Post-print of an Umbrella Review of Psychometric Properties of Neonatal Pain Assessment Scales

Authors: Shen Qiao, Tang Yuman, Leng Hongyao, Lei Ruobing, Zheng Xianlan

Date: 2022-07-09T00:00:00+00:00

Abstract

Background Timely and accurate assessment of neonatal pain is crucial for ensuring the safe and effective implementation of pain management protocols. To date, scholars worldwide have developed and tested more than 40 pain assessment scales for different neonatal populations and pain types. However, substantial differences exist in the reliability and stability of their measurement results. Furthermore, existing individual studies or systematic reviews can only provide fragmented evidence regarding specific scales or particular measurement properties, which is detrimental to clinical decision-making by healthcare professionals.

Objective To conduct an overview of systematic reviews on the measurement properties of neonatal pain assessment scales, and to provide evidence-based support for clinicians and researchers in selecting optimal pain assessment instruments.

Methods Computerized searches of Chinese and English databases were conducted for systematic reviews on the measurement properties of neonatal pain assessment scales published within the past 5 years. Two researchers independently performed literature screening and data extraction. The methodological quality, risk of bias, and reporting quality of included studies were evaluated using the Joanna Briggs Institute (JBI) Critical Appraisal Checklist for Systematic Reviews, the Risk of Bias in Systematic Review (ROBIS) tool, and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement, respectively. The quality of evidence from included studies was further assessed using the Confidence in the Evidence from Reviews of Qualitative research (CERQual) tool.

Results A total of 7 systematic reviews were included. Quality assessment results indicated that 4 studies exhibited high methodological quality with low

risk of bias, while 3 studies had relatively low methodological quality with high risk of bias. PRISMA evaluation revealed that 5 studies had relatively complete reporting (reporting completeness rate >60.00%), 1 study had certain reporting deficiencies (reporting completeness rate of 45.95%), and 1 study had serious information omissions (reporting completeness rate of 10.81%). CERQual results demonstrated that among 22 evidence items regarding measurement properties of neonatal pain assessment scales, there were 2 items of high confidence (9.09%), 8 items of moderate confidence (36.36%), 9 items of low confidence (40.91%), and 3 items of very low confidence (13.64%). Evidence synthesis showed that 25 scales demonstrated good internal consistency, inter-rater reliability, construct validity, and interpretability when used for neonatal pain assessment, and were respectively applicable to acute pain, persistent pain, postoperative pain, or mechanical ventilation-related pain in preterm and/or term infants.

Conclusion No single pain assessment scale is currently available for evaluating all types of pain in neonates. It is recommended to select validated assessment scales based on specific clinical contexts such as patient age and pain type to conduct regular dynamic assessments of infant pain. Additionally, further well-designed, rigorously conducted high-quality studies are needed to examine the reliability and stability of existing scales when measuring various types of pain in Chinese neonatal populations, and to explore the feasibility of expanding their application scope.

Full Text

Preamble

Psychometric Properties of Pain Assessment Scales for Neonatal Infants: An Overview of Systematic Reviews

Shen Qiao[1], Tang Yuman[1], Leng Hongyao[1], Lei Ruobing[2], Zheng Xianlan[1]*

Funding: National Natural Science Foundation of China (72074038); 2019 Chongqing Technology Innovation and Application Development Special Project (cstc2019jscx-msxmX0157)

[1] Department of Nursing, Children' s Hospital of Chongqing Medical University, Chongqing 400014, China; National Clinical Research Center for Child Health and Disorders; Ministry of Education Key Laboratory of Child Development and Disorders; Chongqing Key Laboratory of Pediatrics, Chongqing, China

[2] Chevidence Lab Child & Adolescent Health, Department of Pediatric Research Institute, Children' s Hospital of Chongqing Medical University, Chongqing 400014, China

*Corresponding author: ZHENG Xianlan, Professor/Chief Nurse Specialist, Doctoral Supervisor; E-mail: zhengxianlan@vip.163.com

PROSPERO Registration Number: CRD42021292583

Abstract

Background: Timely and accurate assessment of neonatal pain is essential for the safe and effective implementation of pain management protocols. To date, scholars have developed and tested over 40 pain assessment scales for different neonatal populations and pain types. However, the reliability and stability of these instruments vary considerably. Moreover, existing single studies or systematic reviews provide only fragmented evidence regarding specific scales or isolated psychometric properties, which is inadequate for informed clinical decision-making.

Objective: To conduct an overview of systematic reviews evaluating the psychometric properties of neonatal pain assessment scales, providing evidence-based support for clinicians and researchers in selecting optimal pain assessment instruments.

Methods: We systematically searched Chinese and English databases for systematic reviews on the psychometric properties of neonatal pain assessment scales published within the past five years. Two reviewers independently screened literature and extracted data. Methodological quality, risk of bias, and reporting quality were assessed using the Joanna Briggs Institute (JBI) critical appraisal tool for systematic reviews, the Risk of Bias in Systematic Review (ROBIS) tool, and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement, respectively. Evidence quality was evaluated using the Confidence in the Evidence from Reviews of Qualitative research (CERQual) approach.

Results: Seven systematic reviews were included. Quality assessment revealed that four studies exhibited high methodological quality with low risk of bias, while three had relatively low quality with high risk of bias. PRISMA evaluation showed that five reviews were adequately reported (completion rate >60.00%), one had reporting deficiencies (45.95% completion), and one had serious information gaps (10.81% completion). CERQual assessment of 22 evidence items regarding psychometric properties identified two high-quality items (9.09%), eight moderate-quality items (36.36%), nine low-quality items (40.91%), and three critically low-quality items (13.64%). Evidence synthesis indicated that 25 scales demonstrated good internal consistency, inter-rater reliability, construct validity, and interpretability for assessing acute pain, persistent pain, postoperative pain, or mechanically ventilated pain in preterm and/or term infants.

Conclusions: No single pain assessment scale is suitable for evaluating all types of neonatal pain. We recommend selecting validated scales based on specific clinical contexts, including gestational age and pain type, for regular dynamic assessment. Further well-designed, rigorously conducted high-quality studies are needed to examine the reliability and stability of existing scales in Chinese

neonatal populations across various pain types and to explore the feasibility of expanding their applications.

Keywords: Newborn; Pain measurement; Scale; Psychometric evaluation; Overview of systematic reviews

Introduction

Hospitalized neonates inevitably experience various pain stimuli due to medical procedures. Research indicates that 82.6% of common clinical operations can cause severe pain [1]. The Expert Consensus on Neonatal Pain Assessment and Analgesic Management states that standardized, appropriate pain management can mitigate or prevent various adverse physical and psychological effects of pain and accelerate recovery [2]. Timely and accurate pain assessment is critical for implementing pain management and evaluating analgesic effectiveness. Since neonates cannot verbally express pain, healthcare providers must rely on neurophysiological and behavioral measurements. Pain assessment scales incorporating physiological and/or behavioral indicators are considered the most convenient and cost-effective method, requiring only brief training for clinical staff [3].

The Consensus-Based Standards for the selection of Health Measurement Instruments (COSMIN) emphasize that only scales with robust psychometric properties can serve as reliable clinical tools [4]. Currently, over 40 neonatal pain assessment scales have been developed and tested for different neonatal populations and pain types. However, selecting the most appropriate scale for specific clinical scenarios remains unresolved, primarily because the numerous available scales exhibit substantial variability in reliability and stability [5], and existing single studies or systematic reviews provide only fragmented evidence about particular scales or isolated psychometric properties, hindering clinical decision-making. An overview of systematic reviews, which synthesizes evidence from multiple systematic reviews on a specific topic, can provide comprehensive, high-quality evidence for evidence users [6]. Therefore, this study aims to comprehensively evaluate the psychometric properties of existing neonatal pain assessment scales through an overview of systematic reviews, providing evidence-based support for clinicians and researchers in selecting optimal instruments.

1.1 Inclusion and Exclusion Criteria

Inclusion criteria: (1) Studies investigating the psychometric properties of neonatal pain assessment scales, including reliability (internal consistency, inter-rater reliability, test-retest reliability), validity (content validity, construct validity, criterion validity), responsiveness, and interpretability, feasibility, and practicality (the latter three being important scale characteristics despite not

being psychometric properties); (2) Scales validated for hospitalized neonates; (3) Study designs: systematic reviews, meta-analyses, or meta-syntheses.

Exclusion criteria: (1) Studies on other pain assessment methods such as automated neonatal facial expression recognition systems, near-infrared spectroscopy, or heart rate variability analysis; (2) Studies from which psychometric data could not be extracted; (3) Non-Chinese or non-English publications; (4) Systematic review protocols, conference abstracts, or duplicate publications.

1.2 Literature Search Strategy

On October 31, 2021, we developed search strategies combining free-text terms and subject headings (example strategies for CNKI and PubMed are shown in Boxes 1-2). We searched CNKI, Chinese Biomedical Literature Database, Wanfang, VIP, PubMed, Embase, The Cochrane Library, Web of Science, and CINAHL, supplemented by Google Scholar searches and reference tracking of included studies. Since guideline recommendations should be based on the best available current evidence [7] and systematic review evidence remains valid for approximately 3-5 years [8], we limited our search to the most recent five years.

Box 1. CNKI Search Strategy

```
（主题=疼痛 + 镇痛）AND（主题=新生儿 + 足月儿 + 早产儿 + 低出生体重儿 + 小于胎龄儿 + 婴儿）AN
```

Box 2. PubMed Search Strategy

```
#1 "pain"[Title/Abstract] OR "pain"[MeSH Terms]
#2 (newborn OR neonat* OR "newborn infant"[Title/Abstract] OR premature OR "preterm infant"
#3 (delivery[Title/Abstract] OR childbirth[Title/Abstract] OR labor[Title/Abstract]) OR ("De
#4 #1 AND #2 NOT (#3 AND #1)
#5 (measur*[Title/Abstract] OR assess*[Title/Abstract] OR evaluat*[Title/Abstract] OR rating
#6 (psychometr*[Title/Abstract] OR reliability[Title/Abstract] OR consistency[Title/Abstract
#7 #4 AND #5 AND #6, Filters applied: Meta-Analysis, Systematic Review, in the last 5 years
```

1.3 Literature Screening and Data Extraction

We used EndNote software to remove duplicates from combined database search results. Two researchers independently completed literature screening and data extraction, reaching consensus through cross-checking and discussion. Data were extracted using Excel, including: author, year, title, number of included studies, study population, research topic, and main conclusions.

1.4 Quality Assessment of Included Studies

Two researchers independently evaluated the methodological quality, risk of bias, reporting quality, and evidence quality of included studies, with cross-checking and resolution of discrepancies through group discussion.

1.4.1 Literature Quality Assessment **Methodological quality** was assessed using the Joanna Briggs Institute (JBI) critical appraisal tool for systematic reviews [9], which includes 11 items addressing the evidence-based question, inclusion/exclusion criteria, search strategy, literature quality assessment, data extraction and synthesis, and publication bias. Each item is rated as "yes," "no," "unclear," or "not applicable."

Risk of bias was evaluated using the Risk of Bias in Systematic Review (ROBIS) tool [10]. ROBIS comprises three phases: Phase 1 assesses relevance (primarily for interventional, etiological, diagnostic, and prognostic reviews; therefore not applicable to this study); Phase 2 evaluates bias risk during review development across four domains with 21 signaling questions; Phase 3 assesses overall bias risk with three signaling questions. For Phases 2 and 3, reviewers identified information supporting bias risk judgments, answered signaling questions (yes, probably yes, no, probably no, no information), and rated bias risk as low (all signaling questions answered yes/probably yes), high (any signaling question answered no/probably no), or unclear.

1.4.2 Reporting Quality Assessment Since the PRISMA-COSMIN reporting guideline for systematic reviews of measurement instruments is still under development, we used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [11] to evaluate reporting quality. PRISMA comprises seven sections with 27 items. Each item was rated as "yes" (complete reporting), "no" (partial or no reporting), or "not applicable," with reporting completion rates calculated.

1.4.3 Evidence Quality Assessment Since included systematic reviews did not conduct quantitative synthesis due to clinical and methodological heterogeneity, we used the Confidence in the Evidence from Reviews of Qualitative research (CERQual) approach [12] to evaluate evidence quality. Evidence began as high quality and was downgraded based on four domains: methodological limitations, relevance, coherence, and adequacy of data (no downgrade, -1 serious, -2 very serious), resulting in four levels: high, moderate, low, and critically low.

2.1 Literature Screening Process and Results

The initial search yielded 122 records. After removing 34 duplicates with EndNote, we excluded 76 studies based on title and abstract screening for inappropriate study population, topic, design, or language. After full-text review,

we excluded two conference abstracts and three studies with mismatched populations, ultimately including seven systematic reviews for descriptive analysis. The literature screening flowchart is shown in [Figure 1: see original paper].

2.2 Characteristics of Included Systematic Reviews

The number of studies included in each systematic review ranged from 9 to 352, covering 1 to 40 neonatal pain assessment scales that could be categorized into four pain types: acute pain, persistent pain, postoperative pain, and mechanically ventilated pain. One included review published in *JAMA Pediatrics* [13] provided a comprehensive evaluation of psychometric characteristics of existing neonatal pain scales, while the remaining studies focused on a single scale, specific neonatal population, or particular psychometric property: two reviews [14, 15] evaluated the Neonatal Pain, Agitation, and Sedation Scale (N-PASS) and COMFORT scale, respectively; two reviews [16, 17] focused on mechanically ventilated neonates and neonates undergoing maxillofacial surgery; and two reviews [18, 19] assessed responsiveness and practicality of pain scales. The basic characteristics of included systematic reviews are presented in .

2.3 Literature Quality Assessment Results

Quality assessment revealed that four studies demonstrated high methodological quality with low risk of bias [13, 15, 18, 19], while three had relatively low quality with high risk of bias [14, 16, 17]. Primary factors affecting methodological quality and bias risk included: inappropriate inclusion/exclusion criteria for the review question or unclear specification of included study characteristics; incomplete search strategies or omission of important databases; use of inappropriate quality assessment tools for the review question and study types; and failure to extract and report key outcome measures of scale psychometric properties. Detailed methodological quality and risk of bias assessments are shown in and .

2.4 Reporting Quality Assessment Results

PRISMA evaluation showed that five reviews were adequately reported (>60.00% completion rate) [13-15, 18, 19], one had reporting deficiencies (45.95% completion) [16], and one had serious information gaps (10.81% completion) [17]. Inadequate reporting primarily involved structured abstracts, registration and protocols, search strategies, description of processes and methods for study inclusion in each synthesis, heterogeneity analysis, credibility assessment, funding sources, and availability of data, code, and other materials (reporting compliance <50%). Reporting quality results are detailed in .

2.5 Evidence Quality Assessment Results of Included Studies

Using CERQual, we evaluated the quality of 22 evidence items regarding reliability, validity, and responsiveness of neonatal pain assessment scales. Results showed two high-quality items (9.09%), eight moderate-quality items (36.36%), nine low-quality items (40.91%), and three critically low-quality items (13.64%). Downgrading was primarily due to methodological limitations of primary studies, heterogeneity among study results, indirectness of included studies to the review question, and inadequate data (small sample sizes). Since Shapour et al. [17] did not index included studies or report study characteristics, we could not retrieve relevant information to evaluate evidence quality for psychometric properties. Evidence quality assessment results for other included reviews are presented in .

3.1 Selection and Recommendation of Neonatal Pain Assessment Scales

Psychometric properties are a primary consideration when selecting optimal Patient-Reported Outcome Measures (PROMs) [20]. Scales with validated reliability, validity, and responsiveness can accurately reflect neonatal pain status, guiding clinicians to adjust analgesic regimens for optimal pain prevention and control. This overview employed rigorous evaluation and synthesis of evidence on psychometric properties of neonatal pain assessment scales. Included reviews evaluated over 40 scales, with evidence synthesis () indicating that 25 scales demonstrated good internal consistency, inter-rater reliability, construct validity, and interpretability. Among these, eight scales had low COSMIN bias risk, six had moderate risk, and five had high risk, with applicability to acute pain, persistent pain, postoperative pain, or mechanically ventilated pain in preterm and/or term infants. However, no single scale is supported by sufficient evidence for assessing all types of neonatal pain.

The diversity and limited applicability of existing pain scales have substantially reduced pain assessment rates and constrained pain management implementation [21]. Currently, neonatal units must use at least 3-5 different scales simultaneously, posing significant challenges for users and administrators [22]. Furthermore, lack of clear classification and definition of neonatal pain types directly affects clinicians' ability to correctly identify pain types and select appropriate assessment scales. Therefore, recommendations for scale selection should prioritize, while balancing psychometric performance, general-purpose scales applicable to multiple pain types to avoid underestimation or overestimation of pain due to inappropriate scale selection. Our findings indicate that scales applicable to two or more pain types include: Neonatal Facial Coding System

(NFCS), N-PASS, COMFORT, and COMFORT-B for acute and persistent pain (including mechanically ventilated pain) in preterm and term infants (moderate-quality evidence); Pain Assessment Tool (PAT) for postoperative and persistent pain in preterm and term infants (low-quality evidence); and Pain Observation Scale for Young Children (POCIS) for acute, persistent, and postoperative pain in term infants (low-quality evidence).

Beyond psychometric performance and applicability, clinicians must also consider language and clinical practicality of PROMs. Since these scales originated abroad, standard translation procedures should ensure semantic equivalence and content comprehensibility for local application [23]. Currently, only N-PASS [24] and COMFORT-B [25] have Chinese versions, though the latter requires further validation due to lack of back-translation. The Chinese revised version of N-PASS, introduced by He Biyun et al. [24], has been validated for acute procedural pain with mechanical ventilation [26, 27], mechanical ventilation pain alone [25], and postoperative pain alone [28], consistently demonstrating good reliability, validity, and responsiveness. Regarding clinical practicality, Emma et al.' s [19] review found that 2.80% of randomized trials on neonatal analgesic effectiveness used N-PASS. Domestic surveys show that 4.31% of neonatal units nationwide use N-PASS for routine pain assessment [21]. The scale is also recommended by the American Academy of Pediatrics (AAP) and the Neonatologist Branch of the Chinese Medical Doctor Association [2, 29]. Given N-PASS' s potential to accurately measure various pain types in hospitalized neonates, we recommend that Chinese researchers further expand validation of the Chinese version using standardized psychometric evaluation methods. Until more robust evidence emerges, we recommend selecting validated scales based on specific clinical contexts such as patient age and pain type for regular dynamic assessment [30]. Well-designed, rigorously conducted high-quality studies are needed to examine the reliability and stability of existing scales in Chinese neonatal populations across pain types and to explore feasibility of expanding their applications.

3.2 Overall Quality of Included Systematic Reviews is Suboptimal, Requiring Further Standardization in Conduct and Reporting

The WHO Handbook for Guideline Development states that guideline developers must critically appraise systematic review quality, using only recent high-quality reviews as evidence sources [7]. Developing high-quality systematic reviews is essential for selecting optimal PROMs [31]. The COSMIN working group, comprising psychometric experts, developed the COSMIN guideline to direct systematic reviews of PROMs and aid selection of appropriate health measurement instruments [32]. However, included reviews of neonatal pain assessment scales did not follow COSMIN systematic review standards, instead

following conventional systematic review methods that inadequately addressed reliability, validity, and responsiveness outcomes, preventing judgment of scale recommendations and strength. Additionally, included studies exhibited common problems in systematic reviews/meta-analyses [33]: (1) Unclear structure of the evidence question, inclusion/exclusion criteria, and search strategy: failure to include four key elements (construct, target population, type, and measurement properties) [20], with some inappropriately applying the PICO format for interventional studies [16]; (2) Incomplete search resources: omission of Medline and Embase as recommended by COSMIN [31], or lack of supplementary searches; (3) Inappropriate quality assessment standards: only three reviews [13, 15, 18] used the COSMIN risk of bias checklist for psychometric studies; (4) Inadequate data extraction: missing key outcome data for psychometric properties; (5) Missing important information: incomplete reporting in abstracts (28.57% compliance), methods (48.98%), results (44.44%), and other information (38.10%). These limitations directly affect evidence quality and recommendation grades. We recommend that researchers refer to COSMIN guidelines [31] for defining optimal psychometric evaluation studies, and strictly design, conduct, and report scale psychometric studies. Future systematic reviews of neonatal pain assessment scale psychometric properties should follow COSMIN standards [32] to develop final recommendations guiding clinicians and researchers in selecting optimal pain assessment scales.

3.3 Limitations of This Study

This study has several limitations: findings are based on currently published systematic reviews and may be affected by new publications or updates; meta-analysis was not possible due to clinical and methodological heterogeneity, preventing quantitative comparison of psychometric properties across scales; and lack of direct evidence on scale practicality required using usage frequency as a proxy, which may not reflect true clinical utility.

---

4 Summary

Although current evidence does not support a single pain assessment scale for all neonatal pain types, 25 scales demonstrate good psychometric properties for neonatal pain assessment, applicable to acute, persistent, postoperative, or mechanically ventilated pain in preterm and/or term infants. We recommend that clinicians and researchers select validated scales based on specific clinical contexts such as gestational age and pain type. Additionally, we advocate using COSMIN standards to conduct psychometric evaluation studies of neonatal pain assessment scales in Chinese populations, verifying reliability and stability across pain types, and developing systematic reviews to generate recommendations for optimal pain assessment scales.

Author Contributions: Shen Qiao served as the principal investigator, responsible for study design, implementation, analysis, and reporting. Zheng Xianlan, as corresponding author and expert advisor, provided quality control. Shen Qiao, Tang Yuman, Leng Hongyao, and Lei Ruobing conducted literature screening, data extraction, and quality assessment. Shen Qiao drafted the initial manuscript, with all authors providing critical review and revision.

Conflict of Interest: The authors declare no conflict of interest.

References [1] Wang YJ, Li Y, Sun J, et al. Investigation of procedural pain in neonatal intensive care unit patients [J]. Journal of Nursing Science, 2019, 34(11): 20-23. [2] Cheng R, Yang Y, Shi Y, et al. Expert consensus on neonatal pain assessment and analgesic management (2020 edition) [J]. Chinese Journal of Contemporary Pediatrics, 2020, 22(09): 923-930. [3] Wan L, Zhao Q, Chen J, et al. Chinese expert consensus on pain assessment scales (2020 edition) [J]. Chinese Journal of Pain Medicine, 2020, 16(03): 177-187. [4] Chen YT, Peng J, Shen LJ, et al. Introduction to COSMIN methods: interpretation with a systematic review example [J]. Chinese Nursing Research, 2021, 35(09): 1505-1510. [5] Maxwell LG, Fraga MV, Malavolta CP. Assessment of Pain in the Newborn: An Update [J]. Clinics in Perinatology, 2019, 46(4): 693-707. [6] Higgins JPT, Thomas J, Chandler J, et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021). Cochrane, 2021. Available from www.training.cochrane.org/handbook. [7] World Health Organization. WHO Handbook for Guideline Development, 2nd ed [M]. Geneva: World Health Organization, 2014. https://apps.who.int/iris/handle/10665/145714. [8] Shekelle PG, Ortiz E, Rhodes S, et al. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated? [J]. JAMA, 2001, 286(12): 1461-1467. [9] Gu Y, Zhang HW, Zhou YF, et al. JBI evidence-based healthcare center quality assessment tools for different study types—methodological quality assessment for systematic reviews [J]. Journal of Nurses Training, 2018, 33(08): 701-703. [10] Whiting P, Savović J, Higgins JP, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed [J]. Journal of Clinical Epidemiology, 2016, 69: 225-34. [11] Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews [J]. Revista Española de Cardiología (English Edition), 2021, 74(9): 790-799. [12] Lewin S, Booth A, Glenton C, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings: introduction to the series [J]. Implementation Science, 2018, 13(Suppl 1): 2. [13] Giordano V, Edobor J, Deindl P, et al. Pain and Sedation Scales for Neonatal and Pediatric Patients in a Preverbal Stage of Development: A Systematic Review [J]. JAMA Pediatrics, 2019, 173(12): 1186-1197. [14] Morgan ME, Kukora S, Nemshak M, et al. Neonatal Pain, Agitation, and Sedation Scale' s use, reliability, and validity: a systematic review [J]. Journal of Perinatology, 2020, 40(12): 1753-1763. [15] Maaskant J, Raymakers-Janssen P, Veldhoen E, et al. The clinimetric properties of the COMFORT scale: A

systematic review [J]. European Journal of Pain, 2016, 20(10): 1587-1611. [16] Popowicz H, Kwiecień-Jaguś K, Olszewska J, et al. Pain Scales in Neonates Receiving Mechanical Ventilation in Neonatal Intensive Care Units—Systematic Review [J]. Journal of Pain Research, 2020, 13: 1883-1897. [17] Yaripoor S, Khalili A, Joonbakhsh F, et al. Systematic Review of Pain assessment scales in newborns under maxillofacial surgery admitted to the surgical ward [J]. International Journal of Medical Research and Health Sciences, 2016, 5: 41-44. [18] Meesters N, Dilles T, Simons S, et al. Do Pain Measurement Instruments Detect the Effect of Pain-Reducing Interventions in Neonates? A Systematic Review on Responsiveness [J]. The Journal of Pain, 2019, 20(7): 760-770. [19] Olsson E, Ahl H, Bengtsson K, et al. The use and reporting of neonatal pain scales: a systematic review of randomized trials [J]. Pain, 2021, 162(2): 353-360. [20] Chen YT, Peng J, Shen LJ, et al. Introduction to COSMIN methods: developing systematic reviews of patient-reported outcome measures [J]. Journal of Nurses Training, 2021, 36(8): 699-703. [21] Shen Q, Zheng XL, Lin Z, et al. Survey of pain management practices in 66 medical institutions [J]. Chinese Nursing Management, 2019, 19(02): 187-193. [22] Andersen RD, Munsters JMA, Vederhus BJ, et al. Pain assessment practices in Swedish and Norwegian neonatal care units [J]. Scandinavian Journal of Caring Sciences, 2018, 32(3): 1074-1082. [23] Wei X, Yang H. Cross-cultural translation of medical scales: experience and implications [J]. Overseas English, 2021, 11: 207-208+231. [24] He BY. Translation and psychometric evaluation of the Chinese version of the Neonatal Pain, Agitation and Sedation Scale [D]. Zhejiang University, 2017. [25] Jia YG. Application of three behavioral observation scales in assessing pain in burned infants and young children [D]. Jinan University, [year not specified]. [26] Pang W, Zhang HL, Zhang XB, et al. Value analysis of different scales for assessing acute pain in mechanically ventilated neonates [J]. Progress in Modern Biomedicine, 2019, 19(02): 262-265+283. [27] Huang XZ. Comparison of three scales for assessing acute pain in mechanically ventilated neonates [D]. Southern Medical University, 2017. [28] He XR, Zheng XL, Ke SL, et al. Reliability and validity of three English neonatal pain scales for assessing postoperative pain [J]. Chinese Journal of Evidence-Based Pediatrics, 2021, 16(03): 186-191. [29] COMMITTEE ON FETUS AND NEWBORN and SECTION ON ANESTHESIOLOGY AND PAIN MEDICINE. Prevention and Management of Procedural Pain in the Neonate: An Update [J]. Pediatrics, 2016, 137(2): e20154271. [30] Shen Q, Zheng XL, Li X, et al. Content analysis of clinical practice guidelines for pediatric pain management [J]. Journal of Nursing Science, 2018, 33(07): 50-53. [31] Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures [J]. Quality of Life Research, 2018, 27(5): 1147-1157. [32] Mokkink LB, Prinsen CA, Patrick DL, et al. COSMIN methodology for systematic reviews [EB/OL]. http://www.cosmin.nl/, 2018-02/2022-03-11. [33] Wang ZW, Leng MM, Zhao YJ, et al. Methodological quality analysis of interventional systematic reviews in Chinese nursing literature in recent 5 years [J]. Chinese Nursing Management, 2020, 20(04): 485-489.