The validity of skin conductance for assessing acute pain in mechanically ventilated infants: A cross-sectional observational study

Abstract

Background: Assessing pain in mechanically ventilated infants is challenging. The assessment of skin conductance (SC) is based on the sympathetic nervous system response to stress. This study purpose was to evaluate the validity of SC for assessing pain in mechanically ventilated infants.

Methods: A prospective cross-sectional observational design was used to study SC and its relation to: the category of procedure (i.e., painful or non-painful); the phase of procedure (i.e., before, during and after), and referent pain measurements (i.e., Premature Infant Pain Profile-Revised (PIPP-R) and Neonatal Facial Coding System (NFCS)). Eligible infants were those up to 12 months of age, in intensive care units, who were mechanically ventilated, and required painful and non-painful procedures.

Results: From October 2017 to November 2018, 130 eligible infants were identified, and 55 infants were studied. SC (number of waves per second) during painful procedures (median 0.27, interquartile range 0.2-0.4) was statistically significantly higher than those during non-painful procedures (0, 0-0.09). SC during painful procedures was statistically significantly higher than those before (0, 0-0.07) and after painful procedures (0, 0-0.07). SC showed moderate statistically significant positive correlations with PIPP-R (Spearman’s rho=0.4-0.62) and the four-item NFCS (Spearman’s rho=0.31-0.67) before, during and after painful or non-painful procedures respectively. SC had excellent performance (area under the receiver operator curve=0.979) with excellent sensitivity (92.31%), specificity (95.42%) and negative predictive value (99.21%) but only sufficient positive predictive value (66.67%) when used to discriminate moderate-to-severe pain.

Conclusions: SC showed good validity for assessing pain in critically ill infants requiring mechanical ventilation.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ejp.1816

This article is protected by copyright. All rights reserved
Keywords: Infant, Pain assessment, Skin conductance, Observational research design, Validation
The validity of skin conductance for assessing acute pain in mechanically ventilated infants:  
A cross-sectional observational study

Running head: Skin conductance for assessing pain in sick infants

Jiale Hu, PhD, RN1,2,*, JoAnn Harrold, MD, FRCPC 3,4, Janet E. Squires, RN, PhD2,5, Shokoufeh Modanloo, RN, MScN2, and Denise Harrison, RN, PhD2,6

1 Department of Nurse Anesthesia, College of Health Professions, Virginia Commonwealth University, Richmond, USA
2 School of Nursing, Faculty of Health Sciences, University of Ottawa, Ottawa, Canada
3 Faculty of Medicine, University of Ottawa, Ottawa, Canada
4 Neonatal Intensive Care Unit, Children’s Hospital of Eastern Ontario, Ottawa, Canada
5 Research Institute, Ottawa Hospital, Ottawa, Canada
6 Department of Nursing, Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Melbourne, Australia

* Corresponding Author
Jiale Hu, PhD, RN
Assistant Professor | Director of Research and Global Outreach
Department of Nurse Anesthesia | College of Health Professions
Virginia Commonwealth University
VCU Email: jhu4@vcu.edu

The manuscript is being submitted to the category of original article.

Conflict of interest statement

The authors have no conflicts of interest to declare.

This article is protected by copyright. All rights reserved
Funding

This study was funded by the Canadian Pain Society (CPS) Trainee Research Award (Clinical Science, 2018-2019) and Sigma Theta Tau International/Rosemary Berkel Crisp Research Award (2018-2019). The first author (J. Hu) was supported by the Ontario Trillium Scholarship (2015-2019) and International Doctoral Scholarship.

Significance

Pain assessment in mechanically ventilated infants is challenging. In this study, the validity of skin conductance (SC) for pain assessment is evaluated in the same population of infants during painful and non-painful procedures. SC showed good validity for assessing acute pain in relation to category of procedure, phase of procedure, and referent pain measurements. SC is a promising method, especially with other pain assessment methods and other determinants of pain, in a multi-modal pain assessment approach to understand the complexity of pain in mechanically ventilated infants.

1. Introduction

Despite significant advances in the knowledge of pain in infants, poorly managed pain in this population remains an issue across different settings, especially for sick hospitalized infants (McGrath, Stevens, Walker, & Zempsky, 2014). Infants requiring mechanical ventilation are one of the most vulnerable populations, due to the critical nature of the illness, invasiveness of the initial intubation, and numerous procedures required during ongoing medical and nursing care (Hall, Boyle, & Young, 2007).

Assessing pain in a valid and reliable manner is important for monitoring the effectiveness of pain management in infants, including determining appropriate analgesia dosing (Witt, Coynor, Edwards, & Bradshaw, 2016). Although tremendous advances have been made in understanding pain responses, and more than four dozen pain assessment tools have been developed for infants, assessing pain in infants is still complex and challenging for researchers and care providers (Hatfield & Ely, 2015; van Dijk & Tibboel, 2012). One of the issues in infant pain assessment is
the lack of a “gold standard” for assessing pain, as infants are unable to self-report pain (Pillai Riddell et al., 2016). Behavioral, physiological and multidimensional pain assessment methods are considered as surrogate approaches for self-report (Cong, McGrath, Cusson, & Zhang, 2013; Lee & Stevens, 2014; Stevens, Riddell, Oberlander, & Gibbins, 2007; Stinson, 2009). However, using behavioral indicators or vital signs in mechanically ventilated infants, especially those who are critically ill and receiving neuromuscular blockers or deep sedation, is challenging (Herr, Coyne, McCaffery, Manworren, & Merkel, 2011).

The measurement of skin conductance (SC) is a physiological approach to assessing pain, which is based on the sympathetic nervous system response to stress (Gladman & Chiswick, 1990). When pain occurs, the stress induces sympathetic excitation and leads to emotional sweating at an infant’s plantar (foot) or palmar (hand) regions (Lader & Montagu, 1962). The SC device monitors the increasing and decreasing electrodermal activity, records SC activity peaks and troughs, and calculates SC values (Med-Storm Innovation AS, 2012). Using the SC device is quick and straightforward and requires a short training time, and data are objective. Additionally, the scores of SC can be easily and quickly understood by clinicians and researchers. The number of SC waves per second (NWps) was recommended by many previous studies and the user manual of SC device for pain assessment (Hu, Modanloo, Squires, Harrold, & Harrison, 2019; Med-Storm Innovation AS, 2012).

A recent scoping review, which included twenty-eight studies with 1061 infants, synthesized the validity of SC for assessing acute pain in hospitalized infants (Hu et al., 2019). The results showed that SC might be a promising pain assessment method in infants. However, inconsistent findings across the studies existed. Moreover, most included studies focused on healthy newborns or medically stable hospitalized infants, who were not receiving mechanical ventilation. Only one study reported SC for pain assessment in mechanically ventilated infants (Karpe, Misiolek, Daszkiewicz, & Misiolek, 2013). Karpe et al. (2013) showed that SC was significantly higher during airway suctioning or blood sampling than at rest. However, the study did not compare SC with any referent pain measurements and did not compare SC values during
painful procedures with those during non-painful procedures. Without these comparisons, it is difficult to tell whether the higher SC values during painful procedures compared to baseline and recovery were induced by pain. In addition, the methods regarding the time window of calculating NWps, the method (i.e., maximum, minimum, or average value) and the observation time interval of sampling SC values were reported in insufficient detail to enable replication of the study (Karpe et al., 2013). Therefore, further research transparently reporting SC collection methods and evaluating multiple validity evidence of SC is warranted.

This study purpose was to evaluate the validity of SC for acute pain assessment in mechanically ventilated infants in relation to 1) the category of procedure (i.e., painful and non-painful); 2) the phase of procedure (i.e., before, during and after), and 3) referent pain measurements, which are widely used in research and clinical practice.

2. Methods

2.1 Research design

A prospective cross-sectional observational design was used to study SC and its relation to referent pain measurements in mechanically ventilated infants before, during and after both a painful and a non-painful procedure. Messick’s unitary view of validity was used in this study to examine the validity of SC for acute pain assessment in mechanically ventilated infants (Messick, 1995). The view considers validity as a unitary concept with all validity evidence contributing to construct validity. Thus, validity evidence in relation to multiple variables (i.e., the category of procedure, the phase of procedure, and referent pain measurements) was tested in this study.

2.2 Setting and sample

The setting was a level-three neonatal intensive care unit (NICU) within a tertiary referral pediatric hospital in Canada. In order to maximize sampling, a level-three pediatric intensive care unit (PICU) within the same hospital was added as an additional clinical site after ten months of data collection. Convenience sampling with a consecutive design was used to recruit infants in these two units. Eligibility criteria were infants up to 12 months of age, who were being mechanically ventilated, and who required a painful procedure and non-painful care
Infants were excluded if they had conditions which would have potentially affected SC responses or precluded SC measurement, including being less than corrected gestational age of 25 weeks, and any injury to the skin where the SC probes would be placed. Also, infants with spinal cord malformation (e.g., myelomeningocele and sacral teratoma) were excluded, as these infants’ behavioral pain responses may be different from infants without spinal cord malformations. Finally, as high doses of cholinergic activating/blocking drugs (e.g., neostigmine, glycopyrrolate, and atropine) may influence SC responses (Ledowski, Preuss, & Schug, 2009; Storm, 2008), infants receiving these medications were excluded until a period of 8 hours after these medications were administered.

The two categories of procedure (i.e., painful or non-painful) were defined based on a previous study (Ahn & Jun, 2007). Painful procedures were defined as intrusive interventions that involve skin breaking procedures, such as heel lance or venipuncture, or non-tissue breaking painful procedures, such as naso- or oro-gastric tube insertion (Ahn & Jun, 2007). The non-painful procedures were routine care requiring direct contact between the caregivers and the infants without any invasiveness, such as diaper change or positioning (Ahn & Jun, 2007).

The observation of painful or non-painful procedures was separated into the following three phases: 1) before procedure: the first 30 seconds in the three minutes prior to the painful procedure (i.e., breaking the skin or tube insertion) or non-painful procedure, 2) during procedure: the first 30 seconds following the painful or non-painful procedure, 3) after procedure: the last 30 seconds in the three minutes after completion of the painful or non-painful procedure. The separation of observation periods 1 and 3 from the procedure ensured the included infants were less likely to be handled for preparation before the procedure or settling after the procedure during these periods of observation. Thus, the data collected during these time windows would be more likely to represent the infant’s conditions during baseline and recovery.

2.3 Demographic and medical information

Demographic data and medical information were collected from the patients’ charts.
Demographic data included gestational age (weeks + days), postnatal age (weeks + days), sex, and current weight. The medical information included number of days in NICU or PICU, diagnosis, type of most recent surgery (if applicable), number of days after most recent surgery (if applicable), mechanical ventilation status, and number of days on ventilation. In addition, medication information relating to pain or comfort management administered in the 6 hours period prior to observation was also collected, including analgesics, sedatives, muscle relaxants or anticonvulsants and their route of administration and dosage.

The Neonatal Therapeutic Intervention Scoring System (NTISS) was used as a proxy severity of illness score (Gray, Richardson, McCormick, Workman-Daniels, & Goldmann, 1992; Harrison, 2001). The NTISS is a therapy scoring system consisting of a total of 63 items (Gray et al., 1992). Scores from one to four are assigned to items based on their therapeutic intensity and complexity, with individual items further grouped into eight categories (Canadian Association of Research Ethics Boards, 2010). In this study, the NTISS was scored by the first author (J. Hu) based on therapeutic interventions 24 hours prior to the data collection, which was recorded in the patients’ charts. In the final calculation of scores for NTISS, points were assigned only for the most intense therapeutic intervention (Canadian Association of Research Ethics Boards, 2010). For example, a patient who began a scoring period with a single chest tube in place and subsequently had multiple chest tubes in place would receive final points for multiple chest tubes in place.

2.4 SC measurements

The apparatus and software program used to measure SC responses in this study were commercially developed by Med-Storm Innovations (Med-Storm Innovation AS, 2012). The application of the apparatus and software program in this study followed the product manual (Med-Storm Innovation AS, 2012). The three-electrode system of the apparatus comprises a measuring electrode (black) placed on the sole of the foot, a counter-current electrode (yellow) placed on either side of the ankle, and a reference voltage electrode (blue) placed on the other side of the ankle.
2.5 Referent pain measurements

There is no gold standard for pain assessment in infants (Hatfield & Ely, 2015; Melo, Lélis, Moura, Cardoso, & Silva, 2014). Thus, pain in infants can only be imperfectly measured. Different referent pain measurements could provide different aspects of pain experience in infants and allow researchers to understand infant pain more comprehensively (Pillai Riddell et al., 2016). Thus, two different referent pain measurements were used in this study, including the Premature Infant Pain Profile-Revised (PIPP-R) (Stevens et al., 2014) and the four-item subset of Neonatal Facial Coding System (NFCS) (Grunau, Johnston, & Craig, 1990).

The PIPP is a 7-item multidimensional measurement of pain that has been widely used to measure acute pain in infants (Stevens, Johnston, Petryshen, & Taddio, 1996). In 2014, on the basis of a comprehensive review of the validity of PIPP for measurement of pain in infants, and feedback from researchers and clinicians, the PIPP was revised to PIPP-R to address validity and feasibility concerns in gestational age groups and behavioral states (Gibbins et al., 2014; Stevens et al., 2014). The PIPP-R has the same indicators but a different scoring system from the original version (Stevens et al., 2014). The PIPP-R includes three behavioral responses (facial actions: brow bulge, eye squeeze, nasolabial furrow), two physiological responses (heart rate and oxygen saturation), and two contextual items (gestational age and behavioral state). Each item is numerically scaled and scored on a 4-point scale (0 to 3 points) reflecting changes in each variable from baseline values. The scores obtained for the 7 items are summed for a total pain intensity score, ranging from 0 to 21.

The NFCS is a unidimensional single-domain measurement, which consists of 10 facial expressions; however, a cluster of four facial actions comprised of brow bulge, eye squeeze, naso-labial furrow, and open lips was present in more than 90% of the infants exposed to a painful stimulus (Grunau & Craig, 1987; Grunau et al., 1990). Each facial expression is coded as either 1 or 0 (present or absent), leading to a total score of 0 to 4. Numerous studies suggested using the most frequent facial actions, as the original ