The Validity and Clinical Utility of the COVERS Scale and Pain Assessment Tool for Assessing Pain in Neonates Admitted to an Intensive Care Unit

Anna T. O'Sullivan, MSc,† Simon Rowley, MBChB,† Sharon Ellis, MN,† Kate Faasse, PhD,* and Keith J. Petrie, PhD**

Objectives: Infants admitted to a neonatal intensive care unit (NICU) are routinely subject to a range of painful procedures. However, pain assessments in NICUs are under-utilized due to a lack of a gold standard pain measure. In this study we assessed the psychometric properties and clinical utility of the COVERS and Pain Assessment Tool (PAT), in a neonatal unit.

Methods: We had 72 nurses use the scales to assess pain at baseline and during a heel-lance procedure in 80 NICU infants. An independent research observer and the infant's mother also completed pain ratings. After the study, we assessed nurse preference and clinical utility ratings for both scales.

Results: The COVERS had satisfactory internal consistency at baseline (Cronbach $\alpha = 0.74$) and heel lance ($\alpha = 0.78$), as did the PAT (baseline $\alpha = 0.79$, heel lance $\alpha = 0.85$). Intraclass correlation coefficients demonstrated good inter-rater reliability at baseline and heel lance, respectively, for both the COVERS (0.82 and 0.80) and the PAT (0.83 and 0.86). There were strong associations between total scores on the 2 scales at baseline ($r = 0.81$, $P < 0.001$) and heel lance ($r = 0.91$, $P < 0.001$), between researcher's ratings and total COVERS ($\rho = 0.75$, $P < 0.001$) and PAT scores ($\rho = 0.69$, $P < 0.001$), and between maternal ratings and total COVERS ($r = 0.74$, $P < 0.05$) and PAT scores ($r = 0.65$, $P < 0.05$). Both scales were sensitive to pain and nonpain events. Reliability and validity was mostly upheld across gestational age. Most nurses preferred the COVERS (52%) to the PAT (16%), and 32% had no preference.

Discussion: This study builds on evidence for the COVERS scale and the PAT; both scales were reliable and valid measures of acute pain in neonates as premature as 24-week gestational age.

Key Words: COVERS, Pain Assessment Tool, neonatal pain assessment, neonatal intensive care unit, psychometrics, reliability, validity

(*Clin J Pain* 2016;32:51–57)

The neonatal intensive care unit (NICU) is a clinical setting in which pain is an inevitable consequence of clinical care. Infants admitted to an NICU are subjected to a range of painful procedures between 10 and 12 times in the average day.[1–3] Given the frequency of potentially painful procedures, the NICU represents a clinical setting where regular pain assessment should be a requirement of practice. Currently, pain assessments in NICUs are under-utilized, with estimates indicating that less than a quarter of NICUs or clinicians working in these setting regularly use validated pain assessment measures.[4–9]

It was once thought that newborn infants lacked the capacity to experience pain due to underdeveloped nervous systems. However, research findings in recent decades indicate that even premature neonates are able to experience and interpret pain and display pain behaviors.[10–12] Pain is now considered by some to be the fifth vital sign, and many clinicians agree that pain should be both anticipated and assessed in all clinical settings caring for neonates.[13]

Pain assessment in neonates involves the observation of behavioral and physiological indicators as proxies for pain,[14,15] and is particularly important given that the current gold standard measure of pain—self-report—is unachievable.[16] Scales are constructed using multiple indicators and can be categorized as unidimensional, in which scales assess one indicator of pain, or multidimensional, in which scales assess multiple indicators of pain.[17] The most well-validated neonatal pain assessment scales appear to be those that assess both behavioral and physiological indicators of pain.[16,18] Multidisciplinary guidelines recommend the use of multidimensional scales as these should result in more accurate appraisals of pain experiences.[19,20]

In recent decades there has been a proliferation of pain measures specific to neonates, and there are now over 40 measures available. Despite this, a gold standard or universally preferred scale that differentiates between distress and pain is yet to emerge from published research.[21–23] As well as discriminating between pain and distress, a gold standard scale would demonstrate acceptable degrees of internal consistency, inter-rater reliability, concurrent validity, construct validity, as well as clinical feasibility and utility.[1,24–26] Experts in neonatal pain have outlined deficiencies in pain assessment research, and have recommended a focus on the refinement and further validation of current scales among neonates of varying gestational ages and conditions.[20] The COVERS scale[27] and the Pain Assessment Tool (PAT)[28,29] are 2 commonly used neonatal pain measures in Australia. These scales were chosen for their clinical applicability, which was highlighted by clinicians working in the NICU, and also as there is a lack of literature investigating their psychometric properties. This study assessed the psychometric properties and clinical utility of both scales in a neonatal unit.

Received for publication August 27, 2014; revised March 10, 2015; accepted February 8, 2015.
From the *Department of Psychological Medicine, The University of Auckland; and †Newborn Services, Auckland City Hospital, Auckland, New Zealand.
The authors declare no conflict of interest.
Reprints: Keith J. Petrie, PhD, Department of Psychological Medicine, Faculty of Medical and Health Sciences, University of Auckland, Private Bag 92019, Auckland, New Zealand (e-mail: kj.petrie@auckland.ac.nz).

DOI: 10.1097/AJP.0000000000000228

MATERIALS AND METHODS

Purpose

To assess the reliability, validity, and clinical utility of 2 measures of neonatal pain—the COVERS scale[27] and a modified version of the PAT.[28]

Participants

The study was conducted at The National Women's Newborn Intensive Care Unit (NICU) at Auckland City Hospital from June to August 2013. The sample consisted of 80 neonates and their mothers. Seventy-two NICU nurses assessed pain in neonates using each of the scales. Infants admitted to the NICU were eligible to participate if they required heel lancing for routine blood collection. Infants of all gestational ages, as well as those requiring surgery, were included. Infants receiving oral sucrose for pain relief were also included. With regard to the use of oral sucrose, NICU guidelines allow an oral dose of 0.2 mL of 66.7% sucrose solution to be used as pain relief. Infants must be >1500 g and >31 weeks' postconceptional age to be given sucrose. Up to 4 doses of sucrose can be given in any 24-hour period. Infants were excluded from the study if they were in a critical condition or were expected to die, had severe neurological impairment, were receiving muscle relaxants for ventilation, were pharmacologically paralyzed, or if their sedation was significant enough to reduce their pain responses. The infants' mothers were not asked to participate in the study if they were not fluent in the English language.

Procedure

Nurses working in the Auckland City Hospital NICU were recruited to use the 2 scales to carry out the pain assessments. Before data collection nurses were given verbal training regarding how to score infants using each measure. Nurses' time constraints precluded further education and practice in completing questionnaires. Subsequently, mothers of eligible infants were invited to participate and informed consent was obtained. Ethical approval for the study was granted by The University of Auckland Human Participants Ethics Committee and the Auckland District Health Board Research Review Committee.

Data were collected at 2 points: a nonpain event, or "baseline," and a pain event, a heel-lance procedure. Each infant was assessed only once. The same 2 nurses rated pain during both events. One rater was the nurse in charge of the infant's care while the second rater was drawn from the pool of recruited nurses. Data for the nonpain event were collected in the infant's resting state by nurses who independently and simultaneously scored the neonate's level of pain on both scales. Data for the pain event were then collected as the infant underwent a heel-lance procedure. To control for order effects, scale completion order was counterbalanced. As only 2 nurses were involved in each assessment, 1 nurse performed the heel lance and at the same time made a pain assessment. The researcher, and if present the infant's mother, observed the infant during the pain event and rated the infant's pain. Before the heel lance, mothers were given instruction in how to complete a 0 to 10 pain rating using the 11-point Box-Scale.[30]

Following the completion of data collection from the 80 infants, the researcher approached the 62 nurses that had completed at least one pain assessment and asked them to complete a follow-up questionnaire designed to assess the clinical utility of the scales.

Measures

Clinical data were collected from clinical notes immediately before the heel lance: sex, gestational age at birth, birth weight, reason for admission, current weight, current age, current level of care, any previous surgery, the presence of any breathing support, the use of any muscle relaxants, analgesics or sedatives, and the use of any nursing comfort measure before the heel lance including the administration of oral sucrose, swaddling, kangaroo care, breastfeeding, or the use of a dummy.

The COVERS scale[27] and a modified version of the PAT[28] were used to assess pain (Appendix 1.) The original PAT has 10 undefined response options—1 for each of the 10 items. Therefore, minor additions were made to these items on the scale to help staff to complete the measure and to improve its consistency. Descriptors were included to describe scores of *zero* for "Posture/Tone," "Expression," "Respirations," and "Heart rate" and to describe scores of *one* for "Cry," "Sleep pattern," "Color," "Oxygen saturation," "Blood pressure," and "Nurse perception." The authors of the PAT recommend that scores >5 should result in nursing comfort measures such as soothing, use of a pacifier, and repositioning, and scores >10 should result in analgesia.

The researcher and the mothers (when present) completed an 11-point Box Scale[30] during the heel lance. For the purposes of this study, the scale's anchors were worded so that it was suitable for the mothers and the researcher—"your level of pain" was replaced with "this infant's level of pain."

A questionnaire to evaluate the clinical utility and feasibility of the 2 pain scales was also designed for use in this study. The questionnaire's 20 items (10 relating to each scale) assessed: difficult items, inappropriate items, the clarity of each scale's layout, the clarity of written descriptors, the scale's ease-of-use, the perceived accuracy of the scale, the perceived helpfulness of the scale, how happy nurses would be to use the scale, and an overall rating of the scale. Respondents were also asked to specify their preferred scale. There were 7 continuous items, with a maximum total score of 70, and 3 items with discrete response categories. The order in which the block of questions regarding each scale appeared was counterbalanced to control for order effects.

Data Analyses

Previous studies testing the concurrent validity of neonatal pain scales have evidenced large effect sizes.[27,31] To detect a large effect size ($r = 0.50$) with power of 0.80 and an α-level of 0.05, 28 participants were required. To yield more accurate results, data were collected from 80 infants.[32] Data were analyzed using IBM SPSS Statistics 20. An α-level of 0.05 was used for all tests. Data were assessed for normality, and means and frequencies of clinical data were calculated.

The internal consistency of each scale at each time point was established using mean interitem correlations, corrected item-total correlations, and the Cronbach α reliability coefficients. Inter-rater reliability was established using Kappa Measure of Agreement[32] for categorical data, and intraclass correlations (ICCs) for continuous data.

Copyright © 2015 Wolters Kluwer Health, Inc. All rights reserved.

The Spearman rank order correlations assessed concurrent validity by investigating associations between mean total scores on either scale at baseline, and between mean total scores and mothers' scores. Pearson correlations assessed associations between mean total scores on either scale at heel lance, and between mean total scores and the researcher's score. Construct validity was established by comparing mean total scores during 2 nonpain versus pain events: baseline versus heel lance, and sucrose versus no sucrose. The sensitivity of the scales to heel lance was established using Wilcoxon signed ranks tests, whereas 1-way analyses of covariance (ANCOVAs) established the sensitivity of the scales to the administration of sucrose after controlling for baseline pain scores. Wilcoxon signed ranks tests and McNemar tests were used to detect any statistically significant differences in nurses' attitudes toward each scale.

RESULTS

Infant Characteristics

The majority of the infants were male (60%) and born premature (28 to 37 wk gestation) (61.3%). The remainder were extremely premature ($\leq$28 wk) (26.3%) or term ($\geq$ 37 wk) (12.5%). Gestational age at birth ranged from 23.6 to 41.1 weeks (M = 31.6, SD = 4.7), and birth weights ranged from 580 to 4620 g (M = 1781, SD = 982). Six infants (7.5%) underwent surgery before pain assessment. At the time of assessment, average weight was 1934 g and average age was 14 days. Forty infants (50%) were receiving the highest level of intensive care, 34 (42.5%) were in high-dependency care, and 6 (7.5%) were in low-dependency care. Thirty-four (42.5%) infants were receiving continuous positive airway pressure and 5 (6.3%) were mechanically ventilated. Twenty-five infants (31.3%) received a comfort measure such as swaddling, kangaroo care, breastfeeding, or use of dummy before or during assessment; of these, 60% received oral sucrose.

Reliability

Internal Consistency

The mean interitem correlations on both scales were within the optimal range (M = 0.20 to 0.40).[33] Corrected item-total correlations for COVERS scale items (range = 0.19 to 0.68) and modified PAT items (range = 0.35 to 0.71) were largely acceptable ($r > 0.30$)[34]; the only item that did not perform well was the "Oxygen requirement" item on the COVERS scale. The Cronbach α for the COVERS scale at baseline and heel lance was 0.74 and 0.78, respectively. The Cronbach α for the modified PAT at baseline and heel lance was 0.79 and 0.85, respectively.

Inter-rater Reliability

Kappa Measures of Agreement yielded "fair" (0.21 to 0.40) to "substantial" (0.61 to 0.80) agreement between raters for items on the COVERS scale at baseline (range = 0.29 to 0.78, M = 0.44) and heel lance (range = 0.22 to 0.67, M = 0.39).[35] Items with Fair Kappa values at baseline were Vital Signs, Expression, and Signaling distress, and at heel lance were Vital signs, Expression, Resting, and Signaling distress. With regard to the modified PAT, Kappa statistics ranged from poor (< 0.00) to substantial at baseline (range = −0.07 to 0.60, M = 0.34) and fair to substantial at heel lance (range = 0.25 to 0.69, M = 0.41). At baseline, agreement between nurses was poor on the Nurse perception and Color items, and fair on Oxygen saturation and Posture. At heel lance, the poorest agreement was for the Heart rate item which fell in the fair range. The ICC coefficients for scores on the COVERS scale at baseline and heel lance were 0.82 [95% CI, 0.72-0.88] and 0.80 [95% CI, 0.69-0.87], respectively. The ICC coefficients for scores on the modified PAT at baseline and heel lance were 0.83 [95% CI, 0.73-0.89] and 0.86 [95% CI, 0.78-0.91], respectively. All coefficients indicated high inter-rater agreement ($\rho_I > 0.75$)[36] and were statistically significant ($P < 0.001$).

Validity

The Spearman ρ correlations demonstrated strong associations between the mean total scores on the 2 scales at baseline ($\rho = 0.81$, $P < 0.001$), as well as between the researcher's ratings and mean total scores on both the modified PAT ($\rho = 0.69$, $P < 0.001$) and the COVERS scale ($\rho = 0.75$, $P < 0.001$) at heel lance.[37] The Pearson r correlations also found strong associations between the mean total scores on the 2 scales at heel lance ($r = 0.91$, $P < 0.001$), as well as between maternal ratings and mean total scores on both the modified PAT ($r = 0.65$, $P < 0.05$) and the COVERS scale ($r = 0.74$, $P < 0.05$) at heel lance.[37] These findings support the concurrent validity of both scales.

There were statistically significant increases in pain scores on both the COVERS scale ($Md_{baseline} = 0$; $Md_{heel\ lance} = 3.5$), $z = -7.22$, $P < 0.001$, and the modified PAT scale ($Md_{baseline} = 0.25$; $Md_{heel\ lance} = 5.5$), $z = -7.19$, $P < 0.001$, from baseline to heel lance. Effect sizes for both tests were large ($r = 0.58$). ANCOVAs revealed significant effects of sucrose on pain scores at heel lance on both scales after controlling for the effect of baseline pain score. Infants given sucrose scored lower on the COVERS scale (M = 2.30, SE = 0.55, 95% CI [1.20, 3.39]) than those not given sucrose (M = 3.82, SE = 0.25, 95% CI [3.31, 4.32]), $F_{1,77} = 6.11$, $P = 0.016$, $\eta_p^2 = 0.07$. Infants given sucrose also scored lower on the modified PAT (M = 3.55, SE = 0.80, 95% CI [1.95, 5.14]) than those not given sucrose (M = 5.85, SE = 0.37, 95% CI [5.12, 6.58]), $F_{1,77} = 6.57$, $P = 0.012$, $\eta_p^2 = 0.08$.

Gestational Age

A Kruskal-Wallis test revealed no statistically significant differences in mean total scores at heel lance across the 3 different gestational age groups on either the PAT ($P = 0.15$) or the COVERS ($P = 0.46$). The COVERS scale demonstrated satisfactory internal consistency among extremely preterm neonates, $\alpha = 0.74$, and good internal consistency among preterm neonates, $\alpha = 0.80$, and term neonates, $\alpha = 0.81$.[34] The modified PAT demonstrated good internal consistency among all infants (extremely preterm, $\alpha = 0.87$; preterm, $\alpha = 0.83$; term neonates, $\alpha = 0.88$). Inter-rater reliability for both scales was reasonably high for all neonates aside from term neonates assessed at baseline (Table 1).

Concurrent validity analyses for the 2 scales among the 3 gestational age groups are presented in Table 2. Concurrent validity for term neonates at baseline could not be established for either scale, and correlations between pain scores and researcher ratings for extremely preterm babies at heel lance were moderate.

Copyright © 2015 Wolters Kluwer Health, Inc. All rights reserved.

TABLE 1. Intraclass Correlation Coefficients [and 95% Confidence Intervals] Between Total Scores on the COVERS Scale and Modified PAT According to Gestational Age

	COVERS Scale	Modified PAT
Extremely preterm (n = 21)		
Baseline	0.80** [0.50, 0.92]	0.75* [0.38, 0.90]
Heel lance	0.87** [0.69, 0.95]	0.83** [0.59, 0.93]
Preterm (n = 49)		
Baseline	0.85** [0.74, 0.92]	0.86** [0.75, 0.92]
Heel lance	0.74** [0.53, 0.85]	0.87** [0.77, 0.93]
Term (n = 10)		
Baseline	0.08 [−30.52, 0.78]	0.33 [−0.52, 0.79]
Heel lance	0.89** [0.44, 0.98]	0.88* [0.52, 0.97]

Coefficients represent the reliability between 2 independent raters, on average.
PAT indicates Pain Assessment Tool.
**$P < 0.001$.
*$P < 0.01$.

Clinical Utility and Feasibility

Mean total score for the modified PAT on the measure of clinical utility and feasibility was $M = 49.33$ (SD = 12.16) and for the COVERS Scale $M = 48.95$ (SD = 11.68). There were no significant differences in either scale's format, clarity, ease-of-use, accuracy, helpfulness, overall rating, number of difficult items, or the degree to which the nurse would be happy to use the scale, Ps = 0.28 to 0.72. However, there was a significant difference in the proportion of nurses rating that each scale is appropriate for all infants admitted to the NICU (65.2% for the modified PAT vs. 34.8% for the COVERS scale, $P = 0.016$). The Signaling distress item of the COVERS scale was often cited as inappropriate by nurses as flailing and arching that necessitates the maximum pain score are uncommon in preterm infants.[38,39] The corresponding Posture/Tone item on the modified PAT includes subtler pain responses, for example, clenching of fists, which preterm infants are more capable of displaying. The majority of nurses (52%) preferred the modified PAT; 16% preferred the COVERS Scale; 32% had no preference.

TABLE 2. Correlations Between Mean Total Scores on the Modified PAT, Mean Total Scores on the COVERS Scale, and the Researcher's Rating of Pain, With Data Split by Gestational Age

Extremely preterm (n = 21)	
COVERS scale, modified PAT	
Baseline†	0.79***
Heel lance‡	0.91***
Researcher rating, modified PAT†	0.33
Researcher rating, COVERS scale†	0.47*
Preterm (n = 49)	
COVERS scale, modified PAT	
Baseline†	0.85***
Heel lance‡	0.91***
Researcher rating, modified PAT†	0.77***
Researcher rating, COVERS scale†	0.82***
Term (n = 10)	
COVERS scale, modified PAT	
Baseline†	0.38
Heel lance‡	0.94***
Researcher rating, modified PAT†	0.85**
Researcher rating, COVERS scale†	0.88***

†Spearman rank order correlation.
‡Pearson r correlation.
PAT indicates Pain Assessment Tool.
***$P < 0.001$.
**$P < 0.01$.
*$P < 0.05$.

DISCUSSION

This study investigated the psychometric properties of 2 neonatal pain measures in an intensive care setting, where improved pain assessment practices are urgently needed. This study found that, in general, both the COVERS scale and the modified PAT are reliable and valid measures of pain in neonates. Furthermore, reliability and validity was mostly upheld across the range of gestational ages.

The Cronbach α values obtained in this study are comparable with those of established multidimensional pain scales.[39–41] However, Kappa Measures of Agreement[42] in this study revealed poorer agreement between raters on each scale item than has been found in studies of comparable pain scales. Although agreement between raters in this study was fair to moderate, moderate to substantial agreement has been found between other scales.[43] This study, relative to other studies, recruited a large number of nurse participants with varying levels of experience. As a result, methodological variance may play a part in the findings of low inter-rater agreement for several scale items. Nevertheless, it is our belief that using a larger group of pain assessors is more naturalistic and ecologically valid, given that all staff nurses are required to assess pain in clinical practice. However, more intensive training on the use of the scales is likely to lead to better inter-rater reliability, particularly on "Vital signs," "Expression," "Resting," and "Signaling distress" for the COVERS scale and the "Heart rate" item on the PAT.

ICCs measuring the agreement between raters' total scores are consistent with a previous investigation into the original PAT,[28] as well as with studies of other established pain scales.[44] Lastly, findings which revealed that nurses' ratings were highly correlated irrespective of the degree of pain contradicts the results of a previous investigation which found that as the total PAT score increased the agreement between raters decreased.[45]

Our findings of strong correlations between scale scores at both time points are consistent with previous investigations.[27,45] Furthermore, our findings support the construct validity of both the modified PAT and the COVERS scales and indicate that the scales have construct validity comparable with that of more widely validated neonatal pain measures.[39,44,46–49]

To date, the reliability and validity of the PAT and the COVERS scale has not been investigated in infants younger than 27 weeks' gestational age. The present research extends previous research by finding that, when looking at the entire sample, both scales are reliable and valid among infants born at 24 weeks' gestational age and above. Looking closer there were areas where the scales performed less well, however it is likely that certain findings are a consequence of the small sample of infants, and in particular term infants.

Little research has formally tested the clinical utility of neonatal pain scales, despite these properties being necessary if the measure is to be used successfully in the clinical

Copyright © 2015 Wolters Kluwer Health, Inc. All rights reserved.

setting.[50] Although several established pain scales have been criticized for being cumbersome, in the present study both scales were administered in less than 3 minutes. In this study, both scales received positive evaluations, however, nurses perceived the modified PAT to be more suitable for the range of infants admitted to the NICU. Further work needs to be carried out to establish whether the "Signaling distress" item in the COVERS scale is useful to include in preterm infants.

The PAT and the COVERS have been said to be pertinent to the range of infants requiring intensive care, for example the COVERS scale is unique in that it assesses "Visible crying." Intubated or very preterm infants are often unable to cry audibly, thus the addition of this item improves the relevance of pain assessment to these infants.[27] Distinct from other scales, the modified PAT includes nurses' perceptions of pain, which are important given that these normally determine whether pain relief will be administered.

Several caveats of the research need acknowledging. First, it is important to note that given the similarities between the 2 scales, correlations between pain ratings are highly likely. Second, this study was limited by the design which necessitated 2 nurses assessing pain, meaning that 1 nurse had to perform the heel lance and complete the scales retrospectively. The design of the study also meant that nurse participants were not blind to the administration of nursing comfort measures such as oral sucrose. Both issues might have been resolved by videotaping all infants and having 2 nurses assess pain by viewing the recording. It should be noted that in clinical practice the same nurse would be responsible for performing the heel lance, assessing pain, and administering pain relief, therefore the design of this study increases its ecological validity.

Unlike previous research, this study is strengthened by the inclusion of a broad range of infants as young as 24 weeks' gestational age. In addition, while many previous studies are limited in that they analyze the combined data from infants of varying gestational ages, this study acknowledged the effect of gestational age on pain responses and analyzed data across 3 gestational age groups.

Although this study assessed a broad range of psychometric properties, further research is needed to establish the scales' intrarater reliability, test-retest reliability, predictive validity, and further evidence of construct validity. There is also a need to test whether the scales can discriminate between pain intensities, and between pain and similar constructs such as distress. One of the methods of addressing the issue of whether assessment measures are truly measuring pain is to conduct sensitivity and specificity analyses. Sensitive scale items correctly identify infants experiencing pain and specific items correctly identify infants not experiencing pain.[24] For example, for an item of the PIPP to be considered sensitive it had to be present over 50% of the time during a "painful" procedure, while a specific item could not be present more than 20% of the time during a "nonpainful" procedure.[40] Similar sensitivity and specificity analyses for items on the PAT and the COVERS should be conducted in the future so as to determine whether the scales truly reflect infant pain.

Future research might also consider assessing the scales' construct validity further in situations where pain is expected to vary, such as heel lance versus lumbar puncture. In addition, comparing responses of infants in nonpain situations such as diaper-change versus ultrasound could be investigated, along with "distress" situations such as hunger, tiredness, and agitation. If the scales are assessing pain, a significant difference between heel lance and a more invasive procedure would be anticipated, meanwhile significant differences between variations of distressing situations would suggest that the scales are not truly assessing pain.

There is a drive in the literature to conduct further research so that a gold standard pain assessment scale may eventually arise. To this end, future research should avoid the development of new scales, and should instead focus on the reliability, utility, and most importantly the construct validity of existing scales.[20] Future research should, where possible, also involve analyses for distinct gestational age groups.

Copyright © 2015 Wolters Kluwer Health, Inc. All rights reserved.

APPENDIX 1

TABLE A1. The Modified Pain Assessment Tool and Covers Scale

Parameters	0	1	2
The modified Pain Assessment Tool (PAT)			
Posture/tone	Relaxed Normal Some flexion	Extended Digits widespread Trunk rigid Limbs abducted Shoulders raised off bed	Flexed and/or tense Fists clenched Trunk guarded Limbs drawn to midline Head/shoulders resist posturing
Cry	No	Yes Consolable Can be settled	Yes When disturbed Does not settle after handling Loud Whimpering Whining
Sleep pattern	Relaxed	Easily woken	Agitated or withdrawn Wakes with startle Restless Squirming No clear sleep/wake pattern Eye aversion or "shut out"
Expression	Relaxed Normal	Frown Shallow furrows Eyes lightly closed	Grimace Deep furrows Eyes tightly closed Pupils dilated
Color	Pink, well perfused	Occasionally mottled or pale	Pale/dusky/flushed Palmar sweating
Respirations	Normal baseline rate	Tachypnea At rest	Apnea At rest/with handling
Heart rate	Normal baseline rate	Tachycardia At rest	Fluctuating Spontaneous/at rest
Oxygen saturation	Normal	Fleeting desaturation	Desaturation with/without handling
Blood pressure	Normal	Fluctuates with handling	Hypo-/hypertension at rest
Nurse perception	No pain perceived by me	I think the baby has pain only with handling	I think the baby is in pain
The COVERS scale			
Crying	No	High pitched or visibly crying	Inconsolable or difficult to soothe
Oxygen requirement	None At baseline O_2 Breathing comfortably	Less than 30% Increase less than 20% Change in breathing pattern	More than 30% Increase more than 20% Significant change in breathing pattern
Vital signs	HR and/or BP WNL for age or at baseline No apnea or bradycardia or at baseline	HR and/or BP increase less than 20% of baseline Increase in frequency of apnea and bradycardia	HR and/or BP increase more than 20% of baseline Increase in frequency and severity of apnea and bradycardia
Expression	None Facial muscles relaxed	Grimace Minimal-moderate brow bulge Eye squeeze Nasolabial furrow	Grimace/grunt Moderate-maximum brow bulge Eye squeeze Nasolabial furrow
Resting	Sleeping most of the time	Wakes at frequent intervals—fussy	Constantly awake (even when not disturbed)
Signaling distress	Relaxed	Arms/legs flexed or extended "Time-out signalse	Flailing Arching

Copyright © 2015 Wolters Kluwer Health, Inc. All rights reserved.

REFERENCES

1. Anand K. Consensus statement for the prevention and management of pain in the newborn. *Arch Pediatr Adolesc Med*. 2001;155:173–180.
2. Roofthooft DW, Simons SH, Anand KJ, et al. Eight years later, are we still hurting newborn infants? *Neonatology*. 2014;105:218–226.
3. Carbajal R, Rousset A, Danan C, et al. Epidemiology and treatment of painful procedures in neonates in intensive care units. *JAMA*. 2008;300:60–70.
4. Akuma AO, Jordan S. Pain management in neonates: a survey of nurses and doctors. *J Adv Nurs*. 2012;68:1288–1301.
5. Foster J, Spence K, Henderson-Smart D, et al. Procedural pain in neonates in Australian hospitals: a survey update of practices. *J Paediatr Child Health*. 2013;49:E35–E39.
6. Gharavi B, Schott C, Nelle M, et al. Pain management and the effect of guidelines in neonatal units in Austria, Germany and Switzerland. *Pediatr Int*. 2007;49:652–658.
7. Harrison D, Loughnan P, Johnston L. Pain assessment and procedural pain management practices in neonatal units in Australia. *J Paediatr Child Health*. 2006;42:6–9.
8. Lago P, Guadagni A, Merazzi D, et al. Pain management in the neonatal intensive care unit: a national survey in Italy. *Pediatr Anesth*. 2005;15:925–931.
9. Rohrmeister K, Kretzer V, Berger A, et al. Pain and stress management in the neonatal intensive care unit—a national survey in Austria. *Wien Klin Wochenschr*. 2003;115:715–719.
10. Anand KJS, Hickey PR. Pain and its effects in the human neonate and fetus. *New Engl J Med*. 1987;317:1321–1329.
11. Charnay Y, Paulin C, Chayvialle J, et al. Distribution of substance P-like immunoreactivity in the spinal cord and dorsal root ganglia of the human foetus and infant. *Neuroscience*. 1983;10:41–55.
12. Hall RW, Anand KJS. Physiology of pain and stress in the newborn. *Neoreviews*. 2005;6:61–68.
13. Ranger M, Johnston CC, Anand KJS. Current controversies regarding pain assessment in neonates. *Semin Perinatol*. 2007;31:283–288.
14. Fitzgerald M, Walker SM. Infant pain management: a developmental neurobiological approach. *Nat Clin Pract Neurol*. 2009;5:35–50.
15. Stevens BJ, Pillai-Riddell RR, Oberlander TE, et al. Assessment of pain in neonates and infants. In: Anand KJS, Stevens BJ, McGrath PJ, eds. *Pain in Neonates and Infants*. London: Elsevier; 2007:67–90.
16. Stevens B. Acute pain management in infants in the neonatal intensive care unit. In: Finley GA, McGrath PJ, eds. *Acute and Procedure Pain in Infants and Children*. Seattle: IASP Press; 2001:101–128.
17. Johnston CC, Stevens BJ, Boyer K, et al. Development of psychologic responses to pain and assessment of pain in infants and toddlers. In: Schechter NL, Berde CB, Yaster M, eds. *Pain in Infants, Children, and Adolescents*. Philadelphia: Lippincott Williams & Wilkins; 2003:105–127.
18. Burton J, MacKinnon R. Selection of a tool to assess postoperative pain on a neonatal surgical unit. *Infant*. 2007;3:188–196.
19. American Academy of Pediatrics, Canadian Paediatric Society. Prevention and management of pain in the neonate: an update. *Pediatrics*. 2006;118:2231–2241.
20. Anand KJ, Aranda JV, Berde CB, et al. Summary proceedings from the neonatal pain-control group. *Pediatrics*. 2006; 117:S9–S22.
21. Anand KJS. Pain assessment in preterm neonates. *Pediatrics*. 2007;119:605–607.
22. Stevens BJ, Anand KJS, McGrath PJ. An overview of pain in neonates and infants.. In: Anand KJS, Stevens BJ, McGrath PJ, eds. *Pain in Neonates and Infants*. London: Elsevier; 2007:1–9.
23. Walker SM. Pain in children: recent advances and ongoing challenges. *Br J Anaesth*. 2008;101:101–110.
24. Abu-Saad HH, GJJW Bours, Stevens B, et al. Assessment of pain in the neonate. *Semin Perinatol*. 1998;22:402–416.
25. Franck LS, Greenberg CS, Stevens B. Pain assessment in infants and children. *Pediatr Clin North Am*. 2000;47:487–512.
26. Gibbins S, Stevens B, Asztalos E. Assessment and management of acute pain in high-risk neonates. *Expert Opin Pharmacother*. 2003;4:475–483.
27. Hand IL, Noble L, Geiss D, et al. COVERS neonatal pain scale: development and validation. *Int J Pediatr*. 2010.
28. Hodgkinson K, Bear M, Thorn J, et al. Measuring pain in neonates: evaluating an instrument and developing a common language. *Aust J Adv Nurs*. 1994;12:17–22.
29. Spence K, Henderson-Smart D. Closing the evidence-practice gap for newborn pain using clinical networks. *J Paediatr Child Health*. 2011;47:92–98.
30. Jensen MP, Karoly P, O'Riordan EF, et al. The subjective experience of acute pain an assessment of the utility of 10 indices. *Clin J Pain*. 1989;5:153–160.
31. Cohen J. A power primer. *Psychol Bull*. 1992;112:155.
32. Peat J, Mellis C, Williams K, et al. *Health Science Research: A Handbook of Quantitative Methods*. London: Sage; 2001.
33. Briggs SR, Cheek JM. The role of factor analysis in the development and evaluation of personality scales. *J Pers*. 1986;54:106–148.
34. Nunally JC, Bernstein IH. *Psychometric Theory*. New York: McGraw-Hill; 1994.
35. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–174.
36. Burdock EI, Fleiss JL, Hardesty AS. A new view of inter-observer agreement. *Person Psychol*. 1963;16:373–384.
37. Cohen J. *Statistical Power Analysis for the Behavioural Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
38. Ambuel B, Hamlett KW, Marx CM, et al. Assessing distress in pediatric intensive care environments: the COMFORT scale. *J Pediatr Psychol*. 1992;17:95–109.
39. Hummel P, Puchalski M, Creech S, et al. Clinical reliability and validity of the N-PASS: Neonatal Pain, Agitation and Sedation Scale with prolonged pain. *J Perinatol*. 2008;28: 55–60.
40. Stevens B, Johnston C, Petryshen P, et al. Premature infant pain profile: development and initial validation. *Clin J Pain*. 1996;12:13–22.
41. van Dijk M, de Boer JB, Koot HM, et al. The reliability and validity of the COMFORT scale as a postoperative pain instrument in 0 to 3-year-old infants. *Pain*. 2000;84:367–377.
42. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20:37–46.
43. Peters JW, 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: 353–363.
44. Suraseranivongse S, Kaosaard R, Intakong P, et al. A comparison of postoperative pain scales in neonates. *Br J Anaesth*. 2006;97:540–544.
45. Spence K, Gillies D, Harrison D, et al. A reliable pain assessment tool for clinical assessment in the neonatal intensive care unit. *J Obstet Gynecol Neonatal Nurs*. 2005;33:80–86.
46. Ballantyne M, Stevens B, McAllister M, et al. Validation of the Premature Infant Pain Profile in the clinical setting. *Clin J Pain*. 1999;15:297–303.
47. Debillon T, Zupan V, Ravault N, et al. 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:F36–F41.
48. Grunau R, Craig K. Facial activity as a measure of neonatal pain expression. *Adv Pain Res Ther*. 1990;15:147–155.
49. Lawrence J, Alcock D, McGrath P, et al. The development of a tool to assess neonatal pain. *Neonatal Netw*. 1993;12:59–66.
50. Stevens B, Gibbins S. Clinical utility and clinical significance in the assessment and management of pain in vulnerable infants. *Clin Perinatol*. 2002;29:459–468.

Copyright © 2015 Wolters Kluwer Health, Inc. All rights reserved.