JAMA Pediatrics | Review

Pain and Sedation Scales for Neonatal and Pediatric Patients in a Preverbal Stage of Development
A Systematic Review

Vito Giordano, PhD; Joy Edobor, BSc; Philipp Deindl, MD; Brigitte Wildner, MSc; Katharina Goeral, MD; Philipp Steinbauer, MD; Tobias Werther, MD; Angelika Berger, MD; Monika Olischar, MD

Supplemental content

IMPORTANCE Because children in a preverbal stage of development are unable to voice their feelings, they completely depend on their caregiving team for the interpretation and management of their pain and discomfort. Thus, accurately validated scales to assess pain and sedation levels are crucial.

OBJECTIVE To provide clinicians a complete overview on the validity and reliability of the existing pain and sedation scales for different target populations (preterm infants, term infants, and toddlers) and in different clinical contexts.

EVIDENCE REVIEW BIOSIS Previews, Cumulative Index to Nursing and Allied Health Literature, Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, Embase, MEDLINE, PsycCRITIQUES, PsycINFO, PSYNDEXplus Literature and Audiovisual Media, and PSYNDEXplus Tests were the databases screened from their inception to August 2018. All studies examining the validity or reliability of a given pain or sedation scale for patients in a preverbal stage of development were included in this systematic review. Those scales that were tested for at least construct validity, internal consistency, and interrater reliability were subsequently scored using the consensus-based standards for the selection of health measurement instruments (COSMIN) checklist.

FINDINGS In total, 89 validation articles comprising 65 scales were included. Fifty-seven scales (88%) were useful for assessing pain, 13 scales (20%) for assessing sedation, and 4 scales (6%) for assessing both conditions. Forty-two (65%) were behavioral scales, and 23 (35%) were multidimensional scales. Eleven scales (17%) were validated for infants on mechanical ventilation. Thirty-seven scales (57%) were validated for preterm infants, 24 scales (37%) for term and preterm infants, 7 scales (11%) for term-born children, 7 scales (11%) for preterm infants, term infants, and toddlers, and 17 scales (26%) for term infants and toddlers. Twenty-eight scales (43%) considered construct validity, internal consistency, and interrater reliability.

CONCLUSIONS AND RELEVANCE Clinicians should consider using scales that are validated for at least construct validity, internal consistency, and interrater reliability, combining this information with the population of interest and the construct the scale is intended to measure.

JAMA Pediatr. doi:10.1001/jamapediatrics.2019.3351
Published online October 14, 2019.

Author Affiliations: Division of Neonatology, Pediatric Intensive Care, and Neuropediatrics, Department of Pediatrics and Adolescent Medicine, Medical University of Vienna, Vienna, Austria (Giordano, Edobor, Goeral, Steinbauer, Werther, Berger, Olischar); Department of Neonatology and Pediatric Intensive Care Medicine, University Children's Hospital, University Medical Center Hamburg Eppendorf, Germany (Deindl); Information Retrieval Office, University Library of the Medical University of Vienna, Vienna, Austria (Wildner).

Corresponding Author: Vito Giordano, PhD, Division of Neonatology, Pediatric Intensive Care, and Neuropediatrics, Department of Pediatrics and Adolescent Medicine, Medical University of Vienna, Waehringer Guertel 18-20, 1090 Vienna, Austria (vito.giordano@meduniwien.ac.at).

Neonates and infants admitted to a neonatal intensive care unit (NICU) or a pediatric intensive care unit (PICU) are constantly exposed to painful and stressful stimuli for which they may require sedation or pain treatment.[1] Because neonates and infants cannot verbalize their feelings, they depend on the subjective judgment of their care-providing team. Ensuring a patient's optimal level of comfort through adequate administration of sedative and analgesic drugs as well as reducing stress using nonpharmacological approaches is an important challenge for the entire NICU team.[2]

With particular regard to pain treatment and sedation management, targets should be consensually defined by the care-providing team and maintained within optimal ranges over time.[2,3] Item-based scales could help categorize subjective impressions of several people with various levels of expertise into a more objective score.[3]

To accurately assess pain or sedation, item-based scales need to meet requirements, such as unambiguity and comprehensiveness. The scales need to be efficiently conductible at the patient bedside, easy to interpret, and reproducible. They should also enable consensus reaching among different raters and distinguish various levels of pain and sedation. Finally, such a scale must be validated to become a usable clinical tool.[4]

Despite a large number of scales with a variety of psychometric properties having been published in the last decades, to date, there is no criterion standard when considering the assessment of pain and sedation in patients in preverbal stages of development. The aim of this systematic review was, therefore, to update and summarize the existing literature on this topic to provide clinicians an overview and opportunity to choose the appropriate scale for the target population.

Key Points

Question Which are the best validated scales to assess pain and sedation among children in a preverbal stage of development?

Findings This systematic review evaluated 65 scales that have been used for assessing pain or sedation in preterm infants, term infants, or toddlers. Only 28 of 65 scales (43%) had been tested for construct validity, internal consistency, and interrater reliability.

Meaning Clinicians should consider using well-validated scales when assessing pain or sedation in their target population; construct validity, internal consistency, and interrater reliability are prerequisites all such scales should provide.

Methods

This systematic literature research was officially approved and registered in the international prospective register of systematic reviews PROSPERO. We developed a search strategy for the topics "preterm or infant or newborn or neonate or toddlers" in combination with "pain or sedation or distress" and "scale or assessment or tool or measurement." The strategy was established combining free terms and subject headings. All search terms were transformed into a free-term formulation. Truncation was used to obtain all variations of a root word. To be more specific, proximity operators were used, and the search strategy was restricted to certain database fields. In addition, some of the search terms were transformed into subject headings particular for use in MEDLINE and Embase.

The following 10 scientific literature databases were searched from their inception to August 2018: BIOSIS Previews, Cumulative Index to Nursing and Allied Health Literature, Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, Embase, MEDLINE, PsycCRITIQUES, PsycINFO, PSYNDEXplus Literature and Audiovisual Media, and PSYNDEXplus Tests.

We used the literature management program Endnote, version X7.5, to remove duplicates. Scales that were not suitable for infants, scales that measured constructs other than pain or sedation, scales for which only the Abstract was available, and studies in which the data were in a preliminary stage of validation were excluded. Descriptive information was calculated using SPSS, version 21 (IBM SPSS Statistics), and Excel, version 2016 (Microsoft Cooperation). Eligible full texts were concomitantly screened by 2 raters (V.G. and J.E.) (eFigure 1 in the Supplement). The following aspects of validity were screened:

- Face validity, which represents the extent to which a test is subjectively viewed as covering the concept it purports to measure.[5-7]
- Content validity, which refers to the extent to which a measure represents all facets of a given construct.[5-7]
- Construct validity, which refers to the extent to which the scale scores are in line with the theory the scale is based on.[5-7]
- Convergent validity, which indicates the degree to which 2 measures of the same construct, which theoretically should be related, are in fact related.[5-7]
- Discriminant validity, that is, the degree to which a scale is only measuring the intended construct or also something else.[5-7]
- Criterion/concurrent validity, which provides information about the correlation with an already validated scale or an established clinical criterion.[5-7]

Scales that considered at least construct validity, internal consistency, and interrater reliability were then screened for risk of bias through the consensus-based standards for the selection of health measurement instruments (COSMIN) checklist (http://www.cosmin.nl/).[8] The COSMIN checklist consists of 9 boxes (A to I), including internal consistency, reliability, measurement error, content validity, structural validity, hypotheses testing, cross-cultural validity, criterion validity, and responsiveness. Each box consists of 3 to 35 items; all items deal with study design aspects and statistical methods.[9,10] Each item can be rated as excellent, good, fair, or poor. Each box is rated according to the lowest scored item ("worst score counts").[9,10] The following items were used: 1.1, 1.3 to 1.6, 1.9 to 1.35, 2.22 to 2.31, 3.1 to 3.4, 4.1 to 4.5, 5.1 to 5.4, 6.1 to 6.8, 7.1 to 7.6, 8.1 to 8.3, 9.1 to 9.7, and 10.1 to 10.13. Boxes 4, 6, 9, or 10, depending on the approach used, were particularly relevant when controlling the reliability and the construct approach. Because each scale was validated in different steps, we screened the most relevant articles for the best-achieved scores in the most relevant boxes mentioned above. Scales that in their total screening accumulate more than 1 doubtful or 1 or more inadequate score or the combination of both were reported as having moderate to high risk of bias, respectively. Everything that was not assessed was scored as not applicable.

Assessing risk of bias among included studies was done simultaneously but independently by 2 authors (V.G. and J.E.). Reviewers were not blinded to study authors, institution, or journal when assessing risk of bias. Risk of bias judgments within the included studies were recorded in a Microsoft Excel mask (Microsoft Corporation).

Results

Considering the marked discrepancy in the study design among all eligible studies included in this systematic literature research, a meta-analysis was not possible. Data are therefore descriptively presented in the **Table**[11-91] and in eTables 1 through 5 in the Supplement.

Eight databases were screened using the user interface of OvidSP, 1 database using EBSCO, and 1 database using Embase at embase.com. Initially, after duplicate removal, 1335 records were identified. Sixteen additional articles were found through other sources. After considering all exclusion criteria, only 89 articles comprising 65 scales were deemed relevant for this systematic review (eFigure 1 in the Supplement). All the scales included in this systematic review could be used in a clinical setting to assess pain or sedation in infants and children up to 2 years of age. However, these scales differed from one another in some important measures, such as age of the target population, clinical applicability, validity, and reliability (Table; eTable 1 in the Supplement). An overview of all scales is presented in eFigure 2 in the Supplement.

Of 65 scales[4,11-30,32-35,37-43,45-91] 57 scales (88%) aimed to assess pain,[4,11-22,24-30,32-35,37-43,45-61,63-68,70-76,79,81,84,86,87,89-91] 13 scales (20%) aimed to assess sedation,[22,23,35,48-55,62,69,77,78,80,82,83,85,88] and 4 scales (6%) aimed to assess both conditions (COMFORT, COMFORTneo, COMFORT-Behavior Scale [COMFORT-B] and Neonatal Pain, Agitation and Sedation Scale [N-PASS]) (Table).

Of 65 scales, 32 (49%) were validated to assess acute pain,[4,11,12,14,16-21,24-27,29,32-35,37,38,41-43,45,52-56,60,61,64,67,70-74,76,79,84,89] 14 (22%) to assess prolonged pain,[4,13,15,22,24,28-30,35,37-40,46,48-55,58,59,61,76] and 18 (28%) to assess postoperative pain[15,24,27,40,47,48,55,58,59,61,63,66,68,71-76,81,86,87,91] (Table).

The following 11 scales (17%) were validated for infants on mechanical ventilation: modified Postoperative Comfort Score (PCS), Bernese Pain Assessment Scale (BPSN), Neonatal Facial Coding System (NFCS), shortened NFCS (sNFCS), Nepean Neonatal Intensive Care Unit Pain Assessment Tool (NNICUPAT), Hartwig, State Behavioral Scale (SBS), modified COMFORT scale (mCOMFORT), Richmond Agitation-Sedation Scale (RASS), University of Michigan Sedation Scale (UMSS), and Cardiac Analgesic Assessment Scale (CAAS) (Table). There were scales that assessed pain multidimensionally, considering behavioral and physiological measures, and scales that aimed to assess pain and sedation taking only behavioral measures into account.

Of 65 scales, 42 (65%) were behavioral scales[11,13,15,18-20,22,26,28-30,32,34,35,46,47,52-56,58-64,66,67,70-80,82-88,91] and 23 (35%) were multidimensional scales[4,12,14,16,17,21,23-29,33,35,37-43,45,48-51,56,57,66,68,69,81,89] (Table).

Thirty-seven scales (57%) were validated for preterm infants,[4,11-30,32-35,37-43,45-61] 24 scales (37%) were validated for term and preterm infants,[4,17-30,32-35,37-43,45,46,89] and only 7 scales (11%) were validated for a population of term infants alone.[62-64,66] Finally, 7 scales (11%) were validated for preterm infants, term infants, and toddlers,[15,34,35,47-61] and 17 scales (26%) were validated for term infants and toddlers[48,55,61,67-84] (Table).

Only studies with clinical validation were included. All 65 scales had some kind of validity testing (eTable 1 in the Supplement). All 65 scales (100%) controlled for face validity,[4,11-30,32-35,37-43,45-64,66-78,80-89,91] 28 scales (43%) controlled for content validity,[13,17,22,28,32,33,37,41,46-49,56,59,67,70,72,77,82,84,85,87-89] 9 scales (14%) controlled for convergent validity,[4,18,21,26,38,39,49,62,81,87] 39 scales (60%) controlled for concurrent validity,[11,12,14,17,19,21,22,24-27,29,33,34,40,47,48,54,56,57,59,62,64,67-70,72,77,78,82,84,86,91] and 50 scales (77%) controlled for construct validity.[11,13-15,17-21,24-27,32-35,37-39,41,47,48,54,55,57,63,66,68-72,75,76,80,82,84,86-89,91]

Fifty-seven of 65 scales (88%) had some type of reliability testing (eTable 1 in the Supplement).[4,11-14,17-24,26-28,32-34,37,38,40,41,45-47,49,54,56-59,61-64,66-70,72,75,76,78,80,82,84-87,89,91] Thirty-seven scales (57%) considered internal consistency,[4,11-14,17,18,20,22-24,33,34,37,38,41,45-47,49,54,56,57,61-64,70,75,76,84,85,87,89,91] 55 scales (85%) considered interrater reliability,[4,11-14,17-24,26-28,32-34,37,38,40,41,46,47,49,52-56,58,59,61,62,64,66-70,72,75,76,78,80,82,84-87,89,91] and 11 scales (17%) considered intrarater reliability.[4,18,19,21,33,37,38,41,61,64,82,89] Scales were then filtered according to the flowchart presented in **Figure 1**. Those studies that considered construct validity, internal consistency, and interrater reliability were classified as relevant scales (**Figure 2**). Twenty-eight scales (43%) considered construct validity, internal consistency, and interrater reliability (eTable 1 in the Supplement; Figure 2).

After controlling for the presence of defined cutoff points (eTable 2 in the Supplement), the scales were filtered and are presented according to the population and to the construct of interest in **Figure 3**. Among them, those scales with risk of bias lower than the others were COMFORT, Échelle Douleur Inconfort Nouveau-Né (EDIN), EVENDOL behavioral pain scale, NFCS, N-PASS, and the Premature Infant Pain Profile (PIPP) (**Figure 4**; eFigures 3-5 in the Supplement).

Finally, 19 item categories were used to simplify the large amount of information included in all articles considered in this systematic review (eTable 4 in the Supplement). Specifically, 14 item categories (74%) were used in sedation scales,[22,23,35,48-55,62,69,77,78,80,82,83,85,88] and 17 item categories (89%) were used for infants who were mechanically ventilated.[15,21,23,48,58,59,62,68,78,80,82,83]

The 3 most common item categories used in the scales were mimic (51 scales [78%]), activity/movement (36 scales [55%]), and cry (35 scales [54%]). The least common item category was tolerance to ventilation, which was used in 2 scales (3%).[62,80] The most common combination of items was mimic, cry, and movement/activity, which was used in 18 scales (28%).[20,27,39,47,61,64,66,67,69,70,73,75,79,84] Further information about the tested population and the validation process and optimal cutoff for different levels of pain and sedation are in eTables 1 through 5 in the Supplement.

Discussion

According to the present systematic research, 65 scales were identified for the assessment of pain or sedation in children in a preverbal stage of development. The marked variability in the study

Table. Pain and Sedation Scales for Different Target Populations

Scale	Construct	Age	Type of Scale
Preterm			
Behavioral Indicators of Infant Pain[11]	Acute pain	24-32 GW	Behavioral/multivariable
Adapted COMFORT[12]	Acute pain	28-37 GW	Multidimensional/multivariable
EDIN[13]	Prolonged pain	26-36 GW	Behavioral/multivariable
Faceless Acute Neonatal Pain Scale[14]	Acute pain in infant with covered face	30-35 GW	Multidimensional/multivariable
Modified Postoperative Comfort Score = Clinical Scoring System[15]	Ventilated child; Prolonged pain	29-32 GW	Behavioral/multivariable
Nepean Neonatal Intensive Care Unit Pain Assessment Tool[16]	Acute pain; Ventilated infant	25-36 GW	Multidimensional/multivariable
Pain assessment scale for preterm infants[17]	Acute pain	27-36 GW	Multidimensional/multivariable
Preterm and term Infants			
ABC pain scale[18,19]	Acute pain	32-41 GW	Behavioral/ univariable
Acute Pain in Newborns[20]	Acute pain	25-41 GW	Behavioral/multivariable
Bernese Pain Scale for Neonates[21]	Pain in ventilated or not ventilated; VLBW	27-41 GW	Multidimensional/multivariable
COMFORTneo[22]	Sedation; Prolonged pain; ELBW	24-43 GW	Behavioral/multivariable
Modified COMFORT[23]	Sedation in ventilated infant	23-54 GW	Multidimensional/multivariable
COVERS neonatal pain scale[24,25]	Acute pain	23 GW to 2 mo	Multidimensional/multivariable
CRIES Scale[26,27]	Postoperative pain; Acute pain	32 GW to 1 mo	Multidimensional/multivariable
Modified EDIN[28]	Prolonged pain	31-38 GW	Behavioral/multivariable
Faces Pain Scale-Revised[29,30]	Prolonged pain	25 GW to 3 mo	Behavioral/univariable
Harrison[31]	Acute pain	28-41 GW	Multidimensional/multivariable
Infant Body Coding System[32]	Acute pain	25-41 GW	Behavioral/multivariable
Neonatal Acute Pain Assessment Scale[33]	Acute pain	23-42 GW	Multidimensional/multivariable
Neonatal Infant Pain Scale[26,34-36]	Acute pain	27 GW to 7 mo	Behavioral/multivariable
Neonatal Pain, Agitation and Sedation Scale[4,29,37,38]	Acute pain; Prolonged pain; Sedation	23-30 GW	Multidimensional/multivariable
Objective Pain Scale[27]	Postoperative pain	32-60 GW	Multidimensional/multivariable
Pain Assessment in Neonates[39]	Prolonged pain	26-47 GW	Multidimensional/multivariable
Pain Assessment Tool[24,40]	Postoperative pain; Background pain	23 GW to 6 mo	Multidimensional/multivariable
Premature Infant Pain Profile[17,41-44]	Acute pain	32-40 GW	Multidimensional/multivariable
Premature Infant Pain Profile Revised[45]	Acute pain	From 26 GW	Multidimensional /multivariable
Scale for Use in Newborns[35]	Acute pain	24 GW to 7 mo	Multidimensional /multivariable
Swedish ALPS-Neo[46]	Prolonged pain	23-41 GW	Behavioral/multivariable
Preterm and term infants and toddler			
Children's and Infants' Postoperative Pain Scale[47]	Postoperative pain	35 GW to 5 y	Behavioral/multivariable
COMFORT[35,48-51]	Sedation; Prolonged pain	24 GW to 18 y	Multidimensional/multivariable
COMFORT-Behavior Scale[52-55]	Sedation; Acute pain; Prolonged pain; Child with Down syndrome	35 GW to 4 y	Behavioral/multivariable
Multidimensional Assessment of Pain Scale[56,57]	Postoperative pain	36 GW to 31 mo	Multidimensional/multivariable
Neonatal Facial Coding System[15,58,59]	Ventilated child; Prolonged pain; Postoperative pain	29 GW to 18 mo	Behavioral/multivariable
Shortened NFCS[59]	Ventilated child; Prolonged pain	35 GW to 18 mo	Behavioral/multivariable
Observational visual analog scale[34,56,60,61]	Acute pain	35 GW to 4 y	Behavioral/univariable

(continued)

Table. Pain and Sedation Scales for Different Target Populations (continued)

Scale	Construct	Age	Type of Scale
Term infant			
Hartwig[62]	Sedation in ventilated child	0-10 mo	Behavioral/multivariable
Liverpool Infant Distress Scale[63]	Postoperative pain	Term newborn	Behavioral/multivariable
Modified Behavioral Pain Scale[64]	Acute pain	4-6 mo	Behavioral/multivariable
Modified Behavioral Pain Scale[65]	Acute pain	2-12 mo	Behavioral/univariable
Modified infant pain scale[66]	Postoperative pain	1-7 mo	Multidimensional/multivariable
Neonatal Facial Coding System[65]	Acute pain	2-12 mo	Behavioral/multivariable
Partial modified infant pain scale[66]	Postoperative pain	1-7 mo	Behavioral/multivariable
Infant and Toddler			
Alder Hey Triage Pain Score[67]	Acute pain	0-16 y	Behavioral/multivariable
Cardiac Analgesic Assessment Scale[48,68]	Postoperative pain after cardiac surgery in intubated and ventilated patient	0-16 y	Multidimensional/multivariable
Dartmouth Operative Conditions Scale[69]	Sedation	8 mo to 12 y	Multidimensional/multivariable
EVENDOL behavioral pain scale[70]	Acute and prolonged pain	0-7 y	Behavioral/multivariable
Face, Legs, Activity, Cry and Consolability[36,71-74]	Acute pain	1 d to 7 y	Behavioral/multivariable
	Postoperative pain		
Nursing Assessment of Pain Intensity[75]	Postoperative pain	0-36 mo	Behavioral/multivariable
Pain observation scale for young children[55,61,76]	Burned child	0-4 y	Behavioral/multivariable
	Acute pain		
	Background pain		
	Postoperative pain		
Postoperative pain score[75]	Postoperative pain	0-36 mo	Behavioral/multivariable
Ramsay Sedation Scale[51,77]	Sedation	0-19 y	Behavioral/univariable
Richmond Agitation-Sedation Scale[78]	Sedation, ventilated patient	2 mo to 21 y	Behavioral/univariable
Riley Infant Pain Scale[75]	Postoperative pain	0-36 mo	Behavioral/multivariable
Royal College of Emergency Medicine Composite Pain Scale[79]	Acute Pain	0-16 y	Behavioral/multivariable
State Behavioral Scale[80]	Ventilated patient, sedation	1 mo to 6 y	Behavioral/multivariable
Touch Visual Pain Scale[81]	Acute pain	0-13 y	Multidimensional/multivariable
University of Michigan Sedation Scale[78,82,83]	Sedation, ventilated patient	4 mo to 5 y	Behavioral/univariable
University of Wisconsin Children's Hospital Pain Scale[84]	Acute pain	0-16 y	Behavioral/multivariable
Vancouver Sedative Recovery Scale[85]	Sedation	9 mo to 17 y	Behavioral/multivariable
Toddler			
Behavioral Observational Pain Scale[86]	Postoperative pain	1-7 y	Behavioral/multivariable
Children's Hospital of Eastern Ontario Pain Scale[86,87]	Postoperative pain	1-5 y	Behavioral/multivariable
Pasero Opioid-induced Sedation Scale[88]	Sedation depth	1-17 y	Behavioral/univariable
Preverbal, Early verbal Pediatric Pain Scale[89]	Postoperative pain	12-24 mo	Multidimensional/multivariable
Modified Preverbal, Early Verbal Pediatric Pain Scale[90]	Postoperative pain	12-84 mo	Behavioral/multivariable
	Acute pain		
Toddler-Preschooler Postoperative Pain Score[91]	Postoperative pain	1-5 y	Behavioral/multivariable

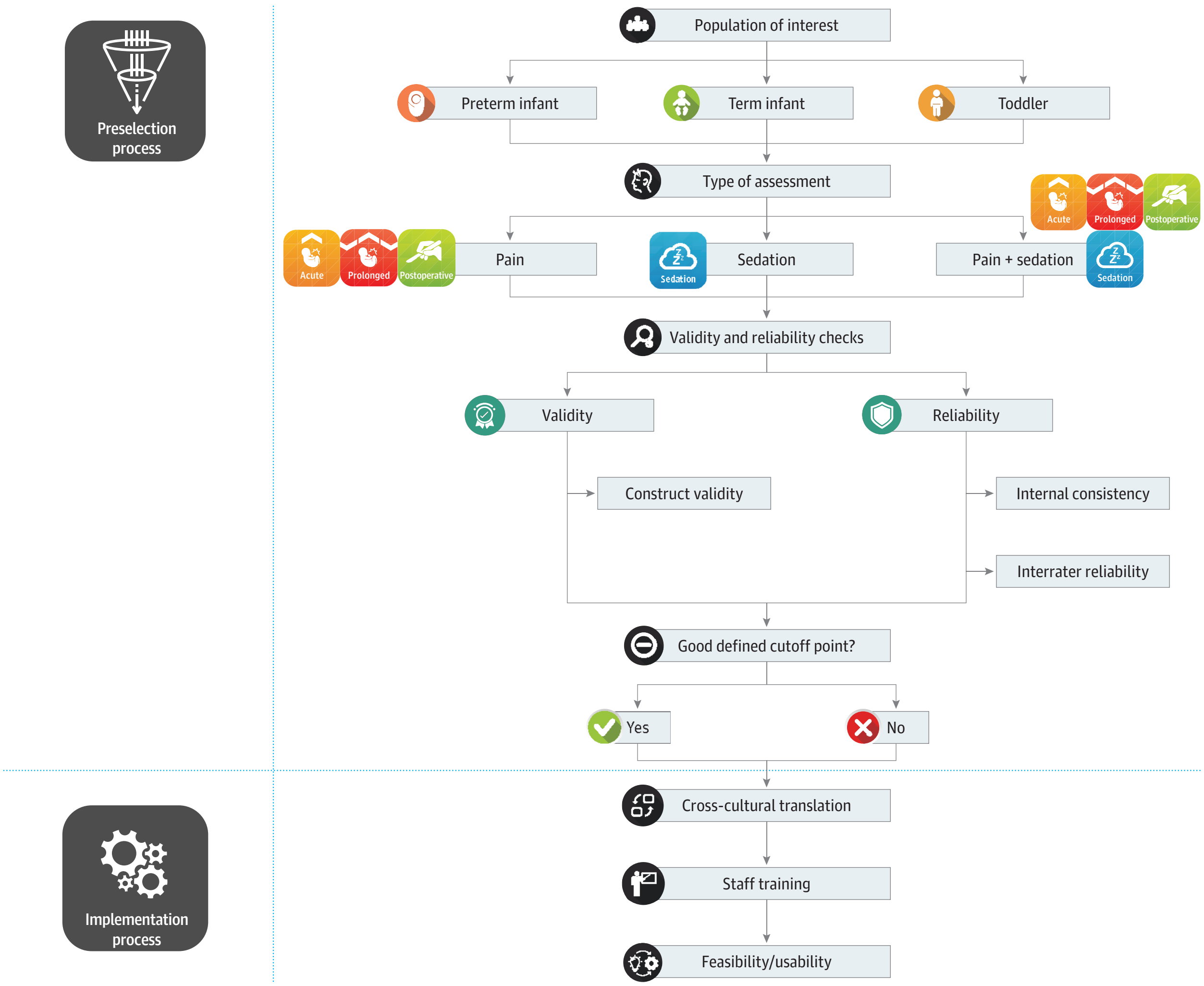
Abbreviations: CRIES, crying requires increased oxygen administration, increased vital signs, expression, sleeplessness: EDIN, Échelle Douleur Inconfort Nouveau-Né; ELBW, extremely low birth weight; GW, gestational weeks; NFCS, Neonatal Facial Coding System; VLBW, very low birth weight.

designs as well as the differences in clinical applicability and population of interest did not allow for a meta-analysis.

According to Slavec and Drnovsek,[92] there are precise steps to consider when validating a scale. These steps include the assessment of dimensionality, validity, and reliability. Assessment of dimensionality is typically conducted in an initial stage of scale development through factor analysis. This step is essential to find the most important factors representing a construct and to control for item homogeneity.[93] Furthermore, it is important to prove both that a given scale is able to measure what it is intended to measure in the context in which it is applied (validity) and that different raters are able to reach consensus using the scale (reliability).

According to our literature research, different statistical approaches were used with the aim to validate the aforementioned scales. None of the included studies tested a scale for all types of validities, whereas all 65 scales were subjectively viewed as covering the concept they were meant to measure. Few studies considered factor analysis or a similar method for the selection of well-defined representative items (COMFORT-B, PIPP-Revised, SBS, pain observation scale for young children (POCIS), PIPP, NFCS, Infant Body Coding System [IBCS], and Children's and Infants' Postoperative Pain Scale [CHIPPS]).[32,37,41,47,54,76,80] Construct validity, which is one of the most important tests to consider when validating a scale, was not always taken into consideration (50 scales

Figure 1. How to Choose a Scale



The process behind scale selection and what to do once a scale has been chosen for the target population.

[77%]).[11,13-15,17-21,24-27,32-35,37-39,41,47,48,54,55,57,63,66,68-72,75,76,80,82,84,86-89,91] The degree to which a scale is only measuring the intended construct (discriminant validity) was also not always considered.

Along with the concept of validity, to control for the reproducibility of a given scale is of great importance from a clinical point of view. A 2011 article by Mandrekar[94] summarizes and clarifies how to easily calculate interrater reliability and when to use κ, weighted κ, and intraclass correlation. In the present study, interrater reliability was tested in most but not all of the scales considered (55 scales [85%]).[4,11-14,17-24,26-28,32-34,37,38,40,41,46,47,49,52-56,58,59,61,62,64,66-70,72,75,76,78-80,82,84-87,89,91]

The American Academy of Pediatrics emphasizes the importance of assessing neonatal pain, especially during and after diagnostic and therapeutic procedures[2] to monitor the effectiveness of pain relief interventions. According to their statement, the following 5 scales are suggested: NFCS, PIPP, N-PASS, Behavioral Indicators of Infant Pain (BIIP) and Acute Pain in Newborns/Douleur Aiguë du Nouveau-né (APN/DAN).

The NFCS[15,58,59] is a behavioral scale that aims to measure acute, prolonged, and postoperative pain in preterm and term neonates by assessing the following "facial actions": brow bulge, eye squeeze, nasolabial furrow, open lips, vertical and horizontal mouth stretch, taut tongue, tongue protrusion, chin quiver, and lip purse.

The PIPP is a very stable scale that has been validated in different steps for various types of validity and has shown good psychometric properties in content and in concurrent and construct validity as well as in interrater reliability.[95,96] The scale aims to assess procedural pain in preterm and term neonates, including items such as gestational age, behavioral state, heart rate, oxygen saturation, brow bulge, eye squeeze, and nasolabial furrow.

The N-PASS was developed as a clinically relevant tool to assess prolonged pain and sedation in infants as well as acute procedural pain.[37,38] Validity and reliability of the N-PASS were provided for acute and prolonged pain and for sedation. Neonates involved in the validation process of the pain subscale had a gestational age at birth between 23 and 40 weeks. The N-PASS includes 5 criteria

Figure 2. Overview of the Relevant Scales

The most relevant scales by clinical applicability as identified by our systematic research.

(crying-irritability, behavior state, facial expression, extremity tone, and vital signs) that are graded 0, 1, or 2 for pain/agitation and 0, −1, or −2 for sedation. The total score is obtained by adding the scores for each criterion.

The BIIP scale measures acute pain in preterm infants (24-32 weeks of gestation) by assessing the following items: "state" (deep sleep, light sleep, drowsy, quite awake, active awake, and agitated/crying), "facial action" (brow bulge, eye squeeze, nasolabial furrow, horizontal mouth stretch, and taut tongue,), and "hand" (hand to mouth, finger splay).[11] Finally, the APN/DAN, which was originally validated in the French language, has shown good psychometric characteristics for the assessment of procedural pain in both preterm and term infants by looking at items such as facial expression, limb movements, and attempts to vocalize.[20]

On the basis of our results, all the scales mentioned above were considered relevant scales (Figures 2-4); however, the NFCS, PIPP, and N-PASS were identified as having lower risk of bias than the BIIP and APN/DAN. Even though important psychometric properties are considered in the BIIP, this scale initially examined only a small percentage of infants assessed at an early gestational age, was validated against the Neonatal Infant Pain Scale (NIPS), which is not considered a criterion standard, and showed moderate correlation to heart rate during painful procedures.[11] In addition, video recording was used to assess pain, which may constitute a bias when assessing some items categories in the clinical setting. Regarding the APN/DAN, it was originally validated in the French language,[20] making it difficult for us to extrapolate all information related to the validation process.

Although the American Academy of Pediatrics statement[2] considers pain assessment in neonates, a recent published scale, the EVENDOL,[70] was meticulously and laudably validated for infants and toddlers. The EVENDOL aims to assess pain by considering the following items: vocal/verbal expression, facial expression, movement, posture, and interaction with the environment.

Another widely used scale, not included within the 5 suggested by the American Academy of Pediatrics, is the Face, Legs, Activity, Cry and Consolability (FLACC) scale. A systematic review of this scale stated that it presents limited and conflicting data for pro-

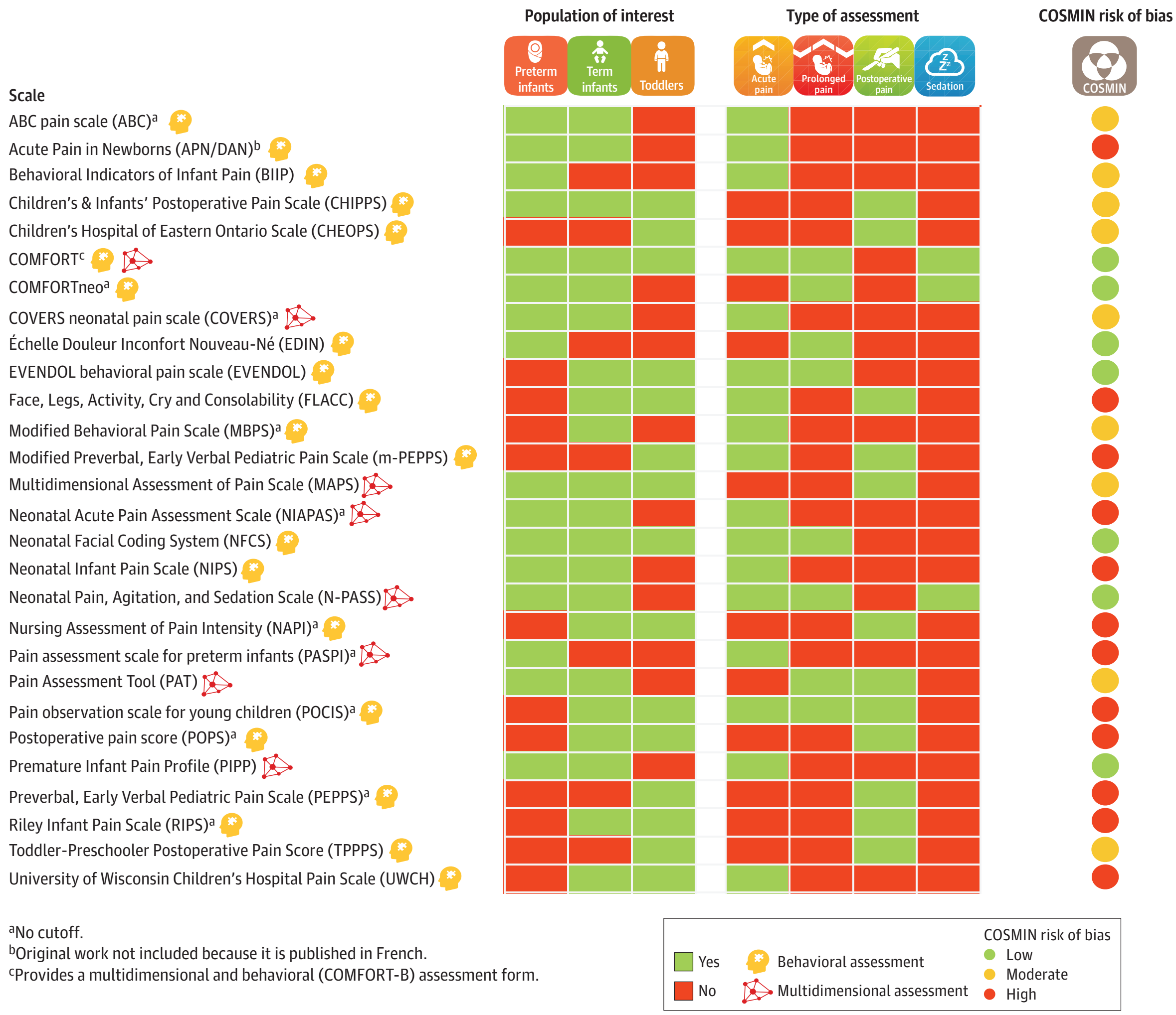
Figure 3. Overview of Relevant Scales With Cutoff

The most relevant scales providing cutoffs, organized by both population of interest and clinical applicability.

cedural pain and that there is insufficient data to support the use of the FLACC scale under all the circumstances and in all the populations in which it is currently being applied.[73]

Procedural pain is not the only condition that can be assessed in a NICU/PICU setting. Postoperative pain, prolonged pain, and sedation represent the daily reality occurring in an intensive care unit as well.

Figure 4. Heatmap of Relevant Scales With COSMIN Risk of Bias



The most relevant scales, organized by population of interest, clinical applicability, and risk of bias.

The EDIN was validated in 76 infants between 25 and 36 weeks of gestation and among infants who were mechanically ventilated.[13] The study assessed prolonged pain through body movements and facial activity.

Another important scale included in our selection (Figure 2 and Figure 3) for the assessment of postoperative pain was the CHIPPS. The CHIPPS was laudably validated, considering a large number (584) of infants.[47] However, it was not clearly stated which correlation coefficient was used in assessing the interrater reliability, which is considered a bias in the COSMIN checklist. Even considering that the stated total interrater reliability score was excellent, the item "cry" obtained a very low agreement between the raters when assessing preterm infants.

Only the N-PASS and the COMFORT scales aimed to assess both pain and sedation. The authors of the COMFORT scale[49] were among the first to show that vital signs are not representative of the assessment of sedation. In fact, the original version of the scale, which includes items such as alertness, calmness, respiratory response, crying, physical movement, muscle tone, and facial tension, was subsequently modified and adapted to the construct of sedation.[12,49,97] This shows how important it is to control for the internal consistency of a scale when applying it in a specific clinical setting. Internal consistency is not the only measure to assess when implementing a scale in a NICU/PICU setting. Construct validity and interrater reliability should also be provided. Figure 2 summarizes all the examined scales that provided information for construct validity, internal consistency, and interrater reliability. However, in some cases these validated aspects were tested at different times and in different populations, sometimes resulting in inadequate to questionable validity and reliability (eTables 1, 3, and 5 in the Supplement). Choosing one among the other scales demonstrating good validity and reliability (Figure 4; eTable 3 in the Supplement) will then depend on their accuracy (ie, ability to distinguish different levels of pain and sedation) but also on the population of interest (eg, preterm vs full-term infants).

Going beyond the concept of validity, it is important from a clinical point of view to concretely define cutoff points for the discrimination of different levels of pain or sedation. Cutoff points could be helpful for the establishment of adequate intervention according to a predefined protocol for the treatment of patients. Rana and collegues,[98] for example, defined NIPS values higher than 3 as a sign that pain relief intervention was needed in their pain and sedation guideline. Deindl and collegues,[3] based their pain and sedation management protocol on well-defined N-PASS cutoff points. The same was successfully done using the COMFORT scales, reducing withdrawal and time on mechanical ventilation among patients postoperatively following cardiac surgery.[99]

Although a given scale per se, if representative of a given construct, can help categorize the opinions of different health personnel, promoting objectivity and building a pain and sedation history of a specific patient on a longitudinal continuum, a well-defined cutoff point can help the clinician *hic et nunc* to achieve a patient's maximal level of comfort (eg, nonpharmacological vs pharmacological intervention, or adapting sedation).

Finally, cross-cultural translation and validation as well as implementation and evaluation of the effects are important steps to perform when applying a given scale in daily clinical practice.[3,100] (Figure 1).

In summary, on the basis of the present systematic research, the most relevant scales are presented in Figure 2 and are presented organized for clinical applicability in Figure 3. All the scales presented in Figure 3 could be used in the clinical setting; however, considering the different validation approaches, they may present some bias. Scales with lower risk of bias compared with the others were COMFORT, EDIN, EVENDOL, NFCS, N-PASS, and PIPP (Figure 4).

Strengths and Limitations

This systematic literature research provides a complete overview of scales used in neonatal setting. However, owing to the different designs and inhomogeneity of the studies, it was not possible to perform a meta-analysis. Consequently, a conclusive ranking of scales was not possible. In addition, some scales were not considered because they were not published in English.

Conclusions

According to the present systematic literature research results, various scales assessing pain or sedation have been published with different levels of validity and reliability. We provided a complete list of scales and their psychometric characteristics suitable for use among patients in a preverbal stage of development. We suggest the use of scales that are validated for construct validity, internal consistency, and interrater reliability and further suggest choosing a particular scale based on the population of interest and the construct intended to measure.

ARTICLE INFORMATION

Accepted for Publication: June 5, 2019.

Published Online: October 14, 2019. doi:10.1001/jamapediatrics.2019.3351

Author Contributions: Dr Giordano had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: Giordano, Edobor, Werther, Berger, Olischar.
Acquisition, analysis, or interpretation of data: Giordano, Edobor, Deindl, Wildner, Goeral, Steinbauer, Olischar.
Drafting of the manuscript: Giordano, Edobor, Wildner, Steinbauer, Olischar.
Critical revision of the manuscript for important intellectual content: Giordano, Deindl, Goeral, Steinbauer, Werther, Berger, Olischar.
Statistical analysis: Giordano, Edobor, Goeral, Steinbauer, Olischar.
Administrative, technical, or material support: Giordano, Edobor, Steinbauer, Olischar.
Supervision: Giordano, Deindl, Werther, Berger, Olischar.

Conflict of Interest Disclosures: None reported.

REFERENCES

1. Carbajal R, Eriksson M, Courtois E, et al; EUROPAIN Survey Working Group. Sedation and analgesia practices in neonatal intensive care units (EUROPAIN): results from a prospective cohort study. *Lancet Respir Med*. 2015;3(10):796-812. doi:10.1016/S2213-2600(15)00331-8

2. Watterberg KL, Cummings JJ, Benitz WE, et al; Committee on Fetus and Newborn and Section on Anesthesiology and Pain Medicine. Prevention and management of procedural pain in the neonate: an update. *Pediatrics*. 2016;137(2):e20154271. doi:10.1542/peds.2015-4271

3. Deindl P, Unterasinger L, Kappler G, et al. Successful implementation of a neonatal pain and sedation protocol at 2 NICUs. *Pediatrics*. 2013;132 (1):e211-e218. doi:10.1542/peds.2012-2346

4. Giordano V, Deindl P, Kuttner S, Waldhör T, Berger A, Olischar M. The Neonatal Pain, Agitation and Sedation Scale reliably detected oversedation but failed to differentiate between other sedation levels. *Acta Paediatr*. 2014;103(12):e515-e521. doi:10.1111/apa.12770

5. Moosbrugger H, Kelava A. Qualitätsanforderungen an einen psychologischen Test (Testgütekriterien). In: Moosbrugger H, Kelava A, eds. *Testtheorie und Fragebogenkonstruktion*. Berlin, Germany: Springer-Lehrbuch; 2012:7-26. doi:10.1007/978-3-642-20072-4_2

6. Kimberlin CL, Winterstein AG. Validity and reliability of measurement instruments used in research. *Am J Health Syst Pharm*. 2008;65(23): 2276-2284. doi:10.2146/ajhp070364

7. Price PC, Jhangiani R, Chiang I-CA. Chapter 5: psychological measurement. In: *Research Methods in Psychology*. BCcampus; 2015:83-102.

8. Playfor S, Jenkins I, Boyles C, et al; United Kingdom Paediatric Intensive Care Society Sedation; Analgesia and Neuromuscular Blockade Working Group. Consensus guidelines on sedation and analgesia in critically ill children. *Intensive Care Med*. 2006;32(8):1125-1136. doi:10.1007/s00134-006-0190-x

9. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010;19(4):539-549. doi:10.1007/s11136-010-9606-8

10. Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, de Vet HCW. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res*. 2012;21(4):651-657. doi:10.1007/s11136-011-9960-1

11. Holsti L, Grunau RE. Initial validation of the Behavioral Indicators of Infant Pain (BIIP). *Pain*. 2007;132(3):264-272. doi:10.1016/j.pain.2007.01.033

12. Caljouw MAA, Kloos MAC, Olivier MY, et al. Measurement of pain in premature infants with a gestational age between 28 to 37 weeks: Validation of the adapted COMFORT scale. *J Neonatal Nurs*. 2007;13(1):13-18. doi:10.1016/j.jnn.2006.11.007

13. Debillon T, Zupan V, Ravault N, Magny JF, Dehan M. Development and initial validation of the EDIN scale, a new tool for assessing prolonged pain in preterm infants. *Arch Dis Child Fetal Neonatal Ed*. 2001;85(1):F36-F41. doi:10.1136/fn.85.1.F36

14. Milesi C, Cambonie G, Jacquot A, et al. Validation of a neonatal pain scale adapted to the new practices in caring for preterm newborns. *Arch Dis Child Fetal Neonatal Ed*. 2010;95(4):F263-F266. doi:10.1136/adc.2008.144758

15. Guinsburg R, Kopelman BI, Anand KJS, de Almeida MFB, Peres CdeA, Miyoshi MH. Physiological, hormonal, and behavioral responses to a single fentanyl dose in intubated and ventilated preterm neonates. *J Pediatr*. 1998;132(6):954-959. doi:10.1016/S0022-3476(98)70390-7

16. Marceau J. Pilot study of a pain assessment tool in the neonatal intensive care unit. *J Paediatr Child Health*. 2003;39(8):598-601. doi:10.1046/j.1440-1754.2003.00215.x

17. Liaw JJ, Yang L, Chou HL, Yin T, Chao SC, Lee TY. Psychometric analysis of a Taiwan-version pain assessment scale for preterm infants. *J Clin Nurs*. 2012;21(1-2):89-100. doi:10.1111/j.1365-2702.2011.03770.x

18. Bellieni CV, Bagnoli F, Sisto R, Neri L, Cordelli D, Buonocore G. Development and validation of the ABC pain scale for healthy full-term babies. *Acta Paediatr*. 2005;94(10):1432-1436. doi:10.1080/08035250510039919

19. Bellieni C, Maffei M, Ancora G, et al. Is the ABC pain scale reliable for premature babies? *Acta Paediatr*. 2007;96(7):1008-1010. doi:10.1111/j.1651-2227.2007.00355.x

20. Carbajal R, Paupe A, Hoenn E, Lenclen R, Olivier-Martin M. APN: a behavioral acute pain rating scale for neonates [in French]. *Arch Pediatr*. 1997;4(7):623-628. doi:10.1016/S0929-693X(97)83360-X

21. Cignacco E, Mueller R, Hamers JPH, Gessler P. Pain assessment in the neonate using the Bernese Pain Scale for Neonates. *Early Hum Dev*. 2004;78(2):125-131. doi:10.1016/j.earlhumdev.2004.04.001

22. van Dijk M, Roofthooft DW, Anand KJ, et al. Taking up the challenge of measuring prolonged pain in (premature) neonates: the COMFORTneo scale seems promising. *Clin J Pain*. 2009;25(7):607-616. doi:10.1097/AJP.0b013e3181a5b52a

23. Lee WK, Young BWY. Measuring the sedation level of mechanically ventilated infants by a modified COMFORT scale. *HK J Paediatr*. 2005;10:189-195. https://pdfs.semanticscholar.org/d594/5d2b17e5ddf5c186d208afa78a7ac2c969c1.pdf. Accessed September 11, 2019.

24. O'Sullivan AT, Rowley S, Ellis S, Faasse K, Petrie KJ. The validity and clinical utility of the COVERS scale and pain assessment tool for assessing pain in neonates admitted to an intensive care unit. *Clin J Pain*. 2016;32(1):51-57. doi:10.1097/AJP.0000000000000228

25. Hand IL, Noble L, Geiss D, Wozniak L, Hall C. COVERS neonatal pain scale: development and validation. *Int J Pediatr*. 2010;2010:496719. doi:10.1155/2010/496719

26. Suraseranivongse S, Kaosaard R, Intakong P, et al. A comparison of postoperative pain scales in neonates. *Br J Anaesth*. 2006;97(4):540-544. doi:10.1093/bja/ael184

27. Krechel SW, Bildner J. CRIES: a new neonatal postoperative pain measurement score: initial testing of validity and reliability. *Paediatr Anaesth*. 1995;5(1):53-61. doi:10.1111/j.1460-9592.1995.tb00242.x

28. Raffaeli G, Cristofori G, Befani B, et al. EDIN scale implemented by gestational age for pain assessment in preterms: a prospective study. *Biomed Res Int*. 2017;2017:9253710. doi:10.1155/2017/9253710

29. Garten L, Deindl P, Schmalisch G, Metze B, Bührer C. Parallel assessment of prolonged neonatal distress by empathy-based and item-based scales. *Eur J Pain*. 2010;14(8):878-881. doi:10.1016/j.ejpain.2010.01.004

30. Hicks CL, von Baeyer CL, Spafford PA, van Korlaar I, Goodenough B. The Faces Pain Scale-Revised: toward a common metric in pediatric pain measurement. *Pain*. 2001;93(2):173-183. doi:10.1016/S0304-3959(01)00314-1

31. Harrison D, Boyce S, Loughnan P, Dargaville P, Storm H, Johnston L. Skin conductance as a measure of pain and stress in hospitalised infants. *Early Hum Dev*. 2006;82(9):603-608. doi:10.1016/j.earlhumdev.2005.12.008

32. Craig KD, Whitfield MF, Grunau RVE, Linton J, Hadjistavropoulos HD. Pain in the preterm neonate: behavioural and physiological indices. *Pain*. 1993;52(3):287-299. doi:10.1016/0304-3959(93)90162-I

33. Pölkki T, Korhonen A, Axelin A, Saarela T, Laukkala H. Development and preliminary validation of the Neonatal Infant Acute Pain Assessment Scale (NIAPAS). *Int J Nurs Stud*. 2014;51(12):1585-1594. doi:10.1016/j.ijnurstu.2014.04.001

34. Lawrence J, Alcock D, McGrath P, Kay J, MacMurray SB, Dulberg C. The development of a tool to assess neonatal pain. *Neonatal Netw*. 1993;12(6):59-66.

35. Blauer T, Gerstmann D. A simultaneous comparison of three neonatal pain scales during common NICU procedures. *Clin J Pain*. 1998;14(1):39-47. doi:10.1097/00002508-199803000-00006

36. Taddio A, Hogan ME, Moyer P, et al. Evaluation of the reliability, validity and practicality of 3 measures of acute pain in infants undergoing immunization injections. *Vaccine*. 2011;29(7):1390-1394. doi:10.1016/j.vaccine.2010.12.051

37. Hummel P, Puchalski M, Creech SD, Weiss MG. Clinical reliability and validity of the N-PASS: Neonatal Pain, Agitation and Sedation Scale with prolonged pain. *J Perinatol*. 2008;28(1):55-60. doi:10.1038/sj.jp.7211861

38. Hummel P, Lawlor-Klean P, Weiss MG. Validity and reliability of the N-PASS assessment tool with acute pain. *J Perinatol*. 2010;30(7):474-478. doi:10.1038/jp.2009.185

39. Hudson-Barr D, Capper-Michel B, Lambert S, Palermo TM, Morbeto K, Lombardo S. Validation of the Pain Assessment in Neonates (PAIN) scale with the Neonatal Infant Pain Scale (NIPS). *Neonatal Netw*. 2002;21(6):15-21. doi:10.1891/0730-0832.21.6.15

40. Spence K, Gillies D, Harrison D, Johnston L, Nagy S. A reliable pain assessment tool for clinical assessment in the neonatal intensive care unit. *J Obstet Gynecol Neonatal Nurs*. 2005;34(1):80-86. doi:10.1177/0884217504272810

41. Stevens B, Johnston C, Petryshen P, Taddio A. Premature Infant Pain Profile: development and initial validation. *Clin J Pain*. 1996;12(1):13-22. doi:10.1097/00002508-199603000-00004

42. Stevens B, Johnston C, Taddio A, Gibbins S, Yamada J. The Premature Infant Pain Profile: evaluation 13 years after development. *Clin J Pain*. 2010;26(9):813-830. doi:10.1097/AJP.0b013e3181ed1070

43. Slater R, Cantarella A, Franck L, Meek J, Fitzgerald M. How well do clinical pain assessment tools reflect pain in infants? *PLoS Med*. 2008;5(6):e129. doi:10.1371/journal.pmed.0050129

44. Nimbalkar SM, Chaudhary NS, Gadhavi KV, Phatak A. Kangaroo Mother Care in reducing pain in preterm neonates on heel prick. *Indian J Pediatr*. 2013;80(1):6-10. doi:10.1007/s12098-012-0760-6

45. Gibbins S, Stevens BJ, Yamada J, et al. Validation of the Premature Infant Pain Profile-Revised (PIPP-R). *Early Hum Dev*. 2014;90(4):189-193. doi:10.1016/j.earlhumdev.2014.01.005

46. Lundqvist P, Kleberg A, Edberg AK, Larsson BA, Hellström-Westas L, Norman E. Development and psychometric properties of the Swedish ALPS-Neo pain and stress assessment scale for newborn infants. *Acta Paediatr*. 2014;103(8):833-839. doi:10.1111/apa.12672

47. Büttner W, Finke W. Analysis of behavioural and physiological parameters for the assessment of postoperative analgesic demand in newborns, infants and young children: a comprehensive report on seven consecutive studies. *Paediatr Anaesth*. 2000;10(3):303-318. doi:10.1046/j.1460-9592.2000.00530.x

48. Cury MR, Martinez FE, Carlotti AP. Pain assessment in neonates and infants in the post-operative period following cardiac surgery. *Postgrad Med J*. 2013;89(1048):63-67. doi:10.1136/postgradmedj-2012-130808

49. Ambuel B, Hamlett KW, Marx CM, Blumer JL. Assessing distress in pediatric intensive care environments: the COMFORT scale. *J Pediatr Psychol*. 1992;17(1):95-109. doi:10.1093/jpepsy/17.1.95

50. Tristão RM, Garcia NVM, de Jesus JAL, Tomaz C. COMFORT behaviour scale and skin conductance activity: what are they really measuring? *Acta Paediatr*. 2013;102(9):e402-e406. doi:10.1111/apa.12325

51. Lamas A, López-Herce J, Sancho L, et al. Assessment of the level of sedation in children after cardiac surgery. *Ann Thorac Surg*. 2009;88(1):144-150. doi:10.1016/j.athoracsur.2009.03.074

52. Valkenburg AJ, Boerlage AA, Ista E, Duivenvoorden HJ, Tibboel D, van Dijk M. The COMFORT-behavior scale is useful to assess pain and distress in 0- to 3-year-old children with Down syndrome. *Pain*. 2011;152(9):2059-2064. doi:10.1016/j.pain.2011.05.001

53. Ista E, van Dijk M, Tibboel D, de Hoog M. Assessment of sedation levels in pediatric intensive care patients can be improved by using the COMFORT "behavior" scale. *Pediatr Crit Care Med*. 2005;6(1):58-63. doi:10.1097/01.PCC.0000149318.40279.1A

54. van Dijk M, de Boer JB, Koot HM, Tibboel D, Passchier J, Duivenvoorden HJ. The reliability and validity of the COMFORT scale as a postoperative pain instrument in 0 to 3-year-old infants. *Pain*. 2000;84(2-3):367-377. doi:10.1016/S0304-3959(99)00239-0

55. de Jong AEEE, Tuinebreijer WE, Bremer M, van Komen R, Middelkoop E, van Loey N. Construct validity of two pain behaviour observation measurement instruments for young children with burns by Rasch analysis. *Pain*. 2012;153(11):2260-2266. doi:10.1016/j.pain.2012.07.021

56. Ramelet A-S, Rees N, McDonald S, Bulsara M, Abu-Saad HH. Development and preliminary psychometric testing of the Multidimensional Assessment of Pain Scale: MAPS. *Paediatr Anaesth*. 2007;17(4):333-340. doi:10.1111/j.1460-9592.2006.02115.x

57. Ramelet AS, Rees NW, McDonald S, Bulsara MK, Huijer Abu-Saad H. Clinical validation of the Multidimensional Assessment of Pain Scale. *Paediatr Anaesth*. 2007;17(12):1156-1165. doi:10.1111/j.1460-9592.2007.02325.x

58. Grunau RVE, Craig KD. Pain expression in neonates: facial action and cry. *Pain*. 1987;28(3):395-410. doi:10.1016/0304-3959(87)90073-X

59. Peters JWB, Koot HM, Grunau RE, et al. Neonatal Facial Coding System for assessing postoperative pain in infants: item reduction is valid and feasible. *Clin J Pain*. 2003;19(6):353-363. doi:10.1097/00002508-200311000-00003

60. van Dijk M, Koot HM, Saad HH, Tibboel D, Passchier J. Observational visual analog scale in pediatric pain assessment: useful tool or good riddance? *Clin J Pain*. 2002;18(5):310-316. doi:10.1097/00002508-200209000-00006

61. de Jong AEE, Bremer M, Schouten M, Tuinebreijer WE, Faber AW. Reliability and validity of the pain observation scale for young children and the visual analogue scale in children with burns. *Burns*. 2005;31(2):198-204. doi:10.1016/j.burns.2004.09.013

62. Hünseler C, Merkt V, Gerloff M, Eifinger F, Kribs A, Roth B. Assessing pain in ventilated newborns and infants: validation of the Hartwig score. *Eur J Pediatr*. 2011;170(7):837-843. doi:10.1007/s00431-010-1354-9

63. Horgan MF, Glenn S, Choonara I. Further development of the Liverpool Infant Distress Scale. *J Child Health Care*. 2002;6(2):96-106. doi:10.1177/136749350200600203

64. Taddio A, Nulman I, Koren BS, Stevens B, Koren G. A revised measure of acute pain in infants. *J Pain Symptom Manage*. 1995;10(6):456-463. doi:10.1016/0885-3924(95)00058-7

65. DiLorenzo MG, Pillai Riddell R, Flora DB, Craig KD. Infant clinical pain assessment: core behavioral cues. *J Pain*. 2018;19(9):1024-1032. doi:10.1016/j.jpain.2018.03.016

66. Buchholz M, Karl HW, Pomietto M, Lynn A. Pain scores in infants: a modified infant pain scale versus visual analogue. *J Pain Symptom Manage*. 1998;15(2):117-124. doi:10.1016/S0885-3924(98)80009-2

67. Stewart B, Lancaster G, Lawson J, Williams K, Daly J. Validation of the Alder Hey Triage Pain Score. *Arch Dis Child*. 2004;89(7):625-630. doi:10.1136/adc.2003.032599

68. Suominen P, Caffin C, Linton S, et al. The Cardiac Analgesic Assessment Scale (CAAS): a pain assessment tool for intubated and ventilated children after cardiac surgery. *Paediatr Anaesth*. 2004;14(4):336-343. doi:10.1046/j.1460-9592.2003.01241.x

69. Chandran V, Jagadisan B, Ganth B. Validation of Adapted Dartmouth Operative Conditions Scale for sedation during pediatric esophagogastroduodenoscopy. *Paediatr Anaesth*. 2017;27(6):621-628. doi:10.1111/pan.13127

70. Fournier-Charrière E, Tourniaire B, Carbajal R, et al. EVENDOL, a new behavioral pain scale for children ages 0 to 7 years in the emergency department: design and validation. *Pain*. 2012;153(8):1573-1582. doi:10.1016/j.pain.2012.02.024

71. Manworren RC, Hynan LS. Clinical validation of FLACC: preverbal patient pain scale. *Pediatr Nurs*. 2003;29(2):140-146.

72. Merkel SI, Voepel-Lewis T, Shayevitz JR, Malviya S. The FLACC: a behavioral scale for scoring postoperative pain in young children. *Pediatr Nurs*. 1997;23(3):293-297. doi:10.1038/nrm2120

73. Crellin DJ, Harrison D, Santamaria N, Babl FE. Systematic review of the Face, Legs, Activity, Cry and Consolability scale for assessing pain in infants and children: is it reliable, valid, and feasible for use? *Pain*. 2015;156(11):2132-2151. doi:10.1097/j.pain.0000000000000305

74. Ranger M, Celeste Johnston C, Rennick JE, Limperopoulos C, Heldt T, du Plessis AJ. A multidimensional approach to pain assessment in critically ill infants during a painful procedure. *Clin J Pain*. 2013;29(7):613-620. doi:10.1097/AJP.0b013e31826dfb13

75. Schade JG, Joyce BA, Gerkensmeyer J, Keck JF. Comparison of three preverbal scales for postoperative pain assessment in a diverse pediatric sample. *J Pain Symptom Manage*. 1996;12(6):348-359. doi:10.1016/S0885-3924(96)00182-0

76. Boelen-van der Loo WJ, Scheffer E, de Haan RJ, de Groot CJ. Clinimetric evaluation of the pain observation scale for young children in children aged between 1 and 4 years after ear, nose, and throat surgery. *J Dev Behav Pediatr*. 1999;20(4):222-227. doi:10.1097/00004703-199908000-00004

77. Lamas A, López-Herce J, Sancho L, et al. Assessing sedation in critically ill children by bispectral index, auditory-evoked potentials and clinical scales. *Intensive Care Med*. 2008;34(11):2092-2099. doi:10.1007/s00134-008-1198-1

78. Kerson AG, DeMaria R, Mauer E, et al. Validity of the Richmond Agitation-Sedation Scale (RASS) in critically ill children. *J Intensive Care*. 2016;4(1):65. doi:10.1186/s40560-016-0189-5

79. James F, Edwards R, James N, Dyer R, Goodwin V. The Royal College of Emergency Medicine Composite Pain Scale for children: level of inter-rater agreement. *Emerg Med J*. 2017;34(6):360-363. doi:10.1136/emermed-2015-205517

80. Curley MA, Harris SK, Fraser KA, Johnson RA, Arnold JH. State Behavioral Scale: a sedation assessment instrument for infants and young children supported on mechanical ventilation. *Pediatr Crit Care Med*. 2006;7(2):107-114. doi:10.1097/01.PCC.0000200955.40962.38

81. Thiadens T, Vervat E, Albertyn R, Van Dijk M, Van As ABS. Evaluation of pain incidence and pain management in a South African paediatric trauma unit. *S Afr Med J*. 2011;101(8):533-536.

82. Malviya S, Voepel-Lewis T, Tait AR, Merkel S, Tremper K, Naughton N. Depth of sedation in children undergoing computed tomography: validity and reliability of the University of Michigan Sedation Scale (UMSS). *Br J Anaesth*. 2002;88(2):241-245. doi:10.1093/bja/88.2.241

83. Malviya S, Voepel-Lewis T, Tait AR. A comparison of observational and objective measures to differentiate depth of sedation in children from birth to 18 years of age. *Anesth Analg*. 2006;102(2):389-394. doi:10.1213/01.ANE.0000184045.01780.73

84. Soetenga D, Frank J, Pellino TA. Assessment of the validity and reliability of the University of Wisconsin Children's Hospital Pain Scale for Preverbal and Nonverbal Children. *Pediatr Nurs*. 1999;25(6):670-676.

85. Macnab AJ, Levine M, Glick N, Susak L, Baker-Brown G. A research tool for measurement of recovery from sedation: the Vancouver Sedative Recovery Scale. *J Pediatr Surg*. 1991;26(11):1263-1267. doi:10.1016/0022-3468(91)90594-J

86. Hesselgard K, Larsson S, Romner B, Strömblad LG, Reinstrup P. Validity and reliability of the Behavioural Observational Pain Scale for postoperative pain measurement in children 1-7 years of age. *Pediatr Crit Care Med*. 2007;8(2):102-108. doi:10.1097/01.PCC.0000257098.32268.AA

87. McGrath PJ, Johnson G, Goodman JT, Dunn J, Chapman J. The CHEOPS: a behavioral scale to measure post-operative pain children. *Adv Pain Res Ther*. 1985;9:395-402.

88. Quinlan-Colwell A, Thear G, Miller-Baldwin E, Smith A. Use of the Pasero Opioid-induced Sedation Scale (POSS) in pediatric patients. *J Pediatr Nurs*. 2017;33:83-87. doi:10.1016/j.pedn.2017.01.006

89. Schultz AA, Murphy E, Morton J, Stempel A, Messenger-Rioux C, Bennett K. Preverbal, Early Verbal Pediatric Pain Scale (PEPPS): development and early psychometric testing. *J Pediatr Nurs*. 1999;14(1):19-27. doi:10.1016/S0882-5963(99)80056-6

90. Strout TD, Baumann MR. Reliability and validity of the modified Preverbal, Early Verbal Pediatric Pain Scale in emergency department pediatric patients. *Int Emerg Nurs*. 2011;19(4):178-185. doi:10.1016/j.ienj.2011.01.003

91. Tarbell SE, Cohen IT, Marsh JL. The Toddler-Preschooler Postoperative Pain Scale: an observational scale for measuring postoperative pain in children aged 1-5: preliminary report. *Pain*. 1992;50(3):273-280. doi:10.1016/0304-3959(92)90031-6

92. Slavec A, Drnovsek M. A perspective on scale development in entrepreneurship research. *Econ Bus Rev Cent South-Eastern Eur.* 2012;14(1):39.

93. Netemeyer RG, Bearden WO, Sharma S. *Scaling Procedures: Issues and Applications*. Sage Research Methods; 2003. doi:10.4135/9781412985772

94. Mandrekar JN. Measures of interrater agreement. *J Thorac Oncol*. 2011;6(1):6-7. doi:10.1097/JTO.0b013e318200f983

95. Johnston CC, Sherrard A, Stevens B, Franck L, Stremler R, Jack A. Do cry features reflect pain intensity in preterm neonates? a preliminary study. *Biol Neonate*. 1999;76(2):120-124. doi:10.1159/000014150

96. Ballantyne M, Stevens B, McAllister M, Dionne K, Jack A. Validation of the premature infant pain profile in the clinical setting. *Clin J Pain*. 1999;15(4):297-303. doi:10.1097/00002508-199912000-00006

97. van Dijk M, Peters JWB, van Deventer P, Tibboel D. The COMFORT Behavior Scale: a tool for assessing pain and sedation in infants. *Am J Nurs*. 2005;105(1):33-36. doi:10.1097/00000446-200501000-00019

98. Rana D, Bellflower B, Sahni J, et al. Reduced narcotic and sedative utilization in a NICU after implementation of pain management guidelines. *J Perinatol*. 2017;37(9):1038-1042. doi:10.1038/jp.2017.88

99. Abou Elella R, Adalaty H, Koay YN, et al. The efficacy of the COMFORT score and pain management protocol in ventilated pediatric patients following cardiac surgery. *Int J Pediatr Adolesc Med*. 2015;2(3):123-127. doi:10.1016/j.ijpam.2015.11.001

100. Olsson E, Anderzén-Carlsson A, Atladóttir SM, et al. Cultural adaptation and harmonization of four Nordic translations of the revised Premature Infant Pain Profile (PIPP-R). *BMC Pediatr*. 2018;18(1):349. doi:10.1186/s12887-018-1322-5