7966	1.2554	0.0641	0.0794	0.3246	0.4905
GLM	0.8789 (0.0504)	0.8095	1.2709	0.0650	0.0810	0.2973	0.4478
MM	0.8841 (0.0661)	0.7854	1.3366	0.0641	0.0803	0.3084	0.5304
TOBIT	0.8767 (0.0473)	0.7966	1.0000	0.0621	0.0770	0.3646	0.4877
BETA	0.8862 (0.0678)	0.7113	0.9998	**0.0576**	0.0719	0.4465	**0.6308**
MLOGIT	0.8787 (0.0603)	0.7751	0.9987	0.0580	**0.0716**	**0.4498**	0.6045
Combination 2							
OLS	0.8787 (0.0573)	0.7968	1.3113	0.0617	0.0778	0.3516	0.5212
GLM	0.8789 (0.0532)	0.8085	1.3626	0.0624	0.0792	0.3269	0.4844
MM	0.8832 (0.0674)	0.7867	1.3904	0.0616	0.0786	0.3385	0.5564
TOBIT	0.8771 (0.0511)	0.7968	1.0000	0.0602	0.0757	0.3862	0.5211
BETA	0.8861 (0.0679)	0.7303	0.9999	**0.0572**	**0.0716**	**0.4504**	**0.6342**
MLOGIT	0.8787 (0.0617)	0.7763	0.9991	0.0577	0.0719	0.4465	0.6073
Combination 3							
OLS	0.8787 (0.0613)	0.7615	1.1514	0.0604	0.0747	0.4023	0.5749
GLM	0.8788 (0.059)	0.7730	1.1871	0.0610	0.0756	0.3869	0.5542
MM	0.8844 (0.0652)	0.7594	1.1799	0.0597	0.0750	0.3969	0.5875
TOBIT	0.8774 (0.0576)	0.7615	1.0000	0.0591	0.0732	0.4250	0.5768
BETA	0.8844 (0.0783)	0.6222	0.9987	0.0584	0.0735	0.4211	**0.6522**
MLOGIT	0.8787 (0.0636)	0.7610	0.9981	**0.0559**	**0.0692**	**0.4863**	0.6424
Combination 4							
OLS	0.8787 (0.0647)	0.7616	1.2422	0.0577	0.0718	0.4478	0.6200
GLM	0.8788 (0.0625)	0.7722	1.2736	0.0584	0.0728	0.4327	0.6012
MM	0.884 (0.0682)	0.7619	1.2772	0.0572	0.0721	0.4431	0.6302
TOBIT	0.8776 (0.0614)	0.7616	1.0000	0.0566	0.0705	0.4671	0.6214
BETA	0.8839 (0.0809)	0.6482	0.9996	0.0575	0.0725	0.4364	**0.6704**
MLOGIT	0.8787 (0.067)	0.7591	0.9994	**0.0543**	**0.0677**	**0.5093**	0.6695
Combination 5							
OLS	0.8787 (0.0699)	0.7471	1.1365	0.0530	0.0669	0.5207	0.6873
GLM	0.8788 (0.0686)	0.7566	1.1429	0.0534	0.0675	0.5114	0.6773
MM	0.8853 (0.0718)	0.7485	1.1651	0.0520	0.0673	0.5140	0.6899
TOBIT	0.878 (0.0682)	0.7471	1.0000	0.0523	0.0661	0.5318	0.6893
BETA	0.8832 (0.0852)	0.6559	0.9994	0.0529	0.0676	0.5100	**0.7265**
MLOGIT	0.8787 (0.0711)	0.7405	0.9994	**0.0502**	**0.0633**	**0.5709**	0.7235
Combination 6							
OLS	0.8787 (0.0702)	0.7548	1.1417	0.0529	0.0665	0.5256	0.6919
GLM	0.8788 (0.069)	0.7644	1.1488	0.0533	0.0672	0.5162	0.6818
MM	0.8852 (0.0718)	0.7511	1.1649	0.0519	0.0670	0.5189	0.6931
TOBIT	0.878 (0.0685)	0.7548	1.0000	0.0521	0.0658	0.5358	0.6934
BETA	0.8831 (0.0853)	0.6608	0.9994	0.0525	0.0670	0.5188	**0.7318**
MLOGIT	0.8787 (0.0715)	0.7420	0.9995	**0.0501**	**0.0630**	**0.5745**	0.7271

Numbers in bold were the best values in each combination. OLS, Ordinary least squares regression; GLM, the general linear regression model; MM, MM-estimator model; Tobit, Tobit regression model; Beta, Beta regression model; MLOGIT, multinomial logistic regression. Mean P, mean of the health utility score predicted; Min P, minimum of the health utility score predicted; Max P, maximum of the health utility score predicted; MAE, mean absolute error; RMSE, root mean square error; adjR^2, adjusted R^2; CCC, concordance correlation coefficient.

10.3389/fpubh.2023.1166760

TABLE 4 Goodness-of-fit results from the 10-fold cross-validation.

Model	Mean P (SD)	Min P	Max P	MAE	RMSE	Adj R^2	CCC
Combination 1							
OLS	0.8788 (0.0023)	0.8027	1.1816	0.0647	0.0798	0.3200	0.4889
GLM	0.8789 (0.0022)	0.8143	1.1874	0.0656	0.0814	0.2919	0.4460
MM	0.8841 (0.0028)	0.7928	1.2471	0.0647	0.0806	0.3042	0.5291
TOBIT	0.8767 (0.0019)	0.8027	1.0000	0.0626	0.0773	0.3615	0.4871
BETA	0.8862 (0.0034)	0.7317	0.9982	**0.0581**	**0.0721**	**0.4435**	**0.6300**
Combination 2							
OLS	0.8787 (0.003)	0.7991	1.1785	0.0629	0.0785	0.3403	0.5190
GLM	0.8789 (0.003)	0.8099	1.1903	0.0636	0.0801	0.3139	0.4820
MM	0.8832 (0.0035)	0.7897	1.2349	0.0628	0.0792	0.3279	0.5547
TOBIT	0.8771 (0.0026)	0.7991	1.0000	0.0613	0.0763	0.3769	0.5197
BETA	0.886 (0.0035)	0.7360	0.9981	**0.0583**	**0.0722**	**0.4416**	**0.6326**
Combination 3							
OLS	0.8788 (0.0031)	0.7849	1.1086	0.0618	0.0756	0.3885	0.5733
GLM	0.8789 (0.003)	0.7926	1.1157	0.0625	0.0766	0.3724	0.5527
MM	0.8845 (0.0034)	0.7835	1.1266	0.0611	0.0759	0.3828	0.5860
TOBIT	0.8774 (0.0031)	0.7849	1.0000	0.0604	**0.0741**	**0.4133**	0.5758
BETA	0.8843 (0.0042)	0.7025	0.9968	**0.0597**	0.0743	0.4092	**0.6507**
Combination 4							
OLS	0.8788 (0.0034)	0.7802	1.1206	0.0597	0.0731	0.4278	0.6176
GLM	0.8789 (0.0033)	0.7881	1.1356	0.0604	0.0742	0.4103	0.5983
MM	0.8884 (0.0101)	0.7841	1.1459	0.0592	0.0744	0.4065	0.6229
TOBIT	0.8775 (0.0035)	0.7802	1.0000	**0.0584**	**0.0717**	**0.4504**	0.6196
BETA	0.8838 (0.0041)	0.6964	0.9959	0.0593	0.0736	0.4194	**0.6686**
Combination 5							
OLS	0.8788 (0.004)	0.7608	1.0938	0.0568	0.0694	0.4833	0.6826
GLM	0.8789 (0.0039)	0.7681	1.1022	0.0572	0.0701	0.4730	0.6725
MM	0.8893 (0.0088)	0.7705	1.1301	0.0560	0.0712	0.4557	0.6739
TOBIT	0.878 (0.0038)	0.7608	1.0000	**0.0559**	**0.0685**	**0.4973**	0.6852
BETA	0.8832 (0.0049)	0.6937	0.9979	0.0566	0.0700	0.4744	**0.7229**
Combination 6							
OLS	0.8788 (0.0040)	0.7633	1.0937	0.0572	0.0695	0.4824	0.6865
GLM	0.8789 (0.0039)	0.7705	1.1023	0.0577	0.0702	0.4716	0.6762
MM	0.8852 (0.0042)	0.7641	1.1107	0.0563	0.0699	0.4759	0.6884
TOBIT	0.8779 (0.0038)	0.7633	1.0000	**0.0563**	**0.0686**	**0.4962**	0.6887
BETA	0.8831 (0.0047)	0.7070	0.9975	0.0568	0.0698	0.4782	**0.7278**

Numbers in bold were the best values in each combination. OLS, Ordinary least squares regression; GLM, the general linear regression model; MM, MM-estimator model; Tobit, Tobit regression model; Beta, Beta regression model; MLOGIT, multinomial logistic regression. Mean P, mean of the health utility score predicted; Min P, minimum of the health utility score predicted; Max P, maximum of the health utility score predicted; MAE, mean absolute error; RMSE, root mean square error; adjR^2, adjusted R^2; CCC, concordance correlation coefficient.

TABLE 5 Regression outputs of the mapping model.

	Combination1	Combination2	Combination3	Combination4	Combination5	Combination6
Intercept	−7.9832 (0.2245)***	−7.6672 (0.2247)***	−4.0199 (0.2942)***	0.5415 (0.0263)***	0.8991 (0.0301)***	0.8652 (0.0313)***
Total score/100	16.7248 (0.3761)***	15.8034 (0.4217)***				
PF score/100			3.2633 (0.3777)***	−0.1221 (0.0465)**		
EF score/100			−5.5249 (0.4787)***	−0.2637 (0.0409)***		
SF score/100			9.2555 (0.3088)***	0.5415 (0.0259)***		
ScF score/100			3.3013 (0.4243)***	0.1888 (0.0373)***		
PF2/100					−0.1566 (0.0283)***	−0.1721 (0.0283)***
PF5/100					0.2031 (0.0123)***	0.1234 (0.0209)***
PF8/100					−0.1554 (0.0339)***	−0.1491 (0.0339)***
EF1/100					−0.1134 (0.0221)***	−0.0996 (0.0222)***
EF4/100					−0.0721 (0.0236)**	−0.0479 (0.0240)*
EF5/100					0.0423 (0.0176)*	0.0338 (0.0175)
SF1/100					0.0605 (0.0162)***	0.0581 (0.0162)***
SF3/100					0.2412 (0.0188)***	0.2432 (0.0187)***
SF4/100					0.1554 (0.0209)***	0.1617 (0.0208)***
ScF1/100					0.1453 (0.0120)***	0.1403 (0.0120)***
ScF2/100					0.0755 (0.0134)***	0.064 (0.0135)***
ScF3/100					−0.1907 (0.0360)***	−0.1996 (0.0360)***
ScF4/100					−0.1386 (0.0349)***	−0.1218 (0.0349)***
Female		−0.1698 (0.0387)***		−0.0058 (0.0033)		0.0016 (0.0030)
Age		0.0446 (0.0155)**		0.0202 (0.0016)***		0.0089 (0.0019)***

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. The Peds QL 4.0 scores used here should be the scaled scores (divided by 100). Total score, Peds QL 4.0 Total score; PF, Physical Functioning; EF, Emotional Functioning; SF, Social Functioning; ScF, School Functioning. PF2, Hard to run; PF5, Hard to take a bath or shower; PF8, Low energy; EF1, Feel afraid or scared; EF4, Trouble sleeping; EF5, Worry about what will happen; SF1, Trouble getting along with others; SF3, Teased; SF4, Not able to do things that others can do; ScF1, Hard to pay attention in class; ScF2, Forget things; ScF3, Trouble keeping up with schoolwork; ScF4, Miss school because of not feeling well.

improve both the efficiency of available clinical data and help decision makers to compare and evaluate relevant interventions, facilitating the development of pharmacoeconomic evaluation.

In accordance with guidance (17), we used six regression models to develop mapping algorithms from Peds QL 4.0 to CHU-9D, including OLS, GLM, MM, Tobit, Beta, and MLOGIT, among which Tobit performed best, followed by Beta. Some studies indicated that indirect mapping could improve the fitting of models compared with direct mapping using linear models (52). In this study, although the sample size was large and the MLOGIT performance was good, we still believed that Tobit or Beta was better. This was mainly due to the lack of dimension level of CHU-9D, which leads to the concentration in some health states. This may lead to an offset in the final result of MLOGIT. In addition, we also found that it could easily lead to fitting abnormalities in MLOGIT fitting process. Therefore, our final results did not use the MLOGIT due to the fact that it requires more comprehensive health state data to prove its accuracy. Furthermore, previous studies had shown that MM was superior to MLOGIT, which was different from our results (19).

In the existing mapping research of Peds QL 4.0, the Peds QL 4.0 total score, the Peds QL 4.0 dimension scores, and the Peds QL 4.0 item scores were often used as independent variables (19–21, 24), as well as some demographic indicators such as age and gender. We screened the independent variables according to the correlation between the variables (see Appendix 1). Some variables that were highly correlated with CHU-9D utility score were directly included as independent variables, for instance, Peds QL 4.0 total score (0.6836), SF (0.6399), and ScF (0.5066). To improve the prediction ability of the model as much as possible, we also screened the variables with low correlation according to the independent variables included in previous studies, such as trying to include gender as an independent variable. In addition, we tried to use the Peds QL 4.0 item scores as independent variables. Due to the large number of Peds QL 4.0 item scores, in order to avoid excessive complexity of the model, we used the stepwise regression method to select Peds QL 4.0 item scores that had a significant correlation with CHU-9D utility value. Finally, the Peds QL 4.0 total score, Peds QL 4.0 dimension scores, Peds QL 4.0 item scores, gender, and age were used as independent variables. We found that when age and gender were used as independent variables, the 10-fold cross-validation and the fitting of the full estimation sample showed different results. In the full estimation sample, the inclusion of age and gender improved the performance of the model. However, in

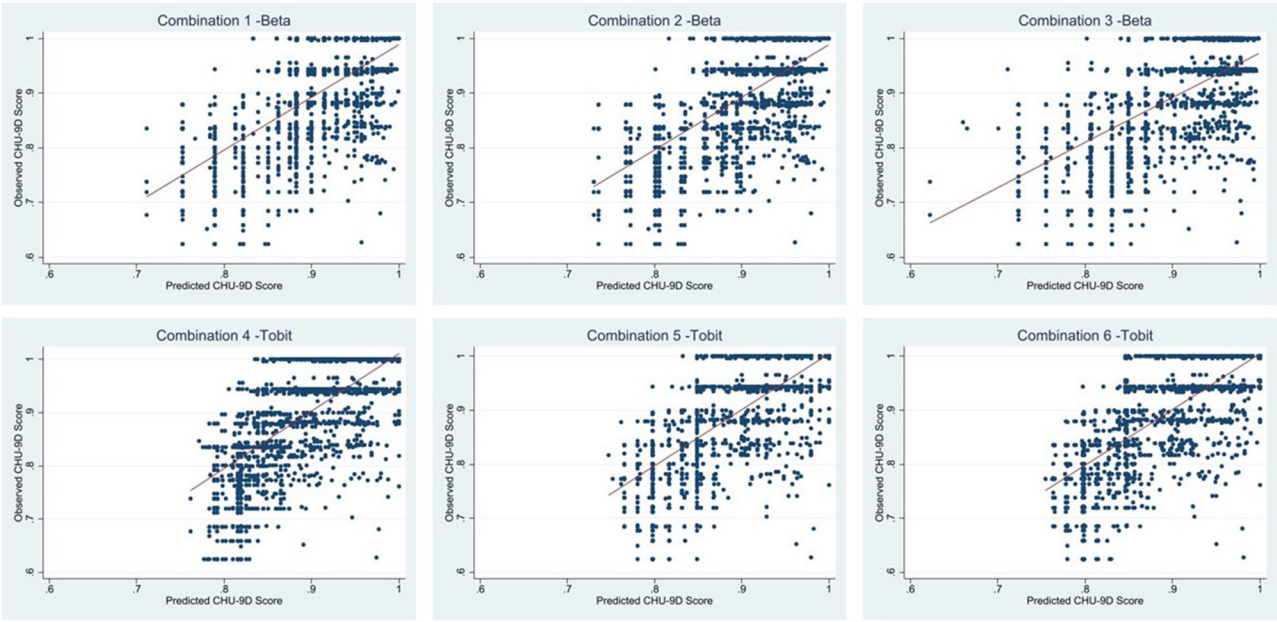


FIGURE 2
The observed and predicted CHU-9D utility score.

the 10-fold cross-validation, only the model performance of the combination of independent variables based on the Peds QL 4.0 dimension scores was improved. The result was similar to Tosin's research (20). Considering the obvious influence of gender and age on the population HUVs, we propose that it is necessary to include age and gender in the construction of the mapping algorithm, but more research is needed to prove the difference.

This study evaluated the prediction performance of different models in different combinations based on full samples and ranked the model performance based on MAE, RMSE, adj R^2, and CCC using the 10-fold cross-validation method. Similar to several other studies, almost all the estimators overestimated the lower bound of the CHU-9D utility score and overpredicted the upper bound of 1 (21, 22). Although overprediction is a difficult problem, some studies have shown that it can be dealt with by truncating or setting the value of the dependent variable between 0 and 1 to the boundary value, such as Tobit model and Beta model (24, 53, 54). Overall, the predictive performance of the Beta model and Tobit model was consistent with the MAE range (0.0408–0.1270) and RMSE range (0.0594–0.1619) observed in the CHU-9D mapping study (19–23).

There are also some limitations in the study. First, although the sample size was large, the sample source was limited to Zhejiang province, which was difficult to represent the level of China. Second, we used the 10-fold cross-validation method to verify the mapping algorithm, but no external validation was performed due to the lack of available data. Third, in the process of data sampling and collection, many samples are concentrated in the lower age group (average age = 7.23), and nearly half of the questionnaires were filled by guardians, which may be related to the factors that children in the lower age group are more likely to receive attention and seek medical treatment in time. However, guardians could hard to understand the real feelings of young people, which may cause agency bias. Fourth, there were some missing CHU-9D dimension levels in the study sample, which may lead to the limited prediction range of the mapping algorithm and the deviation of indirect mapping results (17). Therefore, the model based on indirect mapping was not recommended in the study, and better data were needed to verify in future.

Conclusion

The research first developed a mapping algorithm from Peds QL 4.0 to CHU-9D based on Chinese children and adolescents. We also constructed different mapping algorithms for different combinations of variables, and of all algorithms, the Tobit model with Peds QL 4.0 item scores, gender, and age as the independent variables was the most accurate. However, researchers can reasonably choose mapping algorithms for different combinations of variables based on available data to conduct other studies such as pharmacoeconomic evaluations and provide references for relevant policymakers in China.

Data availability statement

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

Author contributions

QW, CW, and XX made their contributions to the conception and design of the study. QW, YH, and CW made their contributions to the acquisition and analysis of the data. ML and YH made their contributions to the interpretation of data. QW made contributions to the drafting of the study. QW, ML, and XX made their contributions to the revision of the study. All authors of

this study has approved the submitted version and has agreed both to be personally accountable for the author's own contributions and ensure that questions related to the accuracy or integrity of any part of the study.

Acknowledgments

We sincerely express our gratitude to all FD patients and their guardians who participated in the study. Without their help, this study would not have been carried out successfully. In addition, we thank the experts and professionals for their suggestions and guidance during the study.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh.2023.1166760/full#supplementary-material

References

1. Wei Z, Yang X, Xing X, Dong L, Wang J, Qin B. Risk factors associated with functional dyspepsia in Chinese children: a cross-sectional study. *BMC Gastroenterol.* (2021) 21:218. doi: 10.1186/s12876-021-01800-x

2. Korterink JJ, Diederen K, Benninga MA, Tabbers MM. Epidemiology of pediatric functional abdominal pain disorders: a meta-analysis. *PLoS ONE.* (2015) 10:e0126982. doi: 10.1371/journal.pone.0126982

3. Robin SG, Keller C, Zwiener R, Hyman PE, Nurko S, Saps M, et al. Prevalence of pediatric functional gastrointestinal disorders utilizing the rome IV criteria. *J Pediatr.* (2018) 195:134–9. doi: 10.1016/j.jpeds.2017.12.012

4. The Subspecialty Group of Gastroenterology tSoP, Chinese Medical Association; the Subspecialty Group of Gastroenterology, China Association of Traditional Chinese Medicine Children's Health and Drug Research Professional Committee; the Editorial Board, Chinese Journal of Pediatrics. Expert consensus on diagnosis and treatment of functional dyspepsia in Chinese children (2022). *Chin J Pediatrics.* (2022) 60:751–5. doi: 10.3760/cma.j.cn112140-20220417-00339

5. Brook RA, Kleinman NL, Choung RS, Melkonian AK, Smeeding JE, Talley NJ. Functional dyspepsia impacts absenteeism and direct and indirect costs. *Clin Gastroenterol Hepatol.* (2010) 8:498–503. doi: 10.1016/j.cgh.2010.03.003

6. Rippel SW, Acra S, Correa H, Vaezi M, Di Lorenzo C, Walker LS. Pediatric patients with dyspepsia have chronic symptoms, anxiety, and lower quality of life as adolescents and adults. *Gastroenterology.* (2012) 142:754–61. doi: 10.1053/j.gastro.2011.12.043

7. Shelby GD, Shirkey KC, Sherman AL, Beck JE, Haman K, Shears AR, et al. Functional abdominal pain in childhood and long-term vulnerability to anxiety disorders. *Pediatrics.* (2013) 132:475–82. doi: 10.1542/peds.2012-2191

8. Colombo JM, Deacy AD, Schurman JV, Friesen CA. Heartburn in children and adolescents in the presence of functional dyspepsia and/or irritable bowel syndrome correlates with the presence of sleep disturbances, anxiety, and depression. *Medicine.* (2021) 100:e25426. doi: 10.1097/MD.0000000000025426

9. Chen G, Ratcliffe J. A review of the development and application of generic multi-attribute utility instruments for paediatric populations. *Pharmacoeconomics.* (2015) 33:1013–28. doi: 10.1007/s40273-015-0286-7

10. Anle S, Shunguo Z, Nan L, Bin W. Agreement of EQ-5D-Y between children with hematological malignancies and their caregivers on quality of life. *Chin J Pharmacoepidemiol.* (2021) 30:415–20. doi: 10.19960/j.cnki.issn1005-0698.2021.06.011

11. Devanarayana NM, Rajindrajith S, Benninga MA. Quality of life and health care consultation in 13 to 18 year olds with abdominal pain predominant functional gastrointestinal diseases. *BMC Gastroenterol.* (2014) 14:150. doi: 10.1186/1471-230X-14-150

12. Lewis ML, Palsson OS, Whitehead WE, van Tilburg MAL. Prevalence of functional gastrointestinal disorders in children and adolescents. *J Pediatrics.* (2016) 177:39–43. doi: 10.1016/j.jpeds.2016.04.008

13. Ranasinghe N, Devanarayana NM, Rajindrajith S, Perera MS, Nishanthinie S, Warnakulasuriya T, et al. Functional gastrointestinal diseases and psychological maladjustment, personality traits and quality of life. *BMC Gastroenterol.* (2018) 18:33. doi: 10.1186/s12876-018-0760-8

14. Varni JW, Seid M, Knight TS, Uzark K, Szer IS. The PedsQL 4. 0 Generic core scales: sensitivity, responsiveness, and impact on clinical decision-making. *J Behav Med.* (2002) 25:175–93. doi: 10.1023/A:1014836921812

15. Xu F, Chen G, Stevens K, Zhou H, Qi S, Wang Z, et al. Measuring and valuing health-related quality of life among children and adolescents in mainland China–a pilot study. *PLoS ONE.* (2014) 9:e89222. doi: 10.1371/journal.pone.0089222

16. Stevens K. Developing a descriptive system for a new preference-based measure of health-related quality of life for children. *Qual Life Res.* (2009) 18:1105–13. doi: 10.1007/s11136-009-9524-9

17. Wailoo AJ, Hernandez-Alava M, Manca A, Mejia A, Ray J, Crawford B, et al. Mapping to estimate health-state utility from non-preference-based outcome measures: an ISPOR good practices for outcomes research task force report. *Value Health.* (2017) 20:18–27. doi: 10.1016/j.jval.2016.11.006

18. Longworth L, Rowen D. Mapping to obtain EQ-5D utility values for use in NICE health technology assessments. *Value Health.* (2013) 16:202–10. doi: 10.1016/j.jval.2012.10.010

19. Mpundu-Kaambwa C, Chen G, Russo R, Stevens K, Petersen KD, Ratcliffe J. Mapping CHU9D Utility Scores from the PedsQL(TM) 4. 0 SF-15. *Pharmaco Econ.* (2017) 35:453–67. doi: 10.1007/s40273-016-0476-y

20. Lambe T, Frew E, Ives NJ, Woolley RL, Cummins C, Brettell EA, et al. Mapping the paediatric quality of life inventory (PedsQL^{TM}) generic core scales onto the child health utility index-9 dimension (CHU-9D) score for economic evaluation in children. *Pharmacoeconomics.* (2018) 36:451–65. doi: 10.1007/s40273-017-0600-7

21. Sweeney R, Chen G, Gold L, Mensah F, Wake M. Mapping PedsQL(TM) scores onto CHU9D utility scores: estimation, validation and a comparison of alternative instrument versions. *Qual Life Res.* (2020) 29:639–52. doi: 10.1007/s11136-019-02357-9

22. Kelly CB, Soley-Bori M, Lingam R, Forman J, Cecil L, Newham J, et al. Mapping PedsQL^{TM} scores to CHU9D utility weights for children with chronic conditions in a multi-ethnic and deprived metropolitan population. *Qual Life Res.* (2023) 23:1–15. doi: 10.1007/s11136-023-03359-4

23. Robinson T, Oluboyede Y. Estimating CHU-9D utility scores from the WAItE: a mapping algorithm for economic evaluation. *Value Health.* (2019) 22:239–46. doi: 10.1016/j.jval.2018.09.2839

24. Shafie AA, Chhabra IK, Wong JHY, Mohammed NS. Mapping PedsQL^{TM} Generic Core Scales to EQ-5D-3L utility scores in transfusion-dependent thalassemia patients. *Eur J Health Econ.* (2021) 22:735–47. doi: 10.1007/s10198-021-01287-z

25. Canaway AG, Frew EJ. Measuring preference-based quality of life in children aged 6-7 years: a comparison of the performance of the CHU-9D and EQ-5D-Y–the WAVES pilot study. *Qual Life Res.* (2013) 22:173–83. doi: 10.1007/s11136-012-0119-5

26. Rasquin A, Di Lorenzo C, Forbes D, Guiraldes E, Hyams JS, Staiano A, et al. Childhood functional gastrointestinal disorders: child/adolescent. *Gastroenterology.* (2006) 130:1527–37. doi: 10.1053/j.gastro.2005.08.063

27. Chen G, Xu F, Huynh E, Wang Z, Stevens K, Ratcliffe J. Scoring the Child Health Utility 9D instrument: estimation of a Chinese child and adolescent-specific tariff. *Qual Life Res.* (2019) 28:163–76. doi: 10.1007/s11136-018-2032-z

28. Pei W, Yue S, Zhi-Hao Y, Ruo-Yu Z, Bin W, Nan L. Testing measurement properties of two EQ-5D youth versions and KIDSCREEN-10 in China. *Eur J Health Econ.* (2021) 22:1083–93. doi: 10.1007/s10198-021-01307-y

29. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ispor task force for translation and cultural adaptation. *Value Health.* (2005) 8:94–104. doi: 10.1111/j.1524-4733.2005.04054.x

30. Ratcliffe J, Couzner L, Flynn T, Sawyer M, Stevens K, Brazier J, et al. Valuing Child Health Utility 9D health states with a young adolescent sample: a feasibility study to compare best-worst scaling discrete-choice experiment, standard gamble and time trade-off methods. *Appl Health Econ Health Policy.* (2011) 9:15–27. doi: 10.2165/11536960-000000000-00000

31. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol.* (1993) 46:1417–32. doi: 10.1016/0895-4356(93)90142-N

32. Acquadro C, Conway K, Hareendran A, Aaronson N. Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. *Value Health.* (2008) 11:509–21. doi: 10.1111/j.1524-4733.2007.00292.x

33. Yi-Yun L, Qi T, Yuan T, Jin J, Yu-Deng L, Ding-Qun H. Reliability and validity for chinese version of pediatric quality of life inventory PedsQL4. 0. *J Sun Yat-sen Univ.* (2008) 328:31.

34. Wu HH Li H, Gao Q. Psychometric properties of the Chinese version of the pediatric quality of life inventory 4. 0 Generic core scales among children with short stature. *Health Qual Life Outcomes.* (2013) 11:87. doi: 10.1186/1477-7525-11-87

35. Hao Y, Tian Q, Lu Y, Chai Y, Rao S. Psychometric properties of the chinese version of the pediatric quality of life inventory 4. *Int J Quality Life Aspects Treatment Rehab.* (2010) 19:1229–33. doi: 10.1007/s11136-010-9672-y

36. Petrou S, Rivero-Arias O, Dakin H, Longworth L, Oppe M, Froud R, et al. The MAPS reporting statement for studies mapping onto generic preference-based outcome measures: explanation and elaboration. *Pharmacoeconomics.* (2015) 33:993–1011. doi: 10.1007/s40273-015-0312-9

37. Round J, Hawton A. Statistical alchemy: conceptual validity and mapping to generate health state utility values. *PharmacoEconomics.* (2017) 1:233–9. doi: 10.1007/s41669-017-0027-2

38. Vilsbøll AW, Kragh N, Hahn-Pedersen J, Jensen CE. Mapping dermatology life quality index (DLQI) scores to EQ-5D utility scores using data of patients with atopic dermatitis from the national health and wellness study. *Qual Life Res.* (2020) 29:2529–39. doi: 10.1007/s11136-020-02499-1

39. The British Medical Journal. *Correlation and Regression*. (2018). Available online at: https://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/11-correlation-and-regression (accessed February 2,2023).

40. Brazier JE, Yang Y, Tsuchiya A, Rowen DL, A. review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. *Eur J Health Econ.* (2010) 11:215–25. doi: 10.1007/s10198-009-0168-z

41. Hernández Alava M, Wailoo AJ, Ara R. Tails from the peak district: adjusted limited dependent variable mixture models of EQ-5D questionnaire health state utility values. *Value Health.* (2012) 15:550–61. doi: 10.1016/j.jval.2011.12.014

42. Longworth L, Yang Y, Young T, Mulhern B, Hernández Alava M, Mukuria C, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. *Health Technol Assessm.* (2014) 18:1–224. doi: 10.3310/hta18090

43. Yang F, Wong CKH, Luo N, Piercy J, Moon R, Jackson J. Mapping the kidney disease quality of life 36-item short form survey (KDQOL-36) to the EQ-5D-3L and the EQ-5D-5L in patients undergoing dialysis. *Eur J Health Econ.* (2019) 20:1195–206. doi: 10.1007/s10198-019-01088-5

44. Lamu AN, Olsen JA. Testing alternative regression models to predict utilities: mapping the QLQ-C30 onto the EQ-5D-5L and the SF-6D. *Qual Life Res.* (2018) 27:2823–39. doi: 10.1007/s11136-018-1981-6

45. Xu RH, Dong D, Luo N, Wong EL, Yang R, Liu J, et al. Mapping the Haem-A-QoL to the EQ-5D-5L in patients with hemophilia. *Qual Life Res.* (2022) 31:1533–44. doi: 10.1007/s11136-021-03051-5

46. Verardi V, Croux C. Robust regression in stata. *Stata J.* (2008) 9:439–53. doi: 10.1177/1536867X0900900306

47. Hawton A, Green C, Telford C, Zajicek J, Wright D. Using the Multiple Sclerosis Impact Scale to estimate health state utility values: mapping from the MSIS-29, version 2, to the EQ-5D and the SF-6D. *Value Health.* (2012) 15:1084–91. doi: 10.1016/j.jval.2012.07.007

48. Sullivan PW, Ghushchyan V. Mapping the EQ-5D index from the SF-12: US general population preferences in a nationally representative sample. *Med Decis Making.* (2006) 26:401–9. doi: 10.1177/0272989X06290496

49. Basu A, Manca A. Regression estimators for generic health-related quality of life and quality-adjusted life years. *Med Decis Making.* (2012) 32:56–69. doi: 10.1177/0272989X11416988

50. Hunger M, Döring A, Holle R. Longitudinal beta regression models for analyzing health-related quality of life scores over time. *BMC Med Res Methodol.* (2012) 12:144. doi: 10.1186/1471-2288-12-144

51. Shcherbakov M, Brebels A, Shcherbakova NL, Tyukov A, Janovsky TA, Kamaev VA, et al. A survey of forecast error measures. *World Appl Sci J.* (2013) 24:171–6. doi: 10.5829/idosi.wasj.2013.24.itmies.80032

52. Hernández Alava M, Wailoo A, Wolfe F, Michaud K, A. comparison of direct and indirect methods for the estimation of health utilities from clinical outcomes. *Med Decis Making.* (2014) 34:919–30. doi: 10.1177/0272989X13500720

53. Liu T, Li S, Wang M, Sun Q, Chen G. Mapping the Chinese version of the EORTC QLQ-BR53 onto the EQ-5D-5L and SF-6D utility scores. *Patient.* (2020) 13:537–55. doi: 10.1007/s40271-020-00422-x

54. Meunier A, Soare A, Chevrou-Severac H, Myren KJ, Murata T, Longworth L. Indirect and direct mapping of the cancer-specific EORTC QLQ-C30 onto EQ-5D-5L utility scores. *Appl Health Econ Health Policy.* (2022) 20:119–31. doi: 10.1007/s40258-021-00682-0

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 09 June 2023
DOI 10.3389/fpubh.2023.1177317

OPEN ACCESS

EDITED BY
Pei Wang,
Fudan University, China

REVIEWED BY
Zhihao Yang,
Erasmus Medical Center, Netherlands
Anna Rozensztrauch,
Wroclaw Medical University, Poland
Li Cai,
Sun Yat-sen University, China

*CORRESPONDENCE
Chaochun Zou
✉ cc14@zju.edu.cn
Guannan Bai
✉ guannanbai@zju.edu.cn

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

‡These authors have contributed equally to this work and share last authorship

RECEIVED 01 March 2023
ACCEPTED 11 May 2023
PUBLISHED 09 June 2023

CITATION
Chen W, Sun L, He X, Li Z, Ji C, Li F, Shen J, Pan T, Jin X, Dong Y, Hu L, Zou C and Bai G (2023) Health-related quality of life of children with Williams syndrome and caregivers in China.
Front. Public Health 11:1177317.
doi: 10.3389/fpubh.2023.1177317

Health-related quality of life of children with Williams syndrome and caregivers in China

Weijun Chen[1†], Lidan Sun[2†], Xinyu He[2], Ziqiao Li[2], Chai Ji[1], Fangfang Li[1], Jiyang Shen[1], Tianxin Pan[3], Xuejing Jin[4], Yusang Dong[2], Lidan Hu[2], Chaochun Zou[5*‡] and Guannan Bai[2*‡]

[1]Department of Child Health Care, Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Child Health, Hangzhou, Zhejiang, China, [2]Children's Hospital, Zhejiang University School of Medicine, National Children's Regional Medical Center, National Clinical Research Center for Child Health, Hangzhou, Zhejiang, China, [3]Health Economics Unit, Centre for Health Policy, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, VIC, Australia, [4]Centre for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China, [5]Department of Endocrinology, Children's Hospital, Zhejiang University School of Medicine, National Children's Regional Medical Center, National Clinical Research Center for Child Health, Hangzhou, Zhejiang, China

Introduction: Williams syndrome (WS) is a rare genetic disorder that impacts multiple systems and may cause developmental delays. These medical and developmental issues impose a heavy burden on affected children and their families. However, there was no study on children's health-related quality of life (HRQoL) with WS and only two studies about family quality of life globally. Therefore, the primary purpose of this study was to assess the HRQoL of children with WS and their caregivers in China, and the secondary purpose was to identify the potential determinants of children's and caregivers' HRQoL.

Methods: In total, 101 children and caregivers were included. We applied the proxy-reported PedsQL 4.0 Generic Core Module (PedsQL GCM) and PedsQL 3.0 Family Impact Module (FIM) to measure the HRQoL of children and caregivers. Additionally, we collected information on a comprehensive set of social demographic and clinical characteristics. Differences in HRQoL scores across subgroups were assessed by two-independent-samples *t*-tests, one-way ANOVA, and *post hoc* tests. We also calculated effect sizes to indicate clinical relevance. Multivariate linear regression models were applied to assess the potential determinants of HRQoL.

Results: We found that the HRQoL of children with WS and their caregivers was dramatically worse than the norm average scores of the healthy controls of children published in previous studies. Paternal educational level, household income, and the perceived financial burden significantly influenced the HRQoL of both children and families (p-values < 0.05). Multivariate linear regression analysis showed that the perceived financial burden was independently associated with family quality of life (p-values < 0.05)., and the presence of sleeping problem was independently associated with children's HRQoL (p-value=0.01).

Conclusion: We call for attention from policymakers and other stakeholders on the health status and well-being of children with WS and their families. Supports are needed to relieve psychosocial distress and financial burden.

KEYWORDS

Williams syndrome, quality of life, PedsQL, children, family impact

Highlights

The media use some descriptions to introduce children with Williams syndrome (WS), such as "friendly to extreme," "the opposite of autism," and "happy smiling face". Are these children really happy? There has been no data on children's health-related quality of life (HRQoL) with WS. The present study has filled this knowledge gap by providing rich data on the HRQoL of a relatively large sample of Chinese children with WS and their caregivers using reliable and validated instruments. Notably, this study highlighted the significantly worse HRQoL of both children and caregivers than the healthy references. In addition, we must address that the low socioeconomic status indicated by parents' education, household income, and the subjective perception of financial burden impair multiple domains of HRQoL of children and caregivers, highlighting the importance of social support for families with WS.

1. Introduction

Williams syndrome (WS) or Williams-Beuren Syndrome (WBS) is a rare genetic disorder affecting approximately one in every 20,000 live births, with both genders equally involved (1). The prevalence was reported as 1/23,500 in Hong Kong and there is no such data in the mainland China (2). It is a complicated neurodevelopmental disorder caused by a deletion of around 26 genes on chromosome 7q11.23, and these genes are involved in the development of multiple systems (3). WS is found in all racial and ethnic groups. Individuals with WS often display distinctive facial features, including a broad forehead, full cheeks, and a small, upturned nose, known as a "pixie face." Individuals with WS usually have an extremely friendly and outgoing personality and a high degree of verbal ability. In addition, WS is associated with many medical characteristics, such as congenital heart disease, connective tissue abnormalities, and growth deficiency (4). Most WS children may experience developmental delays in early childhood and mild to moderate intellectual disability at school age (5). These medical and developmental issues greatly distress families and individuals with WS, significantly impacting the quality of life of affected children and their families.

Quality of life is a significant concept and a principal endpoint for research and practice in public health and medicine (6). According to the World Health Organization Quality of Life Assessment Group, it is defined as "an individual's perception of their position in the life in the context of the culture in which they live and in relation to their goals, expectations, standards and concerns" (7). Health-related quality of life (HRQoL) is a term referring to the health aspects of quality of life and is generally considered to reflect the impact of disease and treatment on disability and daily functioning; it has also been considered to reflect the impact of perceived health on an individual's ability to live a fulfilling life (8). HRQoL is a subjective and multiple-dimensional concept measuring physical, mental, and social health domains.

Children with WS and their families may face significant challenges to their physical, mental, and social health or well-being. Due to physical, developmental, and behavioral issues, children with WS may require ongoing medical care and treatment. Caregivers may also be exposed to significant emotional, financial, and practical distress when they have to provide ongoing care for their loved ones suffering from WS-related health problems. Based on the literature review, we only found two studies about quality of life-related to WS. Sepúlveda et al. in Spain found that the degree and the presence of intellectual disability did not homogeneously influence the family quality of life, and many variables such as material well-being and interpersonal development could affect the family quality of life (9). Pereira et al. in Brazil found that paternal education, children's cardiopathy, and autonomy in activities of daily living had an impact on the overall family quality of life (10). The studies mentioned above are about family quality of life, and the sample sizes were relatively small. To our knowledge, there was no study on HRQoL of children with WS. This might be explained by the fact that due to the intellectual and cognitive impairments, children with WS cannot evaluate their HRQoL by themselves, so proxy-reported HRQoL is an alternative. In addition, few instruments measure the HRQoL of pediatric patients with a broad age range, for instance, from infants to adolescents. The Pediatric Quality of Life Inventory (PedsQL), developed by Varni et al. in the United States, may be an appropriate measurement to evaluate the HRQoL of healthy and sick children at different ages (11, 12).

Therefore, the primary purpose of this study was to assess the HRQoL of children with WS and their caregivers in China, and the secondary purpose was to identify the potential determinants of children's and caregivers' HRQoL. Our study will benefit from a better understanding of the impacts of WS on children and their families and provide solid evidence for policymakers and other stakeholders, which eventually improves the HRQoL of WS patients.

2. Methods

2.1. Study design and participants

The present study was embedded in the cohort of pediatric patients with WS in China. Since 2009, we have established a WS cohort including over 300 patients who visited the Department of Child Health Care in the Children's Hospital affiliated with Zhejiang University School of Medicine in Hangzhou, China. In 2014, we launched a free Multi-Disciplinary Treatment (MDT) program. To follow up with these families, we had an online patient support group via WeChat (a Chinese social media application) and follow-up by phone. From November 2022 to February 2023, we conducted a questionnaire-based survey among 103 WS patients and their families. Included families were required to meet the two criteria: (1) their child was diagnosed as WS; (2) the age range was from 2 to 18 years. Patients whose caregivers were unwilling to fill in the questionnaire or share the clinical data were excluded from the study. In total, 101 patients and their caregivers completed the questionnaire and were included in analyses. Experienced experts developed the survey questionnaire with research or practice backgrounds in public health, patient-reported outcome measures, and pediatric clinicians in neurodevelopmental and behavioral disorders and rare diseases. The questionnaire includes four parts: social

demographic information, clinical information, caregiver's HRQoL, and proxy-reported HRQoL of children. Several rounds of discussion were organized. A pre-test survey was conducted among 12 parents of children with WS aged from 2 to 18 years, followed by a brief interview by phone to collect suggestions from caregivers. The final version questionnaire was revised accordingly by experts to ensure that this questionnaire was easy to understand and feasible for the online survey, reliable, and validated to measure the HRQoL of children and caregivers. The electronic questionnaire was made by the Wenjuanxing application (a survey platform) that generated a QR code for distribution in the WeChat group, i.e., the online WS patients/families peer-support group. Following the principle of voluntary participation, the study questionnaire was sent by a website to the caregivers with detailed information about the purpose and methods of this study. All eligible participants provided electronically signed consent forms before filling in the questionnaire. If caregivers had any questions, they could contact a research assistant who was a physician for further instruction.

The present study was conducted according to the Declaration of Helsinki and was approved by the Ethics Committee of the Children's Hospital of Zhejiang University School of Medicine (2019-IRB-122).

2.2. Measurements of general characteristics

Information on social demographic characteristics, including child's age, gender, maternal/paternal educational level, marital status, parents living together, household income, residence (urban/rural areas), insurance, cost related to WS treatment, and perceived financial burden due to WS was collected via questionnaire. Education was categorized into three subsequent levels based on the Chinese Standard Classification of Education: high level is a bachelor's degree or above; middle level includes high school, technical school, vocational secondary school, and vocational high school; and low level refers to no education, primary school or middle school. Living together was measured by a single question, "Do the child's parents live together" with three options (always, sometimes, and never). The perceived financial burden was measured by the question, "How much burden does the child's treatment place on the family?" with three options (easily affordable, affordable, hard to afford).

2.3. Measurements of clinical characteristics

We also collected a comprehensive set of clinical variables by a questionnaire including the age of onset of the disease, age of diagnosis, genetic test, rehabilitation, disorders at birth, congenital heart disease and surgery, symptoms of attention deficit and hyperactivity, the presence of health conditions of visual system/endocrine system, / digestive system /urinary system conditions, repeated respiratory tract infection, inguinal hernia, and sleeping problems.

2.4. Health-related quality of life

We applied the proxy-reported PedsQL 4.0 Generic Core Module (PedsQL GCM) and PedsQL 3.0 Family Impact Module (FIM) to measure the HRQOL of children and caregivers. Both instruments were validated in China with good reliability and validity (13, 14). The proxy-reported PedsQL 4.0 GCM is administrated to measure HRQOL in children and adolescents ages 2–18 years in the past month. It contains 23 items and four scales, i.e., physiological (eight items), emotional (five items), social (five items), and school functioning (five items). It is widely used for healthy children in communities and schools and can distinguish pediatric disease patients from healthy controls (15). The PedsQL 3.0 FIM was used to evaluate the impact of WS on the caregiver's QoL and family function in the past month. It contains 28 items and six scales, i.e., physical functioning (six items), emotional functioning (five items), social functioning (four items), cognition functioning (four items), communication (three items), worry (five items), daily activity (three items) and family relationships (five items) (13). Both instruments were filled by caregivers. In the present study, the Cronbach's alpha for PedsQL 3.0 FIM is 0.977 and for GCM is 0.907. The Cronbach's alpha coefficients for each domain of PedsQL 3.0 FIM were 0.938 (physical functioning), 0.941 (emotional functioning), 0.896 (social functioning), 0.959 (cognition functioning), 0.859 (communication), 0.906 (worry), 0.877 (daily activity) and 0.959 (family relationships), respectively. The Cronbach's alpha coefficients for each domain of PedsQL 4.0 GCM were 0.859 (physiological functioning), 0.854 (emotional functioning), 0.880 (social functioning), and 0.762 (school functioning), respectively.

The scoring algorithm of PedsQL 4.0 GCM and PedsQL 3.0 FIM is similar. Response choices were provided on a 0 to 4 scale (0 = never a problem, 1 = seldom a problem, 2 = sometimes a problem, 3 = often a problem, and 4 = always a problem). Items are linearly transformed to a 0–100 scale (0 = 0, 1 = 25, 2 = 50, 3 = 75, 4 = 100). Scale scores and total scores are computed as the sum of the items divided by the number of items answered that accounts for missing data. A higher score indicates better HRQoL.

2.5. Statistical analysis

Firstly, we applied descriptive analysis to calculate means and standard deviations (SDs) for continuous variables and numbers and percentages for categorical variables. Secondly, we compared the averages of the total score and the scale scores of both PedsQL 4.0 GCM and PedsQL 3.0 FIM with those norm average scores of the healthy controls of children that were reported in the published literature (14, 16). Thirdly, we used the two-independent-samples t-tests and one-way ANOVA to assess the differences in mean scores of PedsQL 4.0 GCM and PedsQL 3.0 FIM across groups regarding social demographic and clinical characteristics. In addition, regarding variables with three or more categories, the *post hoc* test was adopted to evaluate the pairwise differences in the averages of HRQoL total and scale scores between two specific subgroups. Tukey HSD correction was used for multi-group differences comparison. Fourth, we selected variables with statistical significance in the last step and included these variables in the multivariate linear regression models; the outcomes were the total score of PedsQL 4.0 GCM and 3.0 FIM. A $p < 0.05$ was used to indicate the statistical significance of all two-sided tests. Despite of statistical significance, we also evaluated the clinical relevance using Cohen's effect size (Cohen's *d*), that was calculated as the absolute value of the difference in average scores divided by the largest SD and was interpreted as follows: $0.2 \leq d < 0.5$, small

difference; $0.5 \leq d < 0.8$, moderate difference; and $d \geq 0.8$, large difference (17). All statistical analyses were conducted using the R software program, version 4.1.

3. Results

3.1. General characteristics of the study population

Table 1 presents the sociodemographic characteristics of the study population. The average age of children was 6.35 years, and 65.3% were boys. 47.5% of children had siblings. Mothers filled out 90.1% of the questionnaire. 73.3% of mothers and 70.3% of fathers had a high educational level, i.e., a bachelor's degree or above. 95% of parents were married, and. 86.1% lived together. 18.8% of households had three or fewer family members living together, 34.7% with four, 24.8% with five, and 21.8% with six or more members. 65.3% of families lived in urban areas. Regarding the annual household income, 31.7% of families earned less than 14,500 US dollars per year, 36.6% earned 14,500 to 36,300 US dollars, 20.8% earned 36,300 to 72,700 US dollars, and 10.9% earned more than 72,700 US dollars. 90.1% of families reported spending less than 14,500 US dollars per year on WS-related treatment, and the cost of 9.9% was between 14,500 to 36,300 US dollars. 79.2% of children with WS had health insurance, and 25.7% received social benefits/bonuses/subsidies. Regarding the perceived financial burden due to taking care of a child with WS, 33.7% considered it a heavy burden they could hardly afford, while 7.9% thought it was easily affordable. 23.8% of children went to a special school or rehabilitation school.

3.2. Clinical characteristics of children with WS

Clinical characteristics of children with Williams syndrome are summarized in Table 2. The average age of disease onset was 0.82 years, and the average age of diagnosis was 1.69 years. 90.1% of children got a genetic test. More specifically, 54.4% had Chromosomal Microarray Analysis (CMA); 15.8% had Whole Exome Sequencing (WES); 5.9% had multiplex ligation-dependent probe amplification (MLPA), 7.9% fluorescence *in situ* hybridization (FISH), and 14.9% whole genome sequencing (WGS). 77.2% of the children had diseases or health problems at birth, 62.4% had congenital heart disease (CHD) with no surgery, and 19.8% had CHD and received surgery. 41.6% presented with symptoms of attention deficit and hyperactivity; 28.7% had health conditions of the visual system; 56.4% had endocrine system diseases; 27.7% had inguinal hernia; 23.8% got repeated respiratory tract infection; 24.8% had digestive system diseases, and 7.9% had urinary system diseases.37.6% of the children had sleeping problems.

3.3. Comparing the mean scores of PedsQL FIM and PedsQL GCM with the norm data

Table 3 presents the mean values and standard deviations of scale scores and total scores of PedsQL 3.0 FIM and PedsQL 4.0 GCM

TABLE 1 General characteristics of children with Williams syndrome (*N*=101).

Variables	Values*
Age [mean (SD)]	6.35 (3.11)
Age group	
2–4 years	39 (38.6)
5–7 years	33 (32.7)
8–12 years	24 (23.8)
13–18 years	5 (5.0)
Gender	
Boy	66 (65.3)
Girl	35 (34.7)
The caregiver who filled in the questionnaire	
Mother	91 (90.1)
Father and others	10 (9.9)
Paternal educational level	
Low	14 (13.9)
Middle	16 (15.8)
High	71 (70.3)
Maternal educational level	
Low	13 (12.9)
Middle	14 (13.9)
High	74 (73.3)
Parental marital status	
Married	96 (95.0)
Divorced or widowed	5 (5.0)
Living together	
Always	87 (86.1)
Sometimes	10 (9.9)
Never	4 (4.0)
Having siblings, Yes	48(47.5)
Number of family members living in the same household	
Three or less	19 (18.8)
Four	35 (34.7)
Five	25 (24.8)
Six and more	22 (21.8)
Annual household income (US dollars)	
Less than 14,500	32 (31.7)
14,500 to 36,300	37 (36.6)
36,300 to 72,700	21 (20.8)
More than 72,700	11 (10.9)
Residence	
Urban	66 (65.3)
Town or rural	35 (34.7)
Health insurance	
No	21 (20.8)
Yes	80 (79.2)

(Continued)

TABLE 1 (Continued)

Variables	Values*
Receiving social benefits/bonuses/subsidies	
No	75 (74.3)
Yes	26 (25.7)
Cost related to WS treatment per year (US dollars)	
Less than 14,500	91 (90.1)
14,500 to 36,300	10 (9.9)
Perceived financial burden	
Easily Affordable	8 (7.9)
Affordable	59 (58.4)
Hard to afford	34 (33.7)
School type	
Normal school	51 (50.5)
Special school/Rehabilitation school	24 (23.8)
Not go to school yet	13 (12.9)
Others	13 (12.9)

*This table presents means, standard deviations, numbers, and percentages. Regarding the educational level, low means no education, primary school or middle school; middle level means high school, technical school, vocational secondary school, and vocational high school; high level means bachelor's degree or above according to the Chinese Standard Classification of Education.

measured in patients with Williams syndrome in our study and those of children as healthy controls extracted from the existing published literature (16, 17). Regarding PedsQL3.0 FIM, the averages of scale scores and the total score were significantly lower than the averages of the healthy controls (all p values <0.001). The range of effect size was from 0.62 to 1.54. Regarding PedsQL 4.0 GCM scale scores and total scores, the means were also significantly lower than those of the healthy controls (all p-values < 0.001). The range of effect size was from 0.88 to 2.03.

3.4. Differences in the mean scores of PedsQL FIM domains and the mean of total score across certain group

Table 4 presents differences in the mean values of PedsQL 3.0 FIM scales and total scores across certain groups. We showed the variables with at least one statistically significant difference (p-value < 0.05) in terms of mean values of PedsQL 3.0 FIM eight scales and total scores. Specifically, boys had lower mean scores in the scale of physical functioning ($p = 0.007$), emotional functioning ($p = 0.041$), social functioning ($p = 0.029$), communication ($p = 0.013$) and worry ($p = 0.021$), as well as the total score ($p = 0.012$). Regarding paternal high educational level, the mean scores were relatively higher in physical functioning ($p = 0.007$), social functioning ($p = 0.034$), communication ($p = 0.005$), family relationships ($p = 0.010$) scales, as well as total scores ($p = 0.005$). The mean scores for maternal high educational level were only higher on the cognitive function scale ($p = 0.023$). The mean scores on physical and cognitive functioning scales were relatively low when parents sometimes lived together ($p = 0.028$ and $p = 0.022$). Caregivers whose residences were in urban areas reported higher mean scores in physical functioning ($p = 0.040$), social functioning ($p = 0.027$), and cognitive functioning ($p = 0.030$)

TABLE 2 Clinical characteristics of children with Williams syndrome (N=101).

Clinical characteristic	Values*
Age of onset, mean (SD), years	0.82 (1.15)
Age of diagnosis, mean (SD), years	1.69 (2.26)
Having a genetic test, yes	91 (90.1)
Chromosomal Microarray Analysis (CMA), Yes	55 (54.5)
Whole Exome Sequencing (WES)	16 (15.8)
Multiplex ligation-dependent probe amplification (MLPA)	6 (5.9)
Fluorescence *in situ* hybridization (FISH)	8 (7.9)
Whole genome sequencing (WGS)	15 (14.9)
Getting rehabilitation, yes	60 (59.4)
Diseases at birth, yes	78 (77.2)
Congenital heart disease and surgery	
No	18 (17.8)
Having CHD but no surgery	63 (62.4)
Having CHD and surgery, receiving	20 (19.8)
Presence of attention deficit and hyperactivity, yes	42 (41.6)
Visual system conditions (%)	
Normal	17 (16.8)
Abnormal	29 (28.7)
Not tested	55 (54.5)
Endocrinologic system conditions, yes	57 (56.4)
Inguinal hernia	
No	51 (50.5)
Yes	28 (27.7)
Not tested	22 (21.8)
Repeated respiratory tract infection, yes	24 (23.8)
Digestive system conditions, yes	25 (24.8)
Urinary system conditions	
No	55 (54.5)
Yes	8 (7.9)
Not tested	38 (37.6)
Sleeping problems, yes	38 (37.6)

*This table presents means, standard deviations, numbers, and percentages.

scales compared with children living in rural areas. Relatively higher mean scores of all scales and the total score were observed across high to low annual household income subgroups (all p-values < 0.001). Caregivers whose children had diseases at birth had higher mean scores in physical functioning ($p = 0.014$), emotional functioning ($p = 0.005$), and communication ($p = 0.014$) scales, as well as for the total score ($p = 0.037$). The mean scores were lower in the family relationships scale of PedsQL FIM ($p = 0.010$) for families who reported their children with sleeping problems. Regarding the variables with three or more categories, we further presented the results of pairwise differences by using the *post hoc* test to show the exact distinction between two subgroups (see Supplementary Table S1). Effect sizes of the differences mentioned above presented in Table 4

TABLE 3 Comparing means of scale scores and total scores of PedsQL FIM and PedsQL GCM between patients with Williams syndrome and the healthy controls*.

	Williams syndrome	Healthy controls	*t* value	*p*-value	Effect size (*d*)
PedsQL FIM[a]					
Physical functioning	58.26 ± 24.52	81.79 ± 15.34	−8.910	<0.001	0.96
Emotional functioning	46.05 ± 25.23	79.98 ± 16.92	−11.969	<0.001	1.34
Social functioning	53.18 ± 24.86	83.48 ± 22.83	−8.717	<0.001	1.22
Cognitive functioning	61.84 ± 25.65	77.80 ± 18.73	−5.241	<0.001	0.62
Communication	50.58 ± 28.60	86.53 ± 16.01	−12.430	<0.001	1.26
Worry	38.95 ± 25.44	78.03 ± 19.52	−12.513	<0.001	1.54
Daily activities	51.32 ± 21.98	73.31 ± 20.54	−7.061	<0.001	1.00
Family relationships	64.91 ± 22.79	79.69 ± 18.23	−5.132	<0.001	0.65
Total score	53.14 ± 19.92	80.07 ± 13.98	−11.688	<0.001	1.35
PedsQL GCM[b]					
Physical functioning	59.56 ± 20.69	82.18 ± 17.6	−10.579	<0.001	1.09
Emotional functioning	57.77 ± 21.75	76.90 ± 16.6	−9.129	<0.001	0.88
Social functioning	39.90 ± 22.35	85.33 ± 15.4	−22.429	<0.001	2.03
School functioning	36.82 ± 26.48	78.03 ± 16.5	−18.146	<0.001	1.56
Total score	50.77 ± 16.59	80.74 ± 16.9	−15.380	<0.001	1.77

*Mean scores and standard deviations (SDs) of scale scores and total scores of PedsQL 3.0 FIM and PedsQL 4.0 GCM of the healthy controls are from the published literature (14, 16).
[a]The sample size of the healthy control group for norm values of PedsQL 3.0 FIM is 210, and we only compared the means and SDs of children aged 5–12 years in our study with the children at the same age in the literature (16).
[b]The sample size of the healthy control group for norm values of PedsQL 4.0 GCM is 284.

are shown in Supplementary Table S2. Additionally, we presented the variables without any statistically significant differences (p-value > 0.05) in terms of mean values of PedsQL 3.0 FIM eight scales and total scores in Supplementary Table S3.

3.5. Differences in mean scores of PedsQL GCM domains and the mean of total score across certain groups

Table 5 presents differences in the means of PedsQL 4.0 GCM total score and scale scores across certain groups. We presented the variables with at least one statistically significant difference (p-value < 0.05) regarding means of total score and scale scores. More specifically, the mean score of physical functioning was observed to be relatively low in children whose families can hardly afford the cost of treatment ($p = 0.001$) and children who had sleeping problems ($p = 0.013$). Regarding the average score of emotional functioning, we observed lower scores in children whose fathers had low education ($p = 0.012$), children whose families had low annual household income ($p = 0.049$) and could hardly afford the cost of treatment ($p = 0.022$), as well as those who had a sleeping problem ($p = 0.005$). The mean score of social functioning was relatively low in children who had symptoms of attention deficit and hyperactivity ($p = 0.032$). Regarding the average score of school functioning, we observed mean scores were significantly different across school types ($p < 0.001$) and was lower in children who had a sleeping problem ($p = 0.011$). The mean of the total score was relatively low in children whose fathers had low education ($p = 0.023$) and whose family could hardly afford the cost of treatment ($p = 0.002$), and children who had sleeping problem ($p = 0.004$); it also was significantly different across household income categories ($p = 0.044$). Regarding the variables with three or more categories, we further presented the results of pairwise differences by using the *post hoc* test to show the exact difference between two subgroups (see Supplementary Table S4). Effect sizes of the differences as mentioned above presented in Table 5 are shown in Supplementary Table S5. Additionally, we presented the variables without any statistically significant differences (p-value > 0.05) in terms of mean values of PedsQL 4.0 GCM scales and total scores in Supplementary Table S6.

Table 6 shows the associations of selected variables with the score of PedsQL GCM and FIM in the multivariate linear regression models. Having sleeping problems was statistically significantly associated with the total score of PedsQL 4.0 GCM (β:−8.43; 95% CI: −14.83, −2.03; $p = 0.010$). Children's gender, paternal education, number of family members in the same household, annual household income, and having diseases at birth were not statistically significantly associated with the total score of PedsQL 4.0 GCM (p-values > 0.05). Perceived financial burden was a statistically significant associated factor of the total score of PedsQL FIM, more specifically, comparing with those who could easily afford the disease-related cost, those who could hardly afford it had significantly lower score of FIM (β:−29.16; 95% CI: −43.79, −14.53; $p < 0.001$); and those who thought the cost affordable had relatively low score of FIM (β:−13.67; 95% CI: −27.25, −0.09; $p = 0.049$). Paternal education, annual household income and perceived financial burden were not statistically significantly associated with the total score of PedsQL GCM (p-values > 0.05).

4. Discussion

The present study assessed proxy-reported HRQoL of children with Williams syndrome and self-reported HRQoL of caregivers in a relatively large sample of Chinese pediatric patients. To our best

TABLE 4 Means and standard deviations of PedsQL 3.0 FIM total score and scale scores (n=101).

	Physical functioning	Emotional functioning	Social functioning	Cognitive functioning	Communication	Worry	Daily activities	Family relationships	Total score
Gender									
Boy (n=66)	53.60 (25.44)	44.32 (25.05)	51.61 (25.62)	58.03 (25.49)	46.59 (29.05)	33.26 (25.11)	50.38 (24.70)	60.91 (23.57)	49.84 (21.15)
Girl (n=35)	68.10 (25.06)	55.43 (26.88)	63.57 (26.14)	67.71 (24.68)	61.43 (26.44)	45.86 (26.75)	59.29 (25.23)	68.43 (28.12)	61.23 (21.54)
p-value	0.007	0.041	0.029	0.069	0.013	0.021	0.090	0.157	0.012
Paternal educational level									
Low (n=14)	39.58 (18.11)	38.93 (22.12)	39.73 (26.59)	38.21 (16.01)	32.14 (23.31)	26.07 (19.03)	40.48 (22.85)	45.00 (24.65)	37.52 (18.88)
Middle (n=16)	55.99 (24.48)	40.62 (24.07)	53.52 (24.26)	59.38 (21.36)	44.27 (29.14)	33.12 (24.89)	56.25 (25.73)	63.44 (19.21)	50.82 (19.28)
High (n=71)	62.97 (26.31)	51.69 (26.77)	59.42 (25.79)	66.41 (25.51)	57.28 (28.13)	40.92 (27.25)	55.40 (24.96)	67.18 (25.38)	57.66 (21.61)
p-value	0.007	0.111	0.034	<0.001	0.005	0.116	0.112	0.010	0.005
Maternal educational level									
Low (n=13)	49.68 (26.49)	41.92 (20.37)	49.04 (26.50)	48.85 (17.81)	42.95 (28.84)	36.92 (27.95)	46.15 (25.60)	51.92 (26.50)	45.93 (21.72)
Middle (n=14)	51.79 (30.17)	37.50 (25.85)	45.98 (18.28)	51.07 (23.87)	45.24 (25.88)	32.50 (22.51)	53.57 (20.34)	68.93 (20.86)	48.32 (18.09)
High (n=74)	61.49 (25.02)	51.28 (26.61)	58.78 (27.15)	65.54 (25.93)	54.50 (29.33)	38.72 (26.82)	54.73 (25.92)	64.53 (25.63)	56.20 (22.29)
p-value	0.186	0.126	0.153	0.023	0.277	0.719	0.529	0.177	0.179
Living together									
Always (n=87)	60.20 (25.86)	49.89 (26.80)	55.96 (26.27)	63.85 (25.09)	53.35 (29.54)	39.60 (26.91)	54.50 (25.25)	65.63 (24.13)	55.37 (21.90)
Sometimes (n=10)	39.17 (17.15)	35.50 (18.33)	48.75 (22.20)	40.50 (18.63)	35.83 (17.59)	24.00 (16.96)	45.00 (14.80)	53.00 (29.08)	40.22 (16.60)
Never (n=4)	72.92 (33.07)	42.50 (21.79)	68.75 (37.50)	60.00 (31.36)	56.25 (31.46)	28.75 (22.87)	52.08 (42.70)	43.75 (34.25)	53.12 (25.94)
p-value	0.028	0.234	0.434	0.022	0.184	0.162	0.528	0.092	0.115
Numbers of family members living in the same household									
Three or less (n=19)	52.85 (25.91)	33.42 (23.46)	50.33 (24.25)	49.47 (23.09)	42.11 (26.86)	29.74 (25.03)	41.23 (22.99)	51.58 (24.84)	43.84 (19.83)
Four (n=35)	62.14 (23.79)	47.71 (21.47)	57.14 (22.80)	69.43 (23.10)	52.62 (27.92)	37.00 (26.60)	55.71 (23.81)	69.86 (23.99)	56.45 (18.83)
Five (n=25)	50.50 (28.52)	45.60 (31.37)	52.50 (31.41)	53.00 (26.85)	49.33 (29.06)	35.80 (25.36)	52.00 (28.59)	60.20 (29.28)	49.87 (25.31)
Six and more (n=22)	67.23 (24.91)	64.55 (20.47)	61.93 (27.20)	68.41 (24.17)	61.36 (30.81)	47.50 (26.49)	62.12 (21.78)	67.50 (19.87)	62.58 (20.82)
p-value	0.093	0.001	0.481	0.006	0.192	0.174	0.055	0.058	0.029
Annual household income (US dollars)									
Less than 14,500 (n=32)	41.41 (22.92)	28.75 (19.01)	38.09 (23.57)	43.12 (22.10)	30.99 (21.09)	23.12 (17.72)	41.15 (23.94)	48.91 (25.93)	36.94 (15.61)

(Continued)

10.3389/fpubh.2023.1177317

TABLE 4 (Continued)

	Physical functioning	Emotional functioning	Social functioning	Cognitive functioning	Communication	Worry	Daily activities	Family relationships	Total score
14,500 to 36,300 ($n = 37$)	61.04 (23.53)	56.76 (24.78)	57.94 (21.58)	67.03 (21.42)	56.98 (28.70)	39.86 (26.91)	52.25 (21.49)	66.08 (22.08)	57.24 (19.38)
36,300 to 72,700 ($n = 21$)	71.23 (23.01)	58.57 (22.92)	73.51 (22.44)	68.57 (24.91)	62.30 (24.95)	49.05 (25.48)	67.86 (24.62)	71.67 (21.76)	65.34 (20.60)
More than 72,700 ($n = 11$)	76.52 (22.07)	55.91 (25.77)	65.91 (26.86)	81.82 (18.74)	74.24 (22.50)	50.45 (29.45)	65.91 (23.70)	81.82 (20.65)	69.07 (18.33)
p-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Residence									
Urban ($n = 66$)	62.50 (26.90)	48.03 (26.77)	59.94 (25.73)	65.38 (26.29)	54.92 (28.98)	40.76 (27.02)	55.30 (26.34)	65.08 (26.28)	56.49 (22.27)
Town or rural ($n = 35$)	51.31 (23.22)	48.43 (25.20)	47.86 (25.90)	53.86 (22.46)	45.71 (28.25)	31.71 (24.01)	50.00 (22.60)	60.57 (23.60)	48.68 (20.43)
p-value	0.040	0.942	0.027	0.030	0.128	0.100	0.315	0.398	0.088
Perceived financial burden									
Easily Affordable ($n = 8$)	86.46 (17.64)	68.12 (16.02)	78.91 (20.85)	90.00 (14.39)	77.08 (23.04)	56.25 (24.89)	78.12 (21.79)	79.38 (16.78)	76.79 (15.72)
Affordable ($n = 59$)	66.24 (21.92)	55.85 (23.75)	63.24 (22.30)	67.29 (21.88)	60.31 (26.05)	44.32 (25.74)	59.60 (22.15)	69.75 (21.42)	60.83 (18.17)
Hard to afford ($n = 34$)	38.85 (21.51)	30.15 (22.00)	37.32 (23.46)	44.41 (23.02)	30.88 (22.24)	21.62 (19.10)	37.01 (21.34)	48.97 (27.07)	36.15 (16.70)
p-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Having diseases at birth									
No ($n = 23$)	46.92 (25.91)	35.00 (22.21)	48.64 (28.39)	54.13 (24.06)	38.77 (24.70)	30.00 (20.50)	46.38 (21.88)	63.91 (24.82)	45.47 (19.92)
Yes ($n = 78$)	62.07 (25.33)	52.05 (26.03)	57.85 (25.47)	63.53 (25.69)	55.56 (29.11)	39.87 (27.43)	55.56 (25.75)	63.40 (25.67)	56.24 (21.93)
p-value	0.014	0.005	0.141	0.121	0.014	0.113	0.124	0.932	0.037
Having sleeping problems									
No ($n = 63$)	61.97 (25.97)	50.24 (25.71)	57.54 (25.60)	62.54 (25.87)	55.42 (28.76)	40.40 (26.67)	57.01 (24.47)	68.49 (22.98)	56.70 (20.99)
Yes ($n = 38$)	53.07 (25.75)	44.74 (26.76)	52.80 (27.52)	59.47 (25.14)	45.61 (28.52)	33.03 (25.22)	47.59 (25.40)	55.26 (27.21)	48.95 (22.71)
p-value	0.097	0.307	0.383	0.561	0.099	0.173	0.068	0.010	0.084

*Values presented in this table are means and standard deviations. Regarding the educational level, low means no education, primary school, or middle school; middle level means high school, technical school, vocational secondary school and vocational high school; high level means bachelor's degree or above according to the Chinese Standard Classification of Education.

10.3389/fpubh.2023.1177317

TABLE 5 Means and standard deviations of PedsQL GCM total score and scale scores (n=101).

	Physical functioning	Emotional functioning	Social functioning	School functioning	Total score
Paternal educational level					
Low (N = 14)	47.99 (22.25)	43.21 (17.61)	32.14 (20.35)	26.19 (25.36)	39.74 (16.13)
Middle (N = 16)	62.11 (21.95)	54.38 (24.96)	40.31 (22.47)	34.48 (26.43)	50.90 (17.96)
High (N = 71)	61.27 (19.63)	61.41 (20.65)	41.34 (22.68)	39.44 (26.48)	52.92 (15.71)
p-value	0.077	0.012	0.374	0.216	0.023
Maternal educational level					
Low (N = 13)	45.43 (17.80)	47.31 (15.49)	43.85 (21.62)	30.00 (25.88)	43.02 (14.61)
Middle (N = 14)	68.53 (19.63)	64.29 (25.10)	41.07 (22.72)	38.69 (26.11)	55.77 (17.79)
High (N = 74)	60.35 (20.35)	58.38 (21.65)	38.99 (22.62)	37.66 (26.81)	51.18 (16.43)
p-value	0.011	0.115	0.757	0.609	0.125
Having siblings					
No (N = 53)	59.91 (16.48)	53.68 (19.98)	37.92 (23.52)	35.72 (27.01)	49.45 (14.29)
Yes (N = 48)	59.18 (24.70)	62.29 (22.92)	42.08 (21.01)	38.02 (26.10)	52.23 (18.86)
p-value	0.861	0.046	0.353	0.665	0.402
Annual household income (US dollars)					
Less than 14,500 (n = 32)	53.03 (21.89)	51.09 (22.49)	37.50 (23.31)	32.92 (26.49)	45.69 (16.10)
14,500 to 36,300 (n = 37)	61.23 (16.38)	56.62 (20.58)	36.76 (21.06)	34.32 (25.25)	49.86 (15.07)
36,300 to 72,700 (n = 21)	60.12 (23.65)	64.29 (20.20)	46.67 (23.47)	44.29 (29.71)	55.43 (19.77)
More than 72,700 (n = 11)	71.88 (19.91)	68.64 (21.34)	44.55 (20.91)	42.27 (23.49)	59.72 (11.48)
p-value	0.059	0.049	0.324	0.372	0.044
Perceived financial burden					
Easily affordable (N = 8)	76.56 (17.76)	68.75 (18.47)	47.50 (18.13)	47.29 (25.38)	62.99 (9.34)
Affordable (N = 59)	62.34 (19.69)	60.76 (20.02)	42.97 (21.56)	36.69 (27.41)	53.20 (16.37)
Hard to afford (N = 34)	50.74 (19.61)	50.00 (23.45)	32.79 (23.33)	34.56 (25.19)	43.69 (15.70)
p-value	0.001	0.022	0.063	0.477	0.002
School type					
Normal school (N = 51)	61.76 (18.08)	60.29 (20.33)	40.78 (22.81)	47.16 (20.55)	53.72 (15.57)
Special school/rehabilitation school (N = 24)	60.81 (23.35)	56.88 (22.98)	33.96 (24.58)	30.97 (26.61)	48.73 (18.76)
Not going to school yet (N = 13)	58.17 (19.20)	60.77 (21.10)	45.77 (14.12)	19.23 (26.22)	50.27 (13.45)
Others (N = 13)	50.00 (25.83)	46.54 (24.27)	41.54 (22.86)	24.62 (31.90)	43.47 (18.03)
p-value	0.324	0.219	0.438	<0.001	0.213
Receiving social benefit/bonuses/subsidies					
No (N = 75)	59.96 (20.59)	58.67 (20.47)	44.20 (20.65)	37.44 (27.43)	52.27 (16.27)
Yes (N = 26)	58.41 (21.34)	55.19 (25.36)	27.50 (22.81)	35.00 (23.92)	46.46 (17.07)
p-value	0.745	0.486	0.001*	0.687	0.125
Having symptoms of attention deficit and/or hyperactivity					
No (N = 59)	59.32 (19.36)	57.97 (22.67)	43.90 (21.01)	34.15 (28.16)	51.26 (16.77)
Yes (N = 42)	59.90 (22.66)	57.50 (20.67)	34.29 (23.21)	40.56 (23.74)	50.09 (16.51)
p-value	0.892	0.916	0.032	0.233	0.729
Sleeping problems					
No (N = 63)	63.49 (18.38)	62.46 (19.79)	41.59 (22.18)	41.96 (23.44)	54.41 (14.38)
Yes (N = 38)	53.04 (22.82)	50.00 (22.87)	37.11 (22.65)	28.29 (29.23)	44.73 (18.38)
p-value	0.013	0.005	0.331	0.011	0.004

*Values presented in this table are means and standard deviations. Regarding the educational level, low means no education, primary school or middle school; middle level means high school, technical school, vocational secondary school and vocational high school; high level means bachelor degree or above according to the Chinese Standard Classification of Education.

TABLE 6 Associations of selected variables with the total score of PedsQL GCM and FIM by multivariate linear regression analyses.

	Total score of PedsQL FIM		Total score of PedsQL GCM	
	β (95% CI)	*p*-value	β (95% CI)	*p*-value
Gender				
Boy ($n = 66$)	Reference		\	\
Girl ($n = 35$)	6.19 (−1.04, 13.42)	0.092	\	\
Paternal educational level				
High level ($n = 71$)	Reference		Reference	
Mid level ($n = 16$)	−3.22 (−12.73, 6.28)	0.502	−1.21 (−9.95, 7.53)	0.784
Low level ($n = 14$)	−9.04 (−20.04, 1.95)	0.106	−9.45 (−19.33, 0.44)	0.061
Numbers of family members living in the same household				
Three or less ($n = 19$)	Reference		\	\
Four ($n = 35$)	5.77 (−3.89, 15.42)	0.239	\	\
Five ($n = 25$)	4.31 (−6.06, 14.69)	0.411	\	\
Six and more ($n = 22$)	9.30 (−1.64, 20.23)	0.095	\	\
Annual household income (US dollars)				
More than 72,700 ($n = 11$)	Reference		Reference	
36,300 to 72,700 ($n = 21$)	−0.28 (−12.84, 12.28)	0.965	−0.11 (−11.67, 11.46)	0.985
14,500 to 36,300 ($n = 37$)	−2.23 (−14.21, 9.74)	0.712	−4.19 (−15.09, 6.70)	0.447
Less than 14,500 ($n = 32$)	−11.14 (−24.94, 2.66)	0.112	−3.20 (−15.76, 9.36)	0.614
Perceived financial burden				
Easily affordable ($n = 8$)	Reference		Reference	
Affordable ($n = 59$)	−13.67 (−27.25, −0.09)	0.049	−5.77 (−17.53, 6.00)	0.333
Hard to afford ($n = 34$)	−29.16 (−43.79, −14.53)	<0.001	−12.71 (−25.97, 0.55)	0.06
Having diseases at birth				
No ($n = 23$)	Reference		\	\
Yes ($n = 78$)	4.71 (−3.52, 12.94)	0.259	\	\
Sleeping problems				
No ($N = 63$)	\	\	Reference	
Yes ($N = 38$)	\	\	−8.43 (−14.83, −2.03)	0.010
Adjusted R^2	0.436		0.161	

CI, confident interval.

knowledge, it was the first study to assess HRQoL in children with WS. In addition, this study demonstrated the associated factors of HRQoL of children with WS and their caregivers. To be noted, having sleeping problems was significantly associated with poor HRQoL of children, and when caregivers perceived the medical cost as hardly unaffordable, their self-reported HRQoL was remarkably worse than those who could easily afford for the cost.

We found that the HRQoL of children with WS and their caregivers were significantly worse than the healthy population as a reference regarding each domain and the overall HRQoL. As indicated by effect sizes, differences in caregiver's HRQoL can be interpreted as large differences in terms of its clinical relevance in all scales and total scores of PedsQL 3.0 FIM except for the scales of cognitive functioning and family relationship that can be interpreted as moderate differences. Differences in proxy-reported HRQoL of children with WS can be interpreted as large differences in our study. Notably, WS can profoundly impact the physical, psychological, social, and role functioning of both children and their caregivers. Due to the nature of this disease, WS can cause impairments in the biological structure and functioning of multiple systems, such as cardiovascular and endocrine systems, which may lead to a poor prognosis. If there were severe damages to organs, in particular, cardiovascular damages, children may have high morbidity and mortality and often die of heart failure. Thus, raising and caring for a child with WS imposes heavy distress and financial burden, and required more involvement of parents' personal time which together may impair the family functioning.

4.1. Health-related quality of life of caregivers of children with WS

Our study showed that the HRQoL of caregivers of male patients was worse than that of female patients regarding physical, emotional, and social functioning, worry domains, and overall HRQoL. Although WS occurs equally in both genders, raising a boy with WS seems more challenging than raising a girl. This finding might be explained by the

gender differences in children's growth and behaviors as well as the role expectations by society. Boys are generally naughtier and more active, and in Chinese social norms, girls are encouraged to be sensible and disciplined. Therefore, raising a boy with WS may cause more distress to caregivers' physical, emotional, and social health and overall HRQoL than raising a girl.

We found the average scores of physical, social, cognitive functioning, communication, and family relationship scales and total score were significantly lower (p-values < 0.05) when the father had a low educational level (i.e., no primary education school and middle school) compared with fathers with high education (i.e., bachelor degree or above); the effect sizes range from 0.48 to 1.10 indicating a moderate to large differences. While regarding maternal educational levels, the statistically significant difference was only observed in the score of the cognitive functioning scale ($p < 0.05$). Our finding was consistent with the Brazilian study showing paternal education positively related to family quality of life (10).

We did not find a significant difference in the caregiver's HRQoL across marital status (i.e., being married vs. being divorced or widowed). However, results in our study showed that the scores of physical and cognitive functioning scales were higher in parents always living together than those only sometimes living together (p-values < 0.05), and effect sizes were 0.81 and 0.93 (Supplementary Table S2), which indicated a large difference. This finding was interesting and might be related to Chinese culture. There have been some couples who got divorced but still lived together for the child's good. Another explanation was that the small sample size of the subgroup of parents who were divorced or widowed might lead to a chance finding.

Our study highlighted significant disparities in caregivers' HRQoL across families with different levels of financial burden as indicated by the annual household income, perceived financial burden due to the disease, and the residence (urban or rural) areas. Especially the effect sizes ranged from 1.12 to 2.43 (see Supplementary Table S2), which indicated large differences when comparing caregivers' HRQoL domain and total scores between the families that can easily afford and those hardly afford. Large differences were observed in all domain scores of caregivers' HRQoL and the overall score between the highest and the lowest subgroup of household income (i.e., more than 72,700 vs. less than 14,500 US dollars per year), with the effect sizes ranging from 1.03 to 2.02. The above finding was consistent with the Brazilian study reporting family income as a determinant of the family's quality of life with WS children (10). They also found that supplemental health insurance and social benefits were associated with better family quality of life, which was not confirmed by our study. It may be explained that China and Brazil have different health insurance systems and policies regarding rare diseases.

Regarding the disease-related characteristics, our study showed that having diseases at birth was related to poorer physical functioning, emotional functioning, and communication within the family, and these differences were interpreted as moderate, with effect sizes ranging from 0.58 to 0.66. Diseases at birth included preterm birth, small for gestational age (SGA), neonatal jaundice, asphyxia, hypothyroidism, congenital heart disease, and indirect inguinal hernia. As mentioned above, children with one or more disorders may have a higher demand for medical care and family care, which may burden the caregiver's physical and emotional health and influence communication among the family members. Children with sleep problems were related to a worse family relationship than those without such an issue; the difference was small according to the effect size ($d = 0.49$). According to a systematic review by Coles et al., poor sleeping quality of children was negatively associated with the quality of family relationships (i.e., between parent and child and within the couple) (18).

Results from the multivariate linear regression analysis showed that the perceived financial burden due to the treatment of WS was significantly independently associated with self-reported HRQoL of caregivers after adjusting other selected variables. This indicated the remarkable influence of perceived financial burden on caregivers' and family's quality of life regardless of the other objective indicators of socioeconomic status (e.g., education and household income) and disease-related characteristics. By now, there was no such financial support feasible in China. Regarding other rare diseases or certain disorders, for instance, autism, the Federation of Disabled Persons and foundations provide certain amount of subsidies for families with sick children to support for medical care and rehabilitation. Therefore, social support, in particular, social benefits, subsidies and health insurances that may cover some costs of WS treatment and care, is warranted.

4.2. Health-related quality of life of children with WS

To the best knowledge, the present study was the first to measure the HRQoL of children with WS. It provided comprehensive data on multiple domains (i.e., physical, emotional, social, and school functioning) of children's HRQoL. Socio-economic status as indicated by education and income was an important factor. Our study showed that the emotional functioning and overall HRQoL of children were poorer in children whose fathers had low educational level than those whose father had high education, and the differences were large (p-values < 0.05; effect sizes: 0.88 and 0.82, see Supplementary Tables S4, S5). The subjective indicator (i.e., perceived financial burden) seemed more sensitive than the objective measurement (i.e., household income) in distinguishing the subgroup differences in children's HRQoL. More specifically, the *post hoc* test did not find significant differences in PedsQL 4.0 GCM scale and total scores between every two subgroups regarding the annual household income (p-values > 0.05, data not shown). We found that children's physical functioning and overall HRQoL were significantly higher in the family who easily afforded the cost of taking care of WS children than those who hardly afforded (p-value < 0.001, effect size = 1.32; p-value < 0.05; effect size = 1.23; see Supplementary Tables S4, S5). Our results suggested that the perception of financial burden related to caring WS children may be a more sensitive indicator of children's HRQoL than the objective measurement such as household income. Having attention deficit and/or hyperactivity was related to the impaired social functioning of children, but the effect size indicated small difference. Sleeping problems may decrease children's overall HRQoL, specifically physical, emotional and school functioning, while the differences were small or moderate. To be noted, the presence of sleeping problems was the

only statistically significant factor in the multivariate regression model, which indicated the very important impact of sleeping on WS children's HRQoL. This finding also inspired us to develop targeted intervention to improve quality of life of these children by improving the sleeping quality.

4.3. Strengths and limitations

Our study has several strengths. The present study was one of the few to assess the HRQoL of children with WS and the family impact by the disease using reliable and valid QoL measurements (i.e., PedsQL 3.0 FIM and PedsQL 4.0 GCM). The sample size was relatively large, with 101 child-caregiver dyads, which, to our knowledge, was the largest sample regarding HRQoL of children with WS. Our study provided affluent data of domain scores and total scores of HRQoL of both caregivers and children, and we compared our data with healthy control at the same age. In addition, a comprehensive set of variables regarding social demographic and clinical characteristics were included in the analyses, which was a preliminary attempt to identify the associated factors with HRQoL. *Post hoc* analyses were adopted to identify the difference between the exact two groups, and effect size (Cohen's *d*) was used to evaluate the clinical relevance.

Several limitations should be taken into account. First, most caregivers participating in this study had a high educational level (i.e., bachelor's degree or higher), which may not represent the population of children with WS and their families. This may be explained by the fact that caregivers with relatively high education may have more resources and higher health literacy and are more likely to seek medical care. Selection bias may occur, so the interpretation of our results should be made with caution. Thirdly, causation cannot be concluded due to the cross-sectional study design.

To summarize, our study has comprehensively assessed the HRQoL of Chinese children with WS and of their caregivers as well as the potential determinants of HRQoL. We highlighted that the HRQoL of children with WS and their caregivers was dramatically worse than the healthy references. Paternal educational level, household income, and the perceived financial burden significantly influence the HRQoL of both children and families. Therefore, we call for attention from policymakers and other stakeholders on the health status and well-being of children with WS and their families. Supports are needed to relieve psychosocial distress and financial burden.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of the Children's Hospital of Zhejiang University School of Medicine (2019-IRB-122). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

WC, CZ, and GB were involved in the study conceptualization and study design. WC, CJ, FL, and JS collected data. XH, ZL, and YD carried out statistical analyses. WC, LS, and GB wrote the first draft of the article. WC, LS, XH, TP, XJ, LH, CZ, and GB provided critical revisions of the article for important intellectual content. All authors contributed to the article and approved the submitted version.

Funding

GB received the starting grant for research from the Children's Hospital, Zhejiang University School of Medicine, National Children's Regional Medical Center, and National Clinical Research Center for Child Health. The funders had no role in the study design; data collection, analysis, or interpretation; in the writing of the report; or in the decision to submit the article for publication.

Acknowledgments

We gratefully thank all the patients with Williams syndrome and their caregivers who would like to participate in the present study and provided valuable data on health-related quality of life of children and caregivers. We also thank all the clinicians, nurses, researchers, and volunteers who contributed to the establishment and maintenance of the cohort of Williams syndrome in China where our study was embedded. In addition, we thank all members of BAI Lab and all members of Pediatric Evidence-based Medical and Clinical Laboratory that is led by GB, for the interesting discussions and significant contributions to this 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/fpubh.2023.1177317/full#supplementary-material

References

1. Strømme P, Bjømstad PG, Ramstad K. Prevalence estimation of Williams syndrome. *J Child Neurol*. (2002) 17:269–71. doi: 10.1177/088307380201700406

2. Yau EKC, Lo IFM, Lam STS. Williams-Beuren syndrome in the Hong Kong Chinese population: retrospective study. *Hong Kong Med J*. (2004) 10:22–7.

3. Hillier LW, Fulton RS, Fulton LA, Graves TA, Pepin KH, Wagner-McPherson C, et al. The DNA sequence of human chromosome 7. *Nature*. (2003) 424:157–64. doi: 10.1038/nature01782

4. Morris CA, Lenhoff HM, Wang PP. *Williams-Beuren syndrome: Research, evaluation, and treatment*. United Kingdom: Johns Hopkins University Press (2006).

5. Mervis CB, John AE. Cognitive and behavioral characteristics of children with Williams syndrome: implications for intervention approaches. *Am J Med Genet*. (2010) 154C:229–48. doi: 10.1002/ajmg.c.30263

6. Fayers PM, Machin D. *Quality of life: The assessment, analysis and reporting of patient-reported outcomes. 3rd* ed. Chichester, West Sussex, UK; Hoboken, NJ: Wiley-Blackwell (2016).

7. The WHOQOL Group. The World Health Organization quality of life assessment (WHOQOL): Position paper from the World Health Organization. *Soc Sci Med*. (1995) 41:1403–9. doi: 10.1016/0277-9536(95)00112-k

8. Ahmed S, Andrich D. *Dictionary of quality of life and health outcomes measurement*. United States: International Society for Quality of Life Research (ISOQOL) (2015).

9. Moraleda Sepúlveda E, López Resa P. Evaluating quality of life in families with Williams syndrome patients. *Health Qual Life Outcomes*. (2021) 19:121. doi: 10.1186/s12955-021-01704-0

10. Pereira RCM, Apis A, dos Santos TR, de Avó LR, Da S, Pilotto RF, et al. Quality of life of Brazilian families who have children with Williams syndrome. *J Intellect Disabil*. (2021):174462952210795. doi: 10.1177/17446295221079583 [Epub ahead of print]

11. Varni JW, Limbers CA, Burwinkle TM. Parent proxy-report of their children's health-related quality of life: an analysis of 13,878 parents' reliability and validity across age subgroups using the PedsQL™ 4.0 generic Core scales. *Health Qual Life Outcomes*. (2007) 5:2. doi: 10.1186/1477-7525-5-2

12. Varni JW, Seid M, Kurtin PS. PedsQL™ 4.0: reliability and validity of the pediatric quality of life inventory™ version 4.0 generic Core scales in healthy and patient populations. *Med Care*. (2001) 39:800–12. doi: 10.1097/00005650-200108000-00006

13. Chen R, Hao Y, Feng L, Zhang Y, Huang Z. The Chinese version of the pediatric quality of life inventory™ (PedsQL™) family impact module: cross-cultural adaptation and psychometric evaluation. *Health Qual Life Outcomes*. (2011) 9:16. doi: 10.1186/1477-7525-9-16

14. Ji Y, Chen S, Li K, Xiao N, Yang X, Zheng S, et al. Measuring health-related quality of life in children with cancer living in mainland China: feasibility, reliability and validity of the Chinese mandarin version of PedsQL 4.0 generic Core scales and 3.0 Cancer module. *Health Qual Life Outcomes*. (2011) 9:103. doi: 10.1186/1477-7525-9-103

15. Varni JW, Seid M, Kurtin PS. PedsQL™ 4.0: reliability and validity of the pediatric quality of life inventory™ version 4.0 generic Core scales in healthy and patient populations. *Med Care*. (2001) 39:800–12. doi: 10.1097/00005650-200108000-00006

16. Yang H, Luo S, Liang X, Lin Q, Cheng T, Zeng L, et al. The association between family impact and health-related quality of life of children with idiopathic central precocious puberty in Chongqing, China. *Health Qual Life Outcomes*. (2021) 19:171. doi: 10.1186/s12955-021-01805-w

17. Cohen J. *Statistical power analysis for the behavioral sciences. 2nd* ed. New York: Routledge (1988).

18. Coles L, Thorpe K, Smith S, Hewitt B, Ruppanner L, Bayliss O, et al. Children's sleep and fathers' health and wellbeing: a systematic review. *Sleep Med Rev*. (2022) 61:101570. doi: 10.1016/j.smrv.2021.101570

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 15 June 2023
DOI 10.3389/fpubh.2023.1129073

OPEN ACCESS

EDITED BY
Jing Wu,
Tianjin University, China

REVIEWED BY
Arlette Setiawan,
Padjadjaran University, Indonesia
Preeti Manmohan Galagali,
Bengaluru Adolescent Care and Counselling Centre, India
Júlia Cassuriaga,
Federal University of Pelotas, Brazil

*CORRESPONDENCE
Ulrike Ravens-Sieberer
✉ ravens-sieberer@uke.de

RECEIVED 21 December 2022
ACCEPTED 16 May 2023
PUBLISHED 15 June 2023

CITATION
Ravens-Sieberer U, Devine J, Napp A-K, Kaman A, Saftig L, Gilbert M, Reiß F, Löffler C, Simon AM, Hurrelmann K, Walper S, Schlack R, Hölling H, Wieler LH and Erhart M (2023) Three years into the pandemic: results of the longitudinal German COPSY study on youth mental health and health-related quality of life. *Front. Public Health* 11:1129073. doi: 10.3389/fpubh.2023.1129073

Three years into the pandemic: results of the longitudinal German COPSY study on youth mental health and health-related quality of life

Ulrike Ravens-Sieberer[1]*, Janine Devine[1], Ann-Kathrin Napp[1], Anne Kaman[1], Lynn Saftig[1], Martha Gilbert[1], Franziska Reiß[1], Constanze Löffler[1], Anja Miriam Simon[2], Klaus Hurrelmann[3], Sabine Walper[4], Robert Schlack[5], Heike Hölling[5], Lothar Heinz Wieler[5,6] and Michael Erhart[1,7,8]

[1]Department of Child and Adolescent Psychiatry, Psychotherapy, and Psychosomatics, University Medical Center Hamburg-Eppendorf, Hamburg, Germany, [2]Infratest dimap, Berlin, Germany, [3]Hertie School, Berlin, Germany, [4]German Youth Institute, Munich, Germany, [5]Department for Epidemiology and Health Monitoring, Robert Koch Institute, Berlin, Germany, [6]Research Cluster Digital Global Public Health, Hasso-Plattner-Institute, University of Potsdam, Potsdam, Germany, [7]Department of Public Health, Alice Salomon University of Applied Sciences, Berlin, Germany, [8]Department of Psychology, Apollon University of Applied Sciences, Bremen, Germany

Purpose: For the past three years, the German longitudinal COPSY (*COVID-19 and PSYchological Health*) study has monitored changes in health-related quality of life (HRQoL) and the mental health of children and adolescents during the COVID-19 pandemic.

Methods: A nationwide, population-based survey was conducted in May–June 2020 (W1), December 2020–January 2021 (W2), September–October 2021 (W3), February 2022 (W4), and September–October 2022 (W5). In total, *n*=2,471 children and adolescents aged 7–17 years (*n*=1,673 aged 11–17 years with self-reports) were assessed using internationally established and validated measures of HRQoL (KIDSCREEN-10), mental health problems (SDQ), anxiety (SCARED), depressive symptoms (CES-DC, PHQ-2), psychosomatic complaints (HBSC-SCL), and fear about the future (DFS-K). Findings were compared to prepandemic population-based data.

Results: While the prevalence of low HRQoL increased from 15% prepandemic to 48% at W2, it improved to 27% at W5. Similarly, overall mental health problems rose from 18% prepandemic to W1 through W2 (30–31%), and since then slowly declined (W3: 27%, W4: 29%, W5: 23%). Anxiety doubled from 15% prepandemic to 30% in W2 and declined to 25% (W5) since then. Depressive symptoms increased from 15%/10% (CES-DC/PHQ-2) prepandemic to 24%/15% in W2, and slowly decreased to 14%/9% in W5. Psychosomatic complaints are across all waves still on the rise. 32–44% of the youth expressed fears related to other current crises.

