

Is It Painful or Not?

Discriminant Validity of the Behavioral Indicators of Infant Pain (BIIP) Scale

Liisa Holsti, PhD, † ‡ § Ruth E. Grunau, PhD,* † ‡ § Tim F. Oberlander, MD, FRCP,* † ‡ § and Horacio Osiovich, MD, FRCP † §*

Objectives: To evaluate the ability of the Behavioral Indicators of Infant Pain (BIIP) scale to discriminate between skin-breaking and nonskin breaking procedures, and to identify sensitized pain responses in preterm infants in the neonatal intensive care unit (NICU).

Methods: Sixty-nine infants born between 24 and 32 weeks gestational age were assessed at 32 weeks postconceptional age during blood collection on one day (procedure A), and then on another day during blood collection preceded by a diaper change (procedure B). Procedure order was randomized. Outcome measures were changes in BIIP coded from continuous bedside video recordings and changes in heart rate (HR).

Results: During blood collection (procedure A), BIIP scores ($P < 0.0001$) and mean HR ($P < 0.0001$) were higher than during the diaper change and higher when the infants had had a preceding diaper change (procedure B vs. procedure A) ($P < 0.03$). HR changed from baseline to the stressors for each procedure. No differences in mean HR were observed during Lance phase between the procedure A and the B blood collection; however, HR remained elevated significantly during the Recovery phase when blood collection was preceded by the diaper change ($P < 0.03$).

Discussion: The BIIP scale is reliable, accurate, and valid assessment for measuring acute pain in preterm infants in the NICU. This assessment combines the relatively most specific,

anatomically based, theoretically derived indicators; and it allows evaluation of behavioral and physiologic pain responses separately.

Key Words: premature infant, pain, assessment

(*Clin J Pain* 2008;24:83–88)

For the preterm infant, particularly those born extremely early, exposure to invasive and intrusive procedures is necessary to ensure their survival. Important recent research shows that infants born as early as 25 weeks gestational age (GA) show activation of the somatosensory cortex in response to acutely painful procedures,¹ and infants born 28 weeks GA and above exhibit specific cortical hemodynamic changes that differentiate tactile from skin-breaking stimuli.² However, for clinicians working in the neonatal intensive care unit (NICU), deciding whether an infant is in pain or not is complex, not only because they must rely on nonverbal behavioral and physiologic cues that are relatively more or less specific to pain, but also because over time, depending on which indicators are chosen to assess the infants, their responses may be diminished or heightened (“sensitized”).^{3,4}

Critical to instituting appropriate pain management strategies is the use of assessments, which are reliable, valid, feasible, and clinically useful. One of the important steps in establishing the validity of a scale is showing that it can differentiate between procedures that are painful and those that are less intrusive particularly because pharmacologic interventions may act differently if pain is or is not present.⁵ In addition, in infants born at extremely early GAs, recommendations for the limited use of some nonpharmacologic interventions (ie, sucrose) increase the need to use assessments that discriminate accurately between painful and less invasive procedures.⁶

Although a number of scales have been developed for the use of measuring acute pain in preterm infants, few have had adequate psychometric testing.⁷ Developed recently, the Behavioral Indicators of Infant Pain (BIIP) is a reliable scale for use in preterm infants in the NICU.⁸ This scale is unique in that it combines relatively specific facial actions, sleep wake states, and 2 theoretically derived, developmentally relevant hand actions.

Received for publication July 20, 2007; revised August 14, 2007; accepted August 15, 2007.

From the *Community Child Health Research, Child and Family Research Institute; Departments of †Occupational Science and Occupational Therapy; ‡Pediatrics, University of British Columbia; and §Children’s and Women’s Health Centre of British Columbia, Vancouver, Canada.

This study was funded by a National Institutes of Health grant HD39783 (R.E.G.), a Canadian Institutes of Health Research grant MOP42469 (R.E.G.), a British Columbia Ministry of Children and Family Development grant through the Human Early Learning Partnership (L.H.); Child and Family Research Institute Establishment Grants (L.H.: 06-2426 and R.E.G.: 02-2403;03-3112), a Canadian Child Health Clinician Scientist Career Award (L.H.), and a Michael Smith Foundation for Health Research Scholar award (R.E.G.).

