## CLINICAL RESEARCH ARTICLE

# The newborn infant parasympathetic evaluation index for acute procedural pain assessment in preterm infants

Julie Gendras[1,2], Pauline Lavenant[1,2], Iona Sicard-Cras[1,2], Maëlys Consigny[3], Laurent Misery[2], Kanwaljeet J. S. Anand[4], Jacques Sizun[1,2] and Jean-Michel Roué[1,2,4]

**BACKGROUND:** Accurate assessments of pain in hospitalized preterm infants present a major challenge in improving the short- and long-term consequences associated with painful experiences. We evaluated the ability of the newborn infant parasympathetic evaluation (NIPE) index to detect acute procedural pain in preterm infants.
**METHODS:** Different painful and stressful interventions were prospectively observed in preterm infants born at 25 + 0 to 35 + 6 weeks gestation. Pain responses were measured using the composite Premature Infant Pain Profile Revised (PIPP-R) scale, the NIPE index, and skin conductance responses (SCR). Outcome measures were correlations between the NIPE index, the PIPP-R score, and the SCR. Sensitivity/specificity analyses tested the accuracy of the NIPE index and SCR.
**RESULTS:** Two hundred and fifty-four procedures were recorded in 90 preterm infants. No significant correlation was found between PIPP-R and the NIPE index. PIPP-R and SCR were positively correlated ($r = 0.27$, $P < 0.001$), with stronger correlations for painful procedures ($r = 0.68$, $P < 0.001$) and especially for skin-breaking procedures ($r = 0.82$, $P < 0.001$). The NIPE index and SCR had high sensitivity and high negative predictive values to predict PIPP-R > 10, especially for skin-breaking painful procedures.
**CONCLUSIONS:** We found no significant correlation between the NIPE index and PIPP-R during routine painful or stressful procedures in preterm infants.

**IMPACT:**

- Exposure to repetitive pain can lead to neurodevelopmental sequelae. Behavior-based pain scales have limited clinical utility, especially for preterm infants.
- New devices for monitoring physiological responses to pain have not been validated sufficiently in preterm infants.
- This study found that the NIPE index was not significantly correlated to the validated PIPP-R scale during acute procedural pain.
- Secondary analysis of this study showed that NIPE index and SCRs may help to exclude severe pain in preterm infants.
- In clinical practice, measurements of physiological parameters should be combined with behavior-based scales for multidimensional pain assessments.

## INTRODUCTION

Hospitalized preterm infants are exposed to significant numbers of painful or stressful procedures during neonatal intensive care unit (NICU) care.[1,2] Repeated pain experiences have been widely associated with short- and long-term sequelae like hyperalgesia, altered brain architecture, or neurodevelopmental impairments.[3–5] This pain burden increases with greater prematurity and many of these procedures are still not treated with analgesia.[1,6]

Some studies have pointed out the limitations of clinical behavior-based scales that may mirror somatic subcortical and motor nervous system activation, rather than the reflection of the cerebral activity.[7–9] Indeed, Slater et al.[10] reported that, in some cases, the pain-specific cortically evoked responses persisted after an acute nociceptive stimulus while the neonate showed no behavioral signs of pain. These results highlight the need for more objective pain assessment in preterm infants, especially when clinical conditions do not allow adequate behavioral assessments.[11]

Alternatives to these pain scales were developed based on the description of nociceptive pathways and neurophysiological responses to pain in newborn infants, such as hormonal, cerebral oxymetry, galvanic skin responses (GSRs), or physiological responses.[12] Their goal was to improve the accuracy of pain assessments using monitoring of reliable and reproducible parameters. Among these methods, two can be routinely and easily applied at the bedside: (1) the GSRs, a reflection of sympathetic activity, which was reported to discriminate pain in neonates as early as 22 weeks of gestational age;[13] and (2) heart rate variability (HRV) to measure the impact of nociception on parasympathetic activity in real time, which was correlated with postoperative pain in newborn infants.[14]

[1]Pôle de Femme, de la Mère et de l'Enfant, Centre Hospitalier Régional Universitaire, Brest, France; [2]Laboratoire Interactions Epithéliums Neurones, EA 4685, Faculté de médecine et des Sciences de la Santé, Université de Bretagne Occidentale, Brest 29200, France; [3]Centre d'Investigation Clinique 0502, Inserm, Brest 29200, France and [4]Pain/Stress Neurobiology Lab, Maternal and Child Health Research Institute, Stanford University School of Medicine, Stanford, CA, USA
Correspondence: Jean-Michel Roué (jean-michel.roue@chu-brest.fr)

The newborn infant parasympathetic evaluation index for acute procedural...
J Gendras et al.

1841

While a study of 29 preterm infants found no correlation between the newborn infant parasympathetic evaluation (NIPE) index and validated pain scales, other studies have shown the ability of HRV analysis to detect pain responses after acute painful procedures in term or preterm neonates, and in children.[15–20] However, the validity of neurophysiological methods must be established before recommending them for routine pain assessment in clinical settings.

Our study aimed to assess the ability of the NIPE index (NIPEi) to detect acute procedural pain and stress responses in preterm infants by concurrently comparing the NIPEi to a validated composite pain scale (PIPP-R) and the skin conductance responses (SCRs).

## METHODS
A prospective, observational study was conducted in the Department of Neonatal Medicine of the University Hospital of Brest, France. Patients were enrolled from January 2017 to January 2018.

### Population
Following signed consents from both parents, hospitalized preterm infants with gestational ages ranging from $25 + 0$ to $35 + 6$ weeks were eligible for this study. Measurements were performed during routine care procedures categorized as stressful or painful according to a previously published classification.[1] Stressful procedures were included to compare painful responses vs. non-painful responses. Exclusion criteria included brain injury (grade 3 or 4 intraventricular hemorrhage, white matter injury), genetic abnormality, severe congenital malformation, administration of neuromuscular blockade, and either anticholinergic or adrenergic antagonists administered during the 48 h prior to recording. A maximum of ten procedures could be recorded for any single patient.

### Study design
Pain and stress responses were measured using (1) the NIPEi (MDMS, Loos, France), (2) the GSR Pain Monitor (Med-Storm, Oslo, Norway), and (3) the PIPP-R score. The NIPE monitor was connected to the cardiorespiratory monitor (Intellivue, MX series, Philips Medical Systems, Eindhoven, The Netherlands) 20 min before the start of the care procedure for signal averaging as required by the manufacturer. The three Pain Monitor leads were placed on palmar or plantar skin. Two analog cameras (Sony HDR-CX740 24.1; Sony Cyber-shot DSC-W830), one focused on the patient's face and the other on physiological parameters (heart rate, $SpO_2$), were used to perform PIPP-R scoring post hoc.

### Measurements
Frequency analysis of HRV was previously validated as a method for pain assessment with decreased parasympathetic responses to painful stimuli during heel lance procedures.[17,21] The NIPE® monitor (MDMS, Loos, France) was designed using this principle, specifically for infants until 2 years old, including preterm infants.[22,14] The instantaneous NIPEi is derived from an algorithm evaluating the short-term HRV in real time.[22] This automated analysis of neonatal HRV in high frequencies (<0.15 Hz) reflects parasympathetic activity. It ranges from 0 to 100 and decreases with the intensity of pain or stress. The NIPEm represents the mean value over 20 min. The instantaneous NIPEi provides a value based on a 64-s moving window analysis, which is updated every 1 s.[20] The NIPE monitor is connected to the cardiac monitor and continuously displays the NIPEm and the NIPEi indexes. The NIPEi seems, therefore, more suitable for acute procedural pain analysis and was used to study acute pain/stress in preterm infants. NIPEi was recorded from 2 min before until 2 min after the procedure, and removed before any other procedure could interfere with the recorded responses.[16,23] Since the NIPEi appeared to change more slowly, on our first data analysis, as compared to the PIPP-R or the SCR, we decided to perform a post hoc analysis for the patients who had the NIPEi measured until 3 min after the procedure and before any other procedure (89% for the first recorded procedure and 84% for the whole sample of procedures).[20] Data were directly exported from the NIPE monitor after recording.

The PIPP-R[24,25] has been validated in term and preterm infants during acute painful procedures. This composite scale includes two physiological parameters (heart rate and oxygen saturation), three behavioral parameters (brow bulge, eye squeeze, and nasolabial furrow), and two contextual factors (GA, behavioral state). The PIPP-R was scored continuously over periods of 30 s starting from 2 min before until 2 min after the procedure. Two trained raters reviewed the video recordings independently and assigned PIPP-R scores; an expert validated these assessments for the first 36% procedures. The calculated intraclass correlation coefficient was high (0.92) showing the inter-rater reliability of the PIPP-R analysis.

The Med-Storm Pain Monitor recorded the SCR integrating a dedicated software using the "premature" mode. Increased number of peaks/s were previously reported to discriminate pain responses[26,27] and was measured in time windows of 30 s from 2 min before until 2 min after the procedure according to previous studies.[26,27]

### Statistical analysis
Power analysis suggested a sample size of 200 patients to estimate a correlation coefficient of 0.8 with an accuracy of 0.05 (half-amplitude of the 95% confidence interval) and $\alpha$-error of 5%. Data were analyzed using the software SAS version 9.4. Correlations between the NIPEi index, the PIPP-R score, and the number of peaks/s (SCR) were calculated using Pearson's or Spearman's correlation coefficients according to the normality of the distribution. In the case of repeated measurements, only the first pair of measurements for each patient was considered for the primary outcome. The secondary outcome integrated the type of stimulation (stressful or painful), the patient's sex and gestational age (<32 vs. 32–36 weeks). A modeling of the change of the different parameters was performed using a generalized linear model for repeated data. Sensitivity and specificity analyses were carried out for the NIPEi and SCR.

### Ethics
The study protocol was approved by the local Ethics committee (Comité de Protection des Personnes (CPP) Ouest 6; No. ID-RCB: 2016-A00012-49) and by French National Agency for Medicines and Health Products Safety. This observational study was registered at www.ClinicalTrials.gov (NCT02885051).

## RESULTS
### Patients and recorded procedures characteristics
Two hundred and fifty-four procedures were recorded for the 90 preterm infants enrolled in this study (Fig. 1). The study was stopped at the end of the planned recruitment period before reaching the calculated sample size. The observed correlation coefficients were much lower than those predicted in our hypothesis. Therefore, the study was stopped because of futility. Patient characteristics and types of procedures are listed in Table 1.

The mean GA was 30.9 weeks with a mean birth weight of 1548 g. Capillary and venipunctures represented 32.7% ($n = 36/110$) of all painful procedures and 65.5% of the skin-breaking procedures. Nursing care represented 45.8% ($n = 66/144$) of the stressful procedures (for details, see Supplementary Table S1).
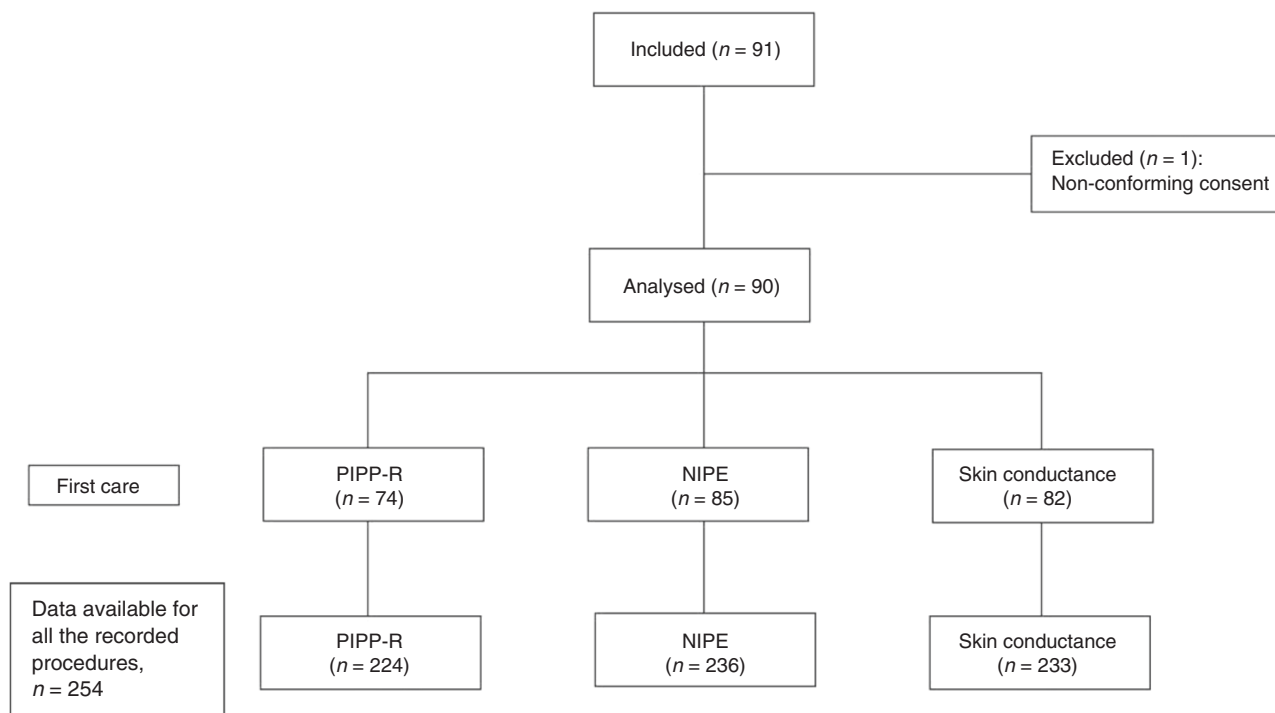
The newborn infant parasympathetic evaluation index for acute procedural...
J Gendras et al.

1842

**Fig. 1** Study progress.

**Table 1.** Characteristics of the patients at birth and procedures studied.

| Patients (n = 90) | | | |
|---|---|---|---|
| Gestational age at birth (weeks)[a] | 30.9 ± 2.8 | Male, n (%) | 42 (46.7) |
| Age at inclusion (days)[a] | 11.9 ± 14.4 | IUGR, n (%) | 18 (20.0) |
| Apgar 5 min | 9.5 ± 0.9, MD = 3 | Singleton, n (%) | 53 (58.9) |
| Birth weight (g)[a] | 1548 ± 488 | C-section, n (%) | 52 (58.4) |
| Type of procedure | | | |
| First recorded procedure (n = 89) | Painful procedures, n (%) | | 42 (47.2) |
| | Skin-breaking procedures, n (%) | | 25 (28.1) |
| | Stressful procedures, n (%) | | 47 (52.8) |
| All recorded procedures (n = 254) | Painful procedures, n (%) | | 110 (43.3) |
| | Skin-breaking procedures, n (%) | | 55 (21.7) |
| | Stressful procedures, n (%) | | 144 (56.7) |

IUGR intrauterine growth restriction, C-section cesarean section.
[a]Mean ± SD (standard deviation).

**Patterns of PIPP-R scores, NIPEi, and SCR over time**
Graphical plots showed peak responses during the procedure for the PIPP-R, over the initial 30 s of the procedure for the SCR, and at 90 s after the procedure for the NIPEi (Fig. 2). The increase in PIPP-R was more pronounced during painful procedures, while the SCR responses were more prominent after stressful procedures. The NIPEi showed no difference between stressful and painful procedures. The peak responses showing the highest PIPP-R scores during the procedure (time = 0 s) were observed in males and in very preterm infants (<32 weeks GA).

**Correlations between NIPEi, PIPP-R, and SCR**
Analysis of the first recorded procedures for each patient showed no significant correlation coefficients (Pearson's or Spearman's) between the NIPEi and PIPP-R scores (baseline, response peaks, and delta) (Table 2). Negative correlation coefficients were expected between the observed decrease in NIPEi responses and the increase in PIPP-R or SCR.

Repeated significant positive correlation coefficients were observed between the PIPP-R and SCR, especially in the 25–32 weeks GA group. Their best correlations ($r > 0.80$) were found for all painful procedures, particularly skin breaking procedures. Analysis of all 254 procedures showed similar trends, but with lower correlation coefficients between PIPP-R and SCR (Table 2). No significant negative correlation coefficients were found between NIPEi and SCR in the whole sample of recorded procedures ($n = 254$). The patients who had NIPEi measurements until 3 min after the procedure ($n = 80$ patients, $n = 213$ procedures) did not show higher correlation coefficients.

**Sensitivity/specificity analyses**
Measurement accuracy parameters are reported in Table 3 and Fig. 3. Based on receiver-operating characteristic (ROC) curve analyses, the highest areas under the curves (AUCs) were reported for skin-breaking procedures with cut-off values of NIPEi <46 and PIPP-R scores >6. With PIPP-R cut-off value >10, sensitivity increased mildly, but the negative predictive values (NPVs) were high (>80%). The NIPEi showed minimal specificity <60%.

The best AUC were reported with the SCR and especially for skin-breaking procedures, associated with high specificity and high NPV. Modeling the relationship between PIPP-R and SCR (peaks/s) associated to NIPEi did not improve the ROC curves significantly (AUC = 0.70 and 0.69, respectively, for PIPP-R > 6 and >10).

**DISCUSSION**
This is the first study that concurrently measures the parasympathetic (NIPEi index) and sympathetic (SCR) autonomic nervous

The newborn infant parasympathetic evaluation index for acute procedural…
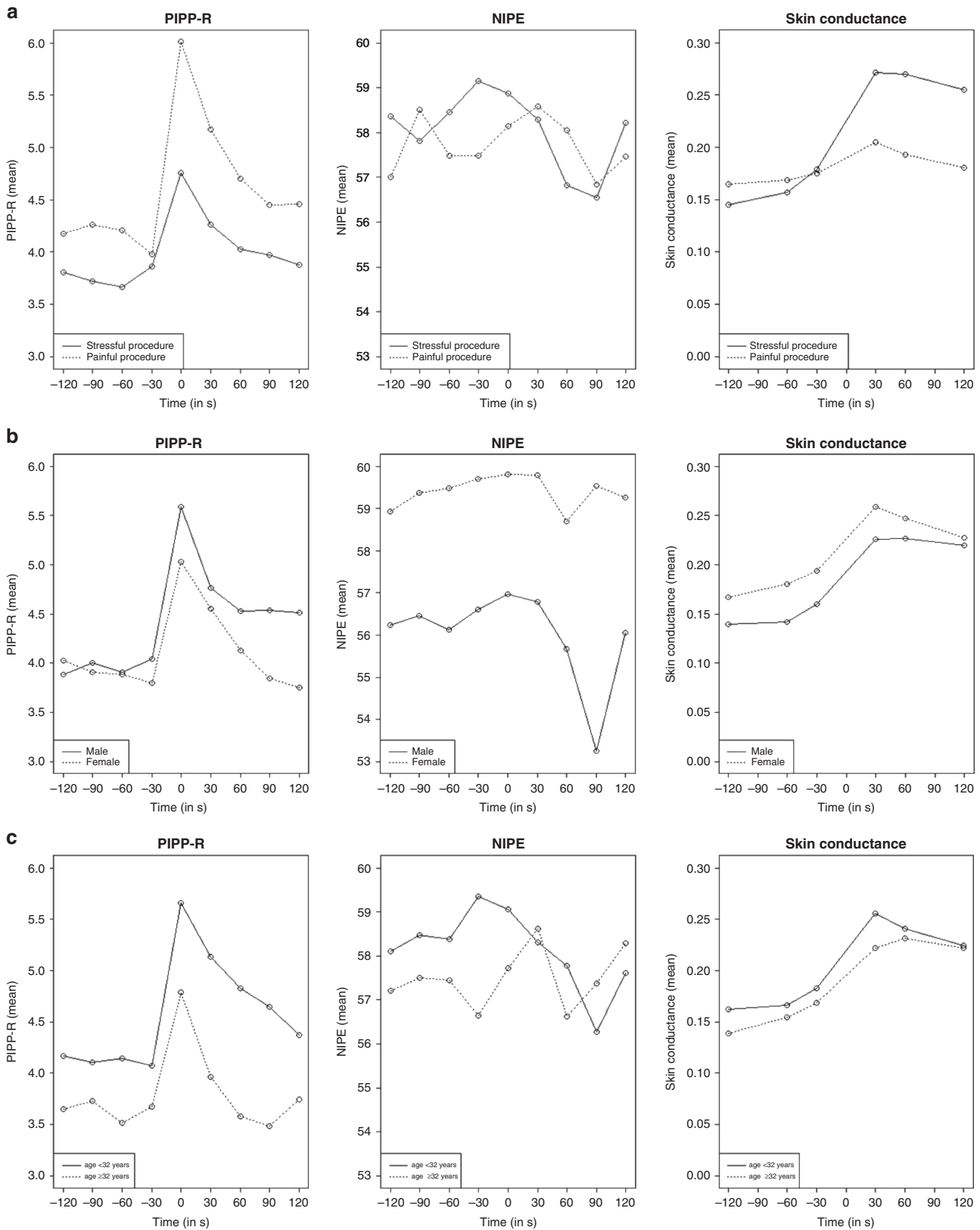J Gendras et al.

1843

**Fig. 2** Patterns of the three parameters (PIPP-R, NIPEi, and skin conductance) over time.

systems to assess the physiological responses to pain and compares these to a validated composite pain scale (PIPP-R). The NIPEi and the PIPP-R scale showed no significant correlations during painful and stressful procedures in preterm infants. The SCR (peaks/s) showed better correlations with the PIPP-R, especially during skin-breaking procedures.

Our findings are consistent with the study published by Cremillieux et al.[15] In a sample of 29 preterm infants, they did

The newborn infant parasympathetic evaluation index for acute procedural...
J Gendras et al.

1844

**Table 2.** Correlation matrix between types of patients, procedures, and responses.

| | All patients | According to GA | | According to gender | | According to the type of procedure | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | | <32 weeks | 32–36 weeks | Male | Female | Painful | Stressful | w/skin breaking | w/o skin breaking |
| **For the first recorded procedure** | | | | | | | | | |
| Baseline before procedure[a] | | | | | | | | | |
| NIPE/PIPP-R (n = 70) | | | | | | | | | |
| PIPP-R/SC (n = 65) | 0.27 [0.02; 0.47] | | | | | **0.63** [0.23; 0.76] | | **0.61** [0.11; 0.80] | |
| NIPE/SC (n = 79) | | | | | 0.44 [0.09; 0.65] | | | | |
| Maximum responses measured[b] | | | | | | | | | |
| NIPE/PIPP-R (n = 71) | | | | | | | | | |
| PIPP-R/SC (n = 69) | 0.28 [0.04; 0.48] | | | | 0.38 [0.04; 0.61] | **0.68** [0.29; 0.78] | | **0.82** [0.30; 0.86] | |
| NIPE/SC (n = 80) | | | | | | | | | |
| Response peaks[c] | | | | | | | | | |
| NIPE/PIPP-R (n = 43) | | | | | | | | | |
| PIPP-R/SC (n = 62) | | **0.56** [0.14; 0.74] | | | | **0.82** [0.40; 0.83] | | **0.98** [0.43; 0.90] | |
| NIPE/SC (n = 51) | | | | | | | | | |
| Delta of responses[d] | | | | | | | | | |
| NIPE/PIPP-R (n = 35) | | | | | | | | | |
| PIPP-R/SC (n = 51) | 0.30 [0.02; 0.53] | **0.77** [0.27; 0.84] | | | | **0.58** [0.13; 0.77] | | **0.92** [0.32; 0.90] | |
| NIPE/SC (n = 48) | | | **0.50** [0.09; 0.71] | | | | | | |
| **For all recorded procedures** | | | | | | | | | |
| Baseline before procedure[a] | | | | | | | | | |
| NIPE/PIPP-R (n = 214) | | | | | | | | | |
| PIPP-R/SC (n = 203) | 0.17 [0.03; 0.30] | | | | | 0.38 [0.16; 0.53] | | 0.38 [0.09; 0.59] | |
| NIPE/SC (n = 219) | | | | | 0.21 [0.02; 0.38] | | | | |
| Maximum responses measured[b] | | | | | | | | | |
| NIPE/PIPP-R (n = 212) | | | | | | | | | |
| PIPP-R/SC (n = 210) | 0.16 [0.02; 0.29] | | 0.23 [0.01; 0.42] | 0.24 [0.03; 0.42] | 0.23 [0.04; 0.39] | 0.31 [0.09; 0.47] | | 0.30 [0.01; 0.53] | 0.19 [0.04; 0.34] |
| NIPE/SC (n = 223) | | | | | | | | | |
| Response peaks[c] | | | | | | | | | |
| NIPE/PIPP-R (n = 131) | | | | | | | | | |
| PIPP-R/SC (n = 189) | 0.15 [0.01; 0.29] | | | | 0.25 [0.05; 0.43] | **0.50** [0.27; 0.62] | | 0.49 [0.18; 0.66] | |
| NIPE/SC (n = 148) | | | | | | | | | |
| Delta of responses[d] (between response peaks and baseline) | | | | | | | | | |
| NIPE/PIPP-R (n = 97) | | | | | | | | | |
| PIPP-R/SC (n = 158) | 0.18 [0.02; 0.32] | | | | 0.34 [0.10; 0.51] | 0.34 [0.09; 0.52] | | | 0.23 [0.04; 0.39] |
| NIPE/SC (n = 129) | 0.23 [0.02; 0.41] | | | | | | | | |

All values: r coefficients and their [95% CI]. Only significant coefficients were described.

[a]NIPEi, PIPP-R, and number of peaks/s (SC, skin conductance) means over the 2-min period before the recorded procedure.

[b]During procedure for PIPP-R, after procedure for NIPEi (over the 2-min period), and mean number of peaks/s (SC) (over the 2-min period).

[c]Peaks of response determined graphically (Fig. 2).

[d]Deltas correspond to the difference between: (1) the value of PIPP-R at the time of the procedure, and the mean PIPP-R score during the 2 min before the procedure (baseline); (2) the value of NIPEi index at 90 s after the procedure, and the value of NIPEi index at the start of the procedure; (3) the mean number of peaks/s (skin conductance) during the 30 s following the procedure, and the mean number of peaks/s during the 2 min before the procedure (baseline).

Bold entries indicate the significant correlations that had coefficients ≥0.50.

The newborn infant parasympathetic evaluation index for acute procedural...
J Gendras et al.

1845

**Table 3.** Diagnostic accuracy of the NIPE Index for all procedures, painful procedures, and skin-breaking procedures.

| Type of procedures | AUC | NIPEi cut-off | Sensitivity | Specificity | PPV | NPV |
|---|---|---|---|---|---|---|
| NIPEi and PIPP-R > 6 | | | | | | |
| All (painful + stressful)[a] | 0.53 | 50 | 0.51 [0.40; 0.62] | 0.62 [0.53; 0.70] | 0.47 [0.36; 0.57] | 0.66 [0.57; 0.74] |
| Painful[b] | 0.56 | 46 | 0.77 [0.57; 0.85] | 0.43 [0.32; 0.62] | 0.56 [0.42; 0.69] | 0.67 [0.47; 0.80] |
| Skin breaking[c] | 0.67 | 46 | 0.86 [0.58; 0.95] | 0.54 [0.37; 0.78] | 0.62 [0.42; 0.81] | 0.81 [0.52; 0.94] |
| NIPEi and PIPP-R > 10 | | | | | | |
| All (painful + stressful) | 0.53 | 54 | 0.67 [0.50; 0.81] | 0.46 [0.39; 0.54] | 0.22 [0.15; 0.31] | 0.86 [0.77; 0.92] |
| Painful | 0.57 | 47 | 0.82 [0.54; 0.89] | 0.45 [0.34; 0.60] | 0.39 [0.24; 0.51] | 0.85 [0.65; 0.92] |
| Skin breaking | 0.67 | 47 | 0.92 [0.46; 0.95] | 0.53 [0.38; 0.74] | 0.44 [0.22; 0.63] | 0.94 [0.64; 0.97] |
| SC (peaks/s) and PIPP-R > 6 | | | | | | |
| All (painful + stressful) | 0.57 | 0.19 | 0.65 [0.54; 0.75] | 0.53 [0.44; 0.62] | 0.48 [0.39; 0.58] | 0.69 [0.58; 0.78] |
| Painful | 0.72 | 0.19 | 0.63 [0.47; 0.77] | 0.82 [0.68; 0.92] | 0.77 [0.60; 0.90] | 0.70 [0.56; 0.82] |
| Skin breaking | 0.78 | 0.08 | 0.91 [0.71; 0.99] | 0.61 [0.39; 0.80] | 0.69 [0.49; 0.85] | 0.88 [0.62; 0.98] |
| SC (peaks/s) and PIPP-R > 10 | | | | | | |
| All (painful + stressful) | 0.60 | 0.12 | 0.89 [0.75; 0.97] | 0.36 [0.29; 0.43] | 0.23 [0.16; 0.31] | 0.94 [0.85; 0.98] |
| Painful | 0.70 | 0.12 | 0.88 [0.69; 0.97] | 0.51 [0.38; 0.64] | 0.42 [0.28; 0.56] | 0.91 [0.77; 0.98] |
| Skin breaking | 0.76 | 0.08 | 1.00 [0.77; 1.00] | 0.52 [0.33; 0.70] | 0.48 [0.29; 0.68] | 1.00 [0.79; 1.00] |

[a] $n = 254$.
[b] $n = 110$.
[c] $n = 55$.

not find any correlation between the NIPEi, the DAN, and the PIPP-R performed during an acute painful procedure, but no sensitivity analyses were performed. However, more recently, Walas et al.,[20] in a sample of 33 infants, reported a significant association between the decrease of the NIPEi and validated pain scales starting from 1 min after noxious stimulation to within 3 min after an acute painful procedure. Their sensitivity analysis for the NIPEi found an area under the ROC curve of 0.79 for the severe pain group. Our secondary analyses reporting a high NPVs and promising area under the ROC curve analyses are consistent with their results and the findings of Chanques et al.[28] in adult patients. The discrepancy between the lack of correlation reported between the NIPEi and pain scales, and the significant results found in the sensitivity analyses needs to be further studied in order to confirm if the NIPEi and SCR could help to identify the highest pain levels in preterm infants.

Our study found the NIPEi to be at its nadir 90 s after the procedure. These results are in line with the findings of Walas et al.,[20] who reported a median nadir of 72 s for the severe pain group. However, the median nadir was, respectively, at 111 and 157 s for the moderate and no/mild pain. When monitoring the NIPEi for pain assessment in preterm infants, longer periods exceeding 90 s could be more appropriate. Faye et al.[14] observed 28 postoperative newborn infants with a mean GA of 37.8 weeks and showed a significant correlation between HRV analysis based on a time analysis of heart rates in the high-frequency domain (high-frequency variability index) and EDIN scores, which reflect prolonged pain.[14]

The low specificity of the NIPEi to acute pain could be partly explained by confounding factors that alter HRV. Postnatal neural maturation may have an influence on HRV responses. The HRV remains lower in preterm newborns compared to term newborns at the same corrected postnatal age, suggesting a delayed maturation of autonomic regulation.[29] Mechanical ventilation could also influence the HRV response as reported by Padhye et al.,[30] who observed decreases in the HRV responses in neonates from 23 to 38 weeks. No patient was mechanically ventilated in our study. A systematic review highlighted the inconsistency of HRV responses for the study of acute pain in infants younger than 1 year and insisted on further research taking into account the possible confounding factors.[31]

The SCR was significantly correlated to PIPP-R scores in preterm newborn infants during painful procedures and especially for those including skin-breaking procedures. These results were consistent with previously published studies. Munsters et al.[13] reported the ability of the GSR measurement (peaks/s) to discriminate sympathetic pain responses as early as 22 weeks of gestational age. This study compared the measurement of SCR to a behavioral pain scale during routine heel lancing in neonates from 22 to 27 weeks of GA. The hemodynamic changes did not influence SCR values, making it an independent tool for pain assessment.[32] One pilot study reported that SCR can be observed in a pain-free state and was correlated to skin temperature, but it included only 11 infants and did not standardize study procedures.[33] Several studies also showed that gestational maturation did not influence the SCR responses in preterm infants,[13,27,34] but these responses were influenced by the postnatal age of term newborns.[35,36]

Strengths of our study include the observation of a wide range of procedures routinely performed in the NICU ($n = 254$), not limited to only heel lances or venipunctures. No correlation was found between NIPEi and PIPP-R scores neither with painful procedures nor with stressful procedures. However, including stressful and painful procedures in the same sample could have underpowered the analysis. Our methodology also ensured the inter-rater reliability of the PIPP-R scoring as showed by the high intraclass correlation coefficient between two independent observers. Despite the early termination of this study, our sample size ($n = 90$) is the largest ever published to date on this specific topic. The calculated sample size was not reached for futility reasons and should not be therefore considered as a limitation. Another limitation is that some observed procedures were still in progress after the end of the recording, which could have interfered with the evolution of some parameters, in particular the NIPEi, whose maximum drop occurred 90 s after the beginning of the procedure. We recommend continued refinement of HRV as a pain assessment method for preterm infants to improve its specificity and positive predictive value.
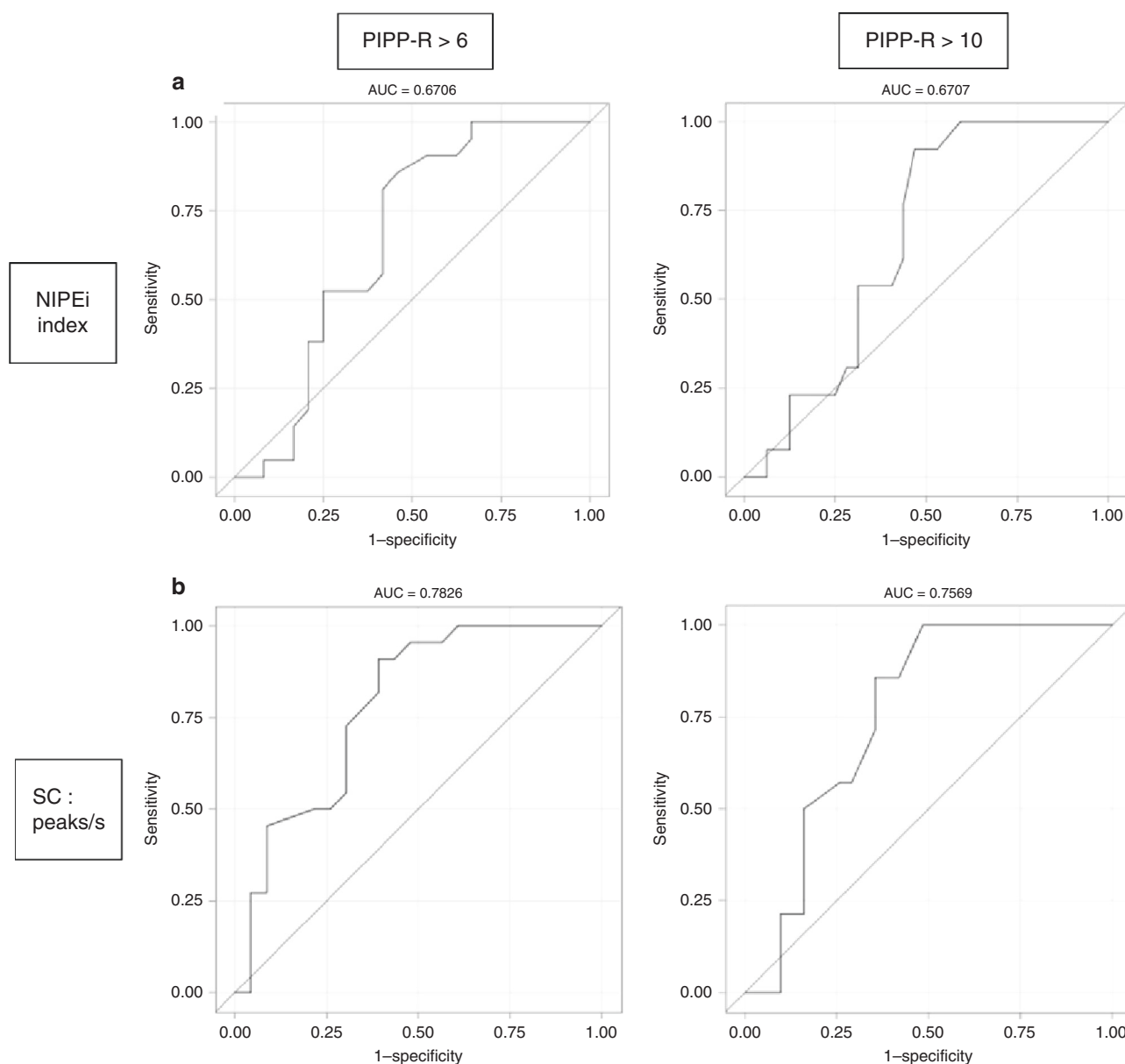
The newborn infant parasympathetic evaluation index for acute procedural...
J Gendras et al.

1846

PIPP-R > 6

PIPP-R > 10

**a**

NIPEi index

AUC = 0.6706

AUC = 0.6707

**b**

SC : peaks/s

AUC = 0.7826

AUC = 0.7569

**Fig. 3** Receiver-operating characteristics (ROC) curves of NIPEi and skin conductance.

No unique method was identified as a gold standard, but SCR seemed to be more reliable than NIPEi to assess pain and stress in preterm infants, probably with a greater accuracy when combined with a validated pain scale. The NIPEi did not show sufficient diagnostic accuracy as a unique method for acute pain assessment, although it may exclude the highest levels of pain as suggested by the high NPVs. In clinical practice, it could be of interest to assess its validity in other neonatal populations with suppressed or absent behavioral responses to pain (e.g., following birth asphyxia, neonatal stroke, severe intraventricular hemorrhage, or use of neuromuscular blockade).

## CONCLUSION
PIPP-R and NIPEi were not correlated during acute painful and stressful procedures in hospitalized preterm infants. SCR (peaks/s) was significantly correlated with PIPP-R scores, especially during painful and skin-breaking procedures. Secondary analysis showed that NIPEi and SCRs could help to exclude severe pain in preterm infants, but this needs to be studied further.

## AUTHOR CONTRIBUTIONS
J.G. contributed to patient enrollment, and collected and analyzed data. She also drafted the manuscript. P.L. contributed to data analysis and drafted the manuscript. I.S.-C. contributed to data analysis and revised the manuscript. M.C. supervised the statistical analysis, interpreted the data, and revised the manuscript. K.J.S.A. interpreted the data and revised the manuscript. J.S. contributed to the study design, the interpretation of the data, and revised the manuscript. J.-M.R. contributed to the study design, supervised the execution of the study, contributed to the data analysis, and drafted and revised the manuscript.

The newborn infant parasympathetic evaluation index for acute procedural...
J Gendras et al.

1847

## REFERENCES

1. Carbajal, R. et al. Epidemiology and treatment of painful procedures in neonates in intensive care units. *JAMA* **300**, 60–70 (2008).
2. Johnston, C., Barrington, K. J., Taddio, A., Carbajal, R. & Filion, F. Pain in Canadian NICUs: have we improved over the past 12 years?. *Clin. J. Pain* **27**, 225–232 (2011).
3. Vinall, J. & Grunau, R. E. Impact of repeated procedural pain-related stress in infants born very preterm. *Pediatr. Res.* **75**, 584–587 (2014).
4. Grunau, R. E. et al. Neonatal pain, parenting stress and interaction, in relation to cognitive and motor development at 8 and 18 months in preterm infants. *Pain* **143**, 138–146 (2009).
5. Ranger, M. et al. Neonatal pain-related stress predicts cortical thickness at age 7 years in children born very preterm. *PLoS ONE* **8**, e76702 (2013).
6. Courtois, E. et al. Epidemiology and neonatal pain management of heelsticks in intensive care units: EPIPPAIN 2, a prospective observational study. *Int. J. Nurs. Stud.* **59**, 79–88 (2016).
7. Slater, R., Cantarella, A., Franck, L., Meek, J. & Fitzgerald, M. How well do clinical pain assessment tools reflect pain in infants? *PLoS Med.* **5**, e129 (2008). https://doi.org/10.1371/journal.pmed.0050129.
8. Gokulu, G. et al. Comparative heel stick study showed that newborn infants who had undergone repeated painful procedures showed increased short-term pain responses. *Acta Paediatr.* **105**, e520–e525 (2016).
9. Treede, R. D., Kenshalo, D. R., Gracely, R. H. & Jones, A. K. The cortical representation of pain. *Pain* **79**, 105–111 (1999).
10. Slater, R. et al. Premature infants display increased noxious-evoked neuronal activity in the brain compared to healthy age-matched term-born infants. *Neuroimage* **52**, 583–589 (2010).
11. Ranger, M., Johnston, C. C. & Anand, K. J. S. Current controversies regarding pain assessment in neonates. *Semin. Perinatol.* **31**, 283–288 (2007).
12. Roué, J.-M. et al. Multi-modal pain assessment: are near-infrared spectroscopy, skin conductance, salivary cortisol, physiologic parameters, and Neonatal Facial Coding System interrelated during venepuncture in healthy, term neonates? *J. Pain Res.* **11**, 2257–2267 (2018).
13. Munsters, J., Wallström, L., Ågren, J., Norsted, T. & Sindelar, R. Skin conductance measurements as pain assessment in newborn infants born at 22–27weeks gestational age at different postnatal age. *Early Hum. Dev.* **88**, 21–26 (2012) .
14. Faye, P. M. et al. Newborn infant pain assessment using heart rate variability analysis. *Clin. J. Pain* **26**, 777–782 (2010).
15. Cremillieux, C., Makhlouf, A., Pichot, V., Trombert, B. & Patural, H. Objective assessment of induced acute pain in neonatology with the Newborn Infant Parasympathetic Evaluation index. *Eur. J. Pain* **22**, 1071–1079 (2018).
16. Zhang, K. et al. Newborn infant parasympathetic evaluation (NIPE) as a predictor of hemodynamic response in children younger than 2 years under general anesthesia: an observational pilot study. *BMC Anesthesiol.* **19**, 98 (2019). https://doi.org/10.1186/s12871-019-0774-y.
17. Lindh, V., Wiklund, U. & Håkansson, S. Heel lancing in term new-born infants: an evaluation of pain by frequency domain analysis of heart rate variability. *Pain* **80**, 143–148 (1999).
18. Lindh, V., Wiklund, U., Sandman, P.-O. & Håkansson, S. Assessment of acute pain in preterm infants by evaluation of facial expression and frequency domain analysis of heart rate variability. *Early Hum. Dev.* **48**, 131–142 (1997) .
19. Lindh, V., Wiklund, U. & Håkansson, S. Assessment of the effect of EMLA® during venipuncture in the newborn by analysis of heart rate variability. *Pain* **86**, 247–254 (2000).
20. Walas, W. et al. Newborn infant parasympathetic evaluation index for the assessment of procedural pain in nonanesthetized infants: a multicenter pilot study. *Am. J. Perinatol.* https://doi.org/10.1055/s-0040-1709458 (2020).
21. Weissman, A., Aranovitch, M., Blazer, S. & Zimmer, E. Z. Heel-lancing in newborns: behavioral and spectral analysis assessment of pain control methods. *Pediatrics* **124**, e921–e926 (2009).
22. Butruille, L. et al. Development of a pain monitoring device focused on newborn infant applications: the NeoDoloris project. *IRBM* **36**, 80–85 (2015).
23. Sabourdin, N. et al. Pain monitoring in anesthetized children: first assessment of skin conductance and analgesia-nociception index at different infusion rates of remifentanil. *Paediatr. Anaesth.* **23**, 149–155 (2013).
24. Gibbins, S. et al. Validation of the Premature Infant Pain Profile-Revised (PIPP-R). *Early Hum. Dev.* **90**, 189–193 (2014).
25. Stevens, B. J. et al. The Premature Infant Pain Profile-Revised (PIPP-R): initial validation and feasibility. *Clin. J. Pain* **30**, 238–243 (2014).
26. Gjerstad, A. C., Wagner, K., Henrichsen, T. & Storm, H. Skin conductance versus the modified COMFORT sedation score as a measure of discomfort in artificially ventilated children. *Pediatrics* **122**, e848–e853 (2008).
27. Storm, H. Skin conductance and the stress response from heel stick in preterm infants. *Arch. Dis. Child Fetal Neonatal Ed.* **83**, F143–F147 (2000).
28. Chanques, G. et al. Analgesia nociception index for the assessment of pain in critically ill patients: a diagnostic accuracy study. *Br. J. Anaesth.* **119**, 812–820 (2017).
29. Oberlander, T. F., Grunau, R. E., Pitfield, S., Whitfield, M. F. & Saul, J. P. The developmental character of cardiac autonomic responses to an acute noxious event in 4- and 8-month-old healthy infants. *Pediatr. Res.* **45**, 519–525 (1999).
30. Padhye, N. S., Williams, A. L., Khattak, A. Z. & Lasky, R. E. Heart rate variability in response to pain stimulus in VLBW infants followed longitudinally during NICU stay. *Dev. Psychobiol.* **51**, 638–649 (2009).
31. Waxman, J. A., Pillai Riddell, R. R., Tablon, P., Schmidt, L. A. & Pinhasov, A. Development of cardiovascular indices of acute pain responding in infants: a systematic review. *Pain Res. Manag.* **2016**, 1–15 (2016).
32. Storm, H. The development of a software program for analyzing skin conductance changes in preterm infants. *Clin. Neurophysiol.* **112**, 1562–1568 (2001).
33. Valkenburg, A. J., Niehof, S. P., van Dijk, M., Verhaar, E. J. M. & Tibboel, D. Skin conductance peaks could result from changes in vital parameters unrelated to pain. *Pediatr. Res.* **71**, 375–379 (2012).
34. Pereira-da-Silva, L. et al. Skin conductance indices discriminate nociceptive responses to acute stimuli from different heel prick procedures in infants. *J. Matern. Fetal Neonatal Med.* **25**, 796–801 (2012).
35. Harrison, D. et al. Skin conductance as a measure of pain and stress in hospitalised infants. *Early Hum. Dev.* **82**, 603–608 (2006).
36. Hernes, K. G. et al. Skin conductance changes during the first year of life in full-term infants. *Pediatr. Res.* **52**, 837–843 (2002).