Conclusion: Mental health of the youth improved in year 3 of the pandemic, but is still lower than before the pandemic.

KEYWORDS

mental health, children and adolescents, longitudinal, COVID-19, pandemic

1. Introduction

Three years into the COVID-19 pandemic, empirical evidence has accumulated that the pandemic has posed a substantial mental health burden on children and adolescents in Germany and worldwide. Systematic reviews summarizing studies from 2020 until spring 2021 conclude that mental health has deteriorated, particularly at the start of the pandemic (1–7). However, one recent meta-analysis showed heterogenous results (8). The nationwide, longitudinal COPSY (***CO****VID-19 and* ***PSY****chological Health*) study (9) has collected data since the very start of the pandemic, i.e., for 3 years, including the latest data collections in 2022. It found that the prevalence of low Health-Related Quality of Life (HRQoL), mental health problems, and anxiety has been elevated over the course of the first three survey waves in 2020 and 2021 compared to prepandemic data (10, 11). In the first wave of the COPSY study, 40% of children and adolescents reported a reduced HRQoL. The second survey wave showed peak levels of reduced HRQoL and of depression and anxiety symptoms, affecting roughly 30% of the participants. Findings of the third wave indicated slight improvements in mental health and HRQoL.

Several studies showed that the prevalence of low HRQoL significantly increased in 2020 (7, 12–15), continued to be low in winter 2020 (9) and spring 2021 (16) and only improved in autumn 2021 (13). Despite this development, some children benefited from more free time and increased time with family during the lockdown in April 2020 (12).

Several longitudinal studies (17, 18) found that mental health problems increased during 2020. Some studies showed that mental health problems peaked in times of lockdowns and school closures [e.g., (19–21)]. An Australian study showed that mental health worsened with the length of lockdowns (22), while a study from Japan reported that the length of school closings was not predictive of mental health (23). Some studies even found an improvement in the well-being of some children during the early lockdown (24, 25). Further, mental health problems slightly improved in autumn 2021 (26).

Two reviews from 2020 described an average doubling of symptoms of anxiety (21%, 26%) and depression (25%, 29%) during the pandemic compared to the prepandemic period (7, 14). Anxiety peaks differ across countries between March to May 2020 (27) (New York, NY) (20) and summer 2021 (United Kingdom, UK) (28). Interestingly, studies from the UK (29) and China (30, 31) report less anxiety during home confinement (29, 31) and a decrease in anxiety from May 2020 to May 2021 (30). Such a recovery effect was also shown in the United States (US) (20) and in a Southern European study (27), but it remained unclear whether the recovery could fully or in part be explained by missing prepandemic data.

Similarly, most studies described an initial increase in depressive symptoms across countries including the Netherlands (17), the US (32), and Australia (21, 33–35). Depressive symptoms either continued to increase through May 2021 (17) or remained stable at a slightly higher level than before the pandemic (26). Only one study from New York reported a recovery effect for depression (20), but again whether full or partial remained unclear.

Another important finding that has been reported by several studies is that specific subgroups of youth (e.g., with previously existing mental health problems, a lower educational level or low socioeconomic status of their parents, with a migration background, limited living space) were particularly vulnerable with respect to their mental health during the pandemic (13, 19, 22, 35–37). On the other hand, children and adolescents with resources like family and social support and an optimistic attitude seem to cope better with the pandemic, i.e., tend to be healthier mentally (9), which has been shown prepandemically as well (38).

From the vastly growing body of research, we learn about mental health trajectories of children and adolescents during the pandemic, although most studies only cover the early years of the pandemic, from 2020 to summer 2021, and study results are mixed. In Germany, there was no lockdown in 2022, the third year of the pandemic, restrictions were loosened further and the COVID-19 infection spread through the country. Most of the 12- to 17-year-olds were either recovered or vaccinated (74.4%) as of November 2022. For 5- to 11-year-olds, the vaccination rate was 22.3% (39, 40). Children's COVID-19 symptoms were generally mild and the physical health of children was on average not impacted severely. However, some studies [e.g., (18)] reported an increase in overall somatic health complaints in a three-wave representative Norwegian longitudinal study of adolescents. In the third year of the pandemic families also faced new crises such as the war in Ukraine, inflation, climate change, and in Germany worries about an energy crisis.

The aims of this study are as follows:

(1) To investigate how childrens' and adolescents' HRQoL and mental health in Germany changed during the past three pandemic years compared to prepandemic data. We hypothesize that the mental burden of the youth declined in the third year of the pandemic due to loosened restrictions and an ongoing adaptation process.
(2) To examine which children are still at risk or more resourceful than others. We expect to replicate our previous results (9) that socially disadvantaged children are still particularly vulnerable and that resourceful children are still mentally healthier in year three.
(3) To explore whether the pandemic still negatively impacts the youth in autumn 2022 or whether new crises (e.g., the climate crisis or the Ukraine war etc.) pose a greater burden. We hypothesize that children and adolescents worry more about new crises than the pandemic now.

2. Materials and methods

2.1. Study design and sample

The German COPSY study is one of the first population-based longitudinal studies to monitor HRQoL and mental health in children and adolescents during the COVID-19 pandemic in Germany and was conducted in five survey waves. Wave 1 (W1, May–June 2020) took place when Germany was under a partial lockdown, Wave 2 (W2, December 2020–January 2021) was conducted during a nationwide lockdown, Wave 3 (W3, September–October 2021) was undertaken after a summer with low infection rates and loosened restrictions, Wave 4 (W4) was at the end of winter (February 2022), when there were still regulations for private gatherings, and Wave 5 (W5) in

autumn 2022 (September–October 2022) with only minimal restrictions in place.

For the COPSY study, families were invited via email to participate in a nationwide online survey which used quota sampling. This method helped to ensure that the sample reflected the sociodemographic characteristics of the German population. The initial participation rate at W1 was 45.8%. The sample sizes of W1-W3, as well as the design and findings of the first three waves, have been described elsewhere (26). Families who had previously participated in the COPSY study were re-invited to each follow-up. To compensate for drop-outs due to aging additional families were recruited, so that sociodemographic representativeness and comparability across all five waves was ensured [please see references (13, 26) for exemplary flowcharts of the sampling of Waves 1 to 3]. Parents provided information on their children and adolescents aged 7 to 17 years, and children and adolescents aged 11 to 17 years were asked themselves. In total, $n = 2{,}471$ families participated in at least one wave of the COPSY study. Samples sizes across all waves ranged between n = 1,586 to n = 1,618 parent-reports and n = 1,040 to n = 1,139 self-reports. Sample sizes and sociodemographic data of Waves 1 to 3 have been described elsewhere (26). Samples of Wave 4 and 5, which are the focus of this manuscript, consisted of n = 1,116 self-reports and n = 1,668 parent-reports (Wave 4) and n = 1,085 self-reports and n = 1,701 parent-reports (Wave 5). The range of response rates across all waves was 80.3 to 86.1% with a response rate for participation in all waves (W1-W5) of 86.1%.

See Figure 1 for the timeline of the COPSY waves in relation to the infection and hospitalization rates in Germany. It should be noted that the survey waves of the COPSY study should not be confused with COVID-19 infection waves, as epidemiologists have recently defined nine pandemic phases based on infection waves (40).

The data of each of the five waves was weighted in order to adjust the samples of the single survey waves to the sociodemographic characteristics of the German population (according to the 2018 Microcensus). Written informed consent to participate in this study was provided by the participants' legal guardian/next kin. The COPSY study was approved by the Local Psychological Ethics Committee (LPEK-0151) and the Commissioner for Data Protection of the University of Hamburg.

For comparison with the HRQoL and mental health of children and adolescents prior to the pandemic, data of $n = 1{,}020$ participants (aged 11 to 17 years) from the nationally representative longitudinal BELLA study (*Behaviour and Well-being of Children and Adolescents in Germany;* (10), conducted between 2014 and 2017 was used. Prepandemic data on psychosomatic complaints of $n = 1{,}073$ children and adolescents, aged 11, 13, and 15 years, was taken from the German HBSC study (*Health Behaviour in School-aged Children;* (41)), which took place between 2017 and 2018.

2.2. Measures

2.2.1. Sociodemographics

The survey covered questions on age, gender, education, living space, single-parenting and migration background.

2.2.2. Health-related quality of life and the burden of COVID-19

For measuring health-related quality of life (HRQoL), the internationally established and widely used self-reported KIDSCREEN-10 Index (42) was used to classify children and adolescents as low, normal or high with respect to their HRQoL, compared to reference data from the national BELLA study. This instrument has proven to be valid and reliable in numerous studies (43–45). Normal HRQoL was defined as M_{BELLA}+/-1SD_{BELLA}.

The survey also included questions about an COVID-19 infection, child and family vaccination status, and whether a family member had died from COVID-19. The perceived burden of the pandemic was surveyed using two questions (13).

2.2.3. Mental health in children and adolescents

Mental health problems were assessed with the internationally well-established Strengths and Difficulties Questionnaire [SDQ; (46)], including the problem scales: emotional problems, conduct problems, hyperactivity, and peer problems with five items each. By summing all

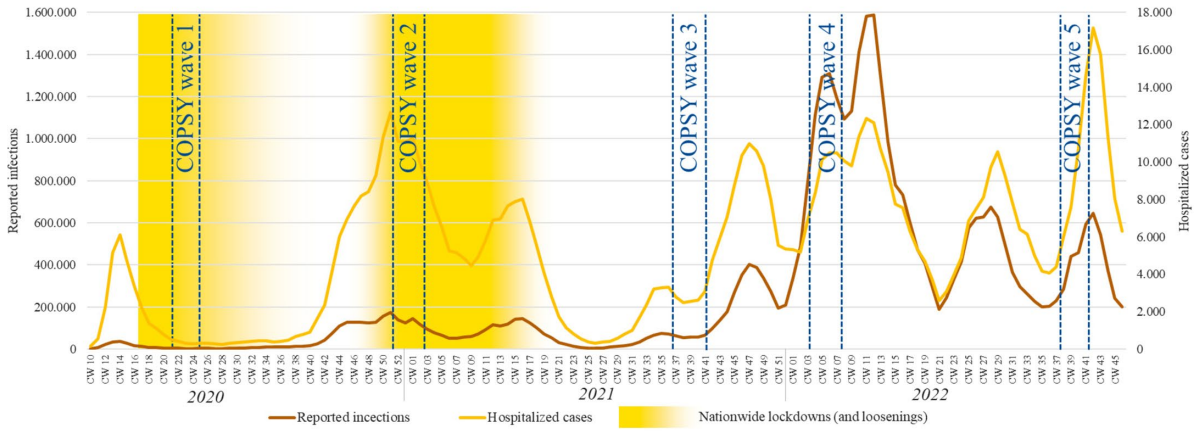


FIGURE 1
Timeline of the survey periods for COPSY W1 to W5 in relation to the infection and hospitalization rates in Germany.

20 items of those subscales a *total difficulties* score can be generated. Higher scores indicate more severe problems. Using established cut-offs (47), participants were divided into two groups: those *with* and those *without* mental health problems (i.e., *abnormal/borderline* vs. *normal*).

Anxiety was assessed using the general anxiety symptom subscale of the Screen for Child Anxiety Related Disorders [SCARED; (48)]. A sum score of the nine items was calculated, with higher scores indicating more severe anxiety. Groups of participants *with* versus *without* anxiety were created using the established cut-off (48).

Depression was measured using the seven symptom items of the German version of the Center for Epidemiological Studies Depression Scale (CES-DC), which have been previously used in the BELLA study (49, 50). In addition, the Patient Health Questionnaire [PHQ-2; (51, 52)] was used. While the CES-DC measures impairment due to depressive symptoms within the past week, the two items of the PHQ-2 ask about core symptoms of major depressive disorder. A higher sum score indicates more severe depressive symptoms according to PHQ-2 and higher impairment by depressive symptoms according to CES-DC. Validated cut-offs were applied to categorize participants *with* and those *without* noticeable depressive symptoms (51, 53).

Further, the self-reported HBSC symptom checklist [HBSC-SCL; (54)], which is a longstanding validated instrument in international studies, was administered to measure the frequencies of psychosomatic complaints. Participants were divided into groups of subjects who experienced each psychosomatic symptom *at least once per week* vs. those who experienced it *less frequently*.

Additional questions were used to assess young people's concerns about other current crises, such as the war in Ukraine, climate change, and the energy crisis. Those were rated on a 5-Point-Likert scale from *not at all* to *very* worried. In addition, a short version of the Dark Future Scale for Children [DFS-K; (55, 56)], measuring the tendency to think about the future with anxiety and uncertainty was administered to assess future anxiety and future worries in the context of different current crises. Regarding to Zaleski (56), future anxiety is defined as a "fear of future events and a feeling that dangerous or adverse changes may occur in the future." A pandemic related version of the DFS-K was first administered in a sample of school students, where acceptable Cronbach's Alpha (0.76), a Retest-correlation of 0.34 and content as well as construct validity are reported (55). All HRQoL and mental health measures (KIDSCREEN-10 Index, SDQ, SCARED, CES-DC, HBSC-SCL PHQ-2) are internationally validated and have good psychometric properties, which have been described previously (26).

2.2.4. Personal, family, and social resources

The Personal Resources Scale [PRS; (57)] was administered to assess personal resources, such as problem-solving skills and optimism of children and adolescents. Perceived social support was measured with four items of the Social Support Scale [SSS; (58, 59)]. Four items from the Cohesion subscale of the Family Climate Scale (FCS) were used to assess family climate (60). Sum scores were calculated for all three scales (PRS, SSS, FCS) with higher values indicating more pronounced resources (57, 58, 60). Respondents with scores below the 20th percentile in the prepandemic BELLA study were considered as having deficits in that particular resource.

2.3. Data analysis – change in HRQoL and mental health

First, the reported burden of the pandemic on children and adolescents was compared across the five COPSY waves. Second, the proportions of youth displaying impaired HRQoL, mental health problems, anxiety, depressive symptoms, and psychosomatic complaints were compared across all measurement points. Age- and gender-adjusted proportions were calculated using logistic regression models for each outcome. Prepandemic population-based data available from the BELLA study (10) and the HBSC study (41) were used to compare HRQoL and mental health outcomes. Chi-square tests and effect sizes (Phi coefficient ϕ, and Cramer's V, resp. with 0.10 indicating a small, 0.30 a medium, and 0.50 a strong effect) were calculated for comparisons across the waves and for group comparisons. Gender differences were expressed as risk ratios (RR) for girls. The chi-square tests as well as the ϕ and V statistics for comparisons across the COPSY waves do not take into account that the majority of measurements represent repeated measures of the same respondents. This led to a lower statistical power in order to detect differences across the COPSY waves. The results of the Dark Future Scale and young people's concerns about other recent crises were reported as descriptive statistics.

Additional logistic regression models were conducted to examine the association between risk group and resource cluster and the main COPSY outcomes (low HRQoL, mental health problems, anxiety, and depressive symptoms). All analyses were controlled for age and gender. Last logistic regression analysis were run to test the association between single-parenting and the above mentioned outcomes.

Alpha-error level was not adjusted because the outcomes were correlated across and within waves, the analyses thus cannot be considered as independent testing of the same hypothesis.

Prior to analyzing the data, a power analysis was conducted using G-Power (Version 3.1). The minimum sample size based on a statistical significance of $p(\alpha) < 0.05$ and a power of $p\ (1-ß) = 0.8$ for a medium effect ($w = 0.3$) between waves according to age groups (7–10 years, 11–13 years, 14–17 years) and female vs. male was $n = 88$.

3. Results

3.1. Sociodemographics

In total, across all five survey waves of the COPSY study, $n = 2{,}471$ families with children and adolescents aged 7–17 years ($M = 13.09$, $SD = 3.83$, 50.1% female) completed parent reports. Self-reports were completed by $n = 1{,}673$ children and adolescents from age 11–17 ($M = 15.23$, $SD = 2.56$, 52.1% female). Sociodemographic characteristics are reported in Table 1 and Figure 2.

Parents who participated in all five surveys were less likely to have a migration background (14.0% vs. 18.4%; $V = 0.06$), were more likely to have a lower level of education (18.7% vs. 12.9%; $V = 0.09$) and were on average 3.3 years older ($d = 0.44$) than those who only responded to 1–4 Waves. No other significant differences in sociodemographic or mental health-related variables were found.

3.2. Changes in HRQoL during the pandemic

After the initial increase in perceived burden of the pandemic (from W1: 70.6% to W2: 82.9%) in children and adolescents aged 11 to 17, the burden decreased between W3 and W4 and W5, i.e., changing from 81.9% in fall 2021 to 80.3% in February 2022 and most recently falling to 73.0% in fall 2022 ($p < 0.001$, $\phi = 0.08$) (see Figure 2).

After the initial significant increases of the prevalence of low HRQoL from prepandemic 15.3% to 40.2% (W1) and 47.7% (W2) in

TABLE 1 Sociodemographic characteristics of the COPSY sample.

	Children and adolescents aged 7–17years (parent reports) (*n*=2,471)		Children and adolescents aged 11–17years (self-reports) (*n*=1,673)	
	n (%)	*M (SD)*	*n* (%)	*M (SD)*
Age[a]		13.09 (3.83)		15.23 (2.56)
7–10 years	775 (31.4)		–	
11–13 years	512 (20.7)		494 (29.5)	
14–17 years[b]	1,184 (47.9)		1,179 (70.5)	
Gender				
Male	1,221 (49.4)		792 (47.3)	
Female	1,239 (50.1)		872 (52.1)	
Other	10 (0.4)		9 (0.5)	
Age of the parent[a]		44.28 (7.57)		46.34 (7.22)
Gender of the parent				
Male	1,076 (43.5)		741 (44.3)	
Female	1,392 (56.3)		930 (55.6)	
Other	3 (0.1)		2 (0.1)	
Migration background				
No	2,033 (82.9)		1,376 (82.9)	
Yes	419 (17.1)		283 (17.1)	
Parental education				
Low	355 (14.4)		272 (16.3)	
Medium	1,397 (56.5)		945 (56.5)	
High	668 (27.0)		417 (24.9)	
No information	51 (2.1)		39 (2.3)	
Single parent				
No	2,018 (81.7)		1,338 (80.0)	
Yes	453 (18.3)		335 (20.0)	
Occupational status				
Full-time employed	1,281 (51.8)		892 (53.3)	
Part-time employed	754 (30.5)		486 (29.0)	
Self-employed	101 (4.1)		71 (4.2)	
Other employment	40 (1.6)		29 (1.7)	
Stay-at-home parent	147 (5.9)		96 (5.7)	
Retiree/pensioner	59 (2.4)		47 (2.8)	
On parental leave	33 (1.3)		13 (0.8)	
Unemployed	56 (2.3)		39 (2.3)	
COVID-19 infection[c]				
A family member has been infected	1,458 (85.7)		946 (84.0)	
The child has been infected	1,056 (62.1)		678 (60.2)	
A relative has died of COVID-19	156 (9.2)		109 (9.7)	

Unweighted data. M = mean; SD = standard deviation. [a]Age at the latest time of participation. [b]Some adolescents had already turned 18 when they participated in the survey but were included in the age group of 14- to 17-year-olds at W5. [c]Any previous confirmed infection with the coronavirus according to parental report at W5.

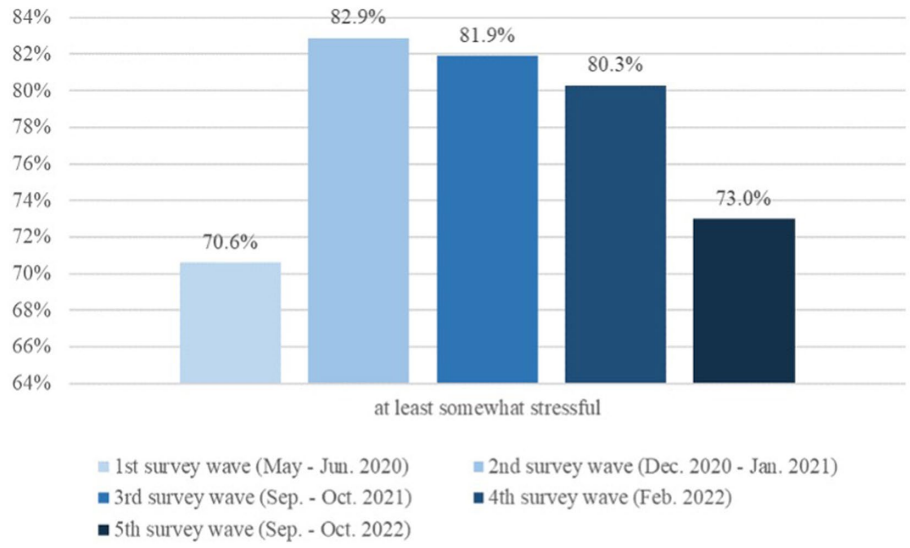


FIGURE 2
Perceived burden of the pandemic among children and adolescents 11–17 years old.

2020, it significantly dropped in spring 2021 to 35.1% (W3), increased in autumn of that year back up to 41.0% (W4), and most recently decreased to 27.0% ($p < 0.001$, $\phi = 0.14$). This latest figure is still higher than prepandemic rates, but lower than at the start of 2020. The prevalence of low HRQoL at W5 is now less than twice as high as the prepandemic rate (see Figure 3). An analysis stratified by gender revealed that girls had a two-fold higher risk (Risk Ratio, RR) of low HRQoL compared to boys in prepandemic times, the RR decreased to 1.2 and 0.9 during the pandemic.

3.3. Changes in mental health during the pandemic

After the initial worsening of overall mental health problems (SDQ) in 2020 (from prepandemic: 17.6% to W1: 30.4%, W2: 30.9%), and a slight but non-significant decrease in mental health problems in Wave 3 (27.0% winter 2020), *via* 28.5% (W4) the downwards trend continued to 22.6% (W5). This is still 5% higher than prepandemic but shows clear improvement in overall mental health over the course of 2022.

Similarly, after an increase in anxiety of children and adolescents aged 11–17 years during 2020 compared to prepandemic data (prepandemic: 14.9%, W1: 24.1%, W2: 30.1%), anxiety decreased in autumn 2021 (W3: 26.8%), slightly increased in Feburary 2022 (W4) to 27.8% and then decreased further to 24.8%, which is about the same level as at the start of the pandemic (W1, see Figure 3).

While depressive symptoms in adolescents aged 11–17 years initially only slightly, non-significantly increased between prepandemic data and W1 (CES-DC: 15.0 to 18.0%; PHQ-2: 10.0 to 11.3%), the increase reached significance in W2 (CES-DC: 24.3%, PHQ-2: 15.0%), followed by a significant decrease in Wave 3 (CES-DC: 19.7%, PHQ-2: 11.1%), a slight increase in W4 (CES-DC: 21.3%, PHQ-2: 12.8%), and finally a significant decrease in W5 (CES-DC: 14.2%, PHQ-2: 8.6%). However, most differences were small.

Girls had a greater risk of reporting both anxiety (RR ranged between 1.0 and 2.1) and depressive symptoms (RR ranged between 1.4 and 1.9).

3.4. Changes in psychosomatic complaints

Self-reported psychosomatic complaints in adolescents aged 11–17 years peaked between W1 and W3, depending on the specific complaint. Irritability, sleeping problems, feeling low and nervousness decreased in autumn 2022, but are still considerably higher than prepandemic. Headaches and stomachaches gradually increased from 2020 to 2022, i.e., about half of all children and adolescents suffered from them at least once during the past week. Back pain has remained on a higher than prepandemic level since W2 (W5: 33.3% vs. prepandemic: 25.9%). Further details of psychosomatic complaints can be found in Figure 4. Girls were at a higher risk of psychosomatic complaints in most complaints and waves, in particular with regard to having headaches (RR between 1.2 and 1.9) and feeling low (RR between 0.9 and 1.9).

Children and adolescents with a reported previous COVID-19 infection self-reported slightly higher prevalences of psychosomatic complaints. However, the differences were only statistically significant for some complaints (i.e., irritability as well as being impaired by these symptoms) ($V = 0.10$–0.14).

3.5. Risks and resources of children and adolescents' HRQoL and mental health

The logistic regression analyses showed that at all five waves, children and adolescents aged 11–17 and belonging to the risk group had a higher risk (Odds Ratio, OR) of experiencing low HRQoL (ORs ranged from 2.4 to 3.3), mental health problems (ORs ranged from 2.7 to 5.1), anxiety symptoms (ORs ranged from 1.2 to 2.2), and depressive

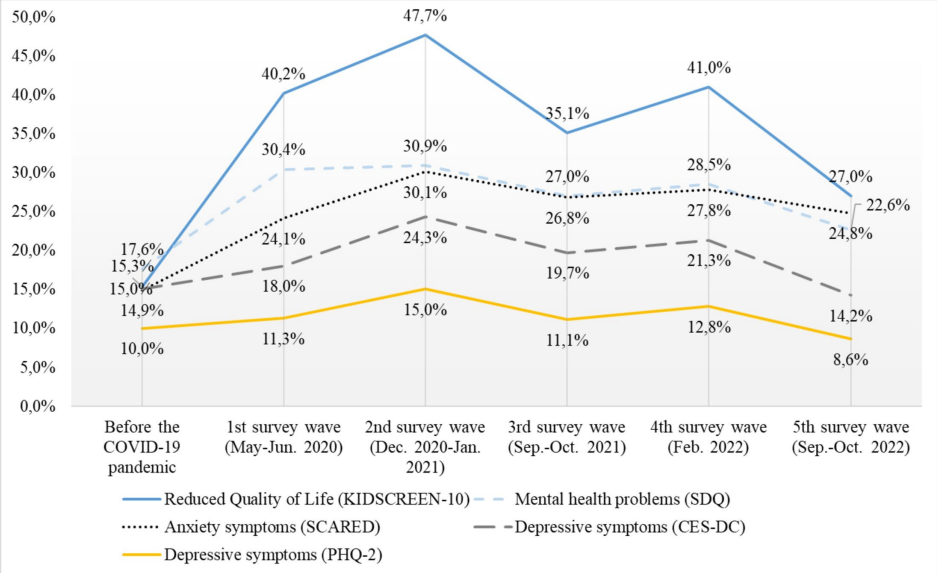


FIGURE 3
Mental health of children and adolescents aged 7–17 years from 2020 to 2022.

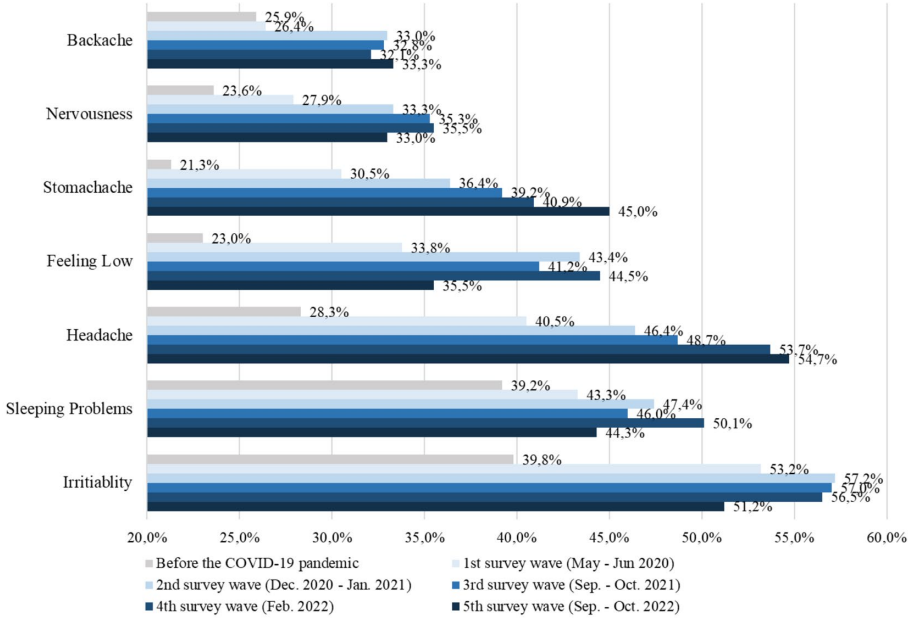


FIGURE 4
Psychosomatic complaints of children and adolescents aged 7–17 years old from 2020 to 2022.

symptoms compared with their peers (ORs ranged from 2.4 to 3.7). The prevalence of having low HRQoL, and overall mental health problems was higher for the risk group across all 5 waves (see Table 2). Children and adolescents belonging to the resource cluster, on the other hand, had a significantly reduced risk (OR) of low HRQoL (OR ranged from 0.1 to 0.2), mental health problems (OR ranged from 0.1 to 0.2), anxiety symptoms (OR ranged from 0.2 to 0.3) and depressive symptoms (OR ranged from 0.1 to 0.2).

Regarding age-specific differences, adolescents aged 14–17 had a lower risk of mental health problems at W3 through W5 compared with 11- to 13-year-olds (OR ranged from 0.4 to 0.6). Further, 14- to 17-year-olds had a lower risk of anxiety at W1 and W2 compared with 11- to 13-year-olds (ORs were 0.5, see Table 2). The logistic regression models fit the data well with the Hosmer and Lemeshow Tests both being non-significant and Nagelkerke's Pseudo R^2 ranging between 0.12 and 0.33.

10.3389/fpubh.2023.1129073

TABLE 2 Logistic regression of risks and resources on low HRQoL, mental health problems and symptoms of anxiety and depression (11–17year olds).

		Low HRQol		Mental health problems		Anxiety symptoms		Depressive symptoms	
		OR	Sig	OR	Sig	OR	Sig	OR	Sig
Risk group	Wave 1	2.35	***	2.67	***	1.98	***	2.84	***
	Wave 2	3.32	***	2.96	***	2.21	***	3.05	***
	Wave 3	2.53	***	2.67	***	1.20		2.35	***
	Wave 4	3.02	***	4.71	***	2.81	***	3.65	***
	Wave 5	3.02	***	5.10	***	2.18	***	2.58	***
Resource cluster	Wave 1	0.15	***	0.21	***	0.28	***	0.18	***
	Wave 2	0.11	***	0.18	***	0.22	***	0.15	***
	Wave 3	0.10	***	0.15	***	0.27	***	0.13	***
	Wave 4	0.18	***	0.16	***	0.24	***	0.16	***
	Wave 5	0.07	***	0.10	***	0.24	***	0.09	***
Female	Wave 1	1.05		0.81		1.03		1.49	
	Wave 2	1.09		0.77		0.72		1.15	
	Wave 3	1.00		0.62		1.59	*	1.46	
	Wave 4	0.96		0.72		1.94	**	1.45	
	Wave 5	0.73		0.51	*	0.88		1.22	
Age 14–17	Wave 1	0.74		0.69		0.49	**	0.90	
	Wave 2	1.14		0.63		0.51	**	1.06	
	Wave 3	0.83		0.43	***	0.71		1.07	
	Wave 4	0.88		0.58	*	0.86		1.53	
	Wave 5	1.06		0.53	*	0.89		2.06	
Female * Age 14–17	Wave 1	1.48		0.96		2.50	**	1.06	
	Wave 2	1.13		1.37		4.19	**	1.48	
	Wave 3	1.60		1.63		1.11		0.98	
	Wave 4	1.01		0.94		0.74		0.77	
	Wave 5	1.37		1.15		1.42		0.76	
Nagelkerke's R^2	Wave 1	0.25		0.14		0.14		0.18	
	Wave 2	0.32		0.17		0.20		0.22	
	Wave 3	0.30		0.19		0.12		0.20	
	Wave 4	0.23		0.22		0.17		0.21	
	Wave 5	0.33		0.26		0.14		0.22	

Significance levels: *$p < 0.05$; **$p < 0.01$, ***$p < 0.001$.

Also, children and adolescents from single parent families had a significantly increased risk (OR) of impaired HRQoL in W1 (OR: 1.55) and W2 (OR: 1.39), for mental health problems (SDQ) in W3 (OR: 1.60), W4 (OR: 1.43) and W5 (OR: 1.41), depressive symptoms (CES-DC) in W2 (OR: 1.50) and W5 (OR: 1.48) and for anxiety symptoms (SCARED) in W1-W5 (OR: 1.64, 1.38, 1.64, 1.40, 1.43).

3.6. New worries of children and adolescents in 2022

Administering novel questions about current worries for the first time in autumn 2022, the COPSY study found that almost half of all children and adolescents aged 11–17 years old were *quite* concerned or *very* worried about the financial and energy crisis during the winter, and nearly as many of them indicated worries about the war in Ukraine. A third of the surveyed youth were also worried about the climate crisis, while only about 10% indicated to still worry about the pandemic (Figure 5).

Items assessing future anxiety in the context of current crises (like the COVID-19 pandemic, climate crisis, the war etc.) revealed that 82.6% of the youth often report the dominating fear that the current crises will remain for a long time. 78.4% reported fear that their life will get worse due to crises and 74.5% were worried that families will be able to afford less in the future. About half of the children (50.6%) were deeply worried about the

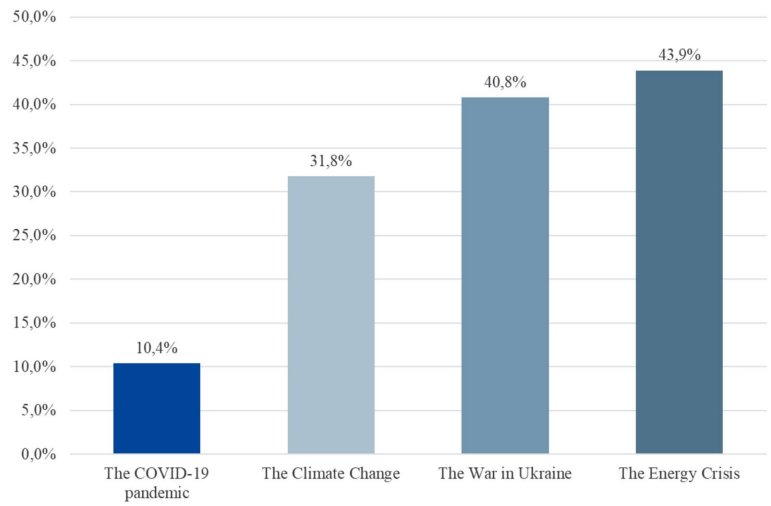


FIGURE 5
Worries of children and adolescents aged 11–17 years in autumn 2022.

uncertainty related to crises, e.g., they were afraid that they will not achieve their school qualification or professional goals in the future.

4. Discussion

The current data (COPSY W5) from autumn 2022 of the longitudinal COPSY study shows a specific pattern (in line with our first hypothesis): After two years of pandemic-related deterioration in child and adolescent mental health, an improvement has finally taken place in 2022. First slight improvements were already visible in autumn 2021, but leveled out in winter 2021/22. The changes between W4 and W5 of the COPSY study provide evidence that this trend of improvement occurred throughout the past year. However, except for depressive symptoms, which show a recovery effect to prepandemic levels, we surprisingly found that in year three of the pandemic most mental health outcomes are still worse than before the pandemic, and some self-reported psychosomatic complaints like headaches and stomachaches are still steadily on the rise.

The finding for the sharp initial increase of mental health problems at the beginning of the pandemic, which then stabilized at a high level from 2020 until spring 2021, is in line with the results of current literature reviews synthesizing the respective evidence during the first 1.5 years of the pandemic (2, 5, 6). Only one recent meta-analysis could not report clear mental health deteriorations during the pandemic, potentially due to comparing studies of young children with university students and comparing studies with small/non-representative samples (8). Most of the to-date published studies either report on prepandemic vs. pandemic effects or were administered between 2020 and spring 2021. To our knowledge, only three other longitudinal studies (20, 61–63) and two health insurance reports (64, 65) have published data on the mental health of children and/or adolescents over the course of 2021 and/or 2022.

Witte et al. (65) published German statutory health insurance data covering the years 2018 until end of 2021. According to their data, diagnosed mental health problems in children and adolescents between the ages of 10 and 17 increased during the pandemic. For girls, there was a particular increase of mood disorders from 18% to 23%, of anxiety disorders from 7% to 24%, and eating disorders from 33% to 54% (65).

In Austria, a four-wave longitudinal study on children aged 3 to 13 years old and using parent reports found that children's mental health (internalizing problems, posttraumatic stress symptoms) was worse in December 2021 than in March 2020 (66). In Italy, a three-wave longitudinal study on 5- to 6-year-old children using parent reports found that between 10/2019 and 03/2020 and 10/2021 the depressive scores of a risk group increased from t1 to t2 and stayed stable until t3, while the depressive scores of the non-risk group first decreased from t1 to t2 and then increased again (61).

In a German cohort study, Schnetzer & Hurrelmann (62, 63) describe, similarly to the COPSY study, slight improvements in mental health in 2022 in 14 to 29-year-olds. From March to October 2022 the percentage of those feeling "stressed out" decreased from 45% to 41% and the percentage of those feeling listless decreased from 35% to 31%. Similarly, small to minor decreases were reported for the proportion of adolescents feeling exhausted (32% to 29%) and depressed (27% to 26%).

Thus far, hardly any other repeated/interval study has reported mental health recovery effects as found by the COPSY and the study from Schnetzer & Hurrelmann (62, 63) for children and adolescents during the pandemic. Small, potentially interim, improvements have been reported by Hawes et al. (20) in relation to the peak of infection rates in New York in 2020 – and by Zuccolo et al. (67) in Brazil, however the improvements found in emotional problems in winter 2020/21 did not continue over the course of 2021.

The COPSY study is one of the first studies to show that the mental health of children and adolescents improved in 2022. These

improvements in mental health could be caused by an ongoing adaptation process during the past pandemic years, e.g., increasing resilience; the resumption of normal social, physical, and entertainment activities due to fewer restrictions; the availability of vaccinations, which make most disease courses less severe; and the fact that most of the children coped with the COVID-19 infection itself well.

While mental health is improving for almost all outcomes in the COPSY study over time, it is important to note that 14% to 27% of the youth still suffer impairments in health-related quality of life, depression, anxiety or other mental problems. Therefore, we cannot yet speak of a full recovery to prepandemic levels.

In terms of our second hypothesis that some children are at a higher mental health risk than others, we replicated our previous results as hypothesized (10): for example that particularly girls show higher HRQoL impairments and more depressive and anxiety symptoms, which was replicated across most COPSY waves. And adolescents aged 14–17 years had a lower risk of mental health problems across most waves compared with 7–10-year-olds. This age and gender effect should be taken into account in prevention and treatment programs for the youth.

The fifth wave of the COPSY study also replicated that children with high family and social support and good personal resources (45%) had a 4 to 14 times higher chance of better mental health outcomes than other children. In contrast, 16% of children and adolescents belonging to a risk group with a higher parental pandemic burden or parental mental health problems or low parental education or restricted living space and migration background had an up to 5 times higher risk of mental health problems. This is in line with other studies (13, 19, 22, 35, 36).

Further, and different from the overall positive trend, the COPSY study surprisingly found an ongoing steady increase in self-reported stomachaches and headaches across the past three pandemic years, with almost half of the children and adolescents reporting them at least once per week. Other studies on somatic complaints in children and adolescents during the pandemic also report increases in somatic complaints. Bantel et al. (68) found an increase in headaches and stomachaches in children at the school entry level in a German town between prepandemic and pandemic assessments, and Hafstadt et al. (18) found an increase in overall somatic complaints in a three-wave representative Norwegian longitudinal study of 12- to 16-year-olds. However, the prevalence of headaches and stomachaches in the COPSY study is higher than the result of a large German school study (69), which found that only 23% of fifth- to tenth- graders had headaches and 14% had stomachaches at least once a week in 2020 and 2021. Whether the higher prevalence of stomachaches and headaches found in COPSY could be related to a higher rate of COVID-19 or other infections with such symptoms in the COPSY sample or could reflect higher strain in the COPSY sample being turned psychosomatic is an open question, which may be answered by future research. The hypothesis that those symptoms are psychosomatic, i.e., children who may not otherwise be able to express their worries or other mental problems may somatize them, may be a more likely hypothesis than those symptoms being directly related to a previous COVID-19 infection. More research should be initiated to further explore psychosomatic and somatic complaints of children during the pandemic.

And finally, as assumed in hypothesis three, while between 2020 and 2021 the COVID-19 pandemic was the topic children were most worried about, in 2022 new crises and future anxieties have emerged. Similarly to a longitudinal study conducted in 2022 in Germany (62) and as expected by us, the COPSY study shows that the pandemic is no longer the topic children and adolescents worry most about. The COPSY W5 found that children aged 11–17 years mostly worry about the energy or financial crisis, Schnetzer & Hurrelmann (62) also found that adolescents between the age of 14 to 29 years worry most about inflation. The second most worrisome topic for youth in both studies was the war in Ukraine, followed by climate change. When comparing the proportions of worried youth in these studies, adolescents aged 14–29 years appear more worried than the 11- to 17-year-olds in the COPSY study. This may be due to age-related differences in political interest and awareness, differences in media use, i.e., older children/adolescents may be more affected by crises, because they may be exposed to more news coverage on that or other factors like different methods of assessment in both studies.

Limitations of the COPSY study are that the sample was drawn by matching data from the German Microcensus (2018), so results may not be generalizable across countries; the survey was administered using an online panel and incentives were deployed, which may have influenced the participation. Lastly, the study does not allow the causal conclusion that the pandemic has caused all the results found; there could be other factors, like response bias, which may have caused them. Strengths of the COPSY study are the longitudinal design comprising almost 3 years, the large population-based sample, the availability of nationally representative, comparative prepandemic data, and the use of validated, internationally-established instruments.

5. Conclusion

Waves 4 and 5 of the COPSY study underline the trend of mental health improvements in 2022, but most mental health problems are still above the prepandemic baseline in year three of the pandemic. Some psychosomatic symptoms like self-reported headaches and stomachaches are on the rise. Only depressive symptoms recovered.

Mental health problems of minors, even if not (yet) clinically diagnosed, should not be overlooked because when untreated they can turn into mental disorders in adulthood causing prolonged suffering and an increase in health care costs.

For those mentally burdened children and adolescents who are still at risk, we recommend nationwide, low-threshold support. Such programs have also been recommended by the German Ethics Committee (70) and the Center for Disease Control (CDC) (71). In line with the German Ethics Committee, we call for support of those children and adolescents and their families to compensate for their many pandemic sacrifices. As an act of transgenerational solidarity, the German Ethics Committee is demanding now that our society focuses on the mental health care of minors (70).

Institutions that offer mental health diagnostics, counseling, treatment, and aids for the participation of youth and support of families need to be financed both reliably and long term. Improving mental health care seems particularly important in neighborhoods with limited living space, low education, financial problems, and a high percentage of migrants. Current severe deficits in mental health

care need to be addressed and remedied. With the increasing demand for mental health care of youth, it is crucial to reduce waiting times for mental health care (72).

In view of current and future crises - like the Ukraine war and the climate crisis - and the demographic change with fewer younger people who have to bear more burdens compared to previous generations, it is crucial that existing burdens are reduced now to avoid long-term mental health impairments.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by Local Psychological Ethics Committee (LPEK-0151). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

UR-S, AK, ME, JD, FR, A-KN, MG, and AS contributed to conception and design of the study. LS organized the database. ME performed the statistical analysis. JD and UR-S drafted the first manuscript. ME, AK, JD, A-KN, MG, and LS wrote sections of the manuscript. All authors contributed substantially to the discussion of results, manuscript revision, read, and approved the submitted version.

Funding

We acknowledge financial support from the Open Access Publication Fund of UKE - Universitätsklinikum Hamburg-Eppendorf and DFG – German Research Foundation. The COPSY study was funded by the Jaekel Foundation and the Foundation "Wissenschaft in Hamburg." Furthermore, the COPSY study is part of the NUM 2.0 project coverCHILD, funded by the Federal Ministry of Education and Research (BMBF). The funders had no role in study design, data collection and analysis, decision to publish, or the preparation of the manuscript.

Acknowledgments

We thank all children, adolescents and their parents who participated in this study for their time and involvement.

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. Bussières EL, Malboeuf-Hurtubise C, Meilleur A, Mastine T, Hérault E, Chadi N, et al. Consequences. of the COVID-19 pandemic on children's mental health: a meta-analysis. *Front. Psychiatry*. (2021) 12:12. doi: 10.3389/fpsyt.2021.691659

2. Elharake JA, Akbar F, Malik AA, Gilliam W, Omer SB. Mental health impact of COVID-19 among children and college students: a systematic review. *Child Psychiatry Hum Dev*. (2022) 54:913–25. doi: 10.1007/s10578-021-01297-1

3. Imran N, Aamer I, Sharif MI, Bodla ZH, Naveed S. Psychological burden of quarantine in children and adolescents: a rapid systematic review and proposed solutions. *Pak J Med Sci*. (2020) 36:1106–16. doi: 10.12669/pjms.36.5.3088

4. Ma L, Mazidi M, Li K, Li Y, Chen S, Kirwan R, et al. Prevalence of mental health problems among children and adolescents during the COVID-19 pandemic: a systematic review and meta-analysis. *J Affect Disord*. (2021) 293:78–89. doi: 10.1016/j.jad.2021.06.021

5. Racine N, McArthur BA, Cooke JE, Eirich R, Zhu J, Madigan S. Global prevalence of depressive and anxiety symptoms in children and adolescents during COVID-19: a meta-analysis. *JAMA Pediatr*. (2021) 175:1142–50. doi: 10.1001/jamapediatrics.2021.2482

6. Samji H, Wu J, Ladak A, Vossen C, Stewart E, Dove N, et al. Mental health impacts of the COVID-19 pandemic on children and youth – a systematic review. *Child Adolesc Ment Health*. (2022) 27:173–89. doi: 10.1111/camh.12501

7. Schlack R, Neuperdt L, Junker S, Eicher S, Heike H, Thom J, et al. Changes in mental health in the German child and adolescent population during the COVID-19 pandemic – Results of a rapid review. *J Health Monit*. (2022) 8:2–72. doi: 10.25646/10761

8. Newlove-Delgado T, Russell AE, Mathews F, Cross L, Bryant E, Gudka R, et al. Annual research review: the impact of Covid-19 on psychopathology in children and young people worldwide: systematic review of studies with pre- and within-pandemic data. *J Child Psychol Psychiatry*. (2022) 64:611–40. doi: 10.1111/jcpp.13716

9. Ravens-Sieberer U, Kaman A, Erhart M, Devine J, Schlack R, Otto C. Impact of the COVID-19 pandemic on quality of life and mental health in children and adolescents in Germany. *Eur Child Adolesc Psychiatry*. (2022) 31:879–89. doi: 10.1007/s00787-021-01726-5

10. Otto C, Reiss F, Voss C, Wüstner A, Meyrose AK, Hölling H, et al. Mental health and well-being from childhood to adulthood: design, methods and results of the 11-year follow-up of the BELLA study. *Eur Child Adolesc Psychiatry*. (2021) 30:1559–77. doi: 10.1007/s00787-020-01630-4

11. Ravens-Sieberer U, Kaman A, Devine J, Löffler C, Reiß F, Napp AK, et al. The mental health and health-related behavior of children and parents during the COVID-19 pandemic. *Dtsch Ärztebl*. (2022) 119:436–7. doi: 10.3238/arztebl.m2022.0173

12. Christner N, Essler S, Hazzam A, Paulus M. Children's psychological well-being and problem behavior during the COVID-19 pandemic: an online study during the lockdown period in Germany. *PLoS One*. (2021) 16:e0253473. doi: 10.1371/journal.pone.0253473

13. Ravens-Sieberer U, Kaman A, Erhart M, Otto C, Devine J, Löffler C, et al. Quality of life and mental health in children and adolescents during the first year of the COVID-19 pandemic: results of a two-wave nationwide population-based study. *Eur Child Adolesc Psychiatry*. (2021) 32:575–88. doi: 10.1007/s00787-021-01889-1

14. Vogel M, Meigen C, Sobek C, Ober P, Igel U, Körner A, et al. Well-being and COVID-19-related worries of German children and adolescents: A longitudinal study from pre-COVID to the end of lockdown in Spring 2020. *JCPP Adv*. (2021) 1:e12004. doi: 10.1111/jcv2.12004

15. Wunsch K, Nigg C, Niessner C, Schmidt SCE, Oriwol D, Hanssen-Doose A, et al. The impact of COVID-19 on the interrelation of physical activity, screen time and health-related quality of life in children and adolescents in Germany: results of the Motorik-Modul Study. *Children*. (2021) 8:98. doi: 10.3390/children8020098

16. Theuring S, van Loon W, Hommes F, Bethke N, Mall MA, Kurth T, et al. Psychosocial wellbeing of schoolchildren during the COVID-19 pandemic in Berlin, Germany, June 2020 to March 2021. *Int J Environ Res Public Health*. (2022) 19:10103. doi: 10.3390/ijerph191610103

17. Green KH, van de Groep S, Sweijen SW, Becht AI, Buijzen M, de Leeuw RNH, et al. Mood and emotional reactivity of adolescents during the COVID-19 pandemic: short-term and long-term effects and the impact of social and socioeconomic stressors. *Sci Rep*. (2021) 11:11563. doi: 10.1038/s41598-021-90851-x

18. Hafstad GS, Sætren SS, Wentzel-Larsen T, Augusti EM. Changes in adolescent mental and somatic health complaints throughout the COVID-19 pandemic: a three-wave prospective longitudinal study. *J Adolesc Health*. (2022) 71:406–13. doi: 10.1016/j.jadohealth.2022.05.009

19. Felfe C, Saurer J, Schneider P, Vornberger J, Klotzbuecher V, Erhart M, et al. The youth mental health crisis during the COVID-19 pandemic: the role of school closures [Preprint]. Available at: https://www.researchsquare.com/article/rs-1304103/v2

20. Hawes MT, Szenczy AK, Olino TM, Nelson BD, Klein DN. Trajectories of depression, anxiety and pandemic experiences; a longitudinal study of youth in New York during the Spring-Summer of 2020. *Psychiatry Res*. (2021) 298:113778. doi: 10.1016/j.psychres.2021.113778

21. Houghton S, Kyron M, Hunter SC, Lawrence D, Hattie J, Carroll A, et al. Adolescents' longitudinal trajectories of mental health and loneliness: the impact of COVID-19 school closures. *J Adolesc*. (2022) 94:191–205. doi: 10.1002/jad.12017

22. Price AM, Measey MA, Hoq M, Rhodes A, Goldfeld S. Child and caregiver mental health during 12 months of the COVID-19 pandemic in Australia: findings from national repeated cross-sectional surveys. *BMJ Paediatr Open*. (2022) 6:e001390. doi: 10.1136/bmjpo-2021-001390

23. Takahashi F, Honda H. Prevalence of clinical-level emotional/behavioral problems in schoolchildren during the coronavirus disease 2019 pandemic in Japan: a prospective cohort study. *JCPP Adv*. (2021) 1:e12007. doi: 10.1111/jcv2.12007

24. Soneson E, Puntis S, Chapman N, Mansfield KL, Jones PB, Fazel M. Happier during lockdown: a descriptive analysis of self-reported wellbeing in 17,000 UK school students during Covid-19 lockdown. *Eur Child Adolesc Psychiatry*. (2022) 32:1–16. doi: 10.1007/s00787-021-01934-z

25. Waite P, Pearcey S, Shum A, Raw JAL, Patalay P, Creswell C. How did the mental health symptoms of children and adolescents change over early lockdown during the COVID-19 pandemic in the UK? *JCPP Adv*. (2021) 1:e12009. doi: 10.1111/jcv2.12009

26. Ravens-Sieberer U, Erhart M, Devine J, Gilbert M, Reiss F, Barkmann C, et al. Child and adolescent mental health during the COVID-19 pandemic: results of the three-wave longitudinal COPSY study. *J Adolesc Health*. (2022) 71:570–8. doi: 10.1016/j.jadohealth.2022.06.022

27. Orgilés M, Francisco R, Delvecchio E, Espada JP, Mazzeschi C, Pedro M, et al. Psychological symptoms in Italian, Spanish and Portuguese youth during the COVID-19 health crisis: a longitudinal study. *Child Psychiatry Hum Dev*. (2022) 53:853–62. doi: 10.1007/s10578-021-01211-9

28. Adegboye D, Lennon J, Batterbee O, Thapar A, Collishaw S, Shelton K, et al. Understanding de novo onset of anxiety during COVID-19: Pre-pandemic socio-emotional functioning in vulnerable children. *JCPP Adv*. (2022) 2:e12076. doi: 10.1002/jcv2.12076

29. Widnall E, Winstone L, Plackett R, Adams EA, Haworth CMA, Mars B, et al. Impact of school and peer connectedness on adolescent mental health and well-being outcomes during the COVID-19 pandemic: a longitudinal panel survey. *Int J Environ Res Public Health*. (2022) 19:6768. doi: 10.3390/ijerph19116768

30. Zhou C, Li R, Yang M, Duan S, Yang C. Psychological status of high school students 1 year after the COVID-19 emergency. *Front. Psychiatry*. (2021) 12:12. doi: 10.3389/fpsyt.2021.729930

31. Xie X, Liu Q, Zhu K, Jiang Q, Feng Y, Xiao P, et al. Psychological symptom progression in school-aged children after COVID-19 home confinement: a longitudinal study. *Front Psychiatry*. (2022) 13:809107. doi: 10.3389/fpsyt.2022.809107

32. Gadassi Polack R, Sened H, Aubé S, Zhang A, Joormann J, Kober H. Connections during crisis: adolescents' social dynamics and mental health during COVID-19. *Dev Psychol*. (2021) 57:1633–47. doi: 10.1037/dev0001211

33. Kostev K, von Vultée C, Weber K, Bohlken J. Zunahme von Depressions- und Angststörungsdiagnosen während der Covid-19-Pandemie bei Kindern und Jugendlichen. *Monit Versorg*. (2021) 14:57–61. doi: 10.24945/MVF.04.21.1866-0533.2334

34. Magson NR, Freeman JYA, Rapee RM, Richardson CE, Oar EL, Fardouly J. Risk and protective factors for prospective changes in adolescent mental health during the COVID-19 pandemic. *J Youth Adolesc*. (2021) 50:44–57. doi: 10.1007/s10964-020-01332-9

35. Naumann E, von den Driesch E, Schumann A, Thönnissen C. Anstieg depressiver Symptome bei Jugendlichen und jungen Erwachsenen während des ersten Lockdowns in Deutschland: Ergebnisse des Beziehungs- und Familienpanels pairfam. Increase of depressive symptoms among adolescents during the first COVID-19 lockdown in Germany: Results from the German family panel pairfam. *Bundesgesundheitsbl Gesundheitsforsch Gesundheitsschutz*. (2021) 64:1533–40. doi: 10.1007/s00103-021-03451-5

36. Cooper K, Hards E, Moltrecht B, Reynolds S, Shum A, McElroy E, et al. Loneliness, social relationships, and mental health in adolescents during the COVID-19 pandemic. *J Affect Disord*. (2021) 289:98–104. doi: 10.1016/j.jad.2021.04.016

37. Lewis J, Jayakumar S, Breaux R, Dvorsky MR, Langberg JM, Becker SP. Prospective examination of psychological trauma among adolescents during the COVID-19 pandemic. *Psychol Trauma Theory Res Pract Policy*. (2022) 15:404–14. doi: 10.1037/tra0001311

38. Kaman A, Otto C, Klasen F, Westenhöfer J, Reiss F, Hölling H, et al. Risk and resource factors for depressive symptoms during adolescence and emerging adulthood – a 5-year follow-up using population-based data of the BELLA study. *J Affect Disord*. (2021) 280:258–66. doi: 10.1016/j.jad.2020.11.036

39. Bundesministerium für Gesundheit, Robert Koch Institut. Wie ist der Fortschritt der COVID-19-Impfung? Aktueller Impfstatus (2022). Available at: https://impfdashboard.de

40. Koch-Institut Robert. *Dritte Aktualisierung der Retrospektiven Phaseneinteilung der COVID-19-Pandemie in Deutschland* (2022). Available at: https://edoc.rki.de/handle/176904/10260 (Accessed November 25, 2022).

41. Moor I, Winter K, Bilz L, Bucksch J, Finne E, John N, et al. The 2017/18 Health Behaviour in School-aged Children (HBSC) study – methodology of the World Health Organization's child and adolescent health study. *J Health Monit*. (2020) 5:88–102. doi: 10.25646/6904

42. Ravens-Sieberer UThe European KIDSCREEN Group. *The European KIDSCREEN questionnaires - quality of life questionnaires for children and adolescent - handbook*. Lengerich: Pabst Science Publisher (2006).

43. Ravens-Sieberer U, Erhart M, Rajmil L, Herdman M, Auquier P, Bruil J, et al. Reliability, construct and criterion validity of the KIDSCREEN-10 score: a short measure for children and adolescents' well-being and health-related quality of life. *Qual Life Res*. (2010) 19:1487–500. doi: 10.1007/s11136-010-9706-5

44. Ravens-Sieberer U, Auquier P, Erhart M, Gosch A, Rajmil L, Bruil J, et al. The KIDSCREEN-27 quality of life measure for children and adolescents: psychometric results from a cross-cultural survey in 13 European countries. *Qual Life Res*. (2007) 16:1347–56. doi: 10.1007/s11136-007-9240-2

45. Ravens-Sieberer U, Gosch A, Rajmil L, Erhart M, Bruil J, Power M, et al. The KIDSCREEN-52 quality of life measure for children and adolescents: psychometric results from a cross-cultural survey in 13 European countries. *Value Health*. (2008) 11:645–58. doi: 10.1111/j.1524-4733.2007.00291.x

46. Goodman R. The strengths and difficulties questionnaire: a research note. *J Child Psychol Psychiatry*. (1997) 38:581–6. doi: 10.1111/j.1469-7610.1997.tb01545.x

47. Woerner W, Becker A, Rothenberger A. Normative data and scale properties of the German parent SDQ. *Eur Child Adolesc Psychiatry*. (2004) 13:II3–10. doi: 10.1007/s00787-004-2002-6

48. Birmaher B, Brent DA, Chiappetta L, Bridge J, Monga S, Baugher M. Psychometric properties of the screen for child anxiety related emotional disorders (SCARED): a replication study. *J Am Acad Child Adolesc Psychiatry*. (1999) 38:1230–6. doi: 10.1097/00004583-199910000-00011

49. Barkmann C, Schulte-Markwort M. Prevalence of emotional and behavioural disorders in German children and adolescents: a meta-analysis. *J Epidemiol Community Health*. (2012) 66:194–203. doi: 10.1136/jech.2009.102467

50. Hautzinger M, Bailer M, Hofmeister D, Keller F. *Allgemeine Depressionsskala (ADS)*. Göttingen: Beltz (2012).

51. Kroenke K, Spitzer RL, Williams JBW. The patient health questionnaire-2: validity of a two-item depression screener. *Med Care*. (2003) 41:1284–92. doi: 10.1097/01.MLR.0000093487.78664.3C

52. Löwe B, Spitzer R, Zipfel S, Herzog W. *Gesundheitsfragebogen für Patienten (PHQ-D). Manual-Komplettversion und Kurzform. Autorisierte deutsche Version des Prime MD Patient Health Questionnaire (PHQ)*. Heidelberg: Pfizer (2002).

53. Fendrich M, Weissman MM, Warner V. Screening for depressive disorder in children and adolescents: validating the Center for Epidemiologic Studies Depression Scale for children. *Am J Epidemiol*. (1990) 131:538–51. doi: 10.1093/oxfordjournals.aje.a115529

54. Haugland S, Wold B. Subjective health complaints in adolescence—reliability and validity of survey methods. *J Adolesc*. (2001) 24:611–24. doi: 10.1006/jado.2000.0393

55. Voltmer K, Von Salisch M. Epidemiebezogene Dark Future Scale für Kinder (eDFS-K) am Beispiel von COVID-19. *Zusammenstellung Sozialwissenschaftlicher Items Skalen ZIS*. (2021). doi: 10.6102/ZIS307

56. Zaleski Z. Future anxiety: concept, measurement, and preliminary research. *Personal Individ Differ*. (1996) 21:165–74. doi: 10.1016/0191-8869(96)00070-0

57. Bettge S, Ravens-Sieberer U. Schutzfaktoren für die psychische Gesundheit von Kindern und Jugendlichen - empirische Ergebnisse zur Validierung eines Konzepts. *Gesundheitswesen*. (2003) 65:167–72. doi: 10.1016/0191-8869(96)00070-0

58. Moser JS, Moran TP, Schroder HS, Donnellan MB, Yeung N. On the relationship between anxiety and error monitoring: a meta-analysis and conceptual framework. *Front Hum Neurosci*. (2013) 7:466. doi: 10.3389/fnhum.2013.00466

59. Sherbourne CD, Stewart AL. The MOS social support survey. *Soc Sci Med*. (1991) 32:705–14. doi: 10.1016/0277-9536(91)90150-B

60. Schneewind KA, Beckmann M, Hecht-Jackl A. Das FK-Testsystem. Das Familienklima aus der Sichtweise der Eltern und der Kinder. Forschungsberichte aus dem Institutsbereich Persönlichkeitspsychologie und Psychodiagnostik (1985).

61. Cimino S, Di Vito P, Cerniglia L. The impact of COVID-19 pandemic on psychopathological symptoms in mothers and their school-age children before, during and after the COVID-19 pandemic peak. *Curr Psychol*. (2022) 41:1–10. doi: 10.1007/s12144-022-03360-z

62. Schnetzer S, Hurrelmann K. Jugend in Deutschland – Trendstudie Winter 2022/23. Die Wohlstandsjahre sind vorbei: Psyche, Finanzen, Verzicht. Datajockey Verl. (2022).

63. Schnetzer S, Hurrelmann K. Trendstudie: Jugend in Deutschland im Dauerkrisenmodus - Klima, Krieg. Datajockey Verl. (2022).

64. Zok K, Roick C. Auswirkungen der COVID-19-Pandemie auf die psychiche Gesundheit von Kindern (2022) 19:1–12. Available at: https://www.wido.de/fileadmin/Dateien/Dokumente/Publikationen_Produkte/WIdOmonitor/wido_monitor_1_2022_pandemiebelastung_kinder.pdf

65. Witte J, Zeitler A, Batram M, Diekmannshemke J, Hasemann L. Kinder- und Jugendreport 2022. Kinder- und Jugendgesundheit in Zeiten der Pandemie. DAK Gesundheit. (2022). Available at: https://www.dak.de/dak/download/dak-kjr22-vand-report-pdf-2572514.pdf

66. Wenter A, Schickl M, Sevecke K, Juen B, Exenberger S. Children's mental health during the first two years of the COVID-19 pandemic: burden, risk factors and posttraumatic growth – a mixed-methods parents' perspective. *Front Psychol*. (2022) 13:901205. doi: 10.3389/fpsyg.2022.901205

67. Zuccolo PF, Casella CB, Fatori D, Shephard E, Sugaya L, Gurgel W, et al. Children and adolescents' emotional problems during the COVID-19 pandemic in Brazil. *Eur Child Adolesc Psychiatry*. (2022) 32:1–13. doi: 10.1007/s00787-022-02006-6

68. Bantel S, Buitkamp M, Wünsch A. Kindergesundheit in der COVID-19-Pandemie: Ergebnisse aus den Schuleingangsuntersuchungen und einer Elternbefragung in der Region Hannover. Child health in the COVID-19 pandemic: results from school entry data and a parent survey in the Hanover region. *Bundesgesundheitsbl Gesundheitsforsch Gesundheitsschutz*. (2021) 64:1541–50. doi: 10.1007/s00103-021-03446-2

69. Hanewinkel R, Hansen J, Neumann C, Petersen FL. *Kinder- und Jugendgesundheit in Schulen. Ergebnisbericht 2020/2021*. Kiel: Institut für Therapie- und Gesundheitsforschung (2021).

70. Ethikrat Deutscher. Ethikrat: Kinder, Jugendliche und junge Erwachsene in gesellschaftlichen Krisen nicht alleinlassen [Internet]. Deutscher Ethikrat (2022). Available at: https://www.ethikrat.org/mitteilungen/mitteilungen/2022/ethikrat-kinder-jugendliche-und-junge-erwachsene-in-gesellschaftlichen-krisen-nicht-alleinlassen/ (Accessed November 15, 2022).

71. Center for Disease Control and Prevention. New reports on health and well-bering of children during COVID-19 pandemic. CDC Newsroom Releases. (2022). Available at: https://www.cdc.gov/media/releases/2022/s0218-children-wellbeing-covid-19.html (Accessed December 15, 2022).

72. Plötner M, Moldt K, In-Albon T, Schmitz J. Einfluss der COVID-19-Pandemie auf die ambulante psychotherapeutische Versorgung von Kindern und Jugendlichen. *Psychother*. (2022) 67:469–77. doi: 10.1007/s00278-022-00604-y

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 20 June 2023
DOI 10.3389/fpubh.2023.1117854

OPEN ACCESS

EDITED BY
Muhammad Shahid Iqbal,
Prince Sattam Bin Abdulaziz University,
Saudi Arabia

REVIEWED BY
Omar Irfan,
Amaris, United Kingdom
Vivek Peche,
Washington University in St. Louis, United States
Umesh Ghimire,
University of Minnesota Twin Cities,
United States

*CORRESPONDENCE
Fernanda Martins
✉ fernanda.martins@hc.fm.usp.br

RECEIVED 14 February 2023
ACCEPTED 16 May 2023
PUBLISHED 20 June 2023

CITATION
Martins F, Gonçalves FT, Imamura M, Barboza DS, Matheus D, Pereira MFB, Marques HHS, Correa-Silva S, Montenegro MM, Fink TT, Lindoso L, Bain V, Ferreira JCOA, Astley C, Matsuo OM, Suguita P, Trindade V, Paula CSY, Litvinov N, Palmeira P, Gualano B, Delgado AF, Carneiro-Sampaio M, Forsait S, Odone-Filho V, Antonangelo L, Battistella LR and Silva CA (2023) Health-related quality of life and functionality in primary caregiver of surviving pediatric COVID-19.
Front. Public Health 11:1117854.
doi: 10.3389/fpubh.2023.1117854

Health-related quality of life and functionality in primary caregiver of surviving pediatric COVID-19

Fernanda Martins[1,2]*, Fernanda T. Gonçalves[3], Marta Imamura[2,4], Daniela S. Barboza[2], Denise Matheus[2], Maria Fernanda B. Pereira[1], Heloisa H. S. Marques[1], Simone Correa-Silva[5], Marilia M. Montenegro[5], Thais T. Fink[1], Livia Lindoso[1], Vera Bain[1], Juliana C. O. A. Ferreira[1], Camilla Astley[6], Olivia M. Matsuo[1], Priscila Suguita[1], Vitor Trindade[1], Camila S. Y. Paula[1], Nadia Litvinov[1], Patricia Palmeira[5], Bruno Gualano[6], Artur F. Delgado[1], Magda Carneiro-Sampaio[1], Silvana Forsait[1], Vicente Odone-Filho[1], Leila Antonangelo[1], Linamara R. Battistella[2,4] and Clovis A. Silva[1]

[1]Faculdade de Medicina, Instituto da Criança e do Adolescente, Hospital das Clínicas HCFMUSP, Universidade de São Paulo, São Paulo, SP, Brazil, [2]Instituto de Medicina Física e Reabilitação, Hospital das Clínicas HCFMUSP, São Paulo, SP, Brazil, [3]Laboratorio de Imunohematologia e Hematologia Forense (LIM-40), Departamento de Medicina Legal, Bioética, Medicina do Trabalho e Medicina Física e Reabilitação, Faculdade de Medicina, Hospital das Clínicas HCFMUSP, Universidade de São Paulo, São Paulo, SP, Brazil, [4]Departamento de Medicina Legal, Bioética, Medicina do Trabalho e Medicina Física e Reabilitação, Faculdade de Medicina FMUSP, Universidade de São Paulo, São Paulo, SP, Brazil, [5]Laboratório de Pediatria Clínica (LIM-36), Departamento de Pediatria, Faculdade de Medicina, Hospital das Clínicas HCFMUSP, Universidade de São Paulo, São Paulo, SP, Brazil, [6]Faculdade de Medicina, Divisão de Reumatologia, Hospital das Clínicas, Universidade de São Paulo (HCFMUSP), São Paulo, SP, Brazil

Objectives: To prospectively assess health-related quality of life (HRQoL), global functionality, and disability in primary caregivers of surviving children and adolescents after COVID-19.

Methods: A longitudinal observational study was carried out on primary caregivers of surviving pediatric post-COVID-19 patients (n=51) and subjects without COVID-19 (n=60). EuroQol five-dimension five-level questionnaire (EQ-5D-5L) and 12-question WHO Disability Assessment Schedule 2.0 (WHODAS 2.0) were answered for both groups. The univariate regression analysis was carried out using SPSS (v 20) and significance was established at 5%.

Results: The median duration between COVID-19 diagnosis in children and adolescents and longitudinal follow-up visits was 4.4months (0.8–10.7). The median age of children and adolescents caregivers with laboratory-confirmed COVID-19 was similar to primary caregivers of subjects without laboratory-confirmed COVID-19 [43.2 (31.6–60.9) vs. 41.5 (21.6–54.8) years, p=0.08], as well as similar female sex (p=1.00), level of schooling (p=0.11), social assistance program (p=0.28), family income/monthU$ (p=0.25) and the number of household's members in the residence (p=0.68). The frequency of slight to extreme problems (level≥2) of the pain/discomfort domain according to EQ-5D-5L score was significantly higher in the former group [74% vs. 52.5%, p=0.03, OR=2.57 (1.14–5.96)]. The frequency of disability according to WHODAS 2.0 total score was similar to those without disability and unknown (p=0.79); however, with a very high disability in both groups (72.5% and 78.3%). Further analysis of

primary caregivers of children and adolescents with post-COVID-19 condition (PCC) [*n*=12/51 (23%)] compared to those without PCC [*n*=39/51(77%)] revealed no differences between demographic data, EQ-5D-5L and WHODAS 2.0 scores in both groups ($p>0.05$).

Conclusion: We longitudinally demonstrated that pain/discomfort were predominantly reported in approximately 75% of primary caregiver of COVID-19 patients, with high disability in approximately three-quarters of both caregiver groups. These data emphasized the prospective and systematic caregiver burden evaluation relevance of pediatric COVID-19.

KEYWORDS

pediatric, children, adolescent, COVID-19, caregiver burden, quality of life, chronic disease, public health

Introduction

Pediatric-coronavirus disease 2019 (COVID-19) survivors may present persistent and disabling conditions after severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), named post-acute COVID-19 condition (PCC) (1–4).

PCC or "long COVID-19" manifestations is generally defined when any symptoms continued after 12 weeks of acute SARS-CoV-2 infection onset and cannot be justified by other conditions (5–8). The prevalence of PCC in pediatric populations ranges from 0.31% to 30% (1, 9, 10) and studies rarely included populations of children and adolescents with chronic preexisting conditions (3, 5).

In this regard, we recently showed that children and adolescents with SARS-CoV-2 infection patients had a prospective impact on health-related quality of life (HRQoL) parameters, specifically in physical and school domains and 23% of them had PCC (3). World Health Organization (WHO) suggest achieving family caregiver health outcomes as post-COVID condition research priorities (11). Only few results appoint a high caregiver burden during the acute phase of children and adolescents with COVID-19, evaluating HRQoL, anxiety, and depression tools, in Iran and Italy cohorts (12, 13).

It has been established that have a child or adolescent with any chronic health condition disrupt the family life in many ways mainly the psychological domain. However, this fact is not assessed and frequently receive little to no assistance or guidance during the medical treatment of the children or adolescent (14, 15).

To our knowledge, there is no report assessing concomitant HRQoL, global functionality, and disability in primary caregivers of children and adolescents with laboratory-confirmed COVID-19 survivors, as well as in PCC subjects. These issues are particularly important in a pediatric population with predominant chronic conditions in a tertiary healthcare facility. To improve care coordination of pediatric population we need to analyze the caregiver burden by mapping their health status and eventually indicating dedicated action plans for caregivers at local services or public policies.

Therefore, the present study aimed to assess longitudinal HRQoL, global functionality, and disability in primary caregivers of children and adolescents with laboratory-confirmed COVID-19 survivors and primary caregivers of children and adolescents without laboratory-confirmed COVID-19 (control group). These parameters were also compared between primary caregivers of COVID-19 patients with and without PCC.

Methods

A longitudinal observational study was carried out on the primary caregiver of surviving pediatric post-COVID-19 patients from Instituto da Crianca e do Adolescente do Hospital das Clinicas da Faculdade de Medicina da Universidade de São Paulo (ICr HCFMUSP), São Paulo, Brazil which was one of the main COVID-19 reference hospitals in the biggest Latin American city. This study included a convenience sample from inpatient and outpatient treated in the hospital since the COVID-19 pandemic beginning. The inclusion criteria were as follows: symptomatic inpatients and outpatients, laboratory-confirmed SARS-CoV-2 infection, and age between <18 years. The exclusion criteria were subjects with asymptomatic disease and those who did not complete the forms as previously reported (3). EuroQol five-dimension five-level questionnaire (EQ-5D-5L) and 12-question World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) instrument were answered by $n=51$ consecutive primary caregivers of surviving pediatric post-COVID-19 patients. The control group of primary caregivers was enrolled at the outpatient clinic of the same tertiary hospital and included $n=60$ primary caregivers of pediatric subjects without laboratory-confirmed COVID-19. Additionally, we point that control group recruitment were paired taking into consideration children or adolescent age, sex and type of chronic health related condition. Data from the present study were collected between October 2020 to October 2021. The primary caregiver was defined by the person directly responsible for all the daily care of children and adolescents.

SARS-CoV-2 infection was measured by real-time reverse transcription-polymerase chain reaction (real-time RT-PCR) and antibody analysis. Real-time RT-PCR to assess SARS-CoV-2 RNA was performed on swab-collected nasopharyngeal and/or oropharyngeal samples in the Molecular Biology Laboratory of our tertiary health-care facility (16). Antibodies against S proteins from the coronavirus spike were carried out by two different methods: an enzyme-linked

immunosorbent assay for anti-SARS-CoV-2 IgG antibodies and a rapid immunochromatographic assay for anti-SARS-CoV-2 IgM and IgG antibodies (17, 18).

Data for all subjects were prospectively inserted in Research Electronic Data Capture (REDCap) database of information at our university hospital (HCFMUSP). This report was approved by the Ethical Committee of our tertiary hospital (CAEE 4.889.659) and the written informed consent was obtained from all the participants and their parents/guardians.

HRQoL, global functionality and disability instruments in primary caregivers

HRQoL was analyzed by EQ-5D-5L Brazilian Portuguese language (19). This self-evaluation tool is a generic, standardized measure of perceived health status. This tool included current information on the day visit and a detailed analysis of five dimensions: mobility, self-care (hygiene and dressing), usual activities (work, study, housework, family, and leisure activities), pain/discomfort, and anxiety/depression. Each dimension comprises a five-level version of the answer: no (level 1), slight (level 2), moderate (level 3), severe (level 4), and extreme problems (level 5).

WHODAS 2.0 global evaluates global functionality and disability in 12-question, covering six domains of functioning in the last 30 days: cognition, mobility, self-care, getting along, life activities, and participation. Each question was recorded from 1 (no difficulty) to 5 (extreme difficulty or cannot do). The sum of the total score ranges from 0 to 48. The results were presented as with disability, without disability, and unknown and categorized as mild (1–4), moderate (5–9), and severe disability (10–48). This tool was validated in the Brazilian Portuguese language (20).

Demographic, anthropometric, and preexisting health condition data in children and adolescents

Demographic data in laboratory-confirmed COVID-19 children and adolescents comprised: current age, sex, and skin color (21). Body mass index was defined by body weight divided by the square of the body height and presented in units of kg/m^2. Skin color was self-reported in a specific form – the options were white, black/brown and yellow. Preexisting pediatric chronic conditions in children and adolescents were categorized according to the duration of signs and symptoms more than 3 months, and diagnosis by physician scientific knowledge, valid methods, or instruments according to specific pediatric diagnostic criteria (22, 23). PASC or long COVID-19 manifestations were defined when clinical abnormalities continue after 12 weeks of the onset of acute COVID-19 diagnosis and cannot be justified by other acute or chronic conditions (3).

Statistical analysis

Statistical analyses were carried out using SPSS software, version 20 (IBM Corporation, Armonk, NY, United States). A non-parametric test (Mann–Whitney test) or parametric test (Student's *t*-test) was used for continuous variables and presented by median (minimum and maximum values) or mean ± standard deviation with an odds ratio (OR), respectively. Fisher's exact or chi-square tests were used for categorical variables. The level of significance was established at 5%.

Results

The median duration between COVID-19 diagnosis in children and adolescents and longitudinal follow-up visits was 4.4 months (0.8–10.7). PASC was observed in 12/51 (23%) of children and adolescents. The most frequently reported symptoms at longitudinal follow-up visits among children and adolescents with long-term PASC were headache in 6/12 (50%), fatigue in 6/12 (50%), dyspnea in 4/12 (30%), anxiety in 2/12 (16%), chest pain in 2/12 (16%), arthralgia in 1/12 (8%), and memory loss in 1/12 (8%).

The median of the current age of the primary caregivers' group of children and adolescents with laboratory-confirmed COVID-19 was similar to primary caregivers of subjects without laboratory-confirmed COVID-19 [43.2 (31.6–60.9) vs. 41.5 (21.6–54.8) years, $p = 0.08$], as well as similar female sex ($p = 1.00$), level of schooling ($p = 0.11$), social assistance program ($p = 0.28$), family income/month U$ ($p = 0.25$) and the number of household's members in the residence ($p = 0.68$) in both groups (Table 1). The median of the current age of children and adolescents in the group of primary caregivers of COVID-19 patients was significantly higher compared to the primary caregiver of subjects without laboratory-confirmed COVID-19 [14.6 (8.1–18.2) vs. 12.7 (2.1–18.9) years, $p = 0.008$], however with low OR = 1.16 (1.04–1.29). The other variables were similar in both groups: female sex ($p = 0.53$), ethnicity ($p = 0.82$), anthropometric data ($p = 0.13$), and preexisting pediatric condition ($p = 0.61$; Table 1).

Table 2 illustrates EQ-5D-5L and WHODAS 2.0 scores in primary caregivers of children and adolescents with laboratory-confirmed COVID-19 compared to primary caregivers of children and adolescents without laboratory-confirmed COVID-19 (control group). The frequency of slight to extreme problems (level ≥ 2) of pain/discomfort domain by EQ-5D-5L score was significantly higher in the primary caregivers of children and adolescents with laboratory-confirmed COVID-19 compared to the other group [74% vs. 52.5%, $p = 0.03$, OR = 2.57 (1.14–5.96)]. No differences were evidenced in the other domains of EQ-5D-5L (mobility, self-care, usual activity, anxiety/depression). The frequency of disability according to WHODAS 2.0 total score was similar to those without disability and unknown ($p = 0.79$); however, with a very high disability in both groups (72.5% and 78.3%). No differences were shown in other WHODAS 2.0 score parameters in both groups ($p > 0.05$). We also performed EQ-5D-5L and WHODAS 2.0 scores evaluations from healthy subjects vs. those with chronic health condition but no difference between these two groups was founded $p > 0.05$ (data not showed).

Further analysis of primary caregivers of children and adolescents with PCC [12/51 (23%)] compared to those without PCC [39/51 (77%)] revealed no differences between demographic data, EQ-5D-5L and WHODAS 2.0 scores in both groups ($p > 0.05$, Table 3).

10.3389/fpubh.2023.1117854

TABLE 1 Primary caregivers of children and adolescents with laboratory-confirmed coronavirus disease 2019 (COVID-19) group compared to primary caregivers of children and adolescents without laboratory-confirmed COVID-19 (control group).

Variables	Primary caregivers of COVID-19 patients (*n*=51)	Primary caregivers of control group (*n*=60)	*p*	OR (CI 95%)
Primary caregiver				
Demographic				
Current age, years	43.2 (31.6–60.9)	41.5 (21.6–54.8)	0.08	1.04 (0.99–1.10)
Female sex	51 (100)	60 (100)	1.00	-
Level of schooling (*n* = 49)				
University	10 (19.6)	20 (33.4)	0.11	1.25 (0.95–1.64)
High school	24 (47.1)	30 (50)		
Middle school	8 (15.7)	4 (6.65)		
Elementary school	6 (11.8)	4 (6.65)		
Unknown	3 (5.9)	2 (3.3)		
Social assistance program	14 (43.8)	17 (32.1)	0.28	1.64 (0.66–4.07)
Family income/month U$	456 (0–2,283)	685 (0–4,035)	0.25	0.91 (0.78–1.07)
Number of household's members in the residence	4 (2–6)	4 (2–8)	0.68	0.93 (0.69–1.27)
Children and adolescents				
Demographic				
Current age, years	14.6 (8.1–18.2)	12.7 (2.1–18.9)	0.008	1.16 (1.04–1.29)
Female sex	31 (61.8)	31 (52.7)	0.53	1.17 (0.71–1.93)
Skin color				
White	27 (52.9)	33 (55)	0.82	1.05 (0.69–1.60)
Black and brown	1 (41.2)	25 (41.7)		
Yellow	2 (3.9)	2 (3.3)		
Unknown	1 (2.0)	0 (0)		
Anthropometric data				
Body mass index, kg/m^2	20.2 (13.7–39.5)	18.9 (13–32.5)	0.13	1.06 (0.98–1.14)
Pediatric preexisting condition			0.61	1.06 (0.85–1.32)
Immunologic	24 (47.1)	22 (36.7)		
Hepatic or renal	6 (11.8)	11 (18.3)		
Oncologic	6 (11.8)	6 (10.0)		
Other chronic condition	7 (13.7)	6 (10.0)		
Healthy	8 (15.7)	15 (25.0)		

Results are presented in *n* (%), median (minimum-maximum values) and mean ± standard deviation.

Discussion

We prospectively demonstrated that pain or discomfort was predominantly reported in 75% of female caregivers after pediatric COVID-19 diagnosis. We also observed a high disability impact in approximately three-quarters of both primary caregivers with and without COVID-19 and anxiety/depression in more than 50% of both groups.

During the COVID-19 pandemic, our pediatric tertiary healthcare facility installed a multidisciplinary and multispecialty outpatient clinic for longitudinal and simultaneous studies for children and adolescents COVID-19 survivors', such as HRQoL, cardiovascular and pulmonary assessments that were recently reported (3, 24, 25). The present study focused on primary caregivers of these children and adolescents, evaluating the caregiver burden in a tertiary health care facility from a low-income country.

The strong point of this observational study was the inclusion of primary caregivers of children and adolescents with laboratory-confirmed SARS-CoV-2 infection (by RT-PCR and sorology), with a predominance of pediatric chronic conditions subjects (22, 26). Another strength was the inclusion of a control group, including primary caregivers of children and adolescents without laboratory-confirmed SARS-CoV-2 infection at study entry, balanced by age, sex, level of schooling, social assistance program, family income/month, and the number of household's members in the residence, since these parameters are a well-known factor that impacts

10.3389/fpubh.2023.1117854

TABLE 2 EuroQol 5-Dimension 5-Level (EQ-5D-5L) and World Hath Organization Schedule (WHODAS) 2.0 scores in primary caregivers of children and adolescents with laboratory-confirmed coronavirus disease 2019 (COVID-19) compared to primary caregivers of children and adolescents without laboratory-confirmed COVID-19 (control group).

Variables of primary caregiver	Primary caregivers of COVID-19 patients (n=51)	Primary caregivers of control group (n=60)	p	OR (CI 95%)
EQ-5D-5L score (n =50, n =59)				
Mobility			0.21	1.48 (0.80–2.75)
Level 1	38 (76.0)	51 (86.4)		
Level 2	6 (12.0)	5 (8.5)		
Level 3	6 (12.0)	2 (3.4)		
Level 4	0 (0)	1 (1.7)		
Level 5	0 (0)	0 (0)		
Slight to extreme problems (level ≥ 2)	12 (24.0)	8 (13.6)	0.21	2.01 (0.74–5.49)
Self-care				
Level 1	46 (92.0)	54 (91.5)	0.83	0.94 (0.53–1.67)
Level 2	2 (4.0)	2 (3.4)		
Level 3	0 (0)	1 (1.7)		
Level 4	2 (4.0)	1 (1.7)		
Level 5	0 (0)	1 (1.7)		
Slight to extreme problems (level ≥ 2)	4 (8.0)	5 (8.5)	1.00	0.94 (0.27–3.41)
Usual activity			0.83	0.94 (0.53–1.67)
Level 1	44 (88.0)	52 (88.1)		
Level 2	5 (10.0)	5 (8.5)		
Level 3	1 (2.0)	1 (1.7)		
Level 4	0 (0)	1 (1.7)		
Level 5	0 (0)	0 (0)		
Slight to extreme problems (level ≥ 2)	6 (12.0)	7 (11.9)	1.00	1.01 (0.30–3.02)
Pain/discomfort			0.12	1.37 (0.92–2.05)
Level 1	13 (26.0)	28 (47.5)		
Level 2	22 (44.0)	17 (28.8)		
Level 3	10 (20.0)	10 (16.9)		
Level 4	5 (10.0)	3 (5.1)		
Level 5	0 (0)	1 (1.7)		
Slight to extreme problems (level ≥ 2)	37 (74.0)	31 (52.5)	0.03	2.57 (1.14–5.96)
Anxiety/depression			0.31	1.21 (0.84–1.74)
Level 1	22 (44.0)	28 (47.5)		
Level 2	14 (28.0)	18 (30.5)		
Level 3	7 (14.0)	9 (15.3)		
Level 4	5 (10.0)	4 (6.8)		
Level 5	2 (4.0)	0 (0)		
Slight to extreme problems (level ≥ 2)	28 (56.0)	29 (52.5)	0.69	1.23 (0.58–2.65)
WHODAS 2.0 12 item total score (0–48) (n = 51, n = 60)	3 (0–20)	2 (0–21)	0.75	1.01 (0.94–1.07)
With disability	37 (72.5)	47 (78.3)	0.79	1.14 (0.43–2.94)
Without disability	9 (17.6)	13 (21.7)		
Unknown	5 (9.8)	0 (0)		
With disability (1–48)			0.53	0.85 (0.51–1.41)
Mild (1–4)	21 (56.8)	25 (41.7)		
Moderate (5–9)	8 (21.6)	8 (13.3)		
Severe (10–48)	8 (21.6)	14 (23.3)		

Results are presented in n (%), median (minimum-maximum values) and mean ± standard deviation.

TABLE 3 Demographic, EuroQol 5-Dimension 5-Level (EQ-5D-5L) and World Hath Organization Schedule (WHODAS) 2.0 scores in primary caregivers of children and adolescents with PCC compared to without PCC.

Variables of primary caregiver	Primary caregivers of COVID-19 patients with PCC (*n*=12)	Primary caregivers of COVID-19 patients without PCC (*n*=39)	*p*	OR (CI 95%)
Demographic				
Mother current age, years	40.22 (32.2–53.4)	44.9 (31.7–60.9)	0.17	0.94 (0.85–1.03)
Mother's level of schooling			0.60	0.87 (0.56–1.40)
University	4 (33.3)	6 (15.4)		
High school	2 (16.7)	22 (56.4)		
Middle school	4 (33.3)	4 (10.3)		
Elementary school	2 (16.7)	4 (10.3)		
Unknown	0 (0)	3 (7.7)		
Social assistance program	3 (37.5)	11 (45.8)	0.71	0.71 (0.13–3.66)
Family minimum wages/month	2.2 (1–10)	2 (0–10)	0.24	1.18 (0.89–1.56)
Number of household's members in the residence	4 (3–6)	4 (2–6)	0.06	1.79 (0.98–3.30)
EQ-5D-5L score				
Mobility			0.19	0.30 (0.05–1.82)
Level 1	10 (90.9)	28 (71.8)		
Level 2	1 (9.1)	5 (12.8)		
Level 3	0 (0)	6 (15.4)		
Level 4	0 (0)	0 (0)		
Level 5	0 (0)	0 (0)		
Slight to extreme problems (level ≥ 2)	1 (9.1)	11 (28.2)	0.25	0.25 (0.02–1.65)
Self-care			1.00	---
Level 1	11 (100)	35 (89.8)		
Level 2	0 (0)	2 (5.1)		
Level 3	0 (0)	2 (5.1)		
Level 4	0 (0)	0 (0)		
Level 5	0 (0)	0 (0)		
Slight to extreme problems (level ≥ 2)	0 (0)	4 (10.2)	0.56	---
Usual activity			0.65	0.63 (0.08–4.66)
Level 1	10 (90.9)	34 (87.2)		
Level 2	1 (9.1)	4 (10.3)		
Level 3	0	1 (2.5)		
Level 4	0	0		
Slight to extreme problems (level ≥ 2)	1 (9.1)	5 (12.8)	1.00	0.68 (0.05–18.9)
Pain/discomfort			0.35	0.69 (0.31–1.51)
Level 1	4 (36.4)	9 (23.1)		
Level 2	4 (36.4)	18 (46.2)		
Level 3	3 (27.3)	7 (17.9)		
Level 4	0 (0)	5 (12.8)		
Level 5	0 (0)	0 (0)		
Slight to extreme problems (level ≥ 2)	7 (63.6)	30 (76.9)	0.44	0.52 (0.14–1.92)
Anxiety/depression			0.22	0.64 (0.32–1.31)
Level 1	7 (63.6)	15 (38.5)		
Level 2	2 (18.2)	12 (30.8)		

(Continued)

TABLE 3 (Continued)

Variables of primary caregiver	Primary caregivers of COVID-19 patients with PCC (*n*=12)	Primary caregivers of COVID-19 patients without PCC (*n*=39)	*p*	OR (CI 95%)
Level 3	1 (9.1)	6 (15.4)		
Level 4	1 (9.1)	4 (10.3)		
Level 5	0	2 (5.1)		
Slight to extreme problems (level ≥ 2)	4 (36.4)	24 (61.5)	0.17	0.36 (0.10–1.51)
WHODAS 2.0 12 item score (0–48)	4 (1–19)	4 (0–21)	0.88	0.99 (0.86–1.13)
With disability	8 (66.7)	29 (74.4)	0.48	2.21 (0.24–20.35)
Without disability	1 (8.3)	8 (20.5)		
Unknown	3 (25.0)	2 (5.1)		
With disability (1–48)			0.93	0.96 (0.36–2.52)
Mild (1–4)	4 (50.0)	17 (58.6)		
Moderate (5–9)	3 (37.5)	5 (17.2)		
Severe (10–48)	1 (12.5)	7 (24.1)		

Results are presented in *n* (%), median (minimum-maximum values) and mean ± standard deviation, PCC, post-COVID-19 condition.

HRQoL in adults (27). Another advantage of the present study was the simultaneous analysis of HRQoL, global functionality, and disability in primary caregivers. Indeed, the same self-reported tools have been used during acute and chronic adult COVID-19 studies (28, 29). However, we did not collect real-time RT-PCR or antibody tests to investigate previous SARS-CoV-2 infections in caregivers.

We extended the previous study and showed that pain or discomfort occurred in most young adults' caregivers of children and adolescents COVID-19 survivors. These findings suggest that these issues might be related to SARS-CoV-2 infection. The marked persistent pain and/or discomfort symptoms in caregivers of children and adolescents with COVID-19 can be subsequent to musculoskeletal involvement and possibly related to muscle weakness, fatigue, chronic pain, and/or arthralgia (30, 31). We hypothesized that those young female caregivers were affected by SARS-CoV-2 and had persistent symptoms probably concomitant with our pediatric patients, as reported by other studies (30, 32).

Of note, disability was observed in more than three-quarters of both primary caregivers from children and adolescents with or without COVID-19. In addition, more than 40% of both our groups with young age had moderate to severe disabilities according to this instrument. Reinforcing the relevant caregiver burden of the present study, normative populational data of total WHODAS 2.0 12 item score showed that the majority of subjects with similar age of the present study (bracket age from 35 to 54 years old) had a mild disability (33).

In addition, anxiety and depression, according to the EQ-5D-5L score, occurred in more than 50% of both groups. This point seems unrelated to SARS-CoV-2 infection and may be explained by the COVID-19 pandemic that increased psychopathology in young people who spent hours per day assessing information about the outbreak (34). The mother's fear of new SARS-CoV-2 infection, underlying disease activity, or complication related to immunosuppressive drugs may also contribute to this finding (27, 35). Furthermore, none of our children and adolescents of the present study were vaccinated at the time of this study which may also have contributed to these mental health abnormalities during this pandemic time.

PCC was not related to any impact on HRQoL, global disability, and functionality in our study. The small sample of PASC observed herein may have contributed to this result.

In conclusion, we longitudinally demonstrated that pain/discomfort were predominantly reported in approximately 75% of female caregivers of COVID-19 patients, with high disability frequency in approximately three-quarters of both caregiver groups. These data emphasized the prospective and systematic caregiver burden evaluation relevance of pediatric COVID-19 suggesting that is important to analyze the caregiver burden by mapping their health status during health care management of pediatric population. Create evidences about caregivers is crucial if we intend to implement a full and integrated lifelong care for chronic health conditions in a pediatric population. We intend to contribute to initial highlights about caregiver consequences in COVID-19 pediatric population mainly for PCC.

Study limitations

The present observational report has some limitations. We did not include clinical and laboratorial data from primary caregivers since we designed the study to assess several outcomes in children and adolescents after COVID-19 diagnosis. This study also included a cohort from only one healthcare pediatric facility at the largest hospital in Latin America and an analysis of the first visit after the median of 4 months of acute COVID-19 diagnosis. Therefore, long-term caregiver burden studies will be required, including additional visit evaluations at 12 months.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by the Ethical Committee of Hospital das Clínicas HCFMUSP, Faculdade de Medicina, Universidade de São Paulo (CAAE 37460620.8.0000.0068). The patients/participants provided their written informed consent to participate in this study.

Author contributions

FM, FG, MI, DB, DM, MP, HM, SC-S, MM, TF, LL, VB, JF, CA, OM, PS, VT, CP, NL, PP, BG, AD, MC-S, SF, VO-F, LA, LB, and CS contributed substantially to the conception and design of the study and to the analysis and interpretation of data. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by grants from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq 304984/2020-5 to CS), Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP 2015/03756-4 to CS), Núcleo de Apoio à Pesquisa "Saúde da Criança e do Adolescente" da USP (NAP-CriAd) to CS, and HCFMUSP with funds donated by NUBANK under the #HCCOMVIDA scheme supported this study (to MM and SC-S).

Acknowledgments

The authors would like to thank HC-FMUSP Pediatric Post-COVID-19 Study Group: Adriana M. E. Sallum, Alberto José S. Duarte, Alexandra V. M. Brentani, Álvaro José S. Neto, Amanda Ihara, Andrea R. Santos, Ana Claudia Latronico, Ana Pinheiro M. Canton, Andreia Watanabe, Angélica C. dos Santos, Antonio C. Pastorino, Bernadette D. G. M. Franco, Bruna Caruzo, Carlos Roberto R. Carvalho, Carlos A. Buchpiguel, Carina Ceneviva, Carolina C. M. F. Martins, Danilo Prado, Deipara M. Abellan, Fabiana B. Benatti, Fabiana Smaria, Fernando D. Penteado, Gabriela S. F. de Castro, Gabriela N. Leal, Guilherme V. Polanczyk, Guilherme S. Gonçalves, Hamilton Roschel, Ilana R. Disi, Isabela G. Marques, Isabela de J. S. Avelar, Inar A. Castro, Izabel M. Buscatti, Jaline Z. Faiad, Jarlei Fiamoncini, Joaquim C. Rodrigues, Jorge D. A. Carneiro, Jose A. Paz, Juliana C. Ferreira, Katia R. Silva, Karina L. M. Bastos, Katia Kozu, Lisa Suzuki, Lilian M. Cristofani, Lucas V. B. Souza, Lucia M. A. Campos, Luiz Vicente R. F. Silva Filho, Marcelo T. Sapienza, Marcos S. Lima, Marlene P. Garanito, Márcia F. A. Santos, Marilia Seelaender, Mayra B. Dorna, Nadia E. Aikawa, Neusa K. Sakita, Paula V. V. Gaiolla, Paula Pasqualucci, Ricardo K. Toma, Simone Correa-Silva, Sofia M. Sieczkowska, Sylvia C. L. Farhat, Silvana Forsait, Vera A. Santos, and Vicente Yingying Zheng.

Conflict of interest

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

Publisher's note

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

References

1. Merzon E, Weiss M, Krone B, Cohen S, Ilani G, Vinker S, et al. Clinical and socio-demographic variables associated with the diagnosis of long COVID syndrome in youth: a population-based study. *Int J Environ Res Public Health*. (2022) 19:5993. doi: 10.3390/ijerph19105993

2. Zimmermann P, Pittet LF, Curtis N. Long covid in children and adolescents. *BMJ*. (2022) 376:o143. doi: 10.1136/bmj.o143

3. Fink TT, Marques HHS, Gualano B, Lindoso L, Bain V, Astley C, et al. Persistent symptoms and decreased health-related quality of life after symptomatic pediatric COVID-19: a prospective study in a Latin American tertiary hospital. *Clinics*. (2021) 76:e3511. doi: 10.6061/clinics/2021/e3511

4. World Health Organization. A clinical case definition for post COVID-19 condition in children and adolescents by expert consensus. (2023);:1–99. Available at: https://www.who.int/publications/i/item/WHO-2019-nCoV-Post-COVID-19-condition-CA-Clinical-case-definition-2023-1

5. Kompaniyets L, Bull-otterson L, Boehmer TK, Baca S, Alvarez P, Hong K, et al. *Post-COVID-19 symptoms and conditions among children and adolescents—United States, March 1, 2020–January 31, 2022*, vol. *71* CDC Morbidity and Mortality Weekly Report (2022) 71:993–999. doi: 10.15585/mmwr.mm7131a3

6. Marques HHS, Pereira MFB, Santos ACD, Fink TT, Paula CSY, Litvinov N, et al. Differences in children and adolescents with SARS-CoV-2 infection: a cohort study in a Brazilian tertiary referral hospital. *Clinics*. (2021) 76:e3488. doi: 10.6061/clinics/2021/e3488

7. World Health Organization. Coronavirus disease (COVID-19): Post COVID-19 condition. (2021). Available at: https://www.who.int/news-room/questions-and-answers/item/coronavirus-disease-(covid-19)-post-covid-19-condition

8. Roessler M, Tesch F, Batram M, Jacob J, Loser F, Weidinger O, et al. Post-COVID-19-associated morbidity in children, adolescents, and adults: a matched cohort study including more than 157,000 individuals with COVID-19 in Germany. *PLoS Med*. (2022) 19:e1004122. doi: 10.1371/journal.pmed.1004122

9. Izquierdo-Pujol J, Moron-Lopez S, Dalmau J, Gonzalez-Aumatell A, Carreras-Abad C, Mendez M, et al. Post COVID-19 condition in children and adolescents: an emerging problem. *Front Pediatr*. (2022) 10:10. doi: 10.3389/fped.2022.894204

10. Qamar MA, Martins RS, Dhillon RA, Tharwani A, Irfan O, Suriya QF, et al. Residual symptoms and the quality of life in individuals recovered from COVID-19 infection: a survey from Pakistan. *Ann Med Surg*. (2022) 75:75. doi: 10.1016/j.amsu.2022.103361

11. World Health Organization. Expanding our understanding of post COVID-19 condition. (2021). 32 p. Available at: https://www.who.int/publications/i/item/9789240025035

12. Mirhosseini S, Grimwood S, Basirinezhad MH, Montazeri R, Ebrahimi H. Psychological distress as a predictor of the burden of care in family caregivers of COVID-19 patients in Iran: a community-based cross-sectional study. *Heal Sci Rep*. (2022) 5:942. doi: 10.1002/hsr2.942

13. Orsini A, Corsi M, Pedrinelli V, Santangelo A, Bertelloni C, Dell'Oste V, et al. Post-traumatic stress, anxiety, and depressive symptoms in caregivers of children tested for COVID-19 in the acute phase of the Italian outbreak. *J Psychiatr Res*. (2021) 135:256–63. Available from:. doi: 10.1016/j.jpsychires.2021.01.024

14. León-Campos Á, García-Mayor S, Martí-García C, Morilla-Herrera JC, Morales-Asencio JM, Lupiáñez-Pérez I, et al. Quality of life, physical and mental health, and economic evaluation of family caregivers of chronic dependent children: INFAPRINT

cohort study protocol. *Int J Environ Res Public Health*. (2023) 20:5081. doi: 10.3390/ijerph20065081

15. Turner J, Clavarino A, Work BS, Yates P, Sci BA, Sci MS, et al. *Development of a resource for parents with advanced cancer: What do parents want*, (2007) 5:135–45. doi: 10.1017/s1478951507070204

16. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DKW, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill*. (2020) 25:45. doi: 10.2807/1560-7917.ES.2020.25.3.2000045

17. Shen B, Zheng Y, Zhang X, Zhang W, Wang D, Jin J, et al. Clinical evaluation of a rapid colloidal gold immunochromatography assay for SARS-Cov-2 IgM/IgG. *Am J Transl Res*. (2020) 12:1348

18. Beavis KG, Matushek SM, Abeleda APF, Bethel C, Hunt C, Gillen S, et al. Evaluation of the EUROIMMUN anti-SARS-CoV-2 ELISA assay for detection of IgA and IgG antibodies. *J Clin Virol*. (2020) 129:104468. doi: 10.1016/j.jcv.2020.104468

19. Craig BM, Monteiro AL, Herdman M, Santos M. Further evidence on EQ-5D-5L preference inversion: a Brazil/U.S. collaboration. *Qual Life Res*. (2017) 26:2489–96. doi: 10.1007/s11136-017-1591-8

20. Silveira C, Parpinelli MA, Pacagnella RC, de Camargo RS, Costa ML, Zanardi DM, et al. Cross-cultural adaptation of the World Health Organization disability assessment schedule (WHODAS 2.0) into Portuguese. *Rev Assoc Med Bras*. (2013) 59:234–40. doi: 10.1016/j.ramb.2012.11.005

21. Durso DF, Bydlowski SP, Hutz MH, Suarez-Kurtz G, Magalhães TR, Pena SDJ. Association of Genetic Variants with self-assessed color categories in Brazilians. *PLoS One*. (2014) 9:e83926. doi: 10.1371/journal.pone.0083926

22. Alveno RA, Miranda CV, Passone CG, Waetge AR, Hojo ES, Farhat SCL, et al. Pediatric chronic patients at outpatient clinics: a study in a Latin American University Hospital. *J Pediatr*. (2018) 94:539–45. doi: 10.1016/j.jped.2017.07.014

23. Buff Passone CG, Grisi SJ, Farhat SC, Della MT, Pastorino AC, Alveno RA, et al. Complexity of pediatric chronic disease: cross-sectional study with 16,237 patients followed by multiple medical specialties. *Rev Paul Pediatr*. (2019) 38:1–8. doi: 10.1590/1984-0462/2020/38/2018101

24. Leal GN, Astley C, Lima MS, Diniz MFR, Lianza AC, Sawamura KSS, et al. Segmental cardiac strain assessment by two-dimensional speckle-tracking echocardiography in surviving MIS-c patients: correlations with myocardial flow reserve (MFR) by 13 N-ammonia PET-CT. *Microcirculation*. (2022) 29:e12750. doi: 10.1111/micc.12750

25. Astley C, Badue Pereira MF, Lima MS, Buchpiguel CA, Carneiro CG, Sapienza MT, et al. In-depth cardiovascular and pulmonary assessments in children with multisystem inflammatory syndrome after SARS-CoV-2 infection: a case series study. *Physiol Rep*. (2022) 10:e15201. doi: 10.14814/phy2.15201

26. Ramos GF, Ribeiro VP, Mercadante MP, Ribeiro MP, Delgado AF, Farhat SCL, et al. Mortality in adolescents and young adults with chronic diseases during 16 years: a study in a Latin American tertiary hospital. *J Pediatr Versão em Port*. (2019) 95:667–73.

27. Zakaria SM, Abdullah N, Md. Akhir N, Amin AS, Mohd Shukry ANA, Abdul Rashid MR, et al. Perceptions of quality of life during the pandemic: a case study on B40 single mothers. *IJERPH*. (2022) 19:12219. doi: 10.3390/ijerph191912219

28. Higgins V, Sohaei D, Diamandis EP, Prassas I. COVID-19: from an acute to chronic disease? Potential long-term health consequences. *Crit Rev Clin Lab Sci*. (2021) 58:297–310. doi: 10.1080/10408363.2020.1860895

29. Battistella LR, Imamura M, De Pretto LR, Van Cauwenbergh SKHAA, Delgado Ramos V, Saemy Tome Uchiyama S, et al. Long-term functioning status of COVID-19 survivors: a prospective observational evaluation of a cohort of patients surviving hospitalisation. *BMJ Open*. (2022) 12:e057246. doi: 10.1136/bmjopen-2021-057246

30. Khoja O, Passadouro BS, Mulvey M, Delis I, Astill S, Tan AL, et al. Clinical characteristics and mechanisms of musculoskeletal pain in long COVID. *J Pain Res*. (2022) 15:1729–48. doi: 10.2147/JPR.S365026

31. CDC. Post-COVID conditions. Centers for Disease Control and Prevention. (2022). Available at: https://www.cdc.gov/coronavirus/2019-ncov/long-term-effects/index.html

32. Seeßle J, Waterboer T, Hippchen T, Simon J, Kirchner M, Lim A, et al. Persistent symptoms in adult patients 1 year after coronavirus disease 2019 (COVID-19): a prospective cohort study. *Clin Infect Dis*. (2022) 74:1191–8. doi: 10.1093/cid/ciab611

33. Andrews G, Kemp A, Sunderland M, von Korff M, Ustun TB. Normative data for the 12 item WHO disability assessment schedule 2.0. *PLoS One*. (2009) 4:e8343–6. doi: 10.1371/journal.pone.0008343

34. Kozu KT, Casella CB, Strabelli CAA, Aikawa NE, Campos LMA, Elias AM, et al. Mental health impact in Latin American pediatric rheumatologists during the COVID-19 pandemic. *J Clin Rheumatol*. (2022) 28:E506–10. doi: 10.1097/RHU.0000000000001782

35. Darlington ASE, Morgan JE, Wagland R, Sodergren SC, Culliford D, Gamble A, et al. COVID-19 and children with cancer: parents' experiences, anxieties and support needs. *Pediatr Blood Cancer*. (2021) 68:1–10. doi: 10.1002/pbc.28790

frontiers | Frontiers in Global Women's Health

TYPE Original Research
PUBLISHED 04 September 2023
DOI 10.3389/fgwh.2023.1138579

OPEN ACCESS

EDITED BY
Pei Wang,
Fudan University, China

REVIEWED BY
Angelo D'Ambrosio,
Freiburg University Medical Center, Germany
Arlette Setiawan,
Padjadjaran University, Indonesia

*CORRESPONDENCE
Desale Bihonegn Asmamaw
desalebihonegn1988@gmail.com

RECEIVED 05 January 2023
ACCEPTED 09 August 2023
PUBLISHED 04 September 2023

CITATION
Asmamaw DB, Negash WD, Aragaw FM, Eshetu HB, Asratie MH and Belachew TB (2023) Spatial distribution and associated factors of poor tetanus toxoid immunization among pregnant women in Ethiopia: spatial and multilevel analysis.
Front. Glob. Womens Health 4:1138579.
doi: 10.3389/fgwh.2023.1138579

Spatial distribution and associated factors of poor tetanus toxoid immunization among pregnant women in Ethiopia: spatial and multilevel analysis

Desale Bihonegn Asmamaw[1]*, Wubshet Debebe Negash[2], Fantu Mamo Aragaw[3], Habitu Birhan Eshetu[4], Melaku Hunie Asratie[5] and Tadele Biresaw Belachew[2]

[1]Department of Reproductive Health, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia, [2]Department of Health Systems and Policy, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia, [3]Department of Epidemiology and Biostatistics, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia, [4]Department of Health Education and Behavioral Sciences, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia, [5]Department of Women's and Family Health, School of Midwifery, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia

Background: Neonatal mortality from tetanus can be reduced by 94% when pregnant women receive at least two doses of tetanus toxoid. In Ethiopia, immunization programs are suboptimal despite their importance. Therefore, the aim of this study was to examine the geographic distribution and associated factors of poor tetanus toxoid (TT) immunization among pregnant women in Ethiopia.

Methods: Secondary data analysis was used using the 2016 Ethiopian Demographic and Health Survey (EDHS). ArcGIS version 10.8 statistical software was used to explore the spatial distribution of poor TT immunization and SaTScan version 9.6 software was used to identify significant hotspot areas of poor TT immunization. For associated factors, a multilevel binary logistic regression model was fitted using STATA version 14 software. In the multivariable multilevel analysis, adjusted OR (AOR) with 95% CI was reported to reveal significantly associated factors of poor TT immunization.

Result: In Ethiopia, the spatial distribution of poor tetanus toxoid immunization was clustered with Global Moran's $I = 0.59$ at p-value of <0.0001. The highest poor TT immunization clusters were observed in the East and South Tigray, the central part of Amhara, West Afar, East Somali, and West Gambella. Pregnant women with no Antenatal care (ANC) visits [Adjusted Odds Ratio (aOR) = 10.46, 95% CI: (8.82, 12.41))], pregnant women with 1–3 ANC visits [aOR = 1.51, 95% CI: (1.31, 1.73)], media exposure [aOR = 1.45, 95% CI: (1.26, 1.67)], poor wealth index [aOR = 1.22; 95% CI: (1.03, 1.45)], middle wealth index [aOR = 1.23; 95% CI: (1.03, 1.47)], family planning use [aOR = 1.28; 95% CI: (1.11, 1.57)] and community level education [aOR = 1.43, 95% CI: (1.14, 1.80)] were significantly associated with poor tetanus toxoid immunization.

Conclusion: Poor tetanus toxoid immunization among pregnant women varies in Ethiopia. It was highest in East and South Tigray, the central part of Amhara, West Afar, East Somali, and West Gambella. Therefore, public health programs should design targeted interventions in identified hot spots to improve tetanus toxoid immunization. Health programmers should be promoting optimal ANC visits, women's education, and family planning use.

KEYWORDS

tetanus toxoid, immunization, spatial, Ethiopia, EDHS

Introduction

Maternal and Neonatal Tetanus (MNT) has been among the most common life-threatening consequences of unclean deliveries and umbilical cord care practices, and are indicators of inequity in access to immunization and other maternal, newborns, and child health services (1, 2). Every year, 3.3 million newborns die from neonatal tetanus, which accounts for a large proportion, particularly in developing countries like Ethiopia where home deliveries are common (3). Approximately 15,000–30,000 maternal mortality is caused by tetanus related to delivery every year (4). The World Health Organization (WHO) report states that the majority (90%) of maternal and neonatal tetanus cases occurred in South East Asian (SEA) and Sub-Saharan African (SSA) countries, and most of them ended with death (5).

The primary strategies for preventing MNT at birth were immunization of mothers with a protective dose of tetanus toxoid (TT) and clean delivery (6, 7). Tetanus toxoid vaccination (TT2+) can reduce neonatal mortality related to tetanus by nearly 94% in women of childbearing age (8). According to the WHO, women of childbearing age should take five successive doses of tetanus toxoid (TT) for lifelong protection (1).

During antenatal care appointments, pregnant women can receive at least two doses of the TT vaccine in different countries. But among expectant mothers, the coverage of TT2+ vaccinations ranged between 27% to 71% (9, 10). It indicates that the majority of the countries were not meeting the WHO's global immunization target of at least 90% national TT vaccine coverage and at least 80% coverage in every district (11). Despite the fact that Ethiopia is expected to reach 86% coverage of national tetanus protection at birth by 2015, only 48% of mothers in 2011% and 49% of mothers in 2016 were protected (12).

Age of the mothers, educational status of the women, marital status, occupational status, joint health decision with husband, distance from health facilities, wealth index, fear of side effects, fear of sterility, ANC follow-up, parity, lack of information about TT vaccination, knowledge, attitude, and low awareness of mothers were significant factors leads to poor TT immunization countries (2, 6, 10, 13–15). TT immunization status also differs from urban to rural mothers as well as from region to region in different countries (14, 16).

Various studies have been conducted in Ethiopia to identify associated factors for TT immunization. The spatial distribution of TT immunization was unclear in regions of Ethiopia. Understanding the level and geographical variation of TT immunization in Ethiopia can help health planners, programmers, partners in the health sector, and policymakers formulate appropriate strategies and interventions and provide quality reproductive health services to increase TT immunization. Hence, this study aimed to assess the spatial distributions of TT immunization and associated factors in Ethiopia using Ethiopian Demographic and Health Survey (EDHS) datasets.

Methods

In Ethiopia, there are nine regions [Afar, Tigray, Amhara, Oromia, Somali, Southern Nations, Nationalities, and People's Region (SNNPR), Benishangul Gumuz, Gambella, and Harari] and two administrative cities (Addis Ababa and Dire Dawa) (17). Based on Worldometer's analysis of the latest United Nations data, Ethiopia has 121,989,792 citizens as of Thursday, December 22, 2022 (18).

A secondary data analysis was undertaken using the 2016 Ethiopian Demographic and Health Survey (EDHS), which was a nationwide representative sample that was carried out from January 18 to June 27, 2016. Once permission had been obtained by securing an online request after explaining the purpose of the study, the EDHS 2016 was accessed from the DHS official database, www.measuredhs.com. Using the 2007 Population and Housing Census (PHC) as a sampling frame, a cross-sectional study design was used with stratified cluster sampling. A stratification was carried out by separating Ethiopia's nine regional states and its two city administrations, then by dividing it into urban and rural areas (17).

A total of 645 Enumeration Areas (EAs) (202 in urban areas and 443 in rural areas) proportional to EA size were selected proportionally to the EA size in the first stage. In the second stage, 28 homes from each cluster (EA) were chosen systematically with an equal chance (12). For this study, the study population was women (aged 15–49 years) who had a pregnancy five years before the survey. We used individual datasets. A total weighted sample of 7,397 pregnant women were included in the present study. Additionally, latitude and longitude coordinates were taken from selected EAs (clusters). Full details of the EDHS sampling system were presented in the report (17, 19).

Study variables

This study's outcome variable was poor TT immunization, which is defined as poor TT immunization if pregnant women did not receive enough immunization (did not receive two and above tetanus toxoid vaccine) and otherwise no. Depending on different literature reviews, individual and community-level variables were included in

the analysis. Age of the women (15–24, 25–34, and 35–49), women's education (no formal education, primary education, and secondary education and above), occupation (employed, not employed), modern family planning use (yes, no), parity (primipara, multipara, grand multipara) and number of antenatal care (ANC) visits (0, 1–3, and ≥4), person who decides on respondents health care (self, jointly, and husband alone) were considered as individual-level variables (20–27). Media exposure; those who read newspapers, listened to the radio, or watched television at least once a week were coded yes and no otherwise (14). The EDHS assessed the wealth index based on 28 data items on household-selected assets, water, and toilet access used to generate the wealth status of the household, and the household wealth index status was done using principal component analysis (PCA). The overall household data items are grouped into nine items: ownership of the house, types of living house floors, walls, roof materials, household properties, types of toilets, types of fuel mainly used for cooking, the main source of water, household animal type and number, farming land ownership, and agricultural product item and amount. To categorize individuals into wealth quintiles (poor, middle, and rich), we used household asset data via principal component analysis (PCA) (28). Wanted last child (wanted no more, wanted later, and wanted then). Wanted last child indicates whether the woman wanted the last child (previous child), wanted to give birth after two years, or did not want to give birth throughout her life. Wanted then means women who wanted to deliver within two years for the last child, wanted later means women who wanted to postpone for at least two years, and wanted no more means women who did not want any more children throughout their life.

Of the community level factors, distance to the health facilities (big problem, not big problem), residence (rural, urban), and in accordance with a previous study conducted in Ethiopia and based on its geopolitical features, region was divided into three regions: small peripheral regions (Afar, Somalia, Benishangul, and Gambela), large central regions (Tigray, Amhara, Oromia, and Southern Nations Nationalities and Peoples Region), and metropolitan (Harari, Dire Dawa, and Addis Ababa) were directly accessed from EDHS data set (29–31). However, the aggregate community-level independent variables (community-level poverty, community-level media exposure, and community-level education) were constructed by aggregating individual-level characteristics at the community (cluster) level. They were classified as high or low based on the distribution of the proportion values calculated for each community after examining the distribution using the histogram. Since the aggregate variable was not normally distributed, the categorization was based on the median value (25, 32, 33).

Data management and analysis

For data analysis, we used STATA 14, ArcGIS 10.8, and SaTScan 9.6 software. For the analysis, sample weights were applied to adjust for the non-proportional sampling of strata and regions during the survey process and to restore representativeness. Text, figures, and tables were used to present descriptive statistics and summary statistics (17, 33).

Spatial analysis

Spatial autocorrelation analysis

The presence of spatial autocorrelation was identified using Moran's index (Moran's I). A Moran's I value close to −1 indicates that disease/events are dispersed, whereas a Moran's I value close to +1 indicates that they are clustered, and a Moran's I value of zero indicates that they are distributed randomly. There was a significant Moran's I ($p < 0.05$), indicating the presence of spatial autocorrelation and rejecting the null hypothesis (poor TT immunization is randomly distributed). Hotspot analysis was conducted using the Getis-Ord Gi* statistic (17).

Spatial interpolation

The spatial interpolation was done to predict poor immunization in unsampled areas of the country based on sampled measurements. Ordinary Kriging (OK) and Empirical Bayesian Kriging (EBK) were done since they statistically optimized the weight to predict the prevalence of poor immunization in the unobserved areas based on the observed measurement. The ordinary Kriging spatial interpolation method was selected for this study for predictions of poor immunization coverage since it had a smaller residual and Root Mean Square Error (RMSE) than EBK.

Spatial scan statistical analysis

Spatial scan statistics applied using Kulldorff's SaTScan software identified statistically significant primary (most likely) and secondary clusters of poor TT immunization. In SaTScan™ works, a window moves across the study areas and the window size needs to be fixed. As the outcome variable was Bernoulli distribution, Kulldorff's method was applied to use a Bernoulli model for a purely spatial analysis. In order to fit the Bernoulli model, respondents with poor TT immunization were considered case, and those with high TT immunization were considered control. Using the default maximum spatial cluster size of 50% of the population as an upper limit, both small and large clusters were detected, and clusters with more than the maximum level were ignored. Poor TT immunization was considered in areas with a high Log Likelihood Ratio and significant *p*-value compared to areas outside the window.

Multilevel analysis

Descriptive statistics were described using frequency and percentage. A variance inflation factor (VIF) was used to test for multicollinearity, and a VIF of less than five was obtained for each independent variable, with a mean VIF of 1.47, indicating there was no significant multicollinearity between independent variables. In the EDHS data, there was a hierarchical structure, which violates the independent observations and equal variance assumptions of a traditional logistic regression model. Therefore, women were nested within households, and households were nested within clusters. Within the cluster, they may have similar characteristics. Hence, multilevel binary logistic regression analysis must take into account the variability between clusters. Since the models were nested.

Intra-class correlation coefficient (ICC) and Proportional Change in Variance (PCV) were computed to measure the variation between clusters. The ICC reveals that the variation of poor TT immunization between clusters is calculated as;. Moreover, the PCV reveals the variation in the prevalence of poor TT immunization among pregnant women explained by factors and calculated as; where; Vnull = variance of the initial model, and VA = area/cluster level variance (34). Model fitness was checked using the deviance test and AIC, and the model with the lowest AIC and lowest deviance was selected as the best-fitted model.

First, bivariable multilevel logistic regression analysis was conducted and those variables with a $p < 0.2$ were considered for multivariable analysis. Finally, multilevel binary logistic regression analysis was done to assess the association between TT immunization and individual and community-level factors. Four models were constructed for multivariable multilevel analysis; the null model (without independent variables), mode I (containing only individual-level factors), mode II (containing only community-level factors), and model III (containing both individual and community-level factors) were fitted. In the multivariable model, variables with an adjusted odds ratio (aOR) with a 95% confidence interval (CI) and a p-value of <0.05 were considered significantly associated factors of poor TT immunization.

Results

Individual and community level factors

A total weighted sample of 7,397 pregnant women were included in this analysis. The median age of the respondents was 28 years (IQR: 24–34) and 87.3% of the women were from rural residence. Of the study participants, 63.1% of the pregnant women had no formal education. More than half (53.9%) of the study participants were not employed and most of (65.6%) the study participants had no media exposure. The majority (90.8%) of the pregnant women were from large central regions. More than two fifths (43.7%) of the pregnant women were from households with poor wealth quantiles (Table 1).

TABLE 1 Individual and community level factors associated with poor TT immunization among pregnant women in Ethiopia.

Variables	Frequency	Percentage (%)
Age in year		
15–24	1,766	23.9
25–34	3,718	50.3
35–49	1,913	25.8
Educational status		
No formal education	4,668	63.1
Primary education	2,095	28.3
Secondary and above	634	8.6
Occupation		
Employed	3,407	46.1
Not employed	3,990	53.9
Household wealth		
Poor	3,230	43.7
Middle	1,552	21
Rich	2,616	35.3
Media exposure		
No	4,849	65.6
Yes	2,548	34.4
Family planning use		
No	4,790	64.8
Yes	2,607	35.2
Number of ANC visits		
No	2,770	37.5
1–3	2,289	30.9
≥4	2,338	31.6
Parity		
Primipara	1,412	19.1
Multipara	3,099	41.9
Grand multipara	2,886	39
Wanted last child		
Wanted then	5,455	73.7
Wanted later	1,277	17.3
Wanted no more	665	9.0
Person who decided on respondents health care		
Self	947	12.8
Joint	4,630	62.6
Husband	1,820	24.6
Resident		
Rural	6,456	87.3
Urban	941	12.7
Community level poverty		
Low	4,484	60.6
High	2,913	39.4
Community-level media exposure		
Low	3,491	47.2
High	3,906	52.8
Community level education		
Low	3,748	50.7
High	3,649	49.3
Region		
Small peripheral	436	5.9
Large central	6,719	90.8
Metropolitan	242	3.3
Distance to the health facilities		
Big problem	4,289	58
Not a big problem	3,108	42

Regional prevalence of poor TT immunization among pregnant women

The prevalence of poor TT immunization during pregnancy varies across the country. The highest and lowest prevalence of poor TT immunization during pregnancy were observed in the Afar (71.7%) and Dire Dawa (34.2%) regions, respectively (Figure 1).

Spatial analysis of poor TT immunization

Spatial autocorrelation and spatial analysis of TT immunization

The spatial autocorrelation analysis revealed that the distribution of poor TT immunization was non-random in Ethiopia, with a Global Moran's Index value of 0.59 ($p < 0.0001$) (Figure 2). A

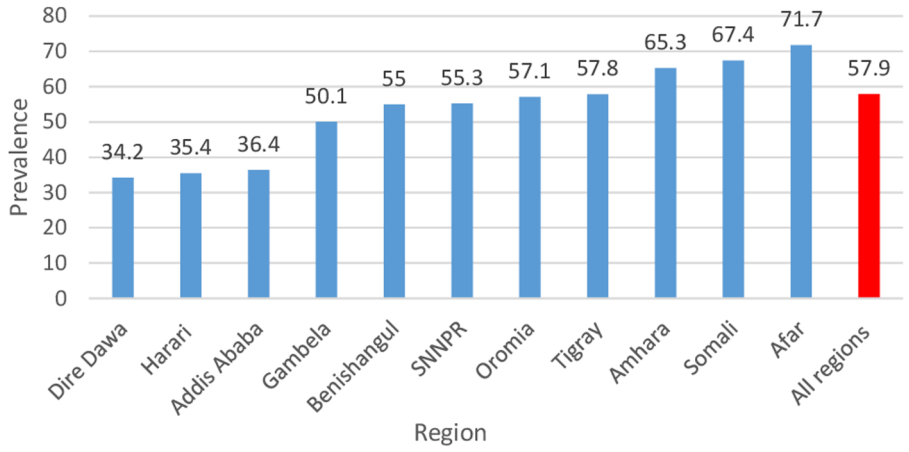


FIGURE 1
Regional prevalence of TT immunization among pregnant women in Ethiopia, 2016.

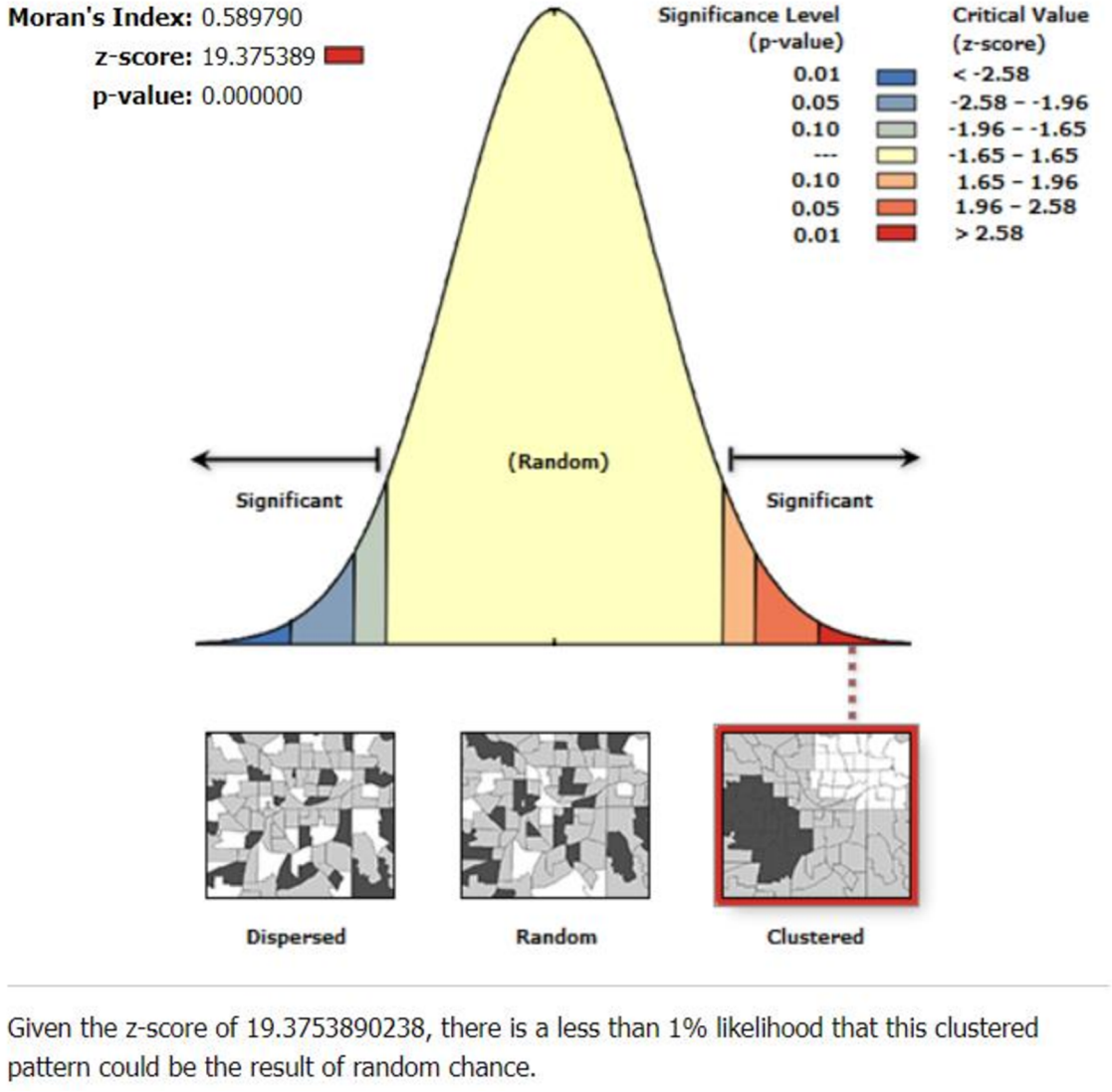


FIGURE 2
Spatial autocorrelation analysis of poor TT immunization among pregnant women in Ethiopia.

higher proportion of poor TT immunization occurred in the east and south of Tigray, the central part of Amhara, West Afar, East Somali, and West Gambella, while low proportions of poor TT immunization were identified in the Addis Ababa, Harari, Dire Dawa, and central Oromia regions (Figure 3).

Getis OrdGi statistical analysis of poor TT immunization

According to the Getis OrdGi statistics analysis, the majority of Tigray and Amhara regions had significant hotspots (areas with poor TT immunization). In addition, it is located in the west of Afar and northeast of the Somali region. Cold spots (areas with high TT immunization) were found in Addis Ababa, Dire Dawa, Hareri, the northern and central parts of Oromia, and the north-eastern part of SNNPR (Figure 4).

Kriging interpolation

In the Kriging interpolation; the predicted poor TT immunization coverage were identified in the Eastern and North-eastern part of Somali, northern part of Afar, central and south western part of Amhara, and Northern part of SNNPR regions whereas, the predicted high coverage of TT immunization were identified in the Northeastern part of Gambella, some part of eastern, western and central Oromia, Addis Ababa, Dire Dawa, and northwest Somali regions (Figure 5).

Spatial Sa Tscan analysis of poor TT immunization

There were 254 significant clusters identified in the spatial Sa Tscan statistics, of which 209 were primary clusters (most likely). As a result of the survey, the primary clusters were found in the Tigray, Amhara, North Eastern Somali region, Benishangul region, and Oromia region. They were located at 14.122895 N, 38.621010 E of geographic location, with a radius of 561.28 km, with a Relative Risk (RR) of 1.3 and Log-Likelihood ratio (LLR) of 77.36, with $p = 0.001$. According to the study, pregnant women within the spatial window were 1.3 times more likely to have poor TT immunization than pregnant women outside it (Figure 6) (Table 2).

Associated factors of TT immunization

Random effect results

Based on the null model, the ICC value for poor TT immunization was 19.6%, which means 19.6% of the

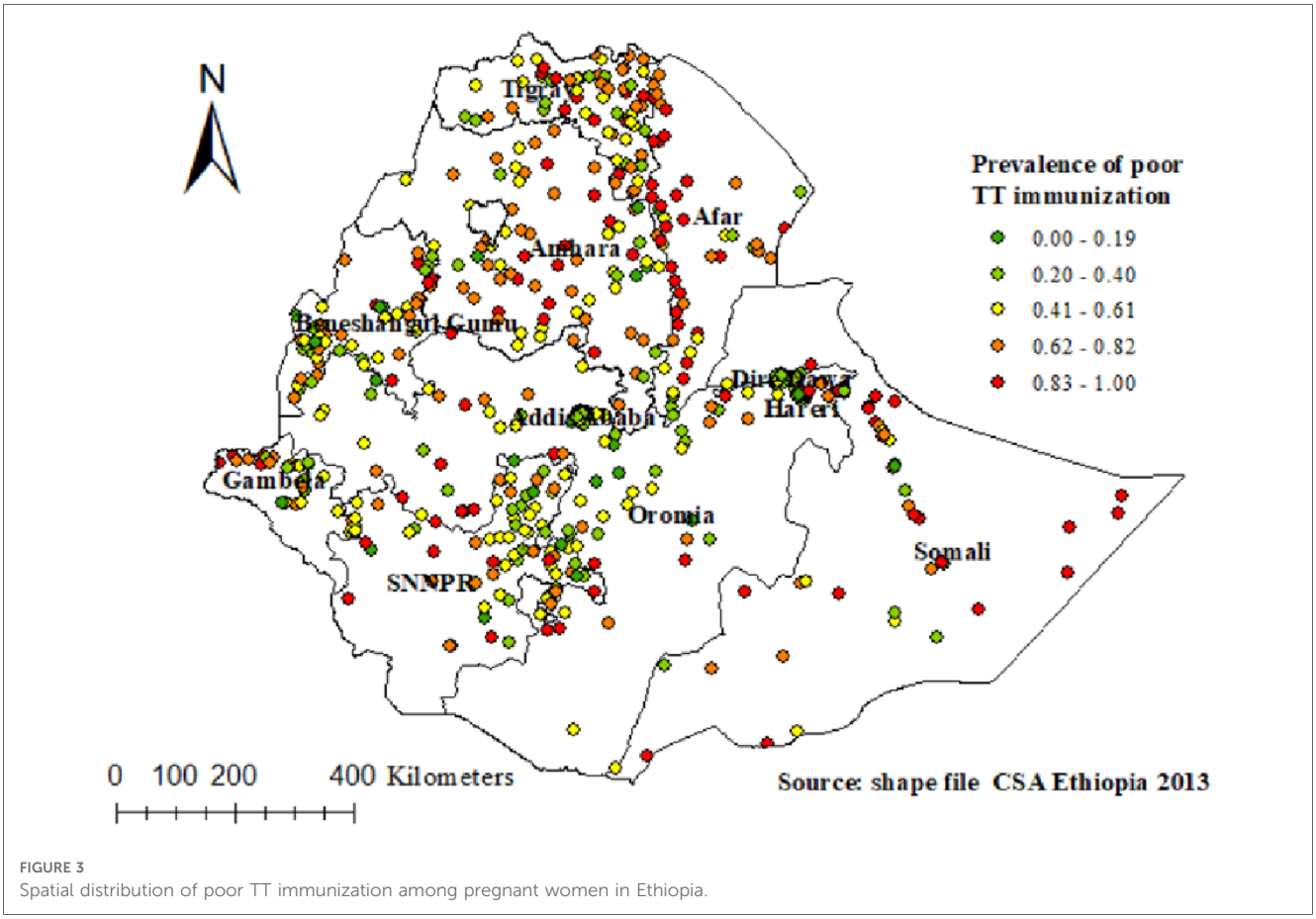


FIGURE 3
Spatial distribution of poor TT immunization among pregnant women in Ethiopia.

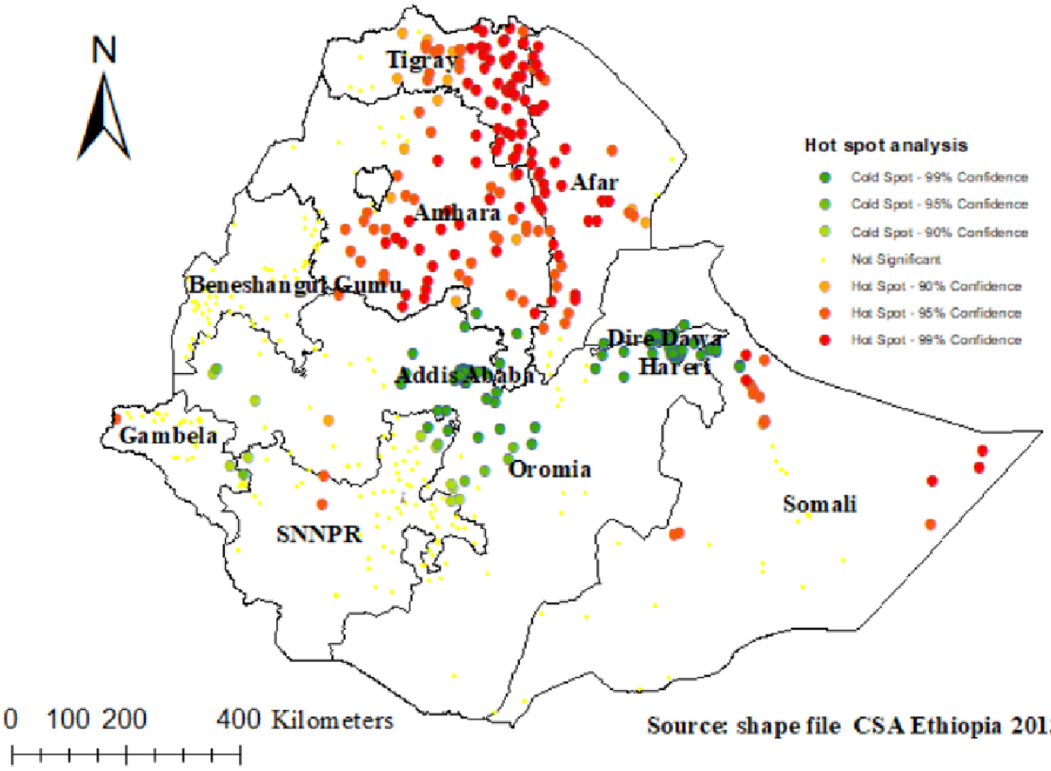


FIGURE 4
Hot spot analysis of poor TT immunization in Ethiopia.

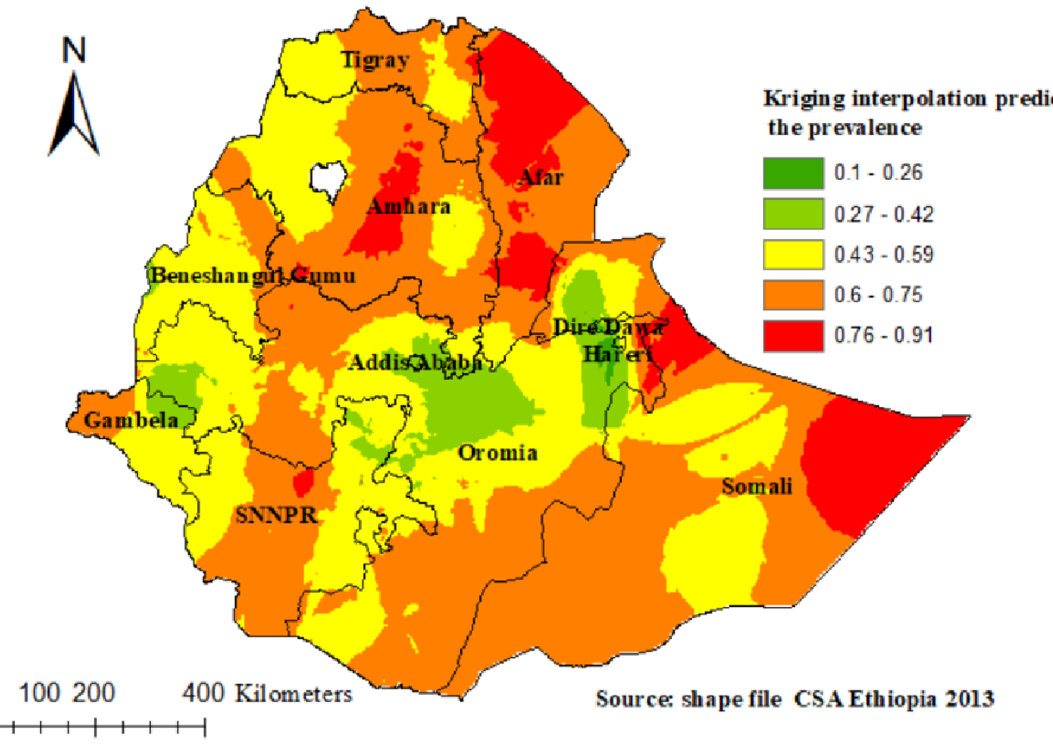


FIGURE 5
Kriging interpolation of poor TT immunization among pregnant women in Ethiopia.

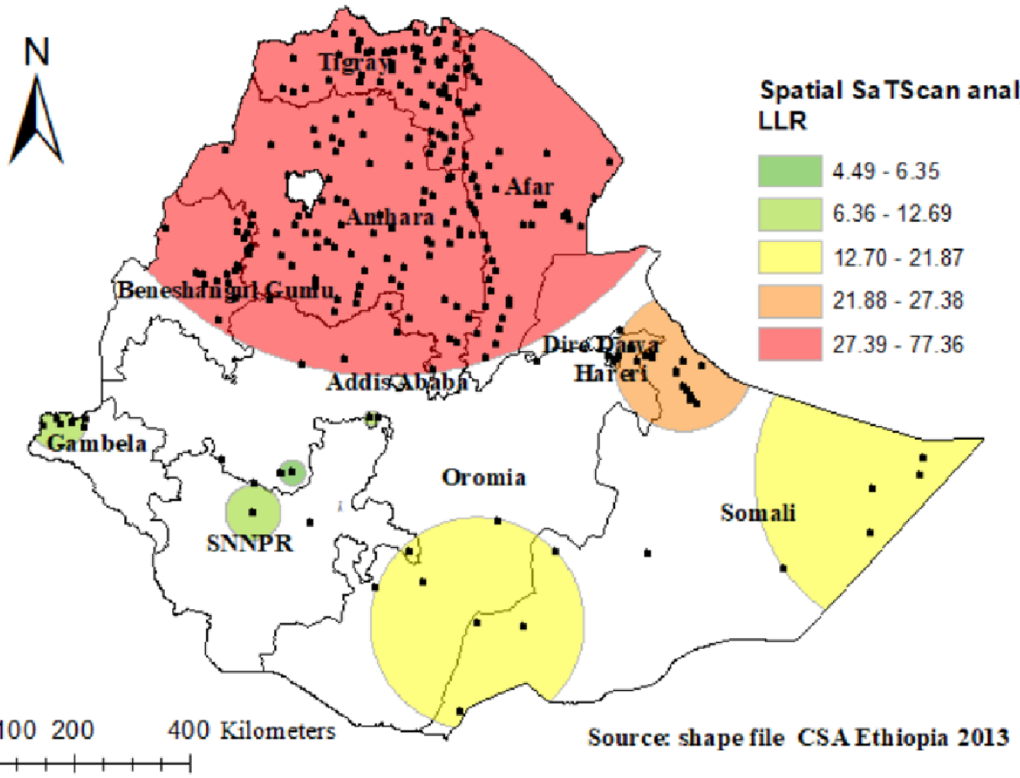


FIGURE 6
SaTScan analysis of poor TT immunization among pregnant women in Ethiopia.

TABLE 2 Sat scan analysis of poor TT immunization among pregnant women in Ethiopia.

Cluster	Enumeration area (cluster) identified	Coordinate/radius	Population	Case	RR	LLR	*P*-value
1 (209)	400, 597, 590, 81, 636, 584, 84, 156, 181, 45, 481, 255, 551, 98, 528, 461, 479, 340, 355, 89, 579, 604, 78, 188, 575, 583, 598, 226, 404, 425, 129, 430, 268, 341, 413, 538, 258, 424, 94, 237, 550, 623, 220, 117, 192, 196, 605, 80, 99, 103, 298, 384, 160, 263, 127, 362, 322, 134, 628, 421, 542, 235, 612, 392, 253, 585, 143, 511, 79, 296, 152, 312, 136, 449, 130, 172, 638, 128, 300, 504, 327, 199, 66, 640, 442, 97, 351, 488, 249, 163, 455, 200, 401, 279, 599, 478, 344, 544, 332, 591, 512, 132, 292, 627, 241, 158, 389, 189, 496, 348, 169, 545, 456, 571, 73, 410, 38, 191, 431, 427, 52, 354, 611, 167, 516, 616, 345, 382, 403, 254, 24, 18, 120, 368, 176, 429, 617, 259, 361, 205, 460, 178, 499, 602, 570, 206, 55, 334, 415, 10, 541, 109, 547, 386, 3, 548, 267, 482, 515, 375, 440, 632, 615, 276, 474, 498, 229, 596, 366, 620, 531, 4, 350, 510, 310, 533, 75, 283, 246, 637, 218, 559, 572, 102, 36, 494, 37, 295, 150, 135, 423, 256, 183, 137, 184, 336, 364, 35, 201, 244, 484, 39, 624, 517, 564, 320, 121, 230, 399	(14.122895 N, 38.621010 E)/561.28 km	2,399	1,605	1.3	77.36	<0.001
2 (23)	527, 277, 568, 33, 22, 116, 439, 64, 57, 239, 210, 573, 8, 251, 214, 186, 566, 1, 622, 436, 212, 501, 307	(9.264633 N, 43.272679 E)/119.44 km	293	226	1.38	27.38	<0.001
3 (5)	269, 630, 629, 77, 146	(7.453674 N, 46.955230 E)/285.37 km	58	55	1.68	21.87	<0.001
4 (8)	377, 394, 422, 7, 34, 289, 480, 398	(5.203234 N, 40.019732 E)/184.64 km	130	106	1.45	18.09	<0.001
5 (6)	309, 536, 370, 618, 233, 69	(8.342595 N, 33.451041 E)/50.46 km	64	55	1.52	12.69	<0.001
6 (2)	338, 76	(6.934084 N, 36.520510 E)/48.00 km	27	26	1.7	11.25	0.012
7 (1)	278	(6.273056 N, 42.688145 E)/0 km	17	171	1.76	9.59	0.032

RR, relative risk; LLR, Log-likelihood Ratio.

variability was explained by group variation, while 80.4% was explained by individual variation. Besides, the higher PCV value (31.63%) in the final model indicates that about 31.63% of the variation of poor TT immunization among pregnant women was attributable to both individual level and community-level factors (Table 3).

TABLE 3 Community-level variability and model fitness for assessment of poor TT immunization among pregnant women in Ethiopia.

Measure of variations	Null model	Model I	Model II	Model III
ICC	19.6	15.2	14.3	13.4
Variance	80.2	63.5	54.7	50.8
PCV (%)	Ref	22.25	27.01	31.63
Model fitness				
Deviance	9,327	8,022.2	9,177.8	8,006.6
AIC	9,331.2	9,102.7	9,013.5	8,207.3

ICC, Intra class correlation coefficient; PCV, proportional change in variance; AIC, Akaike information criterion.

Fixed effect result

Number of ANC visits, media exposure, modern family planning use, wealth index, and community-level education were significantly associated with poor TT immunization in the final model (Model III).

Accordingly, pregnant women who had no and 1–3 ANC visits had 10.46 (aOR: 10.46, 95% CI: 8.82, 12.41) and 1.51 (aOR: 1.51, 95% CI: 1.31, 1.73) times higher odds of poor TT immunization than those who had recommended ANC visits, respectively. Pregnant women who had no media exposure had 1.45 (aOR: 1.45, 95% CI: 1.26, 1.67) times more odds to have poor TT immunization than those who had media exposure. Women who did not use modern family planning had 1.28 (aOR: 1.28, 95% CI: 1.11, 1.57) fold higher odds of poor TT immunization as compared with their counterparts. Compared to study participants in the rich households, pregnant women from households classified as poor and moderate status had 1.22 (aOR: 1.22, 95% CI: 1.03, 1.45) and 1.23 (aOR: 1.23, 95% CI: 1.03, 1.47) fold higher odds of poor TT immunization. Moreover, pregnant women who lived in a community with low education had 1.43 (aOR: 1.43, 95% CI: 1.14, 1.80) times higher odds of poor TT immunization as compared with those who lived in a community with high education (**Table 4**).

Discussion

According to the present study, the spatial distribution of poor TT immunization among pregnant women in Ethiopia was clustered. In most parts of Tigray and Amhara regions, there were significant hotspots of poor TT immunization. It is also found in western Afar and north-eastern Somali region. Whereas the significant "cold spot" areas with high rates of TT immunization were located in Addis Ababa, Dire Dawa, Harari, central and North-eastern parts of Oromia, and north-eastern part of the SNNPR regions. The possible justification for this difference could be due to the different socio-economic and obstetric-related factors of the study participants. For instance, the majority of the respondents in Addis Ababa, Dire Dawa, and Harari were educated, rich, and had the recommended ANC visits. In addition, most participants who lived in these regions were exposed to the media. Which in turn enhances the pregnant woman's understanding of TT immunization (31, 35).

Poor TT immunization coverage was higher among mothers who did not attend ANC visits compared to those who attended the recommended ANC visits. This finding was supported by studies conducted in Ethiopia (15), Kenya (36), and Pakistan (13). The TT immunization was a routine intervention during the ANC visit, and maternal awareness of continuous immunization may be increased through counselling about TT immunization (15).

Similarly, poor TT immunization was more likely to occur among pregnant women from lower-income households. This is in line with studies conducted in Ethiopia (16) and Bangladesh (37). The possible explanation might be due to economic status, which is one of the significant factors that affect maternal healthcare-seeking behavior, including TT immunization services (16, 38, 39).

Similarly, in this study, community-level education was also significantly associated with poor TT immunization. When compared with their counterparts, people who live in a community with low levels of community education are more likely to have poor TT immunization. Scholars showed that maternal health service utilization is low in a community with low community-level education (16, 40).

Poor TT immunization among those who had no media exposure was higher than for those who had media exposure. This is in line with a study conducted in Ethiopia (16). These findings may be explained by the fact that those exposed to the media have a better understanding of reproductive health services, including TT immunizations, and their advantages (41).

Pregnant women who did not use modern family planning methods were more likely to have poor TT immunization as compared with those who used modern family planning. This finding was supported by studies in Ethiopia (42) and Bangladesh (43). As a result of service integration, health information might be provided during family planning sessions (44).

Strength and limitation of the study

The current study's key strength was the use of weighted data that were nationally representative.

As a result, the results of the current study can be applied nationally. In addition, we also identify similar and statistically significant areas with a high cluster of TT immunization by using both ArcGIS and Sat Scan statistical tests. Due to the cross-sectional nature of the study, it does not show the cause-and-effect relationship between the outcome and independent variables. Because the current study relied on secondary data, some essential variables, such as respondents' knowledge of TT vaccines and sociocultural factors, were missing.

TABLE 4 Multi-level mixed-effect logistic regression analysis of individual and community level factors associated with poor TT immunization among pregnant women in Ethiopia.

Variables	Poor TT immunization		COR (95% CI)	Model I AOR (95% CI)	Model II AOR (95% CI)	Model III AOR (95% CI)
Education of the mother	Yes	No				
Secondary education and above	243	391	Ref	Ref		Ref
Primary education	1,039	1,056	1.47 (1.18, 1.83)	1.36 (1.25, 1.48)		0.89 (0.71, 1.13)
No formal education	2,998	1,670	2.31 (1.86, 2.84)	0.99 (0.78, 1.28)		0.97 (0.75, 1.26)
Occupation of the mother						
Employed	1,922	1,485	Ref	Ref		Ref
Not Employed	2,359	1,631	1.12 (0.98, 1.23)	0.98 (0.81, 1.11)		0.95 (0.85, 1.08)
Wealth index						
Rich	1,236	1,379	Ref	Ref		Ref
Middle	936	615	1.57 (1.35, 1.84)	1.06 (0.96, 1.14)		1.23 (1.03, 1.47)*
Poor	2,108	1,222	1.92 (1.67, 2.20)	1.12 (1.03, 1.21)		1.22 (1.03, 1.45)*
Media exposure						
Yes	1,146	1,401	Ref	Ref		Ref
No	3,134	1,715	1.99 (1.75, 2.26)	1.28 (1.19, 1.37)		1.45 (1.26, 1.67)*
Family planning use						
Yes	1,317	1,290	Ref	Ref		Ref
No	2,963	1,827	1.49 (1.33, 1.68)	1,51 (1.38, 1.63)		1.28 (1.11, 1.57)*
Number of ANC visits						
≥4	854	1,485	Ref	Ref		Ref
1–3	1,060	1,229	1.64 (1.43, 1.87)	2.67 (2.35, 2.91)		1.51 (1.31, 1.73)*
None	2,367	403	12.07 (10.27, 14.17)	3.34 (2.98, 3.56)		10.46 (8.82, 12.41)*
Wanted last-child						
Wanted then	3,131	2,324		Ref		Ref
Wanted later	704	573		1.16 (1.05, 1.43)		0.88 (0.76, 1.04)
Wanted no more	445	229		1.23 (1.11, 1.48)		1.19 (0.96, 1.48)
Parity						
Grand multipara	1,806	1,079		Ref		Ref
Multipara	1,751	1,349		1.11 (0.96, 1.27)		1.10 (0.96, 1.26)
Primipara	723	688		1.32 (1.09, 1.58)		1.30 (0.98, 1.56)
Person who decided on respondent health care						
Self	505	443		Ref		Ref
Jointly	2,634	1,996	1.07 (0.91, 1.26)	1.15 (0.96, 1.37)		1.13 (0.95, 1.36)
Husband	1,142	677	1.41 (1.17, 1.69)	1.18 (0.96, 1.44)		1.17 (0.96, 1.43)
Region						
Metropolitan	87	155	Ref		Ref	Ref
Large central	3,910	2,809	2.78 (1.93, 4.01)		1.58 (1.06, 1.36)	1.23 (0.82, 1.84)
Small peripheral	264	152	3.9 (2.53, 6.01)		1.64 (1.04, 1.61)	1.04 (0.64, 1.67)
Residence						
Urban	389	551	Ref		Ref	Ref
Rural	3,890	2,565	2.64 (2.05, 3.40)		1.11 (0.81, 1.51)	0.79 (0.59, 1.05)
Distance to the health facilities						
Not a big problem	1,652	1,457	Ref		Ref	Ref
Big problem	2,628	1,660	1.31 (1.16, 1.48)		1.16 (1.03, 1.31)	0.96 (0.83, 1.11)
Community media exposure						
High	1,729	1,920	Ref		Ref	Ref
Low	2,551	1,196	2.52 (2.05, 3.09)		1.68 (1.34, 12.11)	1.21 (0.96, 1.52)
Community level poverty						
High	1,865	1,048			Ref	Ref
Low	2,416	2,068	0.47 (0.38, 0.59)		0.93 (0.73, 1.18)	1.11 (0.82, 1.36)
Community level education						
High	1,729	1,920	Ref		Ref	Ref
Low	2,551	1,196	2.72 (2.24, 3.32)		1.96 (1.57, 2.45)	1.43 (1.14, 1.80)*

*Statistically significant at p-value <0.05, AOR; Adjusted Odds Ratio, COR; Crude Odds Ratio.

Conclusions

Poor tetanus toxoid immunization among pregnant women varies in Ethiopia. It was highest in the East and South Tigray, the central part of Amhara, West Afar, East Somali, and West Gambella. Number of ANC visits, media exposure, wealth index, family planning use, and community-level education were significantly associated with poor TT immunization. Therefore, TT immunization should be improved in the identified hotspot areas by designing locally targeted public health interventions. Public health interventions like promoting the dissemination of information related to TT immunization in the media, women's education, and modern family planning use have the potential to improve women's awareness of TT vaccines, which in turn will improve TT immunization.

Data availability statement

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

Ethics statement

The studies involving humans were approved by We used secondary data, which is EDHS. The studies were conducted in accordance with the local legislation and institutional requirements. The Ethics Committee/institutional review board waived the requirement of written informed consent for participation from the participants or the participants' legal guardians/next of kin because We used secondary data, which is EDHS.

Author contributions

DA conceived the idea for this study and was involved in the data extraction, analysis, interpretation of the findings, and writing of the original draft. FA, TB, HE, MA, and WN wrote the review and edited the manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgments

We are grateful to the DHS programs, for the permission to use all the relevant DHS data for this 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

References

1. World Health Organization. Weekly epidemiological record. *Wkly Epidemiol Rec.* (2006) 81:325–40.

2. Demicheli V, Barale A, Rivetti A. Vaccines for women for preventing neonatal tetanus. *Cochrane Database Syst Rev.* (2015) (7).

3. Yusuf N, Raza AA, Chang-Blanc D, Ahmed B, Hailegebriel T, Luce RR, et al. Progress and barriers towards maternal and neonatal tetanus elimination in the remaining 12 countries. *Lancet Glob Health.* (2021) 9(11):e1610–7. doi: 10.1016/S2214-109X(21)00338-7

4. World Health Organization. *Global diffusion of eHealth: Making universal health coverage achievable: Report of the third global survey on eHealth.* World Health Organization (2017).

5. World Health Organization: Pertussis: immunization surveillance, assessment and monitoring. (2012). Available at: http:www who int/immunization_monitoring/diseases/pertussis_surveillance/en/index/html

6. World Health Organization. *Protection at birth (PAB) against tetanus.* Global Health Observatory data (2015).

7. Bekele AT, Fiona B, Thomas K, Kassahun A, Kathleen G, Nsubuga P, et al. Factors contributing to routine immunization performance in Ethiopia, 2014. *Pan Afr Med J.* (2017) 27(Suppl 2).

8. Khan REA, Raza MA. Maternal health-care in India: the case of tetanus toxoid vaccination. *Asian Dev Policy Rev.* (2013) 1(1):1–14. doi: 10.18488/journal.107.2013.11.1.14

9. Diamenu S, Bosnu G, Abotsi F, Tweneboa P, Okoh-Owusu M, Amoh P, et al. Introducing protection at birth (Pab) method of monitoring tetanus-diphtheria (Td) vaccination coverage of mothers in Ghana. *Int J Vaccines Immu.* (2015) 1(1).

10. Singh A, Pallikadavath S, Ogollah R, Stones W. Maternal tetanus toxoid vaccination and neonatal mortality in rural north India. *PLoS one.* (2012) 7(11): e48891. doi: 10.1371/journal.pone.0048891

11. Khan R, Vandelaer J, Yakubu A, Raza AA, Zulu F. Maternal and neonatal tetanus elimination: from protecting women and newborns to protecting all. *Int J Women's Health.* (2015) 7:171.

12. Ethiopia CSA, Macro O: *Ethiopia demographic and health survey.* Addis Ababa: Central Statistical Agency (2016).

13. Naeem M, Khan MZ-U-I, Abbas SH, Adil M, Khan A, Naz SM, et al. Coverage and factors associated with tetanus toxoid vaccination among married women of reproductive age: a cross sectional study in Peshawar. *J Ayub Med Coll Abbottabad.* (2010) 22(3):136–40.

14. Tadesse AW, Aychiluhm SB, Mare KU. Individual and community-level determinants of Iron-Folic Acid Intake for the recommended period among pregnant women in Ethiopia: a multilevel analysis. *Heliyon.* (2021) 7(7):e07521. doi: 10.1016/j.heliyon.2021.e07521

15. Nigussie J, Girma B, Molla A, Mareg M. Tetanus toxoid vaccination coverage and associated factors among childbearing women in Ethiopia: a systematic review and meta-analysis. *BioMed Res Int.* (2021) 2021.

16. Liyew AM, Ayalew HG. Individual and community-level determinants of poor tetanus toxoid immunization among pregnant women in Ethiopia using data from 2016 Ethiopian demographic and health survey; multilevel analysis. *Arch Public Health.* (2021) 79(1):1–10. doi: 10.1186/s13690-021-00622-3

17. Tusa BS, Weldesenbet AB, Kebede SA. Spatial distribution and associated factors of underweight in Ethiopia: an analysis of Ethiopian demographic and health survey, 2016. *PLoS One*. (2020) 15(12):e0242744.

18. Woodward M, Reid MA. Cardiovascular disease in the Asia–Pacific region: challenges for health research and policy. *Med J Aust*. (2003) 179(2):71–2. doi: 10.5694/j.1326-5377.2003.tb05438.x

19. Dasa TT, Okunlola MA, Dessie Y. Multilevel analysis of grand multiparity: trend and its determinants in the Sidama National Regional State of Ethiopia: a cross-sectional study design from demographic and health survey 2000–2016. *BMJ Open*. (2022) 12(8):e061697.

20. Zerfu TA, Ayele HT. Micronutrients and pregnancy; effect of supplementation on pregnancy and pregnancy outcomes: a systematic review. *Nutr J*. (2013) 12 (1):1–5. doi: 10.1186/1475-2891-12-20

21. Zegeye B, Ahinkorah BO, Ameyaw EK, Seidu A-A, Yaya S. Utilization of deworming drugs and its individual and community level predictors among pregnant married women in cameroon: a multilevel modeling. *BioMed Res Int*. (2021) 2021.

22. Fereja M, Gebremedhin S, Gebreegziabher T, Girma M, Stoecker BJ. Prevalence of iodine deficiency and associated factors among pregnant women in Ada district, Oromia region, Ethiopia: a cross-sectional study. *BMC Pregnancy Childbirth*. (2018) 18(1):1–8. doi: 10.1186/s12884-018-1905-z

23. Berti C, Biesalski H, Gärtner R, Lapillonne A, Pietrzik K, Poston L, et al. Micronutrients in pregnancy: current knowledge and unresolved questions. *Clin Nutr*. (2011) 30(6):689–701. doi: 10.1016/j.clnu.2011.08.004

24. Boti N, Bekele T, Godana W, Getahun E, Gebremeskel F, Tsegaye B, et al. Adherence to Iron-Folate supplementation and associated factors among Pastoralist's pregnant women in Burji districts, Segen area People's zone, southern Ethiopia: community-based cross-sectional study. *Int J Reprod Med*. (2018) 2018. doi: 10.1155/2018/2365362

25. Asratie MH, Andualem Z. Predictors of early resumption of post-partum sexual intercourse among post-partum period women in Ethiopia: a multilevel analysis based on Ethiopian demographic and health survey 2016. *PLoS One*. (2022) 17(9):e0271372. doi: 10.1371/journal.pone.0271372

26. Gadisa TB, G/Michael MW, Reda MM, Aboma BD. Early resumption of postpartum sexual intercourse and its associated risk factors among married postpartum women who visited public hospitals of Jimma zone, Southwest Ethiopia: a cross-sectional study. *PLoS One*. (2021) 16(3):e0247769. doi: 10.1371/journal.pone.0247769

27. Alum AC, Kizza IB, Osingada CP, Katende G, Kaye DK. Factors associated with early resumption of sexual intercourse among postnatal women in Uganda. *Reprod Health*. (2015) 12(1):1–8. doi: 10.1186/1742-4755-12-1

28. Gebremedhin T, Aschalew AY, Tsehay CT, Dellie E, Atnafu A. Micronutrient intake status and associated factors among children aged 6–23 months in the emerging regions of Ethiopia: a multilevel analysis of the 2016 Ethiopia demographic and health survey. *PLoS One*. (2021) 16(10):e0258954.

29. Abrha S, Shiferaw S, Ahmed KY. Overweight and obesity and its socio-demographic correlates among urban Ethiopian women: evidence from the 2011 EDHS. *BMC Public Health*. (2016) 16(1):1–7. doi: 10.1186/s12889-016-3315-3

30. Ahmed KY, Page A, Arora A, Ogbo FA. Trends and determinants of early initiation of breastfeeding and exclusive breastfeeding in Ethiopia from 2000 to 2016. *Int Breastfeed J*. (2019) 14(1):1–14. doi: 10.1186/s13006-018-0195-4

31. Teshale AB, Tesema GA. Prevalence and associated factors of delayed first antenatal care booking among reproductive age women in Ethiopia; a multilevel analysis of EDHS 2016 data. *PLoS One*. (2020) 15(7):e0235538. doi: 10.1371/journal.pone.0235538

32. Asmamaw DB, Eshetu HB, Negash WD. Individual and community-level factors associated with intention to use contraceptives among reproductive age women in sub-Saharan Africa. *Int J Public Health*. (2022) 107.

33. Asmamaw DB, Negash WD. Magnitude of unmet need for family planning and its predictors among reproductive age women in high fertility regions of Ethiopia: evidence from Ethiopian Demographic and Health Survey. *BMC Women's Health*. (2022) 22(1):1–10.

34. Asmamaw D, Negash W. Unmet need for family planning and associated factors among adolescent girls and young women in Ethiopia: a multilevel analysis of Ethiopian Demographic and Health Survey. *Contracept Reprod Med*. (2023) 8(1):13. doi: 10.1186/s40834-022-00211-x

35. Aychiluhm SB, Tadesse AW, Mare KU, Abdu M, Ketema A: A multilevel analysis of short birth interval and its determinants among reproductive age women in developing regions of Ethiopia. *PLoS One*. (2020) 15(8):e0237602. doi: 10.1371/journal.pone.0237602

36. Haile ZT, Chertok IRA, Teweldeberhan AK. Determinants of utilization of sufficient tetanus toxoid immunization during pregnancy: evidence from the Kenya demographic and health survey, 2008–2009. *J Community Health*. (2013) 38 (3):492–9. doi: 10.1007/s10900-012-9638-9

37. Rahman M. Tetanus toxoid vaccination coverage and differential between urban and rural areas of Bangladesh. *East Afr J Public Health*. (2009) 6(1).

38. Amin R, Shah NM, Becker S. Socioeconomic factors differentiating maternal and child health-seeking behavior in rural Bangladesh: a cross-sectional analysis. *Int J Equity Health*. (2010) 9(1):1–11. doi: 10.1186/1475-9276-9-9

39. Simkhada B, Teijlingen E, Porter M, Simkhada P. Factors affecting the utilization of antenatal care in developing countries: systematic review of the literature. *J Adv Nurs*. (2008) 61(3):244–60. doi: 10.1111/j.1365-2648.2007.04532.x

40. Singh PK, Kumar C, Rai RK, Singh L. Factors associated with maternal healthcare services utilization in nine high focus states in India: a multilevel analysis based on 14 385 communities in 292 districts. *Health Policy Plan*. (2014) 29(5):542–59. doi: 10.1093/heapol/czt039

41. Facha W, Yohannes B, Duressa G. Tetanus toxoid vaccination coverage and associated factors among pregnant women in Duguna Fango District, southern Ethiopia. *Int J Health Sci Res*. (2018) 8(1):148–54.

42. Dubale Mamoro M, Kelbiso Hanfore L. Tetanus toxoid immunization status and associated factors among mothers in Damboya Woreda, Kembata Tembaro zone, SNNP, Ethiopia. *J Nutr Metab*. (2018) 2018. doi: 10.1155/2018/2839579

43. Hashmi FK, Islam M, Khan TA, Tipu MK. Vaccination coverage of mothers during pregnancy with tetanus toxoid and infants after birth. *Pakistan J Pharm*. (2011) 24(2):1–3.

44. Achyut P, Mishra A, Montana L, Sengupta R, Calhoun LM, Nanda P. Integration of family planning with maternal health services: an opportunity to increase postpartum modern contraceptive use in urban Uttar Pradesh, India. *J Fam Plann Reprod Health Care*. (2016) 42(2):107–15. doi: 10.1136/jfprhc-2015-101271

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 03 October 2023
DOI 10.3389/fpsyg.2023.1175402

OPEN ACCESS

EDITED BY
Anam Mehmood,
Harbin Medical University, China

REVIEWED BY
Narimasa Kumagai,
Seinan Gakuin University, Japan
Arlette Setiawan,
Padjadjaran University, Indonesia

*CORRESPONDENCE
Brigitte Essers
✉ brigitte.essers@mumc.nl
Pei Wang
✉ wang_p@fudan.edu.cn

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

RECEIVED 27 February 2023
ACCEPTED 11 September 2023
PUBLISHED 03 October 2023

CITATION
Essers B, Wang P, Stolk E, Jonker MF, Evers S, Joore M and Dirksen C (2023) An investigation of age dependency in Dutch and Chinese values for EQ-5D-Y.
Front. Psychol. 14:1175402.
doi: 10.3389/fpsyg.2023.1175402

An investigation of age dependency in Dutch and Chinese values for EQ-5D-Y

Brigitte Essers[1,2*†], Pei Wang[3*†], Elly Stolk[4], Marcel F. Jonker[5], Silvia Evers[2], Manuela Joore[1,2] and Carmen Dirksen[1,2]

[1]Department of Clinical Epidemiology and Medical Technology Assessment, Maastricht University Medical Centre, Maastricht, Netherlands, [2]Care and Public Health Research Institute (CAPHRI), Maastricht, Netherlands, [3]School of Public Health, Fudan University, Shanghai, China, [4]EuroQol Research Foundation, Rotterdam, Netherlands, [5]Erasmus Choice Modelling Centre, Erasmus University Rotterdam, Rotterdam, Netherlands

Aims: The primary aim was to explore the age dependency of health state values derived via trade-offs between health-related quality of life (HRQoL) and life years in a discrete choice experiment (DCE). The secondary aim was to explore if people weigh life years and HRQoL differently for children, adolescents, adults, and older adults.

Methods: Participants from the general population of the Netherlands and China first completed a series of choice tasks offering choices between two EQ-5D-Y states with a given lifespan. The choice model captured the value of a year in full health, disutility determined by EQ-5D-Y, and a discount rate. Next, they received a slightly different choice task, offering choices between two lives that differed in HRQoL and life expectancy but produced the same number of quality-adjusted life years (QALYs). Participants were randomly assigned to fill out the survey for three or four age frames: a hypothetical person of 10, 15, 40, and 70 years (the last one only applicable to China) to allow the age dependency of the responses to be explored.

Results: A total of 1,234 Dutch and 1,818 Chinese people administered the survey. Controlling for time preferences, we found that the agreement of health state values for different age frames was generally stronger in the Netherlands than in China. We found no clear pattern of differences in the QALY composition in both samples. The probability distribution over response options varied most when levels for lifespan or severity were at the extremes of the spectrum.

Conclusion/discussion: The magnitude and direction of age effects on values seemed dimension- and country specific. In the Netherlands, we found a few differences in dimension-specific weights elicited for 10- and 15-year-olds compared to 40-year-olds, but the overall age dependency of values was limited. A stronger age dependency of values was observed in China, where values for 70-year-olds differed strongly from the values for other ages. The appropriateness of using existing values beyond the age range for which they were measured needs to be evaluated in the local context.

KEYWORDS

health state values, age-dependency, health related quality of life, children, EQ-5D-Y

Introduction

In recent years, the demand for pediatric multi-attribute utility instruments has grown (Chen and Ratcliffe, 2015). One of these utility instruments is the EQ-5D-Youth (EQ-5D-Y), a child-friendly version of the well-known adult questionnaire EQ-5D-3L (Wille et al., 2010). It contains the same five health dimensions, although the wording of three of them (i.e., self-care, anxiety, and usual activities) has been modified in order to fit the needs of the younger respondent. A VAS scale is also included, with endpoints of 0 (the worst health you can imagine) and 100 (the best health you can imagine). The EQ-5D-Y questionnaire can be filled out by children from the age of 8, while for children aged 4–7, a proxy version can be applied. EQ-5D-Y value sets are currently available for nine countries (Devlin et al., 2022).

Key challenges in the area of child health valuation are the impact of different perspectives, i.e., adult, adolescent, or child preferences, and the impact of different health state valuation methods (Rowen et al., 2020). The EQ-5D-Y valuation protocol requires that the general population should be asked to value the EQ-5D-Y health states as proxies for children. People no longer value the health state of a person like themselves but of a 10-year-old hypothetical child. To date, it is unknown whether the obtained values will be sensitive to the specified age of the hypothetical child (e.g., a child aged 10 or an adolescent aged 15), and if so, what framing of age is optimal.

The available evidence about the age dependency of health-related quality of life (HRQoL) values is limited. Kind et al. showed that by using the visual analog scale (VAS), the obtained values were lower for children when respondents were asked to imagine that a health state concerned a 10-year-old child compared to when they valued that state for themselves or another adult (Kind et al., 2015). These results suggest that health problems will affect a child's HRQoL more than an adult's HRQoL. However, the Kind's VAS values were obtained on a scale with the best and worst imaginable health states as the top and bottom anchors and not on the full health-dead scale required for the computation of quality-adjusted life years. Kreimeier et al. (2018) reported that TTO values for children exceed those for adults in the same health state. Shah et al. (2020) found the same result across a range of methods that all produced values on the full health-dead scale.

To better understand the age dependency of HRQoL values, we need to carefully examine the context and meaning of responses given to questions, especially when they involve a time trade-off. Because TTO values are derived from a trade-off between HRQoL and time, HRQoL values are confounded with preferences for time. As a result, differences in TTO values for adults and children have a clear interpretation: Are changes in health affecting children's HRQoL less, or are variations in time preferences impacting the difference as well? This issue needs to be investigated further in order to better understand differences in health state values for children, adolescents, adults, and/or older adults and to advance valuation methods.

The main objective of our research was to examine how age impacts the valuation of EQ-5D-Y health states using a discrete choice experiment (DCE) that included a duration attribute. The second objective was to study if there are cultural differences when valuing health states for children, adolescents, or adults. The third objective was to explore if people attach different relative weights to life years and quality of life for children, adolescents, adults, and older adults.

Methods

Strategy

Respondents in the Netherlands were randomized over three arms that only differed by the framing of the valuation task with respect to the age of the hypothetical person that would experience the health states: 10 years (arm 1), 15 years (arm 2), or 40 years (arm 3), representing a child, an adolescent, and an adult. The study in China adopted the same study design as used in the Netherlands and extended it with a fourth study arm focused on older adults over 70 years. This was done to increase the contrast between arms and increase knowledge of the validity and valuation of the EQ-5D in the elderly population. Respondents in both countries completed two tasks. First, they received a series of questions from a discrete choice experiment featuring EQ-5D-Y health states with an associated duration. Next, respondents received a series of questions asking about their preferences for a "QALY composition". Details of both tasks are provided below. Approval for this study was given by the Ethics Committees of the University of Maastricht and the Institutional Review Board of Fudan University School of Public Health before the start of the study. Data collection took place between August and December 2017 in the Netherlands and between May and July 2019 in China.

EQ-5D-Y

EQ-5D-Y is a five-dimensional measure of health-related quality of life, derived from EQ-5D (Wille et al., 2010). The included dimensions are mobility, looking after myself, doing usual activities, having pain or discomfort, and feeling worried, sad, or unhappy. Each dimension has three levels: no problems, some problems, and a lot of problems.

Sample

In the Netherlands, respondents were recruited from a commercial panel "Panelinzicht". Each respondent received an invitation with a link to participate in the survey. To make sure the sample was representative of the Dutch population, stratified sampling was applied. This means that three strata were defined beforehand: age (with 18 years as a minimum age), gender, and education. Based on the classification as used by Statistics Netherlands (Centraal Bureau voor de Statistiek), the eight levels of education were divided into lower, middle, and higher education. In China, the respondents were enrolled by Survey Engine, and quota sampling was used to generate a representative sample of the general adult population in terms of age and gender.

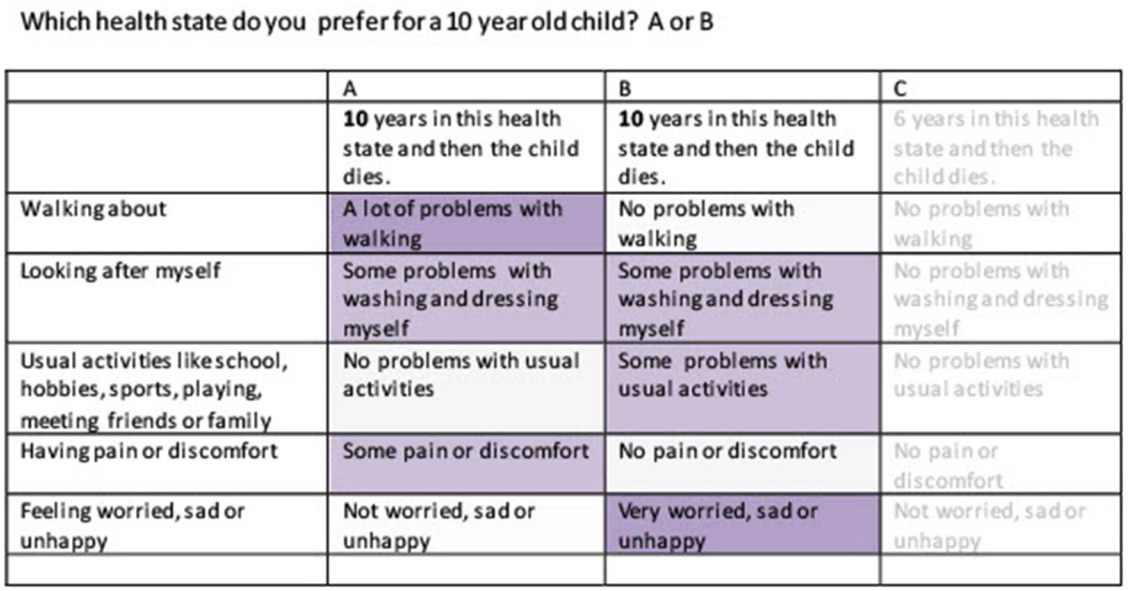


FIGURE 1
Example choice set health state A and B.

The survey

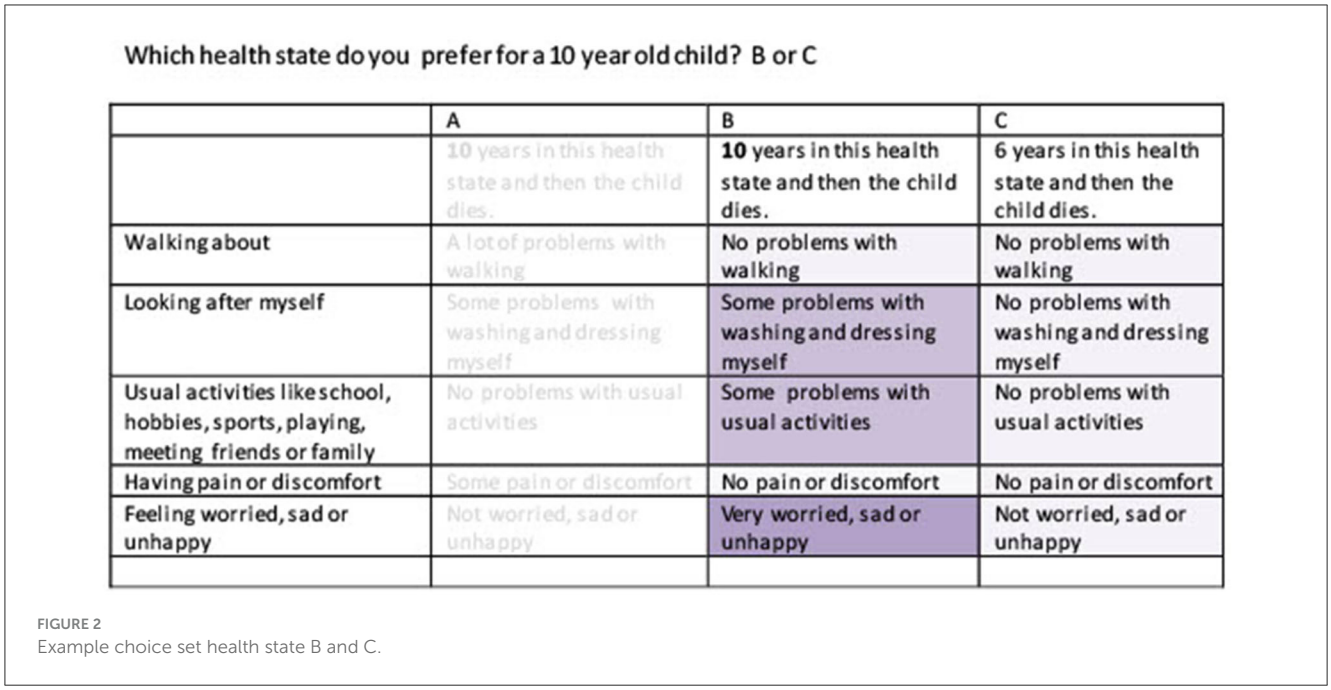
The online survey was developed by Survey Engine in both countries, with the Dutch version being translated into a Chinese version. It started with three questions regarding birth date, gender, and education. Subsequently, respondents were asked to describe their own health based on the EQ-5D-3L and the VAS scale. Then, the objective of the study was explained, and respondents were asked to fill in 15 choice tasks from a discrete choice experiment (DCE). The choice tasks were formatted as matched pairwise choices, following Jonker et al. (2017). This means that they first were asked which of two EQ-5D-Y states, A or B, they preferred for either a 10-year-old child, a 15-year-old adolescent, a 40-year-old adult, or a 70 year old (Chinese version). Both options differed in health but shared an equal life span. Next, they were asked to choose between health states B and C. C represented perfect health, i.e., no problems in any of the five EQ-5D-Y dimensions, but always offered fewer life years compared to B. To make the choice task easier, color coding was applied, with more severe problems darker colored and less severe problems lighter colored (Jonker et al., 2018a). After finishing the DCE, feasibility questions were presented, which means that respondents were asked whether they experienced any difficulties when choosing between A or B and B or C. Examples of both choice sets are presented in Figures 1, 2.

Next, we presented a slightly different choice task, that we dubbed a "QALY composition task". Eight QALY composition tasks were administered. We developed the task to let responses directly tell if people weigh life years and quality of life differently for children, adolescents, adults, or older adults. The QALY composition task involved choices between different ways of achieving a similar QALY total [e.g., life A 2 years in full health (100% QoL) vs. life B 4 years in 50% QoL]. Respondents could indicate their preference for life A or life B on a 5-point Likert scale, varying from a very strong preference for life A to a very strong preference for life B. An example of a QALY task is presented in Figure 3. Eight QALY composition tasks were administered. We developed the task to explore if the relative weights attached to time and HRQoL vary for children, adolescents, adults, or older adults.

At the end of the survey, a number of background questions were asked: employment status, experience in working with children, having children, experience with serious illness of a child, experience with own health during youth, having brother(s) and/or sister(s), experience with serious illness in sibling(s), whether it would have been worse or not if the respondent would have experienced the health states described in the survey instead of the hypothetical 10, 15, 40, or 70-year-olds, what kind of child, adolescent, adult, or older adult they were thinking of when answering the choice tasks, and what kind of religion they belonged to.

Experimental design DCE

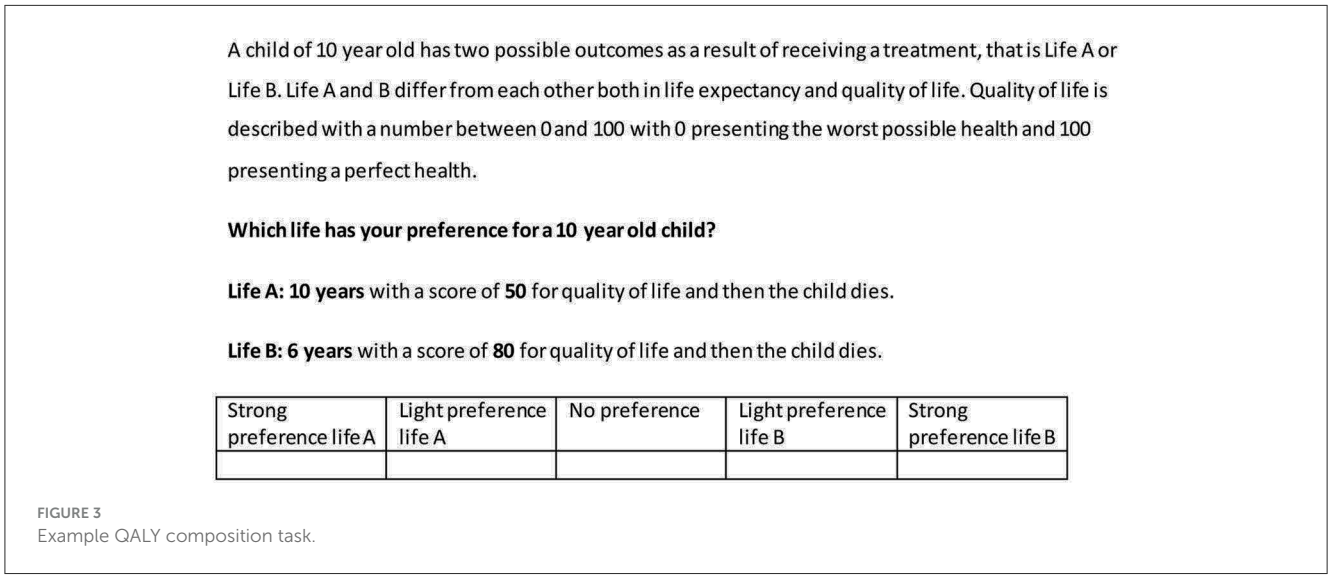
An experimental design with 150 matched pair-wise choice questions was generated using a two-step approach. The EQ-5D-Y states featured as options A and B were selected first, subsequently, option C was added, and in a separate step, the duration levels associated with options A, B, and C were selected. This two-step approach was used to promote consistency with a UK study that used a DCE without duration (Mott et al., 2021 plenary meeting of EuroQoL). Briefly, A and B were selected using an algorithm to create a Bayesian efficient design programmed in Stata. The candidate set was restricted to pairs that had overlapping severity levels in two dimensions. The design accounted for the main effects and two-way interactions. The initial design was created without



Which health state do you prefer for a 10 year old child? B or C

	A	B	C
	10 years in this health state and then the child dies.	**10** years in this health state and then the child dies.	6 years in this health state and then the child dies.
Walking about	A lot of problems with walking	No problems with walking	No problems with walking
Looking after myself	Some problems with washing and dressing myself	Some problems with washing and dressing myself	No problems with washing and dressing myself
Usual activities like school, hobbies, sports, playing, meeting friends or family	No problems with usual activities	Some problems with usual activities	No problems with usual activities
Having pain or discomfort	Some pain or discomfort	No pain or discomfort	No pain or discomfort
Feeling worried, sad or unhappy	Not worried, sad or unhappy	Very worried, sad or unhappy	Not worried, sad or unhappy

FIGURE 2
Example choice set health state B and C.



A child of 10 year old has two possible outcomes as a result of receiving a treatment, that is Life A or Life B. Life A and B differ from each other both in life expectancy and quality of life. Quality of life is described with a number between 0 and 100 with 0 presenting the worst possible health and 100 presenting a perfect health.

Which life has your preference for a 10 year old child?

Life A: 10 years with a score of **50** for quality of life and then the child dies.

Life B: 6 years with a score of **80** for quality of life and then the child dies.

Strong preference life A	Light preference life A	No preference	Light preference life B	Strong preference life B

FIGURE 3
Example QALY composition task.

priors, but data collection was paused two times to allow interim analysis of the data. The obtained coefficients were used as priors to update a design for the next round of data collection. As mentioned above, the C alternative always referred to full health, and hence dominated A and B in terms of quality of life, but was paired with a shorter duration, implying a time trade-off question. The selection of the levels of duration associated with A and B (the same level) and with C (a shorter duration) was also informed by a Bayesian efficient design algorithm (cf.), but this part was programmed in C++ because the utility function accounted for possible non-linearities in preferences for time (i.e., discounting), which standard software packages such as NGENE or STATA could both not handle (Jonker et al., 2018b).

Blocking was applied to divide the 150 matched pairwise choice tasks into 10 blocks, with each block containing 15 pairwise comparisons.

Experimental design QALY composition task

The QALY composition task was constructed on the basis of an orthogonal array. The four variables linked to the orthogonal array were:

1. Life years of a (levels: 2, 4, 6, or 8)
2. Quality of life of a (levels 0.2, 0.4, 0.6, and 0.8)
3. Quality of life of b (levels 0.2, 0.5, and 0.7, 1) and
4. The ratio of total QALYs in a/b (levels 0.8, 1.0, and 1.2).

Together, these four variables were used to define the life years of B, as indicated in Table 1. The scenarios presented to respondents in the QALY composition tasks were defined by the variables in the shaded columns. Three variables were directly obtained from the orthogonal array, and the fourth (life years in option B) was

10.3389/fpsyg.2023.1175402

TABLE 1 Design QALY composition tasks.

Scenarios	Life years A	Qol A	Qol B	QALY multiplier	Life years B*
1	2	0.2	0.25	0.8	1.5
2	4	0.6	1	1	2.5
3	6	0.8	0.5	1	10.0
4	8	0.4	0.7	1.2	5.5
5	4	0.4	0.5	0.8	2.5
6	2	0.8	0.7	1	2.3
7	8	0.6	0.2	1	24.0
8	6	0.2	1	1.2	1.5
9	6	0.6	0.7	0.8	4.0
10	8	0.2	0.5	1	3.0
11	2	0.4	1	1	0.8
12	4	0.8	0.2	1.2	19.0
13	8	0.8	1	0.8	5.0
14	6	0.4	0.2	1	12.0
15	4	0,2	0.7	1	1.2
16	2	0,6	0.5	1.2	3.0

*Numbers are rounded up for convenience.

computed by matching the information of the first three variables with the QALY multiplier. This procedure ensured that decision rules based on longest life, highest quality of life, or maximum number of QALYs would produce different results.

Framing of the survey for the age groups

Exactly, the same DCE design and design of the QALY composition task were used in all arms. The only difference between arms was that respondents were asked to imagine that the health states applied for a different hypothetical person, aged 10, 15, 40, or 70 years. We used the wording of the EQ-5D-Y questionnaire to describe the health states of all arms. Only the examples mentioned between brackets for the dimension usual activities were taken from the adult version of the EQ-5D-3L for the 40-year and 70-year-old arm. For every respondent, randomization was applied per arm, per block, per choice task, and in the left-right order of the health states A and B.

Data analysis

Data quality management

We retained respondents in the sample if they had completed the DCE survey and were not classified as speeders. Speeders were removed from the sample using a speeding threshold set at 530 s for the entire survey. We set this relatively low threshold to account for the fact that choice questions in a DCE repeat much of their content and to avoid undue exclusion of valid responses.