Reprints: Dr Liisa Holsti, PhD, Community Child Health Research, F611, 4480 Oak Street, Vancouver, British Columbia, V6H 3V4, Canada (e-mail: lholsti@cw.bc.ca).

Copyright © 2007 by Lippincott Williams & Wilkins

Interrater reliability, concurrent and construct validity for the BIIP have been established by assessing change in pain scores between periods when the infants were not handled and when they experienced heel lance during routine blood collection.⁸ The purpose of this study was to establish the discriminant validity of the scale. Our primary aim was to determine whether or not the BIIP distinguished between skin-breaking and nonskin-breaking procedures in preterm infants in the NICU. Our secondary aim was to examine whether or not the BIIP was accurate and specific enough to detect sensitized responses during painful procedures.

MATERIALS AND METHODS

In this within subjects, repeated measures, cross-over study, 69 (36 male, 33 female) preterm infants born between 24 and 32 weeks GA were assessed at 32 weeks postconceptional age in the level III to IV NICU at the Children's and Women's Health Center of British Columbia. Ten infants (17%) were small for GA at birth and 47 (68%) were of white ethnicity. Infants who had received analgesics or sedatives within 72 hours of the assessment, who had major congenital anomalies, or whose mothers used illicit drugs during pregnancy were excluded. Infant characteristics are presented in Table 1.

Background Data

A NICU-trained research nurse completed the prospective clinical chart review collecting information from birth to day of testing including the following: birth weight, GA at birth, sex, illness severity on day 1 (Score for Neonatal Acute Physiology-II),¹⁰ cranial ultrasound scan results, daily opioid and other analgesic and sedative exposure, numbers and types of invasive skin breaking procedures, respiratory support, and type and time of last

TABLE 1. Demographic Characteristics to First Study Day 1 (n = 69)

	Mean ± SD	Range
Birth weight (g)	1240 ± 428	500-2525
Gestational age at birth (wk)	29 ± 2	24-32
SNAP-II day 1*	15 ± 9	5-46
Ventilation (d)	10 ± 13	0-50
Other respiratory support	10 ± 9	0-36
Dexamethasone (d)	0.2 ± 0.1	0-8
Pain exposure if Pain only first (procedure A)†	84 ± 55	8-211
Pain exposure if Diaper first (procedure B)†	81 ± 65	11-276
Intravenous morphine exposure if Pain first procedure‡	0.1 ± 0.2	0-0.6
Intravenous morphine exposure if Diaper first procedure‡	0.1 ± 0.2	0-0.1
Maternal age (y)	31 ± 6	19-47

*Score for Neonatal Acute Physiology.¹¹

†Number of invasive (skin breaking) procedures from birth to the first study day.

‡Morphine exposure = (daily average intravenous mg/kg) X days.

handling just before blood collection. Procedural pain exposure was defined as the sum of every skin breaking procedure from birth to the first testing day (eg, heel lance, intramuscular injection, chest tube insertion, central line insertion). Each attempt at a procedure is documented in the medical chart in our nursery; therefore, the total reflected all skin breaks. Total intravenous (IV) morphine exposure was calculated from birth to the test day by multiplying the average daily dose of IV morphine, adjusted for daily weight, by the number of days of IV morphine.^{3,11}

Procedures

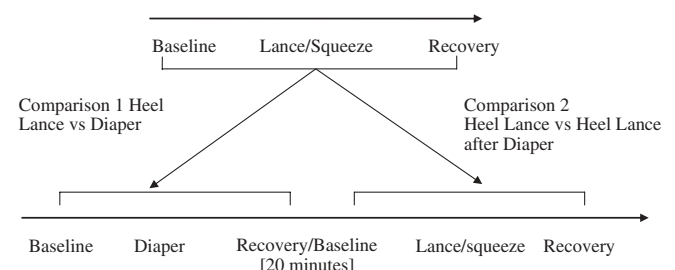
Infants were recruited by a NICU research nurse and written informed consent was obtained from the mother according to a protocol approved by the Clinical Research Ethics Board of the University of British Columbia. Videotaping and physiologic recordings were carried out continuously at bedside; these methods are reported elsewhere.⁸ The infants were studied on separate days in random order. On one day, after at least a 30-minute period in which the infant was not disturbed (average undisturbed time 106 min), behavioral and physiologic responses were recorded before, during, and after blood collection (procedure A). On the other day, after a period of undisturbed time (average undisturbed time 103 min), a fixed series of nursing procedures (diaper change, measuring abdominal girth, temperature, and mouth care) was performed before a scheduled clinical blood collection. Between the nursing procedures and the subsequent blood collection, each infant was left nested and undisturbed for a 20-minute rest period (procedure B). A time-line of the procedures is presented in Figure 1. The clinical status of the infants on the first study day is presented in Table 2.

Measures

BIIP⁸

The BIIP combines into a single score changes in sleep/wake states, 5 facial actions and 2 hand actions. This scale has been shown to have high interrater reliability and moderate concurrent validity with both a multidimensional pain scale (Neonatal Infant Pain Scale)

Procedure A. Blood Collection



Procedure B. Diaper followed by Blood Collection

FIGURE 1. Time line and analysis of procedures.

TABLE 2. Infant Characteristics (n = 69)

	Pain Procedure Only (Procedure A) (n = 31)			Diaper Procedure Before Pain Procedure (Procedure B) (n = 38)		
	Mean ± SD	Range	N (%)	Mean ± SD	Range	N (%)
Postnatal age (d)	21 ± 14	3-59		18 ± 13	3-45	
Receiving mechanical ventilation			7 (23)			5 (13)
Time since last feed (min)*	49 ± 27	0-112		58 ± 34	0-152	
No. painful procedures in 24 h before first study	1 ± 1	0-6		1 ± 2	0-8	
Time since last painful procedure (min)	1635 ± 1793	56-6690		1485 ± 1241	60-5011	

*Four infants were not on oral feeds and 5 were on continuous feeds.

and with changes in mean heart rate (HR).^{8,9} A copy of the scale and the item descriptors can be found in Holsti and Grunau, in press.⁸

HR

Continuous electrocardiographic (ECG) activity was recorded from a single lead of surface ECG and was digitally sampled at 360 Hz off-line using a specially adapted computer acquisition system. Custom physiologic signal processing software was used to acquire, process, and analyze HR.¹² R waves were detected from the sampled ECG, and were used to form a smoothed instantaneous 4-Hz time series.¹³ Epochs of HR (2.2 min each) were selected for Baseline, Lance/squeeze, and Recovery and for Baseline, Diaper, and Recovery. The epoch selection criteria were based on quantitative signal stationarity, the presence of a stable behavioral state, and the absence of gross movement artifact.¹⁴

Video Coding

Infant behaviors were videotaped continuously in real time across 3, 1-minute phases of blood collection (Baseline, Lance/squeeze, and Recovery) and diaper change (Baseline, Diaper, and Recovery), procedures required for the clinical management of the infants. Videotapes were edited for coding in random order of phases; coders were blind to all clinical information about the infants and to the purpose of the study. Interrater reliability coding was carried out on a randomly selected 20% of the sample. Using the 2-way random absolute agreement method of obtaining intraclass correlations, coders achieved reliability ranging between 0.79 (Baseline phase) and 0.92 (Lance/squeeze and Diaper phases).

Data Analysis

To assess changes in BIIP scores and mean HR, repeated measures analysis of variance (ANOVA) was carried out separately across the 3 phases of blood collection (procedure A), the 3 phases of diaper change, and for the 3 phases of blood collection which was preceded by the diaper change (procedure B). Figure 1 shows the separate comparisons. Because we and others have not found sex differences in pain responses of

preterm infants,^{8,11} sex was not used as a between subjects variable. Five infants with significant central nervous system injury (intraventricular hemorrhage, grade III or IV, and/or periventricular leukomalacia) were retained in the analysis because facial and hand actions are not altered in these infants in response to skin-breaking procedures.^{8,15} Bonferroni corrections were used to correct for overall type I error. Repeated measures data were examined for sphericity; Greenhouse-Geisser Epsilon values were used to determine significance. Statistically significant ANOVA was followed by planned Student *t* tests for paired comparisons to identify differences between phases. Paired *t* tests were used to compare BIIP scores and mean HR between the Lance/squeeze phase with the Diaper phase. Paired *t* tests were used to compare the mean BIIP scores of the Lance/squeeze phases between procedure A (blood collection preceded by rest) and procedure B (blood collection preceded by the Diaper procedure); changes in mean HR were analyzed in a similar way.

Finally, because our sample was heterogeneous, that is, we included infants born at varying GAs, illness severity, etc, using ANOVA with GA as a between subjects grouping factor. We compared responses of infants born very early (< 29 wk; n = 29) to infants born at later GAs (29 to 32 wk n = 40). GA at birth, illness severity, days of ventilation are all variables that are highly intercorrelated¹⁶; therefore, we chose GA at birth as a proxy measure. In addition, we used this cut-off because we have shown previously that infants born below 29 weeks are much more likely to require intubation/ventilation and other invasive procedures.¹⁷ In addition, we reanalyzed the data to determine whether including the 10 infants born small for gestational age altered our results.

RESULTS

As our primary aim was to examine whether or not the BIIP discriminated between skin-breaking and non-skin-breaking procedures, sample size estimates were calculated as though we were using a between groups design. GPOWER was used to calculate the estimate,¹⁸ and effect sizes entered into the program were based on

differences in Neonatal Facial Coding System scores in term infants between a painful and nonpainful event.¹⁹ Using this method, we far exceeded the 16 infants which were needed to detect differences between each phase for a power of 0.95 with the statistical significance set at 0.05 and with a large effect size (1.23).

Total handling time for the stressor phases was not statistically significantly different (Lance/squeeze procedure A vs. Diaper: 6.0 and 5.7 min, and Lance/squeeze procedure A vs. Lance/squeeze procedure B: 6.0 and 5.7 min). Significant main effects were found for changes in BIIP scores and mean HR across procedure phases (Baseline, Lance/squeeze, and Recovery; Baseline Diaper, and Recovery; and Baseline, Lance/squeeze, and Recovery—procedure B) (Table 3). Post hoc analyses revealed significant increases in BIIP scores and mean HR from Baseline to the Stressor phases in all instances. Excluding the infants born small for gestational age did not alter any of our results. For procedure A, mean BIIP scores increased significantly from Baseline to the Lance/squeeze phase (procedure A—95% confidence interval (CI): -4.4 to -3.5, $P < 0.001$) and similarly, mean BIIP scores increased from Baseline to the Diaper change (procedure B—95% CI: -2.8 to -1.5, $P < 0.001$). Similar results were found for changes in mean HR [(procedure A) Baseline to Lance, 95% CI: -22.5 to -15.8, $P < 0.0001$; (procedure B) Baseline to Diaper, 95% CI: -15.1 to -10.1, $P < 0.0001$]. In examining changes between the Stressor phases (Lance and Diaper) and the Recovery, the infants' behavioral and physiologic responses returned to Baseline levels except during the following 2 procedures: during Recovery for the Diaper procedure BIIP scores remained elevated significantly (95% CI: -1.5 to -0.1, $P < 0.02$); during the Recovery phase compared with Baseline, HR remained elevated significantly when blood collection was preceded by the Diaper procedure (procedure B; 95% CI: -6.9 to -0.4, $P < 0.03$).

Results from the paired *t* tests indicated that mean BIIP scores were higher during the Lance/squeeze than during the Diaper [5.1 vs. 3.1; $t = 5.1$, $P < 0.0001$ (95% CI: 1.22-2.75)]. BIIP scores did not differ during the Baseline or Recovery phases between procedures. In parallel with BIIP scores, mean HR was also higher during the Lance/squeeze phase than during the Diaper phase [176 vs. 168; $t = 4.7$, $P < 0.0001$ (95% CI: 4.5-11.2)], but did not differ during the Baseline or Recovery phases. Infants who had blood collection after a diaper change showed higher BIIP scores [procedure A—5.1 vs.

procedure B—6.0, $t = 2.3$, $P < 0.03$ (95% CI: 0.12-2.3)], but not mean HR during the subsequent Lance/squeeze phase than had they had been left undisturbed before the blood collection. HR during the Recovery from the blood collection in procedure A did not differ significantly from Recovery from the blood collection in procedure B.

Finally, both groups of infants (< 29 wk GA and 29 to 32-wk GA) showed higher BIIP scores to the Lance/squeeze than to the Diaper; no between groups differences were found. However, infants born at the earlier GA had higher mean HR to the Diaper [$F(1,67) = 3.9$, $P < 0.05$] than infants in the 29 to 32-week GA group. No differences in mean HR were found for the Lance/squeeze phase.

DISCUSSION

Being able to differentiate between procedures that are less intrusive from those that induce pain is important for the appropriate application of pharmacologic and nonpharmacologic interventions.^{5,6} In this study, we have demonstrated that the BIIP discriminates acute skin-breaking from nonskin-breaking procedures in preterm infants assessed at 32 weeks postconceptional age in the NICU, data which supports further the validity of the scale. Although some would suggest that evaluating the validity of this scale should be done on a homogenous population, we would argue that a particular strength of our study is that we used a sample, which was representative of the varying stages of illness/recovery found typically in preterm infants in the NICU.

Previously, graded behavioral responses to procedures have been observed reliably in infants as young as 28 weeks GA; however, body movement indicators were not included in that study.²⁰ In addition, in infants assessed between 28 and 36 weeks GA, wiping a cotton pad on the hand stimulated lower levels of somatosensory oxygenated hemoglobin than did heel lance.² Importantly, the infants born at the earlier GAs showed the greatest changes during the heel lance.² Nevertheless, in that study, the skin-breaking procedure lasted almost twice as long as the tactile procedure; whereas in our study, the time of blood collection and diaper change were equal. Another strength of this study is our analysis of 2.2 minutes of HR data in response to the stressors; longer time periods provide important information on the infants' physiologic regulatory capacities.²¹

TABLE 3. Changes in Mean BIIP Scores and HR Across Procedure Phases (n = 69)

Outcome Measures	Procedure	Baseline (Mean ± SD)	Stressor (Mean ± SD)	Recovery (Mean ± SD)	F	df	Overall Phase P Value
BIIP	Heel lance	1.0 ± 1.7	5.1 ± 2.6	1.6 ± 2.6	75.4	1.68	0.0001
	Diaper	1.0 ± 1.6	3.1 ± 2.5	1.8 ± 2.3	19.7	1.68	0.0001
	Heel lance after diaper	1.0 ± 1.6	6.0 ± 2.7	1.5 ± 2.5	100.5	1.68	0.0001
HR	Heel lance	156.9 ± 10.6	176.0 ± 14.9	159.9 ± 14.5	67.4	1.68	0.0001
	Diaper	155.7 ± 13.2	168.3 ± 14.7	157.8 ± 14.0	51.6	1.68	0.0001
	Heel lance after diaper	156.4 ± 12.6	175.5 ± 14.2	160.1 ± 14.5	90.2	1.68	0.0001

Like others, we found that behavioral and physiologic responses were divergent.^{22,23} During the Recovery phase, after the Diaper procedure, mean HR returned to baseline levels, whereas behavioral responses (BIIP scores) remained elevated. What was once thought to be relatively innocuous, diaper changing induces in some infants' responses that are of greater length and intensity than those during heel lance.²⁴ Indeed, HR responses were higher to the Diaper in infants born at the earlier GAs (< 29 wk). On the other hand, when blood collection was preceded by the Diaper procedure, mean HR remained elevated during the Recovery phase whereas behavioral responses returned to baseline levels. It is difficult to reconcile why these differences in Recovery to the procedures occurred. Divergent results such as these demonstrate the complexities in responses of these infants. Accordingly, unlike with multidimensional scales (those which combine behavioral and physiologic indices into a single score) when used along with changes in physiologic measures, the BIIP allows both researchers and clinicians to evaluate the complex nature of pain responses and the effects of interventions on each response system separately. Moreover, as mentioned previously, the BIIP is a unidimensional scale; that is, it is comprised of behavioral indicators only. From an evolutionary perspective, behavioral indicators are designed specifically to elicit care giving; therefore, they may be more ecologically salient than are physiologic indicators.²⁵

Unique to this study is the evaluation of the BIIP scale for detecting sensitized (heightened) pain responses in these infants; indeed, the BIIP was accurate in detecting differences in pain responses when blood collection after a prolonged rest period was compared with the responses in the same infants after they had been exposed to routine nursing procedures (diaper change) before the blood collection. In term infants, a single major painful event produced sensitization to painful events experienced a few months later.^{26,27} Preterm infants are at great risk for enhanced pain sensitivity. Studies using animal models show that a single skin wound produces a profound and long lasting hyperinnervation of myelinated A-fiber and unmyelinated C-fiber sensory nerve terminals.^{28,29} In addition, neonatal dorsal horn cells in the spinal cord are much larger,³⁰ and natural stimulation of these fields may lower their threshold to additional stimuli (central sensitization).³¹ Furthermore, their descending inhibitory pathways are not developed fully.^{32,33} Accordingly, preterm infants below 35 weeks postconceptional age show altered peripheral nociceptive sensitization with lowered thresholds to tactile stimulation; their thresholds decrease further (primary hyperalgesia) after repeated pain exposure.³⁴ Responses such as these may indicate not only peripheral sensitization, but also allodynia, a condition whereby previously innocuous stimulation is perceived as painful as a result of central sensitization.^{35,36} Thus, changes in peripheral and central pain processing pathways are likely the reason why both skin-breaking and nonskin-breaking procedures induce sensitized behavioral and physiologic responses in these

infants even when a rest period of up to 20 minutes is given between procedures.^{17,37} Further research is needed to determine what the ideal length of time should be between procedures to prevent such heightened responses.

Given the number of assessment tools already available, why prefer the BIIP? First many do not have specifically defined, developmentally appropriate, anatomically based indicators making accurate and reliable scoring less likely. Second, the scale studied and used most extensively to date, the Premature Infant Pain Profile (a multidimensional scale), includes weightings that are applied to account for factors influencing pain responses (behavioral state and GA at birth) in this population.³⁸ As we have argued previously, weightings such as these may conceal important differences in arousal in infants born at varying GAs.⁸ In addition, some have found the PIPP difficult to administer in real time in the clinical setting.³⁹ As an alternative, the BIIP scale has a number of advantages. First, the indicators that comprise the scale have been shown to be the relatively most specific behavioral cues for assessing acute pain.^{11,40,41} Second, among these indicators are 2 hand actions (finger splay and fisting) which are movements that have very specific descriptors and that are derived from a theoretical model developed for assessing levels of stress and stability in preterm and high-risk full-term neonates.⁴² Importantly, these 2 hand actions seem to act as a "counterbalance" for assessing pain in infants whose facial responses may be diminished as a result of being exposed to repeated painful/stressful procedures.^{3,43} Thus, we have achieved accounting for potential differences in facial responses without the need for statistical weightings.

Several potential limitations of our study should be mentioned. First, we did not include infants who were assessed below 32 weeks postconceptional age. Assessing pain in infants born at extremely low GAs remains challenging because many are placed on sedation or analgesic medications. However, preliminary data on a small number of infants show that the BIIP is able to detect pain responses reliably in infants assessed below 28 weeks GA.⁸ In addition, the infants in this study were scored from videotape rather than at bedside; thus, although we can demonstrate reliability and some aspects of validity of the BIIP under controlled conditions, establishing clinical utility and feasibility is still required.

In conclusion, this study provides further evidence that the BIIP scale is valid, reliable, and accurate for measuring acute procedural pain in preterm infants in the NICU. Combining the relatively most specific, developmentally relevant, anatomically based, and theoretically derived indicators, this scale offers significant advantages over others which are available currently. Future work will include evaluating the psychometric properties of the BIIP in infants assessed at earlier GAs. Moreover, further research is required to determine whether or not the BIIP scale is valid and reliable for measuring prolonged pain (eg, postoperative pain) and chronic pain states, conditions for which few measurement tools exist.

ACKNOWLEDGMENTS

The authors thank the staff and families of the Neonatal Intensive Care Unit at the Children's and Women's Health Centre of British Columbia for their participation in this study, Colleen Jantzen from the Early Human Experience Unit, Community Child Health Research, Child and Family Research Institute for carrying out behavioral coding, and the CANDO Research Unit in the Department of Occupational Science and Occupational Therapy, University of British Columbia.

REFERENCES

- Slater R, Cantarella A, Gallella S, et al. Cortical pain responses in human infants. *J Neurosci*. 2006;26:3662–3666.
- Bartocci M, Bergqvist LL, Langercrantz H, et al. Pain activates cortical areas in the preterm newborn brain. *Pain*. 2006;122:109–117.
- Grunau RE, Oberlander TF, Whitfield MF, et al. Demographic and therapeutic determinants of pain reactivity in very low birth weight neonates at 32 weeks' postconceptional age. *Pediatrics*. 2001;107:105–112.
- Morison SJ, Holsti L, Grunau RE, et al. Are there developmentally distinct motor indicators of pain in preterm infants? *Early Hum Dev*. 2003;72:131–146.
- Rahman W, Fitzgerald M, Aynsley-Green A, et al. The effects of neonatal exposure to inflammation and/or morphine on neuronal responses and morphine analgesia in adult rats. In: Jensen TS, Turner JA, Weisenfeld-Halling Z, eds. *Progress in Pain Research and Management*. Vol. 8. Seattle, WA: IASP Press; 1997:738–794.
- Johnston CC, Filion F, Snider L, et al. How much sucrose is too much sucrose? [Letter]. *Pediatrics*. 2007;119:226.
- Duhn LJ, Medves JM. A systematic integrative review of infant pain assessment tools. *Adv Neonatal Care*. 2004;4:126–140.
- Holsti L, Grunau RE. Initial validation of the behavioral indicators of infant pain. *Pain*. 2007. In press.
- Lawrence J, Alcock D, McGrath P, et al. The development of a tool to assess neonatal pain. *Neonatal Network*. 1993;12:59–66.
- Richardson DK, Corcoran JD, Escobar GJ, et al. SNAP-II and SNAPPE-II: simplified newborn illness severity and mortality risk scores. *J Pediatr*. 2001;138:92–100.
- Holsti L, Grunau RE, Oberlander TF, et al. Specific NIDCAP® movements help identify acute pain in preterm infants in the NICU. *Pediatrics*. 2004;114:65–72.
- HR View Software (Computer program). Brighton, MA: Boston Medical Technologies; 1996.
- Berger RD, Saul P, Cohen RJ. Transfer function analysis of autonomic regulation: canine atrial rate response. *Am J Physiol*. 1989;256:H142–H152.
- Oberlander T, Saul JP. Methodological considerations for the use of heart rate variability as a measure of pain reactivity in vulnerable infants. *Clin Perinatol*. 2002;29:427–443.
- Oberlander TF, Grunau RE, Fitzgerald C, et al. Does parenchymal brain injury affect biobehavioral pain responses in very low birth weight infants at 32 weeks' postconceptional age? *Pediatrics*. 2002;110:570–576.
- Grunau RE, Holsti L, Haley D, et al. Neonatal procedural pain exposure predicts lower cortisol and behavioral reactivity in preterm infants in the NICU. *Pain*. 2005;113:293–300.
- Holsti L, Grunau RE, Whitfield MF, et al. Behavioral responses to pain are heightened after clustered care in preterm infants. *Clin J Pain*. 2006;22:757–764.
- Faul F, Erdfelder E. *GPOWER: A Priori-, Post Hoc-, and Compromise Power Analyses for MS-DOS (Computer Program)*. Bonn, Germany: Bonn University; 1998.
- Grunau RE, Johnston CC, Craig KD. Neonatal facial and cry responses to invasive and non-invasive procedures. *Pain*. 1990;42:295–305.
- Johnston CC, Stevens B, Yang F, et al. Developmental changes in response to heel stick in preterm infants: a prospective cohort study. *Dev Med Child Neurol*. 1996;38:438–445.
- Stevens BJ, Riddell RRP, Oberlander TF, et al. Assessment of pain in neonates and infants. In: Anand KJS, Stevens BJ, McGrath PJ, eds. *Pain in Neonates and Infants*. 3rd ed. *Pain Research and Clinical Management*. New York: Elsevier; 2007:67–90.
- Johnston CC, Stevens B, Pinelli J, et al. Kangaroo care is effective in diminishing pain response in preterm neonates. *Arch Pediatr Adolesc Med*. 2003;157:1084–1088.
- Boyer K, Johnston C, Walker CD, et al. Does sucrose analgesia promote physiologic stability in preterm neonates? *Biol Neonate*. 2004;85:26–31.
- Holsti L, Grunau RE, Whitfield MF, et al. Body movements, an important additional factor in discriminating pain from stress in preterm infants. *Clin J Pain*. 2005;21:491–498.
- Barr R. Reflections on measuring pain in infants: dissociation in responsive systems and “honest signaling.” *Arch Dis Child*. 1998;79:F152–F156.
- Taddio A, Coldbach M, Ipp M, et al. Effect of neonatal circumcision on pain responses during vaccination in boys. *Lancet*. 1995;345:291–292.
- Taddio A, Katz J, Ilersich AL, et al. Effect of neonatal circumcision on pain response during subsequent routine vaccination. *Lancet*. 1997;349:599–603.
- Alvarez D, Toursney D, Beland B, et al. Modelling the prolonged effects of neonatal pain. In: Sandkühler J, Bromm B, Gebhart GF, eds. *Nervous System Plasticity and Chronic Pain, Progress in Brain Research*. Amsterdam: Elsevier Science; 2000:365–373.
- Reynolds ML, Fitzgerald M. Long term sensory hyperinnervation following neonatal skin wounds. *J Comp Neurol*. 1995;358:487–498.
- Fitzgerald M. The post-natal development of cutaneous afferent fibre input and receptive field organization in the rat dorsal horn. *J Physiol*. 1985;364:1–18.
- Jennings E, Fitzgerald M. Postnatal changes in responses of rat dorsal horn cells to afferent stimulation: a fibre-induced sensitization. *J Physiol*. 1998;509:859–868.
- Fitzgerald M. The development of nociceptive circuits. *Nat Rev Neurosci*. 2005;6:507–520.
- Beggs S, Fitzgerald M. Development of peripheral and spinal nociceptive systems. In: Anand KJS, Stevens BJ, McGrath PJ, eds. *Pain in Neonates and Infants*. 3rd ed. *Pain Research and Clinical Management*. New York: Elsevier; 2007:11–24.
- Andrews K, Fitzgerald M. Cutaneous flexion reflex in human neonates: a quantitative study of threshold and stimulus-response characteristics after single and repeated stimuli. *Dev Med Child Neurol*. 1999;4:696–703.
- Fitzgerald M, Millard C, McIntosh N. Hyperalgesia in premature infants. *Lancet*. 1988:292.
- Fitzgerald M, Millard C, McIntosh N. Cutaneous hypersensitivity following peripheral tissue damage in newborn infants and its reversal with topical anaesthesia. *Pain*. 1989;39:31–36.
- Holsti L, Grunau RE, Oberlander TF, et al. Prior pain induces heightened motor responses during clustered care in preterm infants in the NICU. *Early Hum Dev*. 2005;81:293–302.
- Stevens B, Johnston CC, Petryshen P, et al. Premature infant pain profile: development and initial validation. *Clin J Pain*. 1996;12:13–22.
- Bellieni CV, Cordelli DM, Caliani C, et al. Inter-observer reliability of two pain scales for newborns. *Early Hum Dev*. 2006;83:549–552.
- Franck LS, Greenburg CS, Stevens B. Pain assessment in infants and children. Acute pain in children. *Pediatr Clin North Am*. 2000;47:487–512.
- Grunau RE, Holsti L, Whitfield MF, et al. Are twitches, startles and body movements pain indicators in extremely low birth weight infants? *Clin J Pain*. 2000;16:37–45.
- Als H. *Manual for the Naturalistic Observation of Newborn Behavior (Preterm and Fullterm)*. Boston: The Children's Hospital; 1984.
- Johnston CC, Stevens BJ. Experience in a neonatal intensive care unit affects pain response. *Pediatrics*. 1996;98:903–925.