Postprint of a Systematic Review of Assessment Tools for Childbirth-Related Psychological Trauma in Perinatal Women

Authors: Yongxia Wen, Sun Hai, Chen Xiaoju, Cai Wanjing, Li Shuni, Guo Honghua, Guo Honghua

Date: 2024-10-17T11:27:11+00:00

Abstract

Background: Childbirth psychological trauma is closely related to maternal mental health; therefore, identifying childbirth psychological trauma and evaluating intervention effects are extremely important. Currently, there are numerous assessment tools for childbirth psychological trauma, but studies evaluating the quality of these assessment tools are still lacking.

Objective: To systematically evaluate the measurement properties and methodological quality of childbirth psychological trauma assessment tools to provide a reference for future research.

Methods: Relevant studies on childbirth psychological trauma assessment tools were searched in China Biology Medicine disc, Wanfang Data Knowledge Service Platform, China National Knowledge Infrastructure, VIP Chinese Journal Service Platform, PubMed, Web of Science, CINAHL, Embase, and APA PsycArticles, with the search period from database inception to November 6, 2023. Literature screening, extraction, and evaluation were independently performed by two researchers and cross-checked.

Results: A total of 6 studies and 6 assessment tools were included: the Birth Trauma Perception Scale (BTPS), the Maternal Childbirth Trauma Scale (MCTS), the Chinese version of the Traumatic Childbirth Perception Scale (TCPS-C), the Birth Trauma Scale (BTS), the Traumatic Childbirth Perception Scale (TCPS), and the Childbirth Trauma Index (CTI). None of the included assessment tools reported measurement error, cross-cultural validity, hypothesis testing, or responsiveness. Regarding content validity, BTPS was "uncertain," CTI was not reported, and the remaining scales were "sufficient." For structural validity, TCPS-C was "sufficient," while the remaining scales were "uncertain." For internal consistency, MCTS was "insufficient," while the

remaining scales were "sufficient." For reliability, BTPS, MCTS, and TCPS-C were "sufficient," while the remaining scales were not reported. TCPS-C, BTS, and TCPS received Grade A recommendation, while BTPS, MCTS, and CTI received Grade B recommendation.

Conclusion: TCPS-C, BTS, and TCPS are effective assessment tools that can be provisionally recommended; however, their measurement properties and methodological quality remain inadequate. Future research should further investigate and continuously optimize these tools to enhance their scientific rigor and precision, thereby advancing the development of related fields.

Full Text

A Systematic Review of Assessment Tools for Maternal Psychological Birth Trauma

WEN Yongxia[1], **SUN Hai**[2], **CHEN Xiaoju**[1], **CAI Wanjing**[1], **LI Shuni**[1], **GUO Honghua**[3]*

[1]International College of Nursing, Hainan Medical University, Haikou 571199, China
[2]International College of Nursing, Hainan Vocational University of Science and Technology, Haikou 571137, China
[3]International College of Nursing, Hainan Medical University, Haikou 571199, China

*Corresponding author: GUO Honghua, Professor; E-mail: 30531391@qq.com
WEN Yongxia and SUN Hai are co-first authors

Abstract

Background
Psychological birth trauma is closely related to maternal mental health, making its identification and the evaluation of intervention effectiveness critically important. While numerous assessment tools for psychological birth trauma exist, research evaluating the quality of these tools is lacking.

Objective
To systematically evaluate the measurement properties and methodological quality of psychological birth trauma assessment tools to provide a reference for future research.

Methods
A systematic search was conducted across Chinese biomedical literature databases, Wanfang Data Knowledge Service Platform, CNKI, VIP Chinese Journal Service Platform, PubMed, Web of Science, CINAHL, Embase, and APA PsycArticles for studies on psychological birth trauma assessment tools, with the search timeframe from database inception to November 6, 2023. Liter-

ature screening, data extraction, and evaluation were performed independently by two researchers and cross-checked.

Results
Six studies involving six assessment tools were included: the Birth Trauma Perception Scale (BTPS), Maternal Childbirth Trauma Scale (MCTS), Chinese version of the Traumatic Childbirth Perception Scale (TCPS-C), Birth Trauma Scale (BTS), Traumatic Childbirth Perception Scale (TCPS), and Childbirth Trauma Index (CTI). None of the included tools reported measurement error, cross-cultural validity, hypothesis testing, or responsiveness. Regarding content validity, BTPS was rated as "uncertain," CTI did not report this property, and the remaining scales were "adequate." For structural validity, TCPS-C was "adequate" while the other scales were "uncertain." For internal consistency, MCTS was "inadequate" while the remaining scales were "adequate." For reliability, BTPS, MCTS, and TCPS-C were "adequate,"while the other tools did not report this property. TCPS-C, BTS, and TCPS received Grade A recommendations, while BTPS, MCTS, and CTI received Grade B recommendations.

Conclusion
TCPS-C, BTS, and TCPS are valid assessment tools that can be provisionally recommended. However, their measurement properties and methodological quality remain inadequate, warranting further research and continuous optimization to improve the scientific rigor and precision of assessment tools and advance the field.

Keywords
Psychological birth trauma; Maternal; Assessment tools; Measurement properties; Systematic review

---

1.1 Literature Search Strategy

A systematic search was conducted using a combination of subject terms and free-text terms across Chinese biomedical literature databases, Wanfang Data Knowledge Service Platform, CNKI, VIP Chinese Journal Service Platform, PubMed, Web of Science, CINAHL, Embase, and APA PsycArticles for studies on psychological birth trauma assessment tools, with the search timeframe from database inception to November 6, 2023. Chinese search terms included "分娩创伤/创伤性分娩," "分娩," "心理创伤," "问卷/量表/工具/信度/效度/反应度/克朗巴赫系数/组内相关系数/内部一致性." English search terms included "Parturition/Birth/*Childbirth*/Labour/Delivery/Labor," "Traumatic/Trauma," "Psychological," "scale/tool/instrument/questionnaire," "assessment/evaluation/measurement/validity/reliability/responsiveness/Cronbach/ICC/EFA/internal consistency." References of included studies were also traced to ensure comprehensive retrieval.

1.2.1 Inclusion Criteria

(1) Studies involving psychological birth trauma assessment tools; (2) Assessment tools validated in maternal populations; (3) Studies reporting at least one measurement property of the assessment tool.

1.2.2 Exclusion Criteria

(1) Secondary studies (reviews, systematic reviews, or Meta-analyses); (2) Duplicate publications or studies where full text was unavailable; (3) Studies where the assessment tool was used solely as an outcome measure (e.g., interventional studies); (4) Conference abstracts, case reports, lectures, or commentaries.

1.3 Literature Screening and Data Extraction

Two researchers trained in systematic evidence-based practice independently conducted literature screening and data extraction, with cross-checking. Disagreements were resolved through discussion with other researchers. Extracted data included: first author, publication year, country/region, sample size, original target population for development, validation population, theoretical model, etc.

1.4 Literature Quality Evaluation

Two researchers trained in systematic evidence-based practice independently performed the evaluation. The process involved: (1) evaluating the methodological quality of included studies using the COSMIN Risk of Bias checklist; (2) evaluating the measurement property quality using COSMIN quality criteria; (3) grading the evidence level using a modified quantitative systematic review evidence grading method; and (4) forming recommendations based on the methodological quality and evidence grading results. All stages were documented and cross-checked, with disagreements resolved through discussion.

1.4.1 Methodological Quality Evaluation Based on the COSMIN Risk of Bias checklist, which includes 10 evaluation dimensions (scale development, content validity, structural validity, internal consistency, cross-cultural validity, reliability, measurement error, criterion validity, hypothesis testing, and responsiveness) comprising 116 items. A 5-point scoring system was used: "very good (V)," "adequate (A)," "doubtful (D)," "inadequate (I)," and "not applicable (NA)," with the "worst score counts" principle applied to evaluate each assessment tool.

1.4.2 Measurement Property Quality Evaluation Based on COSMIN quality criteria, each tool' s content validity, structural validity, internal consistency, cross-cultural validity, reliability, measurement error, criterion validity, hypothesis testing, and responsiveness were rated as "adequate (+),""inadequate (-)," or "uncertain (?)."

1.4.3 Evidence Grading and Recommendation Evidence grading was performed using a modified quantitative systematic review evidence grading method. The initial assumption was that each measurement property of the assessment tools was of high quality, with downgrading considered based on inconsistency, indirectness, imprecision, and risk of bias. Recommendations were formed as follows: (1) Grade A: content validity rated as "adequate (+)" with at least low-quality evidence supporting "adequate (+)"internal consistency; (2) Grade B: tools not meeting Grade A or C criteria; (3) Grade C: high-quality evidence demonstrating "inadequate (-)" for any measurement property.

2.1 Literature Search Results

The initial search yielded 1,239 articles. After removing duplicates, 1,128 articles remained. Following title and abstract screening, 1,111 articles were excluded. After full-text review, 6 articles were ultimately included [17-22], comprising 6 psychological birth trauma assessment tools (Figure 1 [Figure 1: see original paper]).

2.2 Basic Characteristics of Included Assessment Tools

Six studies involving six assessment tools were included: 2 English scales [20,22], 1 Turkish scale [21], and 3 Chinese scales [17-19], with 1 being a translated scale [19]. All scales targeted maternal populations for development or adaptation. Two tools [17-18] reported theoretical foundations, and three tools [17-19] reported retest intervals. Except for the Maternal Childbirth Trauma Scale (MCTS) [18] with a 7-10 day retest interval, the other two tools [17,19] had 2-4 week retest intervals (Table 1).

2.3.1 Scale Development

The Childbirth Trauma Index (CTI) [22] had unclear reporting of content validity and theoretical models, resulting in "inadequate" methodological quality. The remaining scales were rated as "doubtful" due to insufficient detail in qualitative research processes, including: lack of detailed interviewer qualifications [18-19,21], interview outlines [18-21], recording of interviews [17-18,20], and data analysis methods [18-21].

2.3.3 Structural Validity

For structural validity, TCPS-C and BTS both used confirmatory factor analysis with fit indices >0.90, resulting in "very good" methodological quality and "adequate" measurement property ratings. The remaining four assessment tools [17-18,21-22] used exploratory factor analysis without describing fit indices, yielding "adequate" methodological quality and "uncertain" measurement property ratings.

2.3.4 Internal Consistency

All included assessment tools evaluated internal consistency, resulting in "very good" methodological quality ratings. MCTS [18] was rated as "inadequate" for measurement property due to a Cronbach' s α coefficient of 0.65 (<0.70) for the "maternal role adaptation" dimension. The remaining tools demonstrated Cronbach' s α coefficients of 0.70-0.97 across dimensions ($ $0.70), indicating "adequate" measurement properties.

2.3.5 Reliability

Three studies [17-19] reported test-retest reliability but did not report stability during the retest period or whether retest conditions and methods were similar, resulting in "doubtful" methodological quality. However, with intraclass correlation coefficients >0.70, they demonstrated "adequate" measurement properties. The remaining three studies [20-22] did not report ICC or Pearson/Spearman coefficients, preventing evaluation of measurement properties and methodological quality.

2.3.6 Criterion Validity

TCPS, TCPS-C, and CTI all reported criterion validity evaluation but used common scales as gold standards, which does not meet COSMIN guidelines for criterion validity "gold standards" and had correlation coefficients <0.70. Therefore, measurement property quality was rated as "inadequate" and methodological quality as "inadequate." MCTS, TCPS-C, BTS, and TCPS did not clearly describe the qualitative research process, resulting in "doubtful" methodological quality and "adequate" measurement property ratings (Table 2).

2.4 Evidence Grading Evaluation

Regarding risk of bias, CTI did not report content validity, preventing evidence grading. The methodological quality of content validity for the remaining scales was "doubtful," representing substantial risk of bias, resulting in "low" evidence grades. The methodological quality of structural validity for BTPS, MCTS, TCPS, and CTI was "adequate" with "uncertain" measurement properties, possibly introducing bias, resulting in "moderate" evidence grades. The methodological quality of internal consistency for all included scales was "very good" with "adequate" measurement properties, with no downgrading applied. The methodological quality of reliability for BTPS and TCPS-C was "doubtful," resulting in "low" evidence grades. For MCTS, methodological quality was "inadequate," resulting in "very low" evidence grades. The methodological quality of criterion validity for TCPS-C, TCPS, and CTI was "inadequate," resulting in "very low" evidence grades.

Regarding inconsistency, no assessment tools reported inconsistent structural validity and reliability, so no scales were downgraded for this reason. Regard-

ing imprecision, BTPS [17], MCTS [18], and TCPS-C [19] had retest sample sizes <50, resulting in two-level downgrades. Regarding indirectness, TCPS and TCPS-C included not only postpartum women but also pregnant and nulliparous women, introducing indirectness and resulting in one-level downgrades (Table 3).

2.5 Evidence-Based Recommendations

TCPS-C, BTS, and TCPS demonstrated "adequate" reliability in content validity and internal consistency, with internal consistency evidence reaching "low" grade standards, resulting in Grade A recommendations. Although MCTS internal consistency was "inadequate," its evidence grade did not meet high-quality standards, failing to meet Grade C criteria, thus receiving a Grade B recommendation. BTPS and CTI lacked high-quality evidence of "inadequate" measurement properties, failing to meet Grade A or C criteria, thus receiving Grade B recommendations.

3.1.1 Incomplete Content Validity Reporting and Need for Improved Methodological Quality

Content validity measures the objectivity and authenticity of assessment tools in reflecting specific events or concepts and represents a core measurement property, primarily reflected through subjective evaluation of scale items by experts and participants. However, included scales demonstrated multiple deficiencies in content validity: lack of participant evaluation [17,22], unclear research process descriptions [18-20], and failure to include experts from all relevant disciplines [17-18,20]. This may relate to researchers' unclear conceptualization of "psychological birth trauma" and target populations, preventing accurate inclusion of all relevant experts and participants. Similar deficiencies have been identified in other studies (cancer patient resilience assessment tools [24], fear of cancer recurrence assessment tools [25], elderly social frailty assessment tools [26]), reflecting insufficient awareness and emphasis on content validity among scholars. The COSMIN guidelines [23] emphasize that inadequate or missing content validity during scale development prevents scale items from accurately reflecting target population characteristics and core conceptual elements. Conversely, participant interviews [27-28] can significantly improve scale item readability and acceptability while reducing measurement error. Future scale development research should emphasize qualitative research methods to supplement study design, actively incorporate feedback from research participants and professionals for scale item revision and improvement, and provide detailed research process reporting to reduce methodological quality risk bias and promote scale application.

3.1.2 Unreasonable Structural Validity Analysis and Need for Improved Factor Analysis Application

Structural validity examines whether the measured data structure aligns with the target construct, typically demonstrated through exploratory factor analysis (EFA) and confirmatory factor analysis (CFA), though these methods have different applications [29]. Some scholars suggest using EFA in early scale development and CFA during preliminary scale application [30-31]. COSMIN guidelines [23] propose that CFA is preferable to EFA for scales developed based on theoretical models. In this study, two assessment tools [17-18] were developed based on theoretical models but did not employ CFA, while BTS (without a reported theoretical model) used both factor analysis methods, meeting COSMIN guidelines. This aligns with findings from other assessment tool studies [32-33], indicating most scholars have insufficient understanding and application of data analysis methods for scale development, possibly because most scale developers are not statistics professionals [34-35] and lack specialized data analysis knowledge or adequate professional support. Future scale research should emphasize factor analysis method selection and application, seeking statistical expert support during research design and implementation to ensure appropriate statistical methods that improve scale scientific rigor, accuracy, and effectiveness.

3.1.3 Unclear Stability Description Requiring Further Validation

Stability, or reliability, refers to the consistency of results when remeasuring the same participants under identical conditions at different time points, which is crucial for scale reliability and validity over time [36]. Retest intervals that are too short or too long affect reliability [37], as do inconsistent retest conditions and unstable constructs during measurement intervals [25]. In this study, three studies [17-19] reported reliability, with MCTS [18] using a 7-10 day retest interval and TCPS-C [19] using a 4-week interval, both deviating from COSMIN' s recommended 2-week interval without reporting reasons for interval selection or evidence of construct stability or condition similarity during retest periods. Therefore, methodological quality was rated as "doubtful," indicating that the stability of included assessment tools requires further validation. This aligns with other scale studies [35], possibly due to participant attrition during retest periods and objective limitations in maintaining consistent participant states (emotional status, memory, fatigue levels) and test locations. Future scale research should implement measures to reduce such bias, such as establishing standardized testing protocols or using statistical methods (e.g., analysis of covariance) to adjust for potential confounding factors, or carefully considering these biases when reporting stability to enhance result generalizability.

3.1.4 Noncompliant Gold Standard Setting in Criterion Validity per COSMIN Guidelines

COSMIN guidelines [23] emphasize that patient-reported outcome measures typically lack gold standards, though original scales can serve as gold standards when modifying existing scales. Chen [19], YALNIZ et al. [21], and ANDERSON et al. [22] confused criterion validity testing with hypothesis testing, using common scales as gold standards despite different statistical methods [38], raising questions about these tools' accuracy. ZHANG et al. [24] and ZHANG et al. [25] similarly found most scale developers could not distinguish criterion validity from hypothesis testing, likely due to incomplete understanding of these concepts' roles in scale development. Future research should follow COSMIN guidelines to clarify concepts, purposes, and methods, correctly select "gold standards," and employ appropriate statistical methods to improve research quality and reliability for clinical practice support.

3.2 Incomplete Reporting of Measurement Properties for PBT Assessment Tools

Included studies did not report responsiveness, cross-cultural validity, hypothesis testing, or measurement error, representing incomplete measurement property reporting. Cross-cultural validity measures scale stability and consistency across cultural contexts [39]; inadequate consideration of cultural factors during scale development or adaptation raises questions about applicability across cultures. Measurement error reflects sources of variation beyond true variation [25], potentially masking or distorting true results and affecting scale accuracy. Hypothesis testing validates construct validity through specific, verified hypotheses [40]. Responsiveness, or sensitivity, measures a scale' s ability to detect change in the measured concept [41]; higher responsiveness indicates greater sensitivity and reliability. Therefore, the reliability, applicability, and validity of included PBT assessment tools require further validation, consistent with other scale studies [42-43] and reflecting insufficient awareness of cross-cultural validity, measurement error, hypothesis testing, and responsiveness among Chinese scholars. Future scale research should strengthen testing and detailed reporting of all measurement properties to objectively evaluate scale applicability and reliability.

3.3 Selection and Recommendation of PBT Assessment Tools

BTS, TCPS, and TCPS-C received Grade A recommendations and are provisionally recommended. BTS, developed by Chinese scholar ZHANG et al. [20] based on Chinese women' s PBT experiences (neglect, loss of control, physical/emotional reactions, cognitive/behavioral responses), demonstrated "adequate" measurement properties for content validity, structural validity, and internal consistency, indicating accurate reflection of PBT experiences and sta-

ble results across measurements. However, BTS is currently only available in English; Chinese researchers and clinicians should consider language expression issues and contact original authors for a Chinese version before use.

TCPS, developed by Turkish scholar YALNIZ et al. [21] based on Turkish women' s perceptions of physical, emotional, and psychological trauma, demonstrated "adequate" content validity and internal consistency, indicating accurate PBT conceptualization and good logical coherence among items. The moderate-quality evidence grade further demonstrates research and practical value, providing useful reference for Chinese PBT scale development and promoting cross-cultural psychology research.

In 2023, Chinese scholar CHEN [19] translated TCPS into TCPS-C, indirectly validating it in Chinese women. As a Chinese scale with short completion time, TCPS-C offers linguistic and convenience advantages with good clinical applicability and feasibility. Despite "very low or low" evidence grades requiring quality improvement, this does not diminish TCPS-C' s important practical role, suggesting future research should further validate and improve its evidence grade.

CTI, MCTS, and BTPS received Grade B recommendations, indicating these tools have application potential but require cautious use due to immature methodology and requiring further investigation. CTI, though lacking evaluable content validity, had high-quality evidence for "adequate" internal consistency and was developed specifically for adolescents, offering value after further content validity verification. BTPS, with moderate-quality evidence for "adequate" internal consistency, was developed specifically for vaginal delivery mothers, potentially offering greater specificity and applicability for this population. MCTS, despite less obvious advantages in methodological and measurement property quality and evidence grading, is not limited by delivery mode, offering broader applicability and non-negligible value for postpartum PBT assessment. Thus, current PBT assessment tools offer varying research value, suggesting Chinese PBT researchers should carefully select appropriate tools based on study design and clinical context.

This study comprehensively evaluated PBT assessment tools' methodological and measurement property quality using COSMIN guidelines. After careful evaluation, TCPS-C, BTS, and TCPS are provisionally recommended as effective assessment tools. However, all included PBT assessment tools have methodological and measurement property quality limitations, suggesting future research should continue in-depth investigation and optimization to improve accuracy and effectiveness, providing high-quality assessment tools for identifying, preventing, and treating maternal PBT and advancing the field.

Author Contributions: WEN Yongxia proposed the research concept and design, conducted literature search and evaluation, drafted and revised the manuscript, and takes responsibility for the work. SUN Hai conducted literature search and quality evaluation and created figures and tables. CAI Wanjing revised the manuscript and extracted data. LI Shuni collected and organized

data and revised the manuscript. ZHANG Caihong supervised the research process, revised the manuscript, and controlled article quality. GUO Honghua designed the study protocol, revised the final version, and takes responsibility for the work.

Conflict of Interest: None declared.

ORCID IDs:
WEN Yongxia: https://orcid.org/0009-0005-7595-4946
SUN Hai: https://orcid.org/0009-0008-8138-3555
GUO Honghua: https://orcid.org/0000-0001-7126-5100

Funding: Hainan Provincial Natural Science Foundation Project (ZDYY2022SHFZ083)

Citation: WEN YX, SUN H, CHEN XJ, et al. A systematic review of the assessment tools of maternal psychological birth trauma [J]. Chinese General Practice, 2024. DOI: 10.12114/j.issn.1007-9572.2024.0291. [Epub ahead of print].

Editorial Office: Chinese General Practice.

References

[1] PIDD D, NEWTON M, WILSON I, et al. Optimising maternity care for a subsequent pregnancy after a psychologically traumatic birth: a scoping review[J]. Women Birth, 2023, 36(5): e471-480. DOI: 10.1016/j.wombi.2023.03.006.

[2] SCHULLINGER J N. Birth trauma[J]. Pediatr Clin North Am, 1993, 40(6): 1351-1358. DOI: 10.1016/s0031-3955(16)38671-3.

[3] HUANG D Q, DAI L, ZENG T Y, et al. Exploring contributing factors to psychological traumatic childbirth from the perspective of midwives: a qualitative study[J]. Asian Nurs Res, 2019, 13(4): 270-276. DOI: 10.1016/j.anr.2019.10.002.

[4] BECK C T. Birth trauma: in the eye of the beholder[J]. Nurs Res, 2004, 53(1): 28-35. DOI: 10.1097/00006199-200401000-00005.

[5] WANG H, LIANG Y, ZHOU L H, et al. Concept analysis of birth trauma[J]. Chinese Journal of Modern Nursing, 2023, 29(19): 2647-2652. DOI: 10.3760/cma.j.cn115682-20220819-04052.

[6] LI T, GUO X X, HAN T, et al. Investigation of psychological birth trauma and analysis of influencing factors in women after cesarean section[J]. Chinese Journal of Family Planning, 2024, 32(5): 995-1001. DOI: 10.3969/j.issn.1004-8189.2024.05.004.

[7] SUN X Q, FAN X M, CONG S N, et al. Psychological birth trauma: a concept analysis[J]. Front Psychol, 2023, 13: 1065612. DOI: 10.3389/fpsyg.2022.1065612.

[8] ERTAN D, HINGRAY C, BURLACU E, et al. Post-traumatic stress disorder following childbirth[J]. BMC Psychiatry, 2021, 21(1): 155. DOI: 10.1186/s12888-021-03158-6.

[9] BARUT S, UÇAR T, YıLMAZ A N. Comparison of pregnant women' s anxiety, depression and birth satisfaction based on their traumatic childbirth perceptions[J]. J Obstet Gynaecol, 2022, 42(7): 2729-2737. DOI: 10.1080/01443615.2022.2106832.

[10] WEN Y X, SUN H, LI T, et al. Co-word clustering analysis of research hotspots in psychological birth trauma[J]. Chinese Journal of Nursing Education, 2024, 21(4): 492-498. DOI: 10.3761/j.issn.1672-9234.2024.04.018.

[11] VAN SIELEGHEM S, DANCKAERTS M, RIEKEN R, et al. Childbirth related PTSD and its association with infant outcome: a systematic review[J]. Early Hum Dev, 2022, 174: 105667. DOI: 10.1016/j.earlhumdev.2022.105667.

[12] XU L L, MASTERS G A, MOORE SIMAS T A, et al. Labor and delivery clinician perspectives on impact of traumatic clinical experiences and need for systemic supports[J]. Matern Child Health J, 2023, 27(9): 1651-1662. DOI: 10.1007/s10995-023-03700-5.

[13] HU Y, WANG Z W. Nursing Research[M]. 6th ed. Beijing: People's Medical Publishing House, 2022: 81-97.

[14] MOKKINK L B, DE VET H C W, PRINSEN C A C, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures[J]. Qual Life Res, 2018, 27(5): 1171-1179. DOI: 10.1007/s11136-017-1765-4.

[15] TERWEE C B, PRINSEN C A C, CHIAROTTO A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study[J]. Qual Life Res, 2018, 27(5): 1159-1170. DOI: 10.1007/s11136-018-1829-0.

[16] CHEN Y T, SHEN L J, PENG J, et al. Evaluation of patient-reported outcome measures using the modified quantitative systematic review evidence grading method[J]. Nursing Journal of Chinese People' s Liberation Army, 2020, 37(10): 57-60. DOI: 10.3969/j.issn.1008-9993.2020.10.014.

[17] LIAN Z R. Development and application of a perceived birth trauma assessment tool for women with vaginal delivery[D]. Nanchang: Nanchang University, 2022.

[18] WANG H. Development and psychometric testing of the Maternal Childbirth Trauma Scale[D]. Hefei: Anhui Medical University, 2023.

[19] CHEN C F. Translation and psychometric study of the Traumatic Childbirth Perception Scale[D]. Hefei: Anhui Medical University, 2023.

[20] ZHANG K, WU M, ZENG T Y, et al. Development and psychometric testing of a scale for assessing the psychological birth trauma[J]. Front Psychol, 2023, 14: 1071336. DOI: 10.3389/fpsyg.2023.1071336.

[21] YALNIZ H, CANAN F, GENÇ R. Development of a scale of traumatic childbirth perception[J]. Turk Med J, 2016, 8(3): 81-88.

[22] ANDERSON C. Construct validity of the childbirth trauma index for adolescents[J]. J Perinat Educ, 2011, 20(2): 78-90. DOI: 10.1891/1058-1243.20.2.78.

[23] PRINSEN C A C, MOKKINK L B, BOUTER L M, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures[J]. Qual Life Res, 2018, 27(5): 1147-1157. DOI: 10.1007/s11136-018-1798-3.

[24] ZHANG Y S, ZHANG J, XU C, et al. Systematic review of resilience assessment tools for cancer patients based on COSMIN guidelines[J]. Chinese General Practice, 2024, 27(29): 3664-3671. DOI: 10.12114/j.issn.1007-9572.2023.0717.

[25] ZHANG L L, CHEN H, LUO H, et al. Systematic review of fear of cancer recurrence assessment tools based on consensus-based standards for the selection of health measurement instruments[J]. Chinese General Practice, 2023, 26(17): 2138-2146. DOI: 10.12114/j.issn.1007-9572.2022.0810.

[26] LI H, HOU C M, GAO J, et al. Systematic review of social frailty assessment tools for older adults based on COSMIN guidelines[J]. Chinese General Practice, 2024, 27(24): 3067-3068. DOI: 10.12114/j.issn.1007-9572.2023.0647.

[27] ZHI T T, WANG Y B. Application of cognitive interviewing in cultural adaptation of trauma-informed care attitude scale[J]. Chinese Journal of Nursing, 2022, 57(3): 343-347. DOI: 10.3761/j.issn.0254-1769.2022.03.015.

[28] WANG H Y, GAO Y, YE J R, et al. Application of cognitive interviewing in the development of workplace violence management energy scale for psychiatric nurses[J]. Chinese Journal of Nursing, 2023, 58(17): 2126-2133. DOI: 10.3761/j.issn.0254-1769.2023.17.011.

[29] ZHANG Y, YANG R Z, PENG X X, et al. Principles and models of confirmatory factor analysis and its application in evaluating scale structural validity[J]. Chinese Journal of Health Management, 2022, 16(6): 428-432. DOI: 10.3760/cma.j.cn115624-20220510-00427.

[30] WU H Y, DING J, LUO C J, et al. Development of Traditional Chinese Medicine job satisfaction scale based on confirmatory factor analysis[J]. Chinese Journal of Health Statistics, 2019, 36(1): 92-95.

[31] YAN Z T, LIANG Y, WANG F, et al. Systematic review of self-management assessment tools for Chinese hypertensive patients[J]. Chinese Journal of Nursing Education, 2024, 21(1): 73-80. DOI: 10.3761/j.issn.1672-9234.2024.01.013.

[32] JIN J Z, WANG Y, CHEN R, et al. Systematic review of healthy aging assessment tools based on COSMIN guidelines[J]. Journal of Nursing Science, 2023, 38(11): 97-100, 108. DOI: 10.3870/j.issn.1001-4152.2023.11.097.

[33] OPPENHEIMER A, BOITRELLE F, NICOLAS-BOLUDA A, et al. Measurement properties of sexual function assessment questionnaires in women with

endometriosis: a systematic review following COSMIN guidelines[J]. Acta Obstet Gynecol Scand, 2024, 103(5): 799-823. DOI: 10.1111/aogs.14768.

[34] WEI J H, BU M R, BAI Z Y, et al. Systematic review of psychometric properties of clinical efficacy assessment tools for lung cancer in Traditional Chinese Medicine[J]. Journal of Nursing, 2023, 30(22): 56-62. DOI: 10.16460/j.issn1008-9969.2023.22.056.

[35] LI J J, ZHENG G G, WANG Y, et al. Systematic review of symptom assessment tools for heart failure patients[J]. Chinese General Practice, 2024, 27(18): 2272-2278. DOI: 10.12114/j.issn.1007-9572.2023.0786.

[36] PENG J, SHEN L J, CHEN Y T, et al. Interpretation of the COSMIN-RoB checklist for risk of bias in studies on measurement tool stability, measurement error, and criterion validity[J]. Chinese Journal of Evidence-Based Medicine, 2020, 20(11): 1340-1344. DOI: 10.7507/1672-2531.202003164.

[37] MOREIRA R S, BASSOLLI L, COUTINHO E, et al. Reproducibility and reliability of the quality of life questionnaire in patients with atrial fibrillation[J]. Arq Bras Cardiol, 2016, 106(3): 171-181. DOI: 10.5935/abc.20160026.

[38] TERWEE C B, MOKKINK L B, KNOL D L, et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist[J]. Qual Life Res, 2012, 21(4): 651-657. DOI: 10.1007/s11136-011-9960-1.

[39] RATISLAVOVÁ K, HOROVÁ J, MAREK P. Measuring women' s satisfaction with childbirth: a literature review of measurement properties[J]. Zdr Varst, 2024, 63(2): 100-108. DOI: 10.2478/sjph-2024-0014.

[40] CHEN Y T, PENG J, SHEN L J, et al. Introduction to COSMIN methodology: systematic review for developing patient-reported outcome measures[J]. Journal of Nurses' Training, 2021, 36(8): 699-703. DOI: 10.16821/j.cnki.hsjx.2021.08.005.

[41] CHEN Y L, ZHANG S, CHAI S B, et al. Responsiveness study of patient-reported outcome scale for evaluating Traditional Chinese Medicine efficacy in diabetes[J]. China Journal of Traditional Chinese Medicine and Pharmacy, 2020, 35(4): 2031-2035.

[42] ALMEIDA J A B, FLORÊNCIO R B, LEITE J C, et al. Self-efficacy measurement instruments for individuals with coronary artery disease: a systematic review[J]. PLoS One, 2024, 19(3): e0299041. DOI: 10.1371/journal.pone.0299041.

[43] LAW K, HARRIS E, MCCABE M G, et al. Measurement properties of patient-reported outcome measures for adolescent and young adult survivors of a central nervous system tumor: a systematic review[J]. J Adolesc Young Adult Oncol, 2024, 13(1): 40-54. DOI: 10.1089/jayao.2023.0048.

(Received: 2024-05-16; Revised: 2024-09-08)
(Editor: JIA Mengmeng)

Note: Figure translations are in progress. See original paper for figures.