Discrete choice experiment

Logistic regression was used to analyze the respondent's DCE choices (STATA version 14). The parameters of the conditional logit model were estimated using maximum likelihood estimation. Conceptually, the utility that the respondent n obtains from alternative j in a choice task t is computed as the utility obtained from the health state characteristics X_{njt} with their accompanying preference parameters (β_n), multiplied by the net present value (NPV_{njt}) of the number of years T_{njt} associated with that health states, i.e.,

$$U_{njt} = (\beta_n . X_{njt}) . NPV_{njt} + \varepsilon_{njt} \quad (1)$$

An exponential discount function was used to compute NPV (Jonker et al., 2018b), which defines NPV by the discount rate r, i.e.,

$$NPV_{ita} = (1 - \exp(-r\ T_{ita}))/(\exp(r) - 1)\ if\ r \neq 0 \quad (2)$$

Dummy coding was applied for the levels of the EQ-5D-Y with no problems as a reference level. The coefficients from formula 1 that are associated with the dimension severity levels can be converted to the preferred scale for QALY computation, by dividing the relevant β_n by the preference parameter associated with years, based on the Net present value computation.

Feasibility

Feasibility questions for the DCE were analyzed with descriptive statistics in SPSS version 16.

10.3389/fpsyg.2023.1175402

TABLE 2 Characteristics study samples.

	10 year old	15 year old	40 year old	Dutch population	
A. Netherlands					
Age	*N* = 438	*N* = 450	*N* = 346		
18–25	36 (8%)	38 (8%)	28 (8%)	15%	
25–35	49 (11%)	54 (12%)	39 (11%)	15%	
35–45	60 (14%)	55 (12%)	45 (13%)	15%	
45–55	84 (19%)	87 (19%)	63 (18%)	18%	
55–65	80 (18%)	82 (18%)	70 (20%)	16%	
65–75	73 (17%)	77 (17%)	57 (16%)	13%	
>75	56 (13%)	57 (13%)	44 (13%)	9%	
Education					
None	1 (0%)	0 (0%)	2 (1%)		
Lower	95 (22%)	107 (24%)	70 (20%)	31%	
Middle	186 (42%)	185 (41%)	165 (48%)	40%	
High	137 (31%)	139 (31%)	98 (28%)	28%	
Other	19 (4%)	19 (4%)	11 (3%)	1%	
Sex					
Male	192 (44%)	198 (44%)	153 (44%)	51%	
Female	246 (56%)	252 (56%)	193 (56%)	49%	
Having children					
Yes	272 (63%)	278 (63%)	218 (64%)		
No	163 (37%)	164 (7%)	124 (6%)		
	10 year old	15 year old	40 year old	70 year old	Chinese norms
B. China					
Age	*N* = 454	*N* = 455	*N* = 454	*N* = 455	
18–25	83 (18%)	78 (17%)	81 (18%)	81 (18%)	10%
25–35	110 (24%)	111 (24%)	107 (24%)	97 (21%)	17%
35–45	104 (23%)	102 (22%)	106 (23%)	97 (21%)	15%
45–55	64 (14%)	69 (15%)	67 (15%)	79 (17%)	18%
55–65	57 (13%)	55 (12%)	56 (12%)	52 (11%)	11%
65–75	30 (7%)	34 (7%)	35 (8%)	46 (10%)	7%
>75	6 (1%)	6 (1%)	2 (0%)	3 (1%)	4%
Education					
None	2 (0%)	1 (0%)	1 (0%)	1 (0%)	11%
Primary school	3 (1%)	5 (1%)	4 (1%)	4 (1%)	25%
Middle school	15 (3%)	17 (4%)	16 (4%)	20 (4%)	35%
High school	71 (16%)	88 (19%)	80 (18%)	93 (20%)	15%
College and above	363 (80%)	344 (76%)	353 (78%)	337 (74%)	15%
Sex					
Male	225 (50%)	222 (49%)	219 (48%)	239 (53%)	51%
Female	229 (50%)	233 (51%)	235 (52%)	216 (47%)	49%
Having children					
Yes	298 (66%)	281 (62%)	285 (63%)	286 (63%)	
No	79 (17%)	83 (18%)	70 (15%)	93 (20%)	
Unknown/missing	77 (17%)	91 (20%)	99 (22%)	76 (17%)	

QALY composition

The QALY composition task provided ordinal responses on a 5-point Likert scale. By arm, we computed and compared the percentages of responses in each category. We graphically display the results using horizontally stacked bars. Because minimal differences were found, no attempt was made to study differences across arms using non-parametric tests.

Results

Characteristics of the sample

In total, 5,126 Dutch and 4,128 Chinese respondents started the survey, with 1,730 or 2,494 respondents completing it, resulting in a response rate of 34 and 60%, respectively. A total of 496 people were excluded from the Dutch sample as speeders and 676 from the Chinese sample. After these exclusions, the Dutch sample had $N = 438$ respondents in arm 1 (10 years old), $N = 450$ in arm 2 (15 years old), and 346 (40 years old) in arm 3. The final Chinese sample had 454, 455, 454, and 455 respondents in arms 1 (10 years old), 2 (15 years old), 3 (40 years old), and 4 (70 years old), respectively. Sample characteristics are presented in Tables 2A, B. The samples were representative of the populations in terms of sex and age, although the percentage of respondents with lower education in the Netherlands was smaller compared to the population as registered by the Dutch National Bureau Of Statistics (CBS), while in the Chinese sample, the percentage of respondents with college and higher education was much higher compared to Chinese norms (CotSNPCNE, n.d.).

Feasibility

Tables 3A, B shows the answers related to the feasibility questions. In the Netherlands, 53% of the 10-year-old arm felt it was difficult to choose between health states A and B, compared to 45% of the adolescent arm and 34% of the adult arm. In addition, when making a choice between an impaired health state B and perfect health state C but with a shorter life duration, 58% of the respondents in the child arm and 49% in the adolescent arm answered that it was difficult to very difficult compared 43% in the adult arm. On the contrary, respondents across the four arms in China felt the degree of difficulty was similar.

The percentage of respondents answering that their choices would not have been different if they themselves had experienced the health states rather than a hypothetical child, adolescent, adult, or older person, varied across arms in the Netherlands (Table 3A). A total of 62% of the respondents in the adult arm indicated that answering the questions for themselves would have resulted in the same responses, vs. 36% in the child arm and 44% in the adolescent arm. In the child and adolescent arms, 28 and 24% of the people considered health problems or loss of life years *less bad for themselves*, whereas, in the adult arm, respondents more often considered these issues *worse for themselves*. In China, fewer people stated that their responses would have been the same if they were asked about preferences for themselves (11–24% varying across arms), and the majority (varying between 51 and 58%) of the people in all arms state that they would consider health problems or loss of life years worse for themselves (Table 3B).

Results discrete choice experiment

Tables 4A, B shows the results of the regression model on a latent scale for the Netherlands and China. The parameter "years" reflects the additional utility gained from a life year without health problems, before discounting, and is positive—as expected. In both countries, results show that additional life years generate utility. The interaction terms in the Dutch regression model all have the expected negative sign, except mobility level 2, showing that a deviation from full health with no problems is considered negative. The interaction terms for level 2 problems on the dimensions of self-care, usual activities, and pain/discomfort showed unexpected positive signs in China.

The estimated discount rate r varied between 0.22 and 0.25 across the arms in the Netherlands and was ~0.30 in China in all four arms, suggesting strong discounting of future health outcomes.

Figures 4A, B presents the results on a QALY scale (coefficient interaction term divided by coefficient years). Across arms in the Netherlands, we found a high level of agreement on the health state values, except for the dimensions of pain and discomfort and anxiety/depression; respondents traded-off *more* time to avoid these problems for children than for adults. The Chinese results showed that respondents traded-off more time to avoid severe problems in the 70-year arm.

The difference in values for the worst health state (33,333) resulted in −0.630 for children, −0.452 for adolescents, and −0.452 for adults in the Netherlands. On the contrary, older adults in China have a value of −0.870 for the worst state, followed by adolescents (−0.370), children (−0.340), and adults (−0.320).

QALY composition

Figures 5A, B presents the distribution of the Likert responses by QALY composition task. We found no clear pattern of differences across arms in both countries. The distribution over response options varied most when the life years or quality of life were at the extremes of the spectrum. In the Netherlands, the only distinction between arms was that the percentage of responses in the third response category, indicating no preference for A or B, seemed to be the largest when the questions concerned a 10-year-old child. The Chinese results showed a larger percentage of respondents, indicating no preference between life A and life B compared to the Dutch data, with similar or even less clustering in the child's arm on the no preference option.

Discussion

This study examined the impact of framing of age on values for EQ-5D-Y health states in the Netherlands and China. We tested this issue using a DCE duration approach and a task that assessed preferences for QALY composition. The empirical findings indicated that the values derived from the

10.3389/fpsyg.2023.1175402

TABLE 3 Feasibility questions.

	10-year old	15 year old	40 year old	
A. Netherlands				
Difficulty choosing between health state A and B	*N* = 438	*N* = 448	*N* = 346	
Very difficult	55 (13%)	25 (6%)	15 (4%)	
Difficult	176 (40%)	175 (39%)	103 (30%)	
Neutral	134 (31%)	171 (38%)	164 (47%)	
Easy	71 (16%)	67 (15%)	57 (16%)	
Very easy	2 (0%)	10 (2%)	7 (2%)	
Difficulty choosing between health state B and C	*N* = 438	*N* = 448	*N* = 346	
Very difficult	82 (19%)	66 (15%)	40 (12%)	
Difficult	172 (39%)	155 (35%)	110 (32%)	
Neutral	116 (26%)	126 (28%)	93 (27%)	
Easy	56 (13%)	79 (18%)	85 (24%)	
Very easy	12 (3%)	22 (5%)	18 (5%)	
Would your choices have been different if the health problems concerned yourself instead of a hypothetical person?	*N* = 435	*N* = 441	*N* = 341	
Yes, health problems for myself worse	16 (4%)	14 (3%)	23 (7%)	
Yes, loss life years for myself worse	16 (4%)	20 (5%)	25 (7%)	
Yes, health problems for myself less bad	59 (14%)	47 (11%)	10 (3%)	
Yes, loss life years for myself less bad	61 (14%)	57 (13%)	14 (4%)	
No, health problems for myself equally bad	104 (24%)	143 (32%)	163 (48%)	
No, loss life years for myself equally bad	53 (12%)	55 (12%)	49 (14%)	
I do not know	126 (29%)	105 (24%)	57 (17%)	
B. China				
Difficulty choosing between health state A and B	*N* = 390	*N* =3 76	*N* = 368	*N* = 389
Very difficult	21 (5%)	22 (6%)	11 (6%)	19 (5%)
Difficult	94 (24%)	87 (23%)	88 (23%)	86 (22%)
Neutral	126 (32%)	113 (30%)	114 (30%)	119 (31%)
Easy	113 (29%)	122 (32%)	118 (32%)	127 (33%)
Very easy	36 (9%)	32 (9%)	37 (9%)	38 (10%)
Difficulty choosing between health state B and C	*N* = 390	*N* = 376	*N* = 368	*N* = 389
Very difficult	23 (6%)	28 (7%)	27 (7%)	27 (7%)
Difficult	97 (25%)	81 (22%)	66 (22%)	86 (22%)
Neutral	104 (27%)	99 (26%)	101 (26%)	85 (22%)
Easy	126 (32%)	121 (32%)	140 (32%)	138 (35%)
Very easy	40 (10%)	47 (13%)	34 (13%)	53 (14%)
Would your choices have been different if the health problems concerned yourself instead of a hypothetical person?	*N* = 373	*N* = 361	*N* = 355	*N* = 374

(Continued)

10.3389/fpsyg.2023.1175402

TABLE 3 (Continued)

	10 year old	15 year old	40 year old	70 year old
B. China				
Yes, health problems for myself worse	88 (24%)	96 (27%)	27% (104)	98 (26%)
Yes, loss life years for myself worse	111 (30%)	95 (26%)	26% (101)	119 (32%)
Yes, health problems for myself less worse	39 (10%)	34 (9%)	9% (44)	39 (10%)
Yes, loss life years for myself less worse	32 (9%)	31 (9%)	9% (27)	40 (11%)
No, health problems for myself equally worse	60 (16%)	56 (16%)	16% (43)	28 (7%)
No, loss life years for myself equally worse	22 (6%)	28 (8%)	8% (17)	16 (4%)
I do not know	21 (6%)	21 (6%)	6% (19)	34 (9%)

TABLE 4A Results non-linear preferences on a latent scale Dutch population.

	10 year old		15 year old		40 year old	
	Coefficient	95% CI	Coefficient	95% CI	Coefficient	95% CI
Years	1.04	0.89; 1.19	1.12	0.98; 1.25	1.05	0.90; 1.20
Mo2*years	0.05	0.02; 0.08	0.06	0.04; 0.09	0.07	0.04; 0.11
Mo3*years	−0.07	−0.10; −0.05	−0.06	−0.09; −0.04	−0.09	−0.12; −0.06
Sc2*years*	−0.01	−0.04; 0.01	0.00	−0.02; 0.03	−0.01	−0.04; 0.02
Sc3*years	−0.12	−0.15; −0.09	−0.13	−0.15; −0.10	−0.14	−0.17; −0.11
Ua2*years	−0.06	−0.09; −0.04	−0.07	−0.10; −0.05	−0.05	−0.08; −0.03
Ua3*years	−0.32	−0.37; −0.28	−0.32	−0.36; −0.28	−0.28	−0.32; −0.24
Pd2*years	−0.14	−0.18; −0.11	−0.10	−0.13; −0.08	−0.08	−0.11; −0.06
Pd3*years	−0.54	−0.61; −0.47	−0.49	−0.54; −0.43	−0.44	−0.50; −0.38
Ad2*years	−0.19	−0.22; −0.15	−0.17	−0.20; −0.14	−0.16	−0.19; −0.12
Ad3*years	−0.64	−0.72; −0.56	−0.63	−0.69; −0.56	−0.57	−0.64; −0.49
Discount rate	0.25	0.22; 0.28	0.23	0.20; 0.25	0.22	0.19; 0.25

*Indicate that this is an interaction between the domain like mobility and years as described under the heading results discrete choice experiment.

TABLE 4B Results non-linear preferences on a latent scale Chinese population.

	10 year old		15 year old		40 year old		70 year old	
	Coefficient	95% CI	Coefficient	95% CI	Coefficient	95% CI	Coefficient	95% CI
Years	0.27	0.23; 0.31	0.29	0.25; 0.33	0.33	0.29; 0.37	0.23	0.19; 0.27
Mo2*years	−0.01	−0.03; 0.02	−0.01	−0.03; 0.02	0.01	−0.02; 0.03	−0.01	−0.04; 0.02
Mo3*years	−0.08	−0.10; 0.05	−0.10	−0.14; 0.06	−0.11	−0.15; −0.07	−0.09	−0.13; −0.05
Sc2*years	0.04	0.01; 0.08	0.03	0.02; 0.08	0.06	0.03; 0.09	0.02	−0.01; 0.05
Sc3*years	−0.02	−0.04; 0.03	−0.02	−0.05; 0.01	−0.05	−0.08; −0.02	−0.06	−0.10; −0.02
Ua2*years	0.01	−0.02; 0.03	0.01	−0.02; 0.05	0.03	0.00; 0.06	0.03	−0.00; 0.06
Ua3*years	−0.04	−0.08; 0.00	−0.05	−0.08; 0.00	−0.05	−0.09; −0.01	−0.05	−0.09; −0.02
Pd2*years	0.01	−0.02; 0.03	0.01	−0.02; −0.03	0.00	−0.02; 0.03	0.02	−0.01; 0.05
Pd3*years	−0.08	−0.13; 0.00	−0.10	−0.14; 0.04	−0.10	−0.14; −0.06	−0.10	−0.12; −0.06
Ad2*years	−0.00	−0.03; 0.02	−0.03	−0.06; 0.00	−0.00	−0.03; 0.02	−0.03	−0.04; 0.00
Ad3*years	−0.14	−0.17; −0.11	−0.13	−0.17; −0.09	−0.13	−0.18; −0.08	−0.13	−0.17; −0.10
Discount rate	0.29	0.25; 0.33	0.33	0.28; 0.38	0.32	0.28; 0.36	0.32	0.27; 0.37

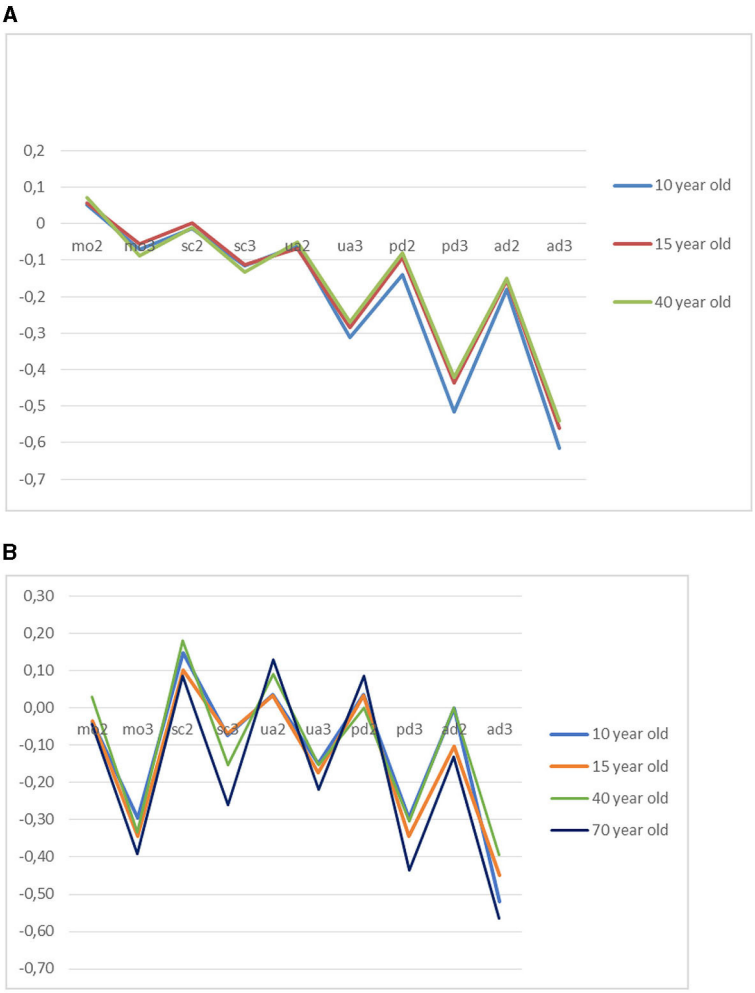


FIGURE 4
(A) Utility decrements per EQ-5D-Y dimension severity level in the Netherlands. **(B)** Utility decrements per EQ-5D-Y dimension severity level in China.

DCE duration task were similar in the Netherlands for children, adolescents, and adults (except for "pain") and varied more for children, adolescents, adults, and older adults in China, where the lowest values were found in the group stating preferences for a 70-year-old person. Dutch people comparatively traded-off more time to avoid pain in children than for adults, resulting in lower values, while Chinese people were more willing to trade-off time to avoid any type of severe problem in the elderly compared to the other arms. The QALY composition task showed no clear differences in values for health across age groups.

No evidence for age dependency of health state values was found in the Netherlands. Our results for the 10-year-old arm are consistent with Kreimeier's TTO results (Kreimeier et al., 2018). Based on international results, Kreimeier reported that TTO values applied to children generally were higher compared to values of adults, but in that study, the Dutch results were an exception. In the Netherlands, people gave a lower TTO value to a health state when it concerned a 10 years old compared to themselves (Kreimeier et al., 2018). This indicates that Dutch respondents are prepared to trade-off life years against the quality of life for children. In our research, the results also showed that respondents were prepared to trade-off more time to avoid pain in children than in adults, resulting in lower values, although generally, the agreement of health state values for different ages was quite strong. While the congruence between studies supports the validity of our findings, care should still be taken when generalizing our results to other countries. Stronger evidence for age dependency of values was found in China, where the inclusion of the 70-year arm increased the contrast between groups.

Our estimation of health state utilities followed a state-of-the-art DCE duration approach, requiring a multiplicative utility function that involves a non-linear discount function. The estimated discount rates indicated that respondents valued quality of life in the short term more compared to the long term, which was anticipated, and as argued by Jonker and Bliemer (2019), valid health state utility values can only be obtained if the model adequately accounts for such time preferences. The estimated discount rates were, however, relatively high when compared to the standard rates usually applied in economic evaluations, especially in China. While the discount rates were still within the range of previously estimated discount rates for health-related outcomes

A

1=children,2=adolescents,3=adults

B

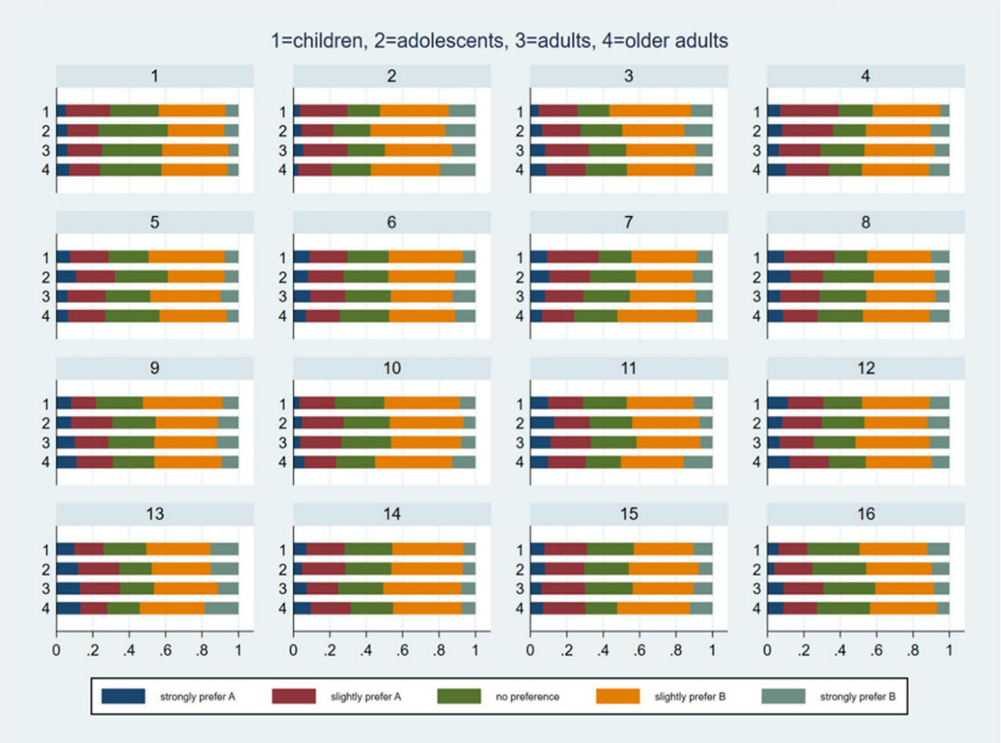


FIGURE 5
(A) Distribution of likert responses by scenario in the Netherlands. **(B)** Distribution of likert responses by scenario in China.

(Attema et al., 2018), their reliability needs to be established in the future research. A limitation of the DCE duration method is that the best way to account for time preferences, especially in the presence of discounting, has not been identified. Discount rates can be computed in different ways. Models that account for non-linear time preferences are complex and have not been implemented yet in the standard software that we used for choice modeling, and this limits the modeling options (e.g., we cannot simultaneously account for preference heterogeneity and for non-linear time preferences). Furthermore, this way of assessing preferences places high demands on the design, necessitating interim design updates to ensure that the design is based on adequate priors, and the end results may still depend on the data quality obtained along the way. We excluded speeders *post-hoc*, not before design updates.

If we are examining preferences for a subject like a trade-off between life years and quality of life, we also need to carefully consider what advantages and disadvantages different valuation methods may have when used in such a context. We consider it possible that the use of TTO poses even greater challenges than the DCE of the required accuracy in rating health states and direct assessment. A specific result that may be worth noting is the larger clustering of responses in the child arm vs. the other arms on the no preference answer option in the QALY composition task in the Netherlands. This might indicate that a larger fraction of respondents in the child arm feel uncertain when trading-off quality of life and life years. However, it is also possible that respondents are neutral about their preference for either one of the options and consider them equivalent. Either way, it shows that more respondents in the child's arm were reserved when making a choice. However, it appears that the Chinese results showed a reversed pattern, with more respondents in the child's arms who were more certain to make a decision. The possible explanation may be a cultural difference: paternalism is more prevalent in China.

The findings of this study may be taken into consideration for future updates of the EQ-5D-Y valuation protocol. EQ-5D-Y values are currently elicited from adults who value health states accruing to a 10-year-old child (Ramos-Goñi et al., 2020). This study reflects on the appropriateness of using a specified age (here, 10 years of age) in the elicitation of values that are used across a wider age group by varying the specified age. Age dependency of values was limited in the Netherlands, suggesting that values elicited for a 10-year-old child may also be validly applied for a 15-year-old. However, in China, the values for 70-year-olds differed strongly from the values for other ages, suggesting that the appropriateness of using a fixed, specified age may be questioned. Moreover, many respondents indicated that their choices would have been different if the health state had been experienced by themselves rather than by someone else. This finding is in line with results from other studies (Lipman et al., 2021; Reckers-Droog et al., 2022). More research on the sensitivity of values to age and perspective is warranted.

Conclusion

Age dependency was observed in the stated preferences for hypothetical health states. The magnitude and direction of age effects in values seemed dimension- and country-specific. In the Netherlands, we found a few differences in dimension-specific weights elicited for 10- and 15-year-olds compared to 40-year-olds, but the overall age dependency of values was limited. A stronger age dependency of values was observed in China, where values for 70-year-olds differed strongly from the values for other ages. The appropriateness of using existing values beyond the age range for which they were measured needs to be evaluated in the local context.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committees of the University Maastricht and Institutional Review Board of Fudan University School of Public Health. Written informed consent from the participants was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

BE, PW, ES, MFJ, SE, MJ, and CD: conceptualization and writing and critical review. BE, PW, MFJ, and ES: design and analysis. All authors contributed to the article and approved the submitted version.

Funding

This study was made possible through funding provided by the EuroQol Research Foundation (project number: EQ-2016740).

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

Attema, A. E., Brouwer, W. B. F., and Claxton, K. (2018). Discounting in economic evaluations. *Pharmacoeconomics*. 36, 745–758. doi: 10.1007/s40273-018-0672-z

Chen, G., and Ratcliffe, J. A. (2015). Review of the development and application of generic multi-attribute utility instruments for paediatric populations. *PharmacoEconomics*. 33, 1013–1028. doi: 10.1007/s40273-015-0286-7

CotSNPCNE (n.d.). *Attainment CotSNPCNE*. Available online at: https://www.gov.cn/guoqing/2021-05/13/content_5606149.htm (accessed February 16, 2023).

Devlin, N., Roudijk, B., Viney, R., and Stolk, E. (2022). EQ-5D-Y-3L value sets, valuation methods and conceptual questions. *PharmacoEconomics*. 40, 123–127. doi: 10.1007/s40273-022-01226-7

Jonker, M. F., Attema, A. E., Donkers, B., Stolk, E. A., and Versteegh, M. M. (2017). Are health state valuations from the general public biased? A test of health state reference dependency using self-assessed health and an efficient discrete choice experiment. *Health Econ*. 26, 1534–1547. doi: 10.1002/hec.3445

Jonker, M. F., and Bliemer, M. C. J. (2019). On the optimization of Bayesian D-efficient discrete choice experiment designs for the estimation of QALY tariffs that are corrected for nonlinear time preferences. *Value Health J. Int. Soc. Pharmacoecon. Outcomes Res*. 22, 1162–1169. doi: 10.1016/j.jval.2019.05.014

Jonker, M. F., Donkers, B., de Bekker-Grob, E. W., and Stolk, E. A. (2018a). Effect of level overlap and color coding on attribute non-attendance in discrete choice experiments. *Value Health J. Int. Soc. Pharmacoecon. Outcomes Res*. 21, 767–771. doi: 10.1016/j.jval.2017.10.002

Jonker, M. F., Donkers, B., de Bekker-Grob, E. W., and Stolk, E. A. (2018b). Advocating a paradigm shift in health-state valuations: the estimation of time-preference corrected QALY tariffs. *Value Health J. Int. Soc. Pharmacoecon. Outcomes Res*. 21, 993–1001. doi: 10.1016/j.jval.2018.01.016

Kind, P., Klose, K., Gusi, N., Olivares, P. R., and Greiner, W. (2015). Can adult weights be used to value child health states? Testing the influence of perspective in valuing EQ-5D-Y. *Qual. Life Res. Int. J. Qual. Life Aspects Treat. Care Rehabil*. 24, 2519–2539. doi: 10.1007/s11136-015-0971-1

Kreimeier, S., Oppe, M., Ramos-Goñi, J. M., Cole, A., Devlin, N., Herdman, M., et al. (2018). Valuation of EuroQol five-dimensional questionnaire, youth version (EQ-5D-Y) and EuroQol five-dimensional questionnaire, three-level version (EQ-5D-3L) health states: the impact of wording and perspective. *Value Health*. 21, 1291–1298. doi: 10.1016/j.jval.2018.05.002

Lipman, S. A., Reckers-Droog, V. T., Karimi, M., Jakubczyk, M., and Attema, A. E. (2021). Self vs. other, child vs. adult. An experimental comparison of valuation perspectives for valuation of EQ-5D-Y-3L health states. *Eur. J. Health Econ. HEPAC Health Econ. Prevent. Care*. 22, 1507–1518. doi: 10.1007/s10198-021-01377-y

Mott, D. J., Shah, K. K., Ramos-Goñi, J. M., Devlin, N. J., and Rivero-Arias, O. (2021). Valuing EQ-5D-Y-3L health states using a discrete choice experiment: do adult and adolescent preferences differ? *Med. Decision Making* 41, 584–596.

Ramos-Goñi, J. M., Oppe, M., Stolk, E., Shah, K., Kreimeier, S., Rivero-Arias, O., et al. (2020). International valuation protocol for the EQ-5D-Y-3L. *PharmacoEconomics*. 38, 653–663. doi: 10.1007/s40273-020-00909-3

Reckers-Droog, V., Karimi, M., Lipman, S., and Verstraete, J. (2022). Why do adults value EQ-5D-Y-3L health states differently for themselves than for children and adolescents: a think-aloud study. *Value Health J. Int. Soc. Pharmacoecon. Outcomes Res*. 25, 1174–1184. doi: 10.1016/j.jval.2021.12.014

Rowen, D., Rivero-Arias, O., Devlin, N., and Ratcliffe, J. (2020). Review of valuation methods of preference-based measures of health for economic evaluation in child and adolescent populations: where are we now and where are we going? *PharmacoEconomics*. 38, 325–340. doi: 10.1007/s40273-019-00873-7

Shah, K. K., Ramos-Goñi, J. M., Kreimeier, S., and Devlin, N. J. (2020). An exploration of methods for obtaining 0 = dead anchors for latent scale EQ-5D-Y values. *Eur. J. Health Econ. HEPAC Health Econ. Prevent. Care*. 21, 1091–1103. doi: 10.1007/s10198-020-01205-9

Wille, N., Badia, X., Bonsel, G., Burstrom, K., Cavrini, G., Devlin, N., et al. (2010). Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. *Qual. Life Res. Int. J. Qual. Life Aspects Treat. Care Rehabil*. 19, 875–886. doi: 10.1007/s11136-010-9648-y

frontiers | Frontiers in Public Health

TYPE Original Research
PUBLISHED 23 November 2023
DOI 10.3389/fpubh.2023.1184756

OPEN ACCESS

EDITED BY
Pei Wang,
Fudan University, China

REVIEWED BY
Martin Musalek,
Charles University, Czechia
Kenneth Land,
Duke University, United States

*CORRESPONDENCE
Jinmei Fu
✉ fujinmei1976@163.com

RECEIVED 12 March 2023
ACCEPTED 08 November 2023
PUBLISHED 23 November 2023

CITATION
Lv W, Fu J, Zhao G, He Z, Sun S, Huang T, Wang R, Chen D and Chen R (2023) A cohort study of factors influencing the physical fitness of preschool children: a decision tree analysis. *Front. Public Health* 11:1184756. doi: 10.3389/fpubh.2023.1184756

A cohort study of factors influencing the physical fitness of preschool children: a decision tree analysis

Wendi Lv[1,2], Jinmei Fu[3]*, Guanggao Zhao[2], Zihao He[4], Shunli Sun[3], Ting Huang[5], Runze Wang[6], Delong Chen[2] and Ruiming Chen[2]

[1]College of Physical and Health, Jiangxi University of Chinese Medicine, Nanchang, China, [2]School of Physical Education, Nanchang University, Nanchang, China, [3]Jiangxi Sports Science and Medicine Center, Nanchang, China, [4]School of Sport Science, Beijing Sport University, Beijing, China, [5]School of Physical Education and Sport Science, Fujian Normal University, Fuzhou, China, [6]PLA Army Academy of Artillery and Air Defense, Nanjing, China

Objective: Based on the decision tree model, to explore the key influencing factors of children's physical fitness, rank the key influencing factors, and explain the complex interaction between the influencing factors.

Methods: A cohort study design was adopted. 1,276 children (ages 3–6) from 23 kindergartens in Nanchang, China, were chosen for the study to measure the children's physical fitness at baseline and a year later and to compare the physical fitness scores at the two stages. The study was conducted following the Chinese National Physical Fitness Testing Standard (Children Part); To identify the primary influencing factors of changes in physical fitness, a decision tree model was developed, and a questionnaire survey on birth information, feeding patterns, SB, PA, dietary nutrition, sleep, parental factors, and other relevant information was conducted.

Results: The levels of physical fitness indicators among preschool children showed a significant increase after 1 year. The accuracy of the CHAID model is 84.17%. It showed that 7 variables were strongly correlated with the physical changes of children's fitness, the order of importance of each variable was weekend PA, weekend MVPA, mother's BMI, mother's sports frequency, father's education, mother's education, and school day PA. Three factors are related to PA. Four factors are related to parental circumstances. In addition to the seven important variables mentioned, variables such as breakfast frequency on school day, puffed food, frequency of outing, school day MVPA, parental feeling of sports, father's occupation, and weekend breakfast frequency are all statistically significant leaf node variables.

Conclusion: PA, especially weekend PA, is the most critical factor in children's physical fitness improvement and the weekend MVPA should be increased to more than 30 min/d based on the improvement of weekend PA. In addition, parental factors and school day PA are also important in making decisions about changes in fitness for children. The mother's efforts to maintain a healthy BMI and engage in regular physical activity are crucial for enhancing the physical fitness of children. Additionally, other parental factors, such as the parents' educational levels and the father's occupation, can indirectly impact the level of physical fitness in children.

KEYWORDS
decision tree, preschool, physical fitness, physical activity, cohort study

1 Introduction

Early childhood motor development is very active, especially in spatial orientation and coordination. Therefore preschool age is usually named the golden age of motoric. The behavioral habits formed during this period, especially the habit of physical activity (PA) started during this period, have an essential influence on the habit patterns in adulthood (1). Moreover, the development of physical fitness in preschoolers will not only affect the level of fitness in the majority (2, 3) and is even closely related to the risk of disease development in adulthood (4, 5). Therefore, a targeted approach to physical fitness promotion in preschoolers is significant.

However, five times the National Physical Fitness Testing Standard (Children Part) data show that the proportion of children aged 3 to 6 years old who do not reach the qualified level is as high as 6.4 ~ 14.3%. In addition, the physical development of preschoolers shows unevenness in different indicators. Compared with the National Physical Fitness Testing Standard (Children Part) data in 2014, the average levels of height, sitting height, weight, chest circumference, and balance beam walking for both boys and girls in 2020 have increased. Still, the double-leg timed hoping, sit and reach test, and standing long jump are also decreasing. Although the report did not provide a correlation statistical description, the results described are sufficient to attract attention from society and the country (6). To develop a scientific strategy for fostering early childhood fitness and to target the work on early childhood fitness, the prerequisite is to grasp the main influencing factors of early childhood fitness development.

Studies have shown many factors are associated with physical health. PA has a positive effect in enhancing the physical fitness levels of young children (7–10), The researchers discovered a substantial positive association between VPA, MVPA, and physical fitness. These findings imply that promoting MVPA may yield enduring advantages for body composition and physical fitness. Compared to PA, SB has negative effects on health (11, 12), Sedentary behavior is commonly observed among young children, and prolonged sedentary time is associated with higher body fat composition. On the contrary, maintaining a healthy dietary nutrition during early childhood has a positive impact on reducing the risk of obesity and promoting the physical fitness of young children (13, 14). Research substantiated that genetic factors not only play a significant role in determining body type and physical adaptability but also have a crucial impact on the level of physical health and development (15, 16). Some studies have found that sleep plays a crucial role in children's health, and there is a close association between obesity, sleep, and physical activity in preschool children.

Apart from sleep, several studies have indicated that breastfeeding reduces the risk of childhood overweight and obesity compared to non-breastfed children (17, 18). In the process of child development, the influence of parents on their children cannot be overlooked. There is a correlation between parental-reported physical activity levels of preschool children and objectively measured physical health and body composition. The environment and support provided by parents play a crucial role in promoting the development of children's motor skills (19, 20). Among the many factors, identifying the key factors influencing changes in physical fitness and grasping their potential interactions have become the focus of academic attention.

Existing studies have mostly used ANOVA, correlation analysis, and logistic regression to explore the degree of influence of different factors on the physical fitness of preschoolers. However, traditional models can be challenging to implement for individuals without a strong statistical background. Furthermore, interpreting the results of these regression models can be complex. as these methods lack deep data mining and decision analysis and are vulnerable to covariance problems (21, 22). In recent years, scholars have used decision trees in data mining techniques to conduct functional explorations in many fields (23–32). As an important classification technique in data mining, decision trees are capable of effectively identifying the main influencing factors of physical changes. They can optimally partition samples based on various types of variables, automatically determining the classification based on the significance of tests, and overcoming collinearity issues. Additionally, decision trees utilize tree diagrams to illustrate the interaction between variables at different levels (21, 28). Decision trees have been applied in beneficial explorations in the fields of physical fitness and exercise training, such as the decision-making of human muscle strength indicators (24), analysis of influencing factors in practical teaching of public physical education courses in universities (26), key influencing factor analysis of overweight and obesity in young children (28), as well as the analysis of factors influencing physical activity in children and adolescents during the COVID-19 pandemic (29).

Given the advantages of decision trees in terms of optimal sample partitioning, overcoming covariance, and mining interaction relationships. We conducted an accurate and visually representative decision tree model, which garnered attention and recognition from scholars. This not only further enhanced the research framework for studying children's physical fitness but also provided important decision-making references for comprehensive promotion of their physical health (33). However, the decision tree model established in this study was based on cross-sectional surveys, where data was collected by measuring different indicators for individuals at the same time point. Consequently, it cannot establish explicit causal relationships between variables. To further improve the research framework on factors influencing young children's physical fitness, this study obtained physical fitness data of young children over two consecutive years through a cohort study. The overall physical fitness scores of the same child over the two-year period were compared using age-based scoring. A decision tree was then employed to identify key factors contributing to changes in young children's physical fitness, offering important evidence-based recommendations for the development of tailored physical fitness promotion strategies suited to the needs of families, kindergartens, and society.

2 Materials and methods

2.1 Participants

The study was a cohort study, using a whole-group stratified random sampling method. The sampling of participants was based on the public welfare activity of "Concerning the physical fitness of preschoolers and taking care of children's health" of Nanchang Sports Bureau, China. A total of 23 kindergartens were selected in each of the 6 districts and 3 counties in Nanchang, with a total of 5,870 participants aged 3–6 years old, with an average age of 4.05 ± 0.86 years. All participants were healthy and had no motor impairment. All participants had parental consent to participate in the study and signed an informed consent form. All participants underwent two early childhood physical fitness tests at baseline and 1 year later, during which questionnaires were administered. Daily habits were maintained during the testing period. At the end of the test, those who did not complete the two stages of the physical fitness test and those who did not complete or completed the questionnaire were excluded, and the remaining 1,276 children who participated effectively throughout the test were used as participants for statistical analysis of the data in this cohort study (Table 1).

2.2 Physical fitness test

The physical fitness test and comprehensive rating were conducted according to the test method and scoring standard of "National Physical Fitness Testing Standard (Children Part)" from China using the designated equipment for national physical fitness testing. The test indexes included physical form indexes and physical quality indexes, including standard weight (reflecting body development and nutritional status), height (reflecting longitudinal skeletal growth), 10-m shuttle run (reflecting body agility), standing long jump (reflecting lower body explosive power and body coordination), tennis throw (reflecting upper body and core muscle strength), double-leg timed hop (reflecting coordination and lower limb muscle strength), sit and reach testing (reflecting trunk and lower limb flexibility), and balance beam walking (reflecting balance ability), with timing accuracy of 0.1 s, measurement accuracy of 0.1 cm, and weight accuracy of 0.1 kg. These 8 indicators were scored separately based on different age groups, with each item having a maximum score of 5 points, totaling 40 points. The comprehensive grading criteria were as follows: excellent (>31 points), good (28–31 points), pass (20–27 points), fail (<20 points). Subsequently, the difference in the overall physical fitness score between the two stages was calculated to determine and analyze the changes in children's physical fitness, specifically whether there was improvement, no change, or decrease. Taking into account the actual sample conditions and the requirements of the decision tree model used in this study, the target variable (difference in physical fitness scores) was categorized into 2 levels of: "improvement (including no change)" (≥0 points) and "decrease" (<0 points).

TABLE 1 The basic information of participants.

Gender	Population/n	Height/cm	Weight/kg
Male	707	110.26 ± 5.98	19.77 ± 3.37
Female	569	108.87 ± 5.68	18.72 ± 2.98

2.3 Questionnaires

The questionnaire referred to the content of the Questionnaire on Health Behavior of parental and Preschool Children from the Excellent Academic Leaders Program of Shanghai, China (12XD1404500), and used the Delphi method to screen the indicator variables by integrating the opinions of 10 expert teachers from Shanghai Institute of Physical Education, Nanchang University, and Jiangxi Institute of Sports Science. The questionnaire included six primary indicators of young children's birth information (gender, birth length, birth weight), feeding patterns (feeding mode within 4 months after birth, caregiver, attend early childhood classes), PA (school day SB, weekend SB, school day PA, weekend PA, school day MVPA, weekend MVPA, frequency of outing, frequency of participating in sports organization activities every week), dietary nutrition [fruits, vegetables, candy, carbonated soft drinks, milk, fish and shrimp, eggs, hamburgers or hot dogs, pizza, instant noodles, puffed food (d/w)], breakfast frequency on school day, weekend breakfast frequency [vitamin D, calcium, cod liver oil, zinc (d/w)], sleep (total sleep time, sleep time on school day, weekend Sleep Time, time to fall asleep, frequency of getting enough sleep in a week), and parental factors (age of father and mother, BMI of father and mother, type of registered residence of father and mother, education background of father and mother, income of father and mother, parental occupation, parental working days SB, parents' weekend days SB, parental frequency of vigorous activity, parental sports frequency in every weekend, parental feeling of sports, marriage and upbringing of biological parents), which were divided into 59 secondary indicators (Table 2). The questionnaire was validity scored by five early childhood health experts, with a mean score of 96.20. The questionnaire was tested for reliability using the retest method, and 812 children tested in three kindergartens were selected to repeat the survey with a 2-week interval, and the intra-group correlation coefficient (ICC) was used to test, yielding a reliability coefficient of 0.86, which met the statistical requirements.

Questionnaires were distributed to parental of preschool children with the assistance of the Jiangxi Institute of Sports Science, Nanchang Sports Bureau and the kindergartens where the

TABLE 2 The basic information of participants (baseline).

Age	Boy		Girl		Total	
	n	%	*n*	%	*n*	%
3	162	12.7%	142	11.1%	304	23.8%
3.5	167	13%	132	10.3%	299	23.4%
4	193	15.1%	149	11.7%	342	26.8%
4.5	154	12.1%	125	9.8%	279	21.8%
5	28	2.2%	16	1.3%	44	3.4%
5.5	3	0.2%	5	0.4%	8	0.63%
Total	707	55.4%	569	44.5%	1,276	100%

participants were enrolled, and were brought back by parents to be filled out and collected the next day. A professional data entry company processed the collected questionnaires, and the questionnaire results were entered into the database using EpiData software. Double-entry and logical error checking were performed to eliminate those who did not complete the questionnaires or had invalid questionnaire options. The number of valid questionnaires was counted after processing.

2.4 Statistical analyses

The results of the physical fitness and questionnaire survey data of preschool children were entered using EpiData 3.1 software with dual input. With the help of SPSS 22.0 data software, the information from physical fitness test and questionnaire survey were matched and integrated to eliminate invalid data. The software IBM SPSS modeler was used to create a decision tree model (parameters to be set as follows: the maximum depth of the tree structure is 5, the minimum number of cases of nodes influencing the change in the physical fitness of toddlers is 100, the minimum number of instances of sub-nodes is 50, and the minimum change value of Gini coefficient is 0.0001), and a 10-level crossover was used to verify the accuracy of the model identification. The Chi-squared Automatic Interaction Detector (CHAID), also known as the chi-square automatic interaction detection algorithm, is a decision tree technique based on adjusted significance testing (Bonferroni test), uses the chi-squared test to evaluate the significance of variable grouping on the target variable and selects the splitting point that maximizes the chi-squared value as the optimal splitting point (21–33). Due to the characteristics of large sample size, multiple indicators, this study used the significance of chi-square tests to automatically determine the grouping variables and split values of the multivariate contingency table, thereby quickly and effectively identifying the main influencing factors. The target variables in the model, which is the change in physical fitness, were divided into 2 categories according to the research needs, with children with improved (or unchanged) physical fitness (≥0) marked as "1" and children with decreased physical fitness marked as "2" (<0). The categorical variables were assigned values, and the best cut-off points for the remaining continuous variable types were identified and split by the decision tree model.

3 Results

3.1 Results of the physical fitness test

This study collected 1,276 valid samples, of which 707 were boys, accounting for 55.4%; there are 569 girls, accounting for 44.5%. There are 304, 299, 342, 279, 44, and 8 children aged 3, 3.5, 4, 4.5, 5, and 5.5, respectively. The average age is 3.8 ± 0.6 years, respectively, ranging from 3 to 5.5 years. In this study, data comparison revealed that the levels of all eight physical fitness indicators were significantly higher ($p < 0.01$) in preschool children in the second stage compared with the first stage (Table 3), and the scores of sit and reach testing, tennis throw, standing long jump, double-leg timed hop, and balance beam walking were significantly higher (Table 4).

3.2 Construction of decision tree model

The decision tree model created in this study has 5 layers and 41 leaf nodes (Figure 1), and the decision tree model shows that the root node variable is weekend PA. Therefore weekend PA is the most important influencing factor on the change in physical fitness of preschool children. The 1st layer of the tree structure splits the nodes according to weekend PA. This layer has three nodes ">1.5 h/d," "30 min ~ 1.5 h/d," and "≤30 min/d," where the node weekend PA ">1.5 h/d" (85.1%) had the highest rate of physical fitness improvement for preschoolers, followed by "30 min ~ 1.5 h/d" (77.0%), while "≤30 min/d" (30.4%) was the lowest. Weekend PA ">1.5 h/d" (85.1%) was very significantly different from that of children with weekend PA "30 min ~ 1.5 h/d" (77.0%) and "≤30 min/d" (30.4%; $p < 0.001$). It is clear that preschoolers with longer weekend PA have better physical fitness improvement.

The three split nodes in the layer 2 leaf nodes were all weekend MVPA. Combined the analysis with the layer 1 nodes, it was found that in the group with weekend PA ">1.5 h/d," Weekend MVPA node ">5" (>1 h/d; 86.3%) and node "(3–5)" (30 min ~ ≤1 h/d; 94.1%) had a significantly higher rate of physical fitness improvement than node "≤3" (≤30 min/d) toddlers (68.6%; $p < 0.001$).

Branching downward from the second layer of leaf nodes, the branching characteristics of different nodes are quite different. The third level leaf node variables include mother's sports frequency, mother's BMI, father's education, and school day MVPA. The branch weekend MVPA nodes ">4" (>1.5 h/d) were classified based on the frequency of mother's PA, and the weekly frequency of mother's PA nodes "≥2" (1 and more times per week; 71.6%) had a significantly higher rate of fitness improvement than those with node "1" (0 times per week; 23.1%; $p = 0.001$). The downward branch of "≥2" (1 and more times per week) on frequency of mother's PA nodes is the mother's educational background, and the last branch is weekend breakfast. The branch weekend MVPA node "(3–5)" (30 min ~ ≤1 h/d) with weekend PA node ">4" (>1.5 h/d) was split into "1" and "2" according to mother' s BMI, children with mother' s BMI node "2" (<24.17; 98.0%) had a significantly higher rate of fitness improvement than those with maternal BMI node "1" (≥24.17; 63.2%; $p < 0.001$). Mother BMI node "1" (≥24.17) branches down to buffered food.

The branch with weekend PA node ">4" (>1.5 h/d) was the weekend MVPA node ">5" (>1 h/d), which was classified according to the father's education, and the father's education "≤2" (vocational high school/junior high school/technical high school or junior high school and below) were significantly higher (95.1%) than those with ">2" (bachelor's/master's/doctoral degree; 82.2%). Among children whose fathers' education was node ">2," preschoolers with school day PA ">4" (>1.5 h/d; 91.0%) had a significantly higher rate of physical fitness improvement than preschoolers with school day PA "≤4" (≤1.5 h/d; 76.4%). The downward branches of school day PA "≤4" (≤1.5 h/d) and school day PA ">4" (>1.5 h/d) are father's love of sport and father's occupation, respectively.

Combining the nodes in stratum 1, it was found that among preschoolers with weekend PA "30 min ~ ≤1.5 h/d," stratum 2 weekend MVPA node ">5" (>1 h/d; 76.2%) and node "(3–5)" (30 min ~ ≤1 h/d; 94.7%) were significantly higher ($p < 0.001$) than preschoolers with node "≤3" (≤30 min/d; 65.8%). In addition to weekend PA, weekend MVPA, mother's BMI, frequency of mother's sports, father's education,

TABLE 3 Influencing factors and assignment table.

Level I indicators	Secondary indicators	Type	Assign value
Birth information	gender	Cla	1 = male, 2 = female
	birth length (cm)	Con	
	birth weight (kg)	Con	
Feeding patterns	feeding mode within 4 months after birth	Cla	1 = breast feeding, 2 = artificial feeding, 3 = mixed feeding
	caregiver	Cla	1 = independent care by parents, 2 = independent care by grandparents, 3 = joint care by parents and grandparents, 4 = joint care by parents and nannies, 5 = joint care by grandparents and nannies, 6 = joint care by parents, grandparents and nannies, 7 = others
	attend early childhood classes	Cla	1 = Yes, 2 = No
PA and SB	school day SB (h/d)	Cla	1 = ≈0, 2 = 0.5, 3 = 1, 4 = 2, 5 = 3, 6 = 4, 7 = 5, 8 = 6, 9 = ≥7
	weekend SB (h/d)	Cla	
	school day PA (h/d)	Cla	1 = ≈0, 2 = 0.5, 3 = 1, 4 = 1.5, 5 = 2, 6 = 2.5, 7 = 3, 8 = > 3
	weekend PA (h/d)	Cla	
	school day MVPA (min/d)	Cla	1 = ≈0, 2 = 15, 3 = 30, 4 = 45, 5 = 60, 6 = > 60
	weekend MVPA (min/d)	Cla	
	frequency of outing (d/w)	Cla	1 = ≈0, 2 = 1, 3 = 2, 4 = 3, 5 = 4, 6 = 5, 7 = 6, 8 = 7
	frequency of participating in sports organization activities every week	Cla	1 = 0, 2 = 1 ~ 3, 3 = ≥4
Dietary nutrition	fruits, vegetables, candy, carbonated soft drinks, milk, fish and shrimp, eggs, hamburgers or hot dogs, pizza, instant noodles, puffed food (d/w)	Cla	1 = 0, 2 = 1 ~ 2, 3 = 3 ~ 4, 4 = 5 ~ 6, 5 = 7
	breakfast frequency on school day (d/w)	Cla	1 = 0, 2 = 1, 3 = 2, 4 = 3, 5 = 4, 6 = 5
	weekend breakfast frequency (d/w)	Cla	1 = 0, 2 = 1, 3 = 2
	vitamin D, calcium, cod liver oil, zinc (d/w)	Cla	1 = 0, 2 = ≥1
Sleep	total sleep time (h)	Con	
	sleep time on school day	Cla	1 = ≤20:30, 2 = 21:00, 3 = 21:30, 4 = 22:00, 5 = 22:30, 6 = 23:00, 7 = ≥23:00
	weekend Sleep Time	Cla	
	time to fall asleep	Con	
	frequency of getting enough sleep in a week (d/w)	Cla	1 = 4 ~ 7, 2 = 3 ~ 5, 3 = 1 ~ 2, 4 = 0

(Continued)

10.3389/fpubh.2023.1184756

TABLE 3 (Continued)

Level I indicators	Secondary indicators	Type	Assign value
Parental factors	age of father and mother (y)	Con	
	BMI of father and mother (kg/m^2)	Con	
	type of registered residence of father and mother	Cla	1 = urban hukou in Nanchang, 2 = agricultural hukou in Nanchang, 3 = urban hukou in other provinces and cities, 4 = agricultural hukou in other provinces and cities, 5 = Hong Kong, Macao, Taiwan and overseas
	education background of father and mother	Cla	1 = junior high school or below, 2 = senior high school or technical secondary school, 3 = junior college, 4 = undergraduate, 5 = master's degree, 6 = doctor's degree
	income of father and mother (cny/M)	Cla	1 = ≈0， 2 = ≤2000， 3 = 2001 ~ 4,000， 4 = 4,001 ~ 8,000， 5 = 8,001 ~ 15,000 ， 6 = ≥15,000
	parental occupation	Cla	1 = person in charge of state organs, party mass organizations, enterprises and public institutions, 2 = professional technicians, 3 = staff and relevant personnel, 4 = business and service personnel, 5 = production personnel in agriculture, forestry, animal husbandry, fishery and water conservancy, 6 = production and transportation equipment operators and relevant personnel, 7 = military personnel, 8 = other employees
	parental working days SB (h/d) parents' weekend days SB (h/d)	Cla	1 = ≈0， 2 = 0.5， 3 = 1， 4 = 2， 5 = 3， 6 = 4， 7 = 5， 8 = 6， 9 = ≥7
	parental frequency of vigorous activity (d/w)	Cla	1 = ≈0， 2 = 1， 3 = 2， 4 = 3， 5 = 4， 6 = 5， 7 = 6， 8 = 7
	parental sports frequency in every weekend	Cla	1 = 0， 2 = 1 ~ 2， 3 = 3 ~ 5， 4 = ≥6
	parental feeling of sports	Cla	1 = very like, 2 = quite like, 3 = average, 4 = quite dislike, 5 = very dislike
	marriage and upbringing of biological parents	Cla	1 = normal marriage, 2 = divorced and unmarried, supported by father, 3 = divorced and unmarried, supported by mother, 4 = divorced and remarried, supported by father, 5 = divorced and remarried, supported by mother, 6 = other

Cla: Variable of classification, Con: Variable of continuous, y: year old, M: month, w: week, d: day, h: hour, min: minute, SB: sedentary behavior, PA: physical activity, MVPA: medium and high intensity physical activity, BMI: body mass index, cny: RMB. The "school day" activities only include "outside the school."

10.3389/fpubh.2023.1184756

TABLE 4 Physical fitness of preschoolers.

Stage	Height (cm)	Weight (kg)	Sit and reach (cm)	Tennis throw (m)	Standing long jump (cm)	10-m shuttle run (s)	Double-leg timed hop (s)	Balance beam walking (s)
Baseline	103.5 ± 5.8	17.7 ± 3.0	10.7 ± 4.1	3.0 ± 1.2	63.9 ± 21.8	9.3 ± 2.0	9.2 ± 4.5	14.2 ± 9.6
12 months	109.6 ± 5.9*	19.3 ± 3.2*	11.7 ± 4.6*	4.2 ± 1.5*	82.2 ± 19.0*	8.2 ± 1.3*	6.7 ± 2.5*	11.9 ± 8.1*

*$p < 0.01$.

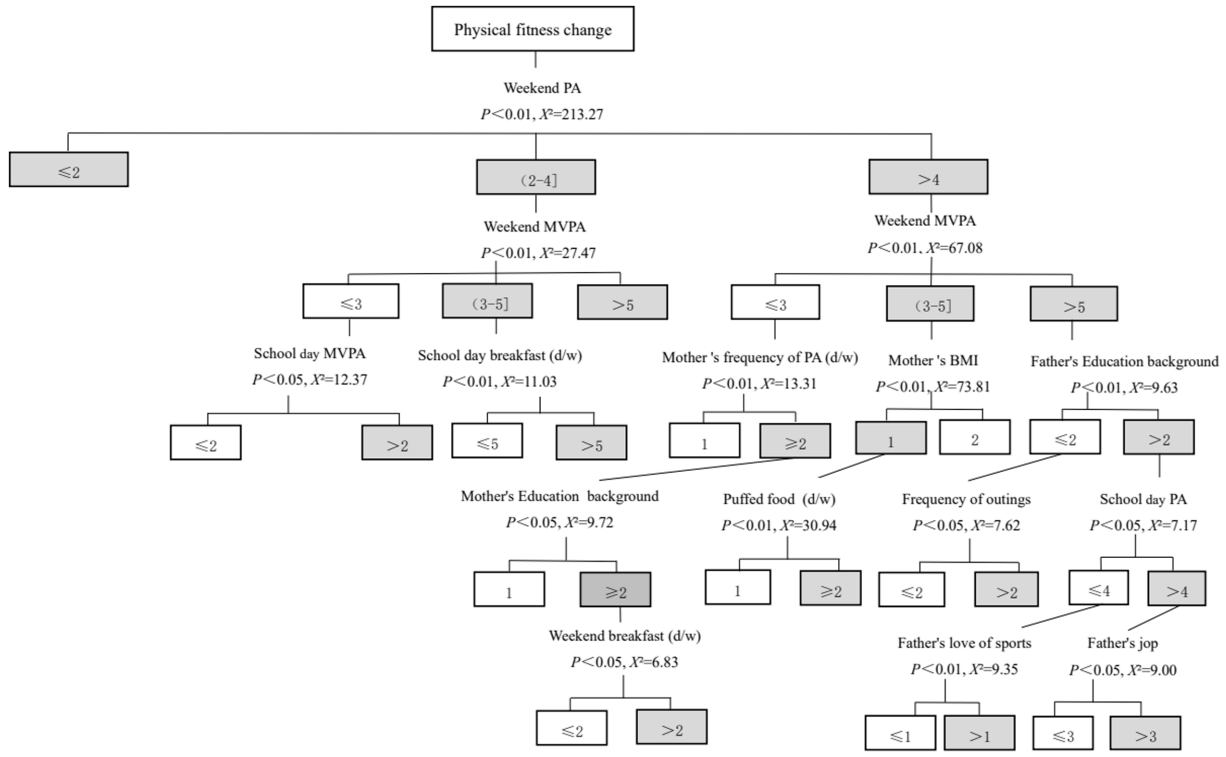


FIGURE 1
The model of CHAID. Layer 1 – weekend PA: 2 = 30 min, 3 = 1 h, 4 = 1.5 h; Layer 2 – weekend MVPA: 3 = 30 min, 4 = 45 min, 5 = 1 h; Layer 3 - mother's BMI: 1 = <24.17, 2 = ≥24.17; Layer 3 - frequency of mother's sports: 1 = 0 times; 2 = 1 time; Layer 3 – school PA: 2 = 30 min; Level 3 - Father's education background: 2 = vocational high school/technical secondary school/technical high school or technical secondary school; Level 4 – Education background of mother: 1 = junior high school or below.

mother's education, and school day PA as the main influencing factors, school day breakfast, puffed food, days out with children, school day MVPA, father's love of sports, father's occupation, and weekend breakfast were all statistically significant leaf node variables.

3.3 Ranking of key influencing factors

The CHAID model generated by IBM SPSS modeler software was used to rank the importance of each variable based on the decision tree node positions as well as the values. A total of seven variables that were important for the change in physical fitness of preschool children were screened. Each variable was ranked in order of importance: weekend PA, weekend MVPA, mother's BMI, mother's sports frequency, father's education, mother's education, and school day PA (Table 5).

3.4 Evaluation of decision tree model

A 10-layer cross-validation model was used to identify the accuracy, and the model's accuracy was 84.17%, indicating that the model was effective (Table 6).

4 Discussion

In this study, by tracking the measurement of preschool children's fitness levels and changes. It was found that overall, various aspects of preschool children's physical fitness improve with age. Through in-depth analysis using decision trees, the seven most significant influencing factors were identified. These factors exert their influence through both direct and indirect pathways. Among them, PA and parental factors were identified as key factors influencing the changes

TABLE 5 Physical fitness score of preschoolers.

Stage	Height weight score	Sit and reach (cm)	Tennis throw (m)	Standing long jump (cm)	10-m shuttle run (s)	Double-leg timed hop(s)	Balance beam walking (s)
Baseline	4.0 ± 1.5	3.1 ± 1.2	2.2 ± 0.9	2.9 ± 1.2	2.7 ± 1.0	1.6 ± 0.9	2.7 ± 1.1
12 months	4.1 ± 1.5	3.4 ± 1.2*	2.3 ± 1.0#	3.3 ± 1.0*	2.8 ± 1.0	3.6 ± 1.2*	2.4 ± 1.0*

*$p < 0.01$.
#$p < 0.05$.

TABLE 6 Recognition accuracy of the model.

Name	*n*	%
Model accuracy	1,074	84.17
Model error	202	15.83
Total	1,276	100

in preschool children's physical fitness. This finding is consistent with the results of most relevant studies, and once again suggests that the importance of promoting PA in children's 24-h behavior should be emphasized, and the role of families in promoting PA should be played. This result also proves the effectiveness of the decision tree model in identifying the influencing factors of preschool children's physical fitness.

4.1 Construction of decision tree model

Many previous studies have analyzed the factors affecting children's physical fitness from different perspectives, mainly from personal factors (gender, age, physical activity, race, etc.) (34, 35), family factors (parents' physical activity, children's nutrition, etc.) (17, 18, 36), community factors (built environment, community support, etc.) (37, 38) and other aspects. These studies mainly analyze the effects of single or multiple factors, and it is easy to overlook the impact of other factors on the physical fitness of young children. Indeed, there is a relationship between the effects of various factors that overlap or cancel out, that is, the interaction between various factors has an interactive effect on the physical fitness of preschool children. In fact, the improvement or decrease of preschool children's physical fitness and the influencing factors do not show a simple linear relationship. After segmenting the samples for different types of variables, the close relationship between independent variables and dependent variables in a specific type of sample can be explored. For example, in a study by Zhenya et al. (39), the dose effect of 24-h isochronous replacement of motor behavior time and changes in body mass in preschool children was investigated using component data analysis. On the contrary, as the time of isochronous substitution of LPA, SB and sleep for MVPA increased, the level of body mass decreased rapidly and the decrease was greater than the increase in the other three behaviors of MVPA isochronous, and these findings corroborate the existence of complex interactions among the independent variables. Decision tree, as an important classification method in data mining, is characterized by optimal segmentation of samples based on different types of variables, and automatic classification of samples based on the significance of the test.

The CHAID model in decision tree is used in this study, and the constructed CHAID model is 5 levels high, divided into 41 leaf nodes, and has a branching tree. The accuracy of the model reaches 84.17%, which meets the needs of this study.

4.2 Decision tree analysis of factors influencing physical fitness

Weekend PA, weekend MVPA, mother's BMI, mother's frequency of PA, father's education, mother's education and school day PA were the seven most important influencing factors. Besides, school day breakfast, puffed food, taking the child out for a few days, school day MVPA, father's love for sports, father's occupation, and weekend breakfast all influenced different nodes of body mass change in preschool children. From the above-influencing factors, it was found that PA and parental factors are the key to influence the evolution of body mass in preschool children.

4.2.1 Physical activity

Many studies have confirmed that PA (including exercise) is an important factor influencing changes in fitness levels. A systematic review study exploring school-based PA interventions found (40) that the included high-quality RCTs confirmed that PA interventions are effective in improving physical fitness and reducing skin-fold thickness in children and adolescents aged 5–18 years. Not coincidentally, another meta-analysis found that school-based PA interventions were associated with a significant small increase in cardiorespiratory fitness (CRF) in children aged 3–12 years (Hedges' g = 0.22; 95% CI 0.14 to 0.30; $p < 0.001$). According to subgroup analysis, the increase in CRF was significant in girls (Hedges' g = 0.25; 95% CI 0.13 to 0.37; $p > 0.001$) (41). The findings of these studies undoubtedly provide an important reference for young children's physical fitness promotion and further confirm the important role of PA in enhancing young children's physical fitness, consistent with the results of the present study. Investigating the mechanism, cardiorespiratory fitness, regular PA induces molecular adaptations in the cardiovascular system. The increased demand for oxygen during exercise activates signaling pathways that promote angiogenesis, or the growth of new blood vessels. This process involves the release of factors such as vascular endothelial growth factor and fibroblast growth factor, which stimulate the formation of new capillaries, enhancing blood flow to the muscles. Additionally, exercise activates pathways that signal the heart to undergo hypertrophy, leading to increased heart muscle mass and improved cardiac function. In addition, other indicators of physical fitness, such as height, weight, muscle strength, and muscle endurance, have gradually been discovered through relevant mechanisms. However, it is worth noting that most of these mechanism studies are based on adults or children and adolescents, which can provide ideas for the possible mechanisms of PA improving physical fitness in young children. However, further

exploration is still needed for the relevant mechanism research in preschooler.

This study found that weekend PA and weekend MVPA were two key factors affecting preschoolers' physical fitness, highlighting the importance of performing activities on weekends. This is because most preschoolers follow their teachers on school days as required in the school, and therefore their PA levels are all within a certain range. On the other hand, preschoolers have relatively more choices during the weekend, which may lead to more screen time and possibly more time spent outdoors. The behavior of young children in kindergarten is basically carried out under the supervision of teachers, and whether it is SB or PA, it is mainly structured. However, with the trend of using electronic products at a younger age, more and more young children are addicted to electronic products. Once parents fail to supervise during weekends, young children may use electronic products for a long time. On the contrary, parents consciously taking their children out for activities and limiting screen time will be beneficial for the development of children's physical fitness. The difference in PA between school days and weekends was also found in Meimei ji 's research (42) on Chinese young children. Similarly, the findings of this study are consistent with that of scholar Zhao, who found that weekend PA is the key influencing factor for children's physical development ranking first (33). The study also mentioned that it may be because during the semester of this study, compared to the more regular lifestyle on school days, weekend children had more freedom in choosing physical activities and sedentary behavior, which had a greater impact on their physical fitness. This shows that adequate attention should be paid to preschoolers' weekend PA, which includes weekend MVPA. In addition, the 6th influencing factor is school day PA, which appears in both the 2nd and 3rd levels of the branch. Therefore, the PA factors of fitness changes in preschool children include not only weekend PA, but also school day PA is an important factor. Both of the previous studies were conducted at kindergarten sites. As the place where children spend most of their day, the kindergarten curriculum on school days is critical, and adequate PA is critical for preschoolers' physical fitness. This evidence further confirms and adds to the significance of PA in the physical fitness of children and adolescents, including preschoolers.

In addition, the performance characteristics of PA in preschool children have been a hot topic of interest for researchers. The decision tree model showed that preschoolers with weekend PA ">1.5 h/d" had significantly higher fitness levels than those with "30 min ~ 1.5 h/d" and "≤30 min/d. In addition, preschoolers with MVPA ">1 h/d" on the downward branching weekend also improved significantly more than preschoolers with MVPA "≤30 min/d." The findings of this study are generally consistent with the recommendations of national PA guidelines for young children, with the Australian (43), Canadian (44) and WHO PA guidelines for children referring to no less than "1 h/d" of active play (45), and similarly the Chinese Guidelines for Exercise in Preschool Children (3–6 years) recommending a specific duration of no less than "1 h/d" of MVPA cumulatively (46). These guidelines were developed through a comprehensive search of literature related to preschool children in various fields, combined with the physiological and psychological characteristics, and repeatedly discussed by scholars in various fields, which have essential reference and practical significance. The splitting point value at the root node of the decision tree model is consistent with the present study, that is, MVPA is a vital factor influencing the fitness changes of preschool children, and at least "1 h/d" is a critical splitting point. The difference is that this study found that weekend PA ">1.5 h/d" is a critical point for improving the physical fitness of preschool children, which is lower than the recommended amount of PA in each guideline. The possible reason is that the physical activity level of Chinese children is generally low (42), and children who exceed this threshold can have a significant improvement in their physical fitness.

4.2.2 Parental factors

From the tree diagram of the decision tree, it is clear that both level 1 and level 2 are related to PA, and from level 3 parental factors appear. Parental factors, as indirect influences, may impact changes in preschool children's physical fitness through direct factors. This study found that mother's BMI, fequency of mother's PA, father's education and mother's education exerted significant influence in the change of body mass of preschool children. Preschoolers whose mothers' BMI was greater than 24.17 had a significantly higher decrease in body mass than preschoolers whose mothers' BMI was less than 24.17. Studies have shown that the effect of mothers' BMI on preschool children is mainly in the areas of preschool children's growth and development, diet and exercise habits. On the other hand, overweight and obese mothers have influence on preschoolers' lifestyle in terms of diet and PA, especially mothers as the main caregivers of preschoolers, and their lifestyle behavioral habits are obvious on the habit formation of preschoolers. A study by Li et al. (47), which explored the effect of parental overweight/obesity on overweight/ obesity in children and adolescents, found that children and adolescents whose mothers were overweight or obese were 2.74 times more likely to be overweight and obese than children whose parents were neither overweight/ obese. As the primary caregiver of preschool children's eating behavior, the mother's BMI status is related to her eating habits and health perceptions. These differences in habits and perceptions can lead mothers to adopt different feeding practices, affecting the eating behavior habits of their preschool children and, consequently, their body shape. In addition to the influence of acquired factors, some studies have suggested that parental genes are an important factor affecting the physical fitness of young children. These studies have found that innate genes have a certain impact on the physical fitness of young children (48), but some studies have also suggested that the degree of genetic influence is limited (49). However, in the process of developing children's physical fitness, the acquired influence can be changed and has a profound impact (50). Therefore, it is necessary to pay attention to the impact of family environment on the physical fitness of young children.

The results of this study also reaffirm the positive effect of mothers' frequency of physical activity participation on the development of PA habits and physical fitness of preschool children. The present study also found that preschool mothers' participation in sports once a week or more was significantly higher than preschool children's physical fitness who participated in sports 0 times a week. The influence of mothers' sports habits on their children is undoubtedly subtle. Mothers who participate in sports regularly can create a good PA atmosphere for their children and act as role models in developing their children's sports habits. These sports habits play an important role in the promotion of physical fitness in preschool children. Zhao et al. (51) explored the differences in physical fitness and PA between overweight and normal weight preschoolers and the relationship between their parental and children's PA levels, and found that

maternal PA was positively correlated with their children's PA levels, and overweight parents were positively correlated with their children's overweight. Scholars Rhodes et al. (52) explored the effect of a family-based intervention on improving preschool children's MVPA and fitness levels through a randomized controlled trial. The study population was divided into a Family PA Planning Plus Information/Education and an Education Information Only group, and it was found that preschoolers in the intervention group had significantly improved MVPA at both the 6 and 13-week measures compared to the control group, and also had significantly improved cardiorespiratory fitness. Thus, family-based facilitation strategies, such as the involvement of mothers in their children's activities, may be an effective measure to improve preschoolers' fitness levels.

In addition, parental education and father's occupation influenced changes in physical fitness in preschool children. The results of the present study are similar to those of a foreign study on adolescents. Finger et al. (53) investigated 5,251 German adolescents aged 11–17 years and found that higher levels of parental education were associated with better aerobic fitness (PA). Higher levels of parental education were found to be associated with higher levels of PA in girls. The reason for this is that parents whose work status is predominantly SB, especially those with higher education, may exercise more in their leisure time and may have stronger expectations to motivate their children to exercise with them (54); in contrast, parents whose work is physically demanding (a group that is mostly less educated) may be less active in their leisure time, as they usually stay at night or on weekends at home, using media for entertainment and relaxation. Parental education does not directly affect the level of fitness of preschoolers, but may in turn affect the change in fitness of preschoolers by increasing PA. In a study that explored the relationship between parental education and PA from another perspective, Xiu et al. (55) found that the total number of parents with educational attainment of junior high school and below had only 6.2% of their children participating in various special classes, while the total number of parents with educational attainment higher than junior high school was as high as 93.8%. In addition, Wang et al. (56) explored the relationship between fathers' occupation and children's PA and SB. They found that students whose fathers were agricultural workers reported lower rates of physical activity.

In contrast, fathers with occupations such as clerks, police officers, and military personnel reported higher rates of PA in their children. In addition, fathers whose occupations were casual workers/layoffs and urban farmers reported higher rates of "watching TV/video," while fathers whose occupations were police and military, and professionals and technicians reported lower rates. The present study also found that preschool children whose fathers' occupations were state agencies, professionals, clerks, and related personnel had significantly higher physical fitness than preschool children whose fathers were employed in agriculture, animal husbandry, or fishing. The reason for this may be related to family lifestyle and environmental impacts. Engaging in professions such as agriculture, animal husbandry, and fishing may mean more physical labor and fewer opportunities to participate in aerobic or physical exercise. The occupation and nature of work of parents may affect their time arrangement and lifestyle, affecting the opportunities and quality of their children's participation in sports activities. In contrast, fathers who are engaged in state organs, professional technicians and clerks may have more time and opportunities to encourage their children to engage in sports activities and provide a better sports environment. In summary, based on current research analysis, parents' educational background and occupation may affect the physical fitness level of young children by influencing their PA.

5 Strength and limitations

Through a one-year cohort study, it illustrates the causal relationship between the characteristics of physical change and development and various factors. The decision tree model is used to find the main influencing factors, and the complex interaction between influencing factors is discussed. This provides more practical reference for children's physical fitness promotion work, and also provides research basis for future research. The limitation of this study is that the measurement of PA and SB time in young children is a subjective questionnaire, which may have some bias compared to the objective measure.

6 Conclusion

Through the decision tree model, a total of 7 critical influencing factors for changes in preschool children's physical fitness were screened. The top 2 factors in order of their importance were weekend PA and weekend MVPA, followed by mother's BMI, mother' s sports frequency, paternal education, mother' education, and school day PA, of which 3 factors were related to preschool children's PA and 4 factors were related to preschool children's parental factors. The parental sports habits, body shape, and Feeding patterns have important impacts on the changes in preschool children's physical fitness.

The results of the tree diagram suggest that to improve the physical fitness level of preschool children, the weekend MVPA should be increased to more than 30 min/d on improving the weekend PA. Mothers with appropriate BMI and higher sports frequency also play an important role in improving the physical fitness of preschool children. Other parental factors such as parental education and father's occupation are also influencing factors for changes in children's physical fitness. In addition, parental factors and school day PA have decisional implications for changes in physical fitness in preschool children of different level types.

Through the decision tree model, a total of 7 critical influencing factors for changes in preschool children's physical fitness were screened. The top 2 factors in order of their importance were weekend PA and weekend MVPA, followed by mother's BMI, mother's sports frequency, father's education, mother' education, and school day PA, of which 3 factors were related to preschool children's PA and 4 factors were related to preschool children's parental factors.

The duration of the present cohort study was 1 year. Although the decision tree model could find the main influencing factors, a follow-up study with a longer period might explain more accurately the influence of different factors on the change of physical fitness of preschool children, so the follow-up study period of the subsequent design could be extended. In addition, the decision tree model demonstrates the complex interaction between different influencing factors through tree diagrams, future studies can conduct comprehensive intervention experiments based on the results of this study to integrate many factors affecting preschool children's physical

fitness for empirical interventions to clarify further and identify key influencing factors, so that more research results will be extended to kindergartens and educational institutions to improve and refine preschool, this will improve the evaluation and service system of children's physical health. At the same time, the decision tree model is suitable for large-scale, large-group, and large-sample research, and has some prospects for application not only in the field of preschool children's physical fitness research, but also in the whole field of physical fitness testing, so that more decision tree results can provide scientific and reasonable health advice for different age groups, and target health problems to ultimately achieve the purpose of improving physical health and human quality of life.

Data availability statement

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

Ethics statement

The studies involving humans were approved by Jiangxi University of Chinese Medicine Ethics Commit/Jiangxi University of Chinese Medicine. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

WL and GZ were responsible for the design of the study. JF and SS provided policy support for the study. ZH, TH, RW, DC, and RC collected the data. WL and ZH has conducted all the analysis and written the first draft of the paper. All authors contributed to the article and approved the submitted version.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was funded by the National Social Science Fund of China (Grant No. 21BTY088), Social Science "Fourteenth Five-Year Plan" Fund of Jiangxi Province (Grant No. 22TY20D), Sports scientific research project of Jiangxi Provincial Bureau of Sports Research (Grant No. 202122), and Teaching Reform Project of Jiangxi University of Chinese Medicine (Grant No. 2022jzqn-5).

Acknowledgments

We thank kindergarten staff, children and their parents who participated in the survey.

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. Pate RR, O'Neill JR, Brown WH, McIver KL, Howie EK, Dowda M. Top 10 research questions related to physical activity in preschool children. *Res Q Exerc Sport*. (2013) 84:448–55. doi: 10.1080/02701367.2014.897592

2. Malina RM. Tracking of physical activity and physical fitness across the lifespan. *Res Q Exerc Sport*. (1996) 67:S48–57. doi: 10.1080/02701367.1996.10608853

3. Dwyer T, Magnussen CG, Schmidt MD, Ukoumunne OC, Ponsonby AL, Raitakari OT, et al. Decline in physical fitness from childhood to adulthood associated with increased obesity and insulin resistance in adults. *Diabetes Care*. (2009) 32:683–7. doi: 10.2337/dc08-1638

4. Twisk JW, Kemper HC, van Mechelen W. Tracking of activity and fitness and the relationship with cardiovascular disease risk factors. *Med Sci Sports Exerc*. (2000) 32:1455–61. doi: 10.1097/00005768-200008000-00014

5. Thomas NE, Baker JS, Davies B. Established and recently identified coronary heart disease risk factors in young people: the influence of physical activity and physical fitness. *Sports Med*. (2003) 33:633–50. doi: 10.2165/00007256-200333090-00001

6. Yue Y, Juan H, Jing C, Hang C, Heng Z. Analysis of the Physical Health Level of 3-6 Year Old Children in Jiangsu Province [J]. *China School Health*. 43:1372–1375. doi: 10.16835/j.cnki.1000-9817.2022.09.022

7. Leppänen MH, Henriksson P, Delisle Nyström C, Henriksson H, Ortega FB, Pomeroy J, et al. Longitudinal physical activity, body composition, and physical fitness in preschoolers. *Med Sci Sports Exerc*. (2017) 49:2078–85. doi: 10.1249/MSS.0000000000001313

8. Huang T, Zhao GG, Tan H, Wu H, Fu J, Sun S, et al. Effects of family intervention on physical activity and sedentary behavior in children aged 2.5-12 years: a meta-analysis. *Front Pediatr*. (2021) 9:720830. doi: 10.3389/fped.2021.720830

9. Williams HG, Pfeiffer KA, O'Neill JR, Dowda M, McIver KL, Brown WH, et al. Motor skill performance and physical activity in preschool children. *Obesity (Silver Spring)*. (2008) 16:1421–6. doi: 10.1038/oby.2008.214

10. Guanggao Z, Wendi L, Jinmei F, Shunli S, Wujin H Dinghong M, et al. The preliminary study on the relationship between gut microbiota, inflammation and physical activity and BMI in preschool children. *China Sport Sci*. (2019) 39:29–37. doi: 10.16469/j.css.201908004

11. Kariippanon KE, Chong KH, Janssen X, Tomaz SA, Ribeiro EHC, Munambah N, et al. Levels and correlates of objectively measured sedentary behavior in young children: SUNRISE study results from 19 countries. *Med Sci Sports Exerc*. (2022) 54:1123–30. doi: 10.1249/MSS.0000000000002886

12. Collings PJ, Brage S, Ridgway CL, Harvey NC, Godfrey KM, Inskip HM, et al. Physical activity intensity, sedentary time, and body composition in preschoolers. *Am J Clin Nutr*. (2013) 97:1020–8. doi: 10.3945/ajcn.112.045088

13. Skouteris H, McCabe M, Swinburn B, Hill B. Healthy eating and obesity prevention for preschoolers: a randomised controlled trial. *BMC Public Health*. (2010) 10:220. doi: 10.1186/1471-2458-10-220

14. Evaristo OS, Moreira C, Lopes L, Abreu S, Agostinis-Sobrinho C, Oliveira-Santos J, et al. Associations between physical fitness and adherence to the Mediterranean diet

with health-related quality of life in adolescents: results from the LabMed physical activity study. *Eur J Pub Health*. (2018) 28:631–5. doi: 10.1093/eurpub/cky043

15. Silventoinen K, Maia J, Jelenkovic A, Pereira S, Gouveia É, Antunes A, et al. Genetics of somatotype and physical fitness in children and adolescents. *Am J Hum Biol*. (2020) 33:e23470. doi: 10.1002/ajhb.23470

16. Cieśla E, Zaręba M, Kozieł S. The level of physical fitness in children aged 6-7years with low birthweight. *Early Hum Dev*. (2017) 111:23–9. doi: 10.1016/j.earlhumdev.2017.05.008

17. Liu F, Lv D, Wang L, Feng X, Zhang R, Liu W, et al. Breastfeeding and overweight/obesity among children and adolescents: a cross-sectional study. *BMC Pediatr*. (2022) 22:347–8. doi: 10.1186/s12887-022-03394-z

18. Huang H, Gao Y, Zhu N, Yuan G, Li X, Feng Y, et al. The effects of breastfeeding for four months on thinness, overweight, and obesity in children aged 3 to 6 years: a retrospective cohort Study from National Physical Fitness Surveillance of Jiangsu Province, China. *Nutrients*. (2022) 14:4154. doi: 10.3390/nu14194154

19. Escolano-Pérez E, Sánchez-López CR, Herrero-Nivela ML. Early environmental and biological influences on preschool motor skills: implications for early childhood care and education. *Front Psychol*. (2021) 12:725832. doi: 10.3389/fpsyg.2021.725832

20. Palou P, Muntaner-Mas A, Cantallops J, Borràs PA, Labayen I, Jiménez-Pavón D, et al. A single question of parent-reported physical activity levels estimates objectively measured physical fitness and body composition in preschool children: the PREFIT project. *Front Psychol*. (2019) 10:1585. doi: 10.3389/fpsyg.2019.01585

21. Henrard S, Speybroeck N, Hermans C. Classification and regression tree analysis vs. multivariable linear and logistic regression methods as statistical tools for studying haemophilia. *Haemophilia*. (2015) 21:715–22. doi: 10.1111/hae.12778

22. Xie JY, Yan XQ, Zeng J, Qian Z, Zhou J, Xia C, et al. Analysis of physical fitness status of preschool children in Zhongshan City. *Chin J Sch Health*. (2020) 41:604–7. doi: 10.16835/j.cnki.1000-9817.2020.04.033

23. Famularo S, Milana F, Cimino M, Franchi E, Giuffrida M, Costa G, et al. Upfront surgery versus neoadjuvant perioperative chemotherapy for Resectable colorectal liver metastases: a machine-learning decision tree to identify the best potential candidates under a parenchyma-sparing policy. *Cancers*. (2023) 15:613. doi: 10.3390/cancers15030613

24. Yu D, Zhong Y, Yu Y. Application of data mining technology in the analysis of human muscle strength data—a case study of hand grip muscle strength test data. *Sports Sci*. (2010) 30:70–4. doi: 10.16469/j.css.2010.02.012

25. Chen W, Yang G, Qi D. Comprehensive evaluation of college Students' physical health and sports mode recommendation model based on decision tree classification model. *Comput Intell Neurosci*. (2022) 2022:1–8. doi: 10.1155/2022/5504850

26. Che Y, Che K, Li Q. Application of decision tree in PE teaching analysis and management under the background of big data. *Comput Intell Neurosci*. (2022) 2022:1–9. doi: 10.1155/2022/8091838

27. Shephard RJ. Qualified fitness and exercise as professionals and exercise prescription: evolution of the PAR-Q and Canadian aerobic fitness test. *J Phys Act Health*. (2015) 12:454–61. doi: 10.1123/jpah.2013-0473

28. Wendi L, Guanggao Z, Jinmei F, Shunli S, Jing C, Zihao H, et al. Exploring key factors of overweight in preschool children: a decision tree analysis. *J Chengdu Sport Univ*. (2020) 46:86–93. doi: 10.15942/j.jcsu.2020.01.014

29. Guerrero MD, Vanderloo LM, Rhodes RE, Faulkner G, Moore SA, Tremblay MS. Canadian children's and youth's adherence to the 24-h movement guidelines during the COVID-19 pandemic: a decision tree analysis. *J Sport Health Sci*. (2020) 9:313–21. doi: 10.1016/j.jshs.2020.06.005

30. Chen H, Li X, Meng F, Ai W, Lin Q, Cai K. A decision tree network with semi-supervised entropy learning strategy for spectroscopy aided detection of blood hemoglobin. *Spectrochim Acta A Mol Biomol Spectrosc*. (2023) 291:122354. doi: 10.1016/j.saa.2023.122354

31. Weerackody SC, Clutterbuck GL, Johnston LM. Measuring psychological, cognitive, and social domains of physical literacy in school-aged children with neurodevelopmental disabilities: a systematic review and decision tree. *Disabil Rehabil*. (2022) 45:3456–75. doi: 10.1080/09638288.2022.2131004

32. Quan B, Li M, Lu S, Li J, Liu W, Zhang F, et al. Predicting disease-specific survival for patients with primary cholangiocarcinoma undergoing curative resection by using a decision tree model. *Front Oncol*. (2022) 12:824541. doi: 10.3389/fonc.2022.824541

33. Guanggao Z, Wendi L, Jinmei F, Shunli S, Wujin H, Dinghong M, et al. Influence factors of physical fitness in preschool children: a decision tree analysis. *China Sport Sci*. (2020) 40:32–9. doi: 10.16469/j.css.202004

34. Minghui Q, Chunyi F, Tang Z, Longkai L, Peijie C. Dose-response relationship between bouts of physical activity and physical health in preschool children. *China Sport Sci*. (2020) 40:39–45. doi: 10.16469/j.css.202003004

35. Schmutz EA, Leeger-Aschmann CS, Radtke T, Muff S, Kakebeeke TH, Zysset AE, et al. Correlates of preschool children's objectively measured physical activity and sedentary behavior: a cross-sectional analysis of the SPLASHY study. *Int J Behav Nutr Phys Act*. (2017) 14:1. doi: 10.1186/s12966-016-0456-9

36. Hesketh KR, Goodfellow L, Ekelund U, McMinn AM, Godfrey KM, Inskip HM, et al. Activity levels in mothers and their preschool children. *Pediatrics*. (2014) 133:e973–80. doi: 10.1542/peds.2013-3153

37. Carson V, Rosu A, Janssen I. A cross-sectional study of the environment, physical activity, and screen time among young children and their parents. *BMC Public Health*. (2014) 14:61. doi: 10.1186/1471-2458-14-61

38. Barboza M, Marttila A, Burström B, Kulane A. Contributions of preventive social services in early childhood home visiting in a disadvantaged area of Sweden: the practice of the parental advisor. *Qual Health Res*. (2021) 31:1380–91. doi: 10.1177/1049732321994538

39. Zhenya C, Shuming W. Study on the isochronous substitution effect of movement behavior on physical health of preschool children. *China Sport Sci*. (2020) 40:50–7. doi: 10.16469/j.css.202010006

40. Sun C, Pezic A, Tikellis G, Ponsonby AL, Wake M, Carlin JB, et al. Effects of school-based interventions for direct delivery of physical activity on fitness and cardiometabolic markers in children and adolescents: a systematic review of randomized controlled trials. *Obes Rev*. (2013) 14:818–38. doi: 10.1111/obr.12047

41. Pozuelo-Carrascosa DP, García-Hermoso A, Álvarez-Bueno C, Sánchez-López M, Martinez-Vizcaino V. Effectiveness of school-based physical activity programmes on cardiorespiratory fitness in children: a meta-analysis of randomised controlled trials. *Br J Sports Med*. (2018) 52:1234–40. doi: 10.1136/bjsports-2017-097600

42. Ji M, Tang A, Zhang Y, Zou J, Zhou G, Deng J, et al. The relationship between obesity, sleep and physical activity in Chinese preschool children. *Int J Environ Res Public Health*. (2018) 15:527. doi: 10.3390/ijerph15030527

43. Okely AD, Ghersi D, Hesketh KD, Santos R, Loughran SP, Cliff DP, et al. A collaborative approach to adopting/adapting guidelines - the Australian 24-hour movement guidelines for the early years (birth to 5 years): an integration of physical activity, sedentary behavior, and sleep. *BMC Public Health*. (2017) 17:869. doi: 10.1186/s12889-017-4867-6

44. Tremblay MS, Chaput JP, Adamo KB, Aubert S, Barnes JD, Choquette L, et al. Canadian 24-hour movement guidelines for the early years (0-4 years): an integration of physical activity, sedentary behaviour, and sleep. *BMC Public Health*. (2017) 17:874. doi: 10.1186/s12889-017-4859-6

45. WHO. *Guidelines on physical activity, sedentary behaviour and sleep for children under 5 years of age*. Geneva: World Health Organization (2019) Licence: CC BY-NC-SA 3.0 IGO.

46. Hongyan G, Xing Z, Sha Q, Jianxin W, Yufeng Y, Jianjun G, et al. Physical activity guideline for Chinese preschoolers aged 3-6 years. *Chin J Child Health Care*. (2020) 28:714–20. doi: 10.11852/ZGETBJZZ2020-0982

47. Li X-h, Ma J, Wang H-j, Dong B, Fu LG, Song Y, et al. Analysis of influencing factors on physical endurance of Chinese primary and middle school students aged 10 to 17. *Chin J Prevent Med*. (2013) 47:700–6. doi: 10.3760/cma.j.issn.0253-9624.2013.08.007

48. Bost KK, Teran-Garcia M, Donovan SM, Fiese BHSTRONG Kids Team. Child body mass index, genotype and parenting in the prediction of restrictive feeding. *Pediatr Obes*. (2018) 13:239–46. doi: 10.1111/ijpo.12219

49. Li A, Robiou-du-Pont S, Anand SS, Morrison KM, McDonald SD, Atkinson SA, et al. Parental and child genetic contributions to obesity traits in early life based on 83 loci validated in adults: the FAMILY study. *Pediatr Obes*. (2018) 13:133–40. doi: 10.1111/ijpo.12205

50. Masip G, Foraita R, Silventoinen K, Adan RAH, Ahrens W, de Henauw S, et al. The temporal relationship between parental concern of overeating and childhood obesity considering genetic susceptibility: longitudinal results from the IDEFICS/I.Family study. *Int J Behav Nutr Phys Act*. (2021) 18:139. doi: 10.1186/s12966-021-01205-9

51. Zhao F. The influence of physical fitness activities on overweight children and their parents. *Chin J Sch Health*. (2015) 36:1848–50. doi: 10.16835/j.cnki.1000-9817.2015.12.030

52. Rhodes RE, Blanchard CM, Quinlan A, Naylor PJ, Warburton DER. Family physical activity planning and child physical activity outcomes: a randomized trial. *Am J Prev Med*. (2019) 57:135–44. doi: 10.1016/j.amepre.2019.03.007

53. Finger JD, Mensink GB, Banzer W, Lampert T, Tylleskär T. Physical activity, aerobic fitness and parental socio-economic position among adolescents: the German health interview and examination survey for children and adolescents 2003-2006 (KiGGS). *Int J Behav Nutr Phys Act*. (2014) 11:43. doi: 10.1186/1479-5868-11-43

54. Finger JD, Tylleskär T, Lampert T, Mensink GBM. Physical activity patterns and socioeconomic position: the German National Health Interview and examination survey 1998 (GNHIES98). *BMC Public Health*. (2012) 12:1079. doi: 10.1186/1471-2458-12-1079

55. Xiu Y, Yong-wei S. Family factors influencing the constitution of 3-6Years old children of Liaoning Province. *J Shenyang Sport Univ*. (2008) 2:11–4. doi: 10.3969/j.issn.1004-0560.2008.02.004

56. Zhirong W, Zhiyong AYinan H, Yu Y, Shi R, Hongwei W, et al. Analysis of the relationship between students' static activity behavior and fathers occupation. *Chin J School Health*. (2012) 33:884–5. doi: CNKI:SUN:XIWS.0.2012-07-048

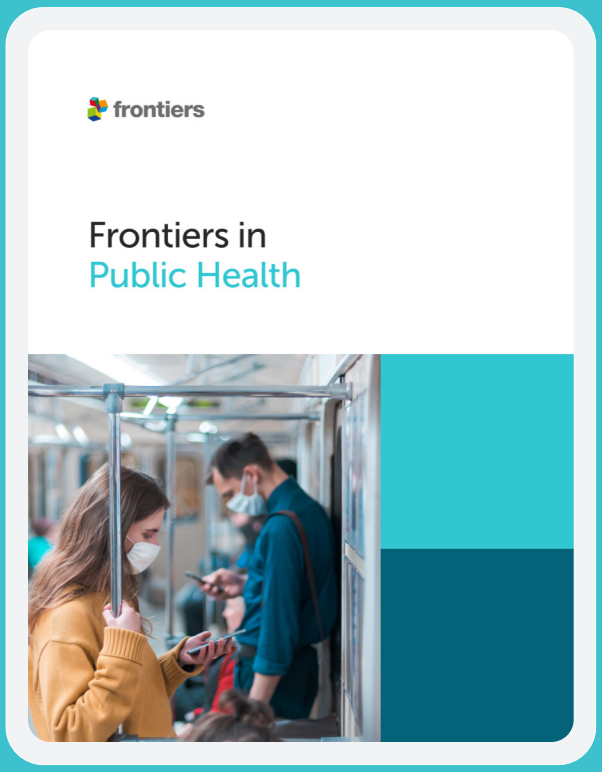
Frontiers in Public Health

Explores and addresses today's fast-moving healthcare challenges

One of the most cited journals in its field, which promotes discussion around inter-sectoral public health challenges spanning health promotion to climate change, transportation, environmental change and even species diversity.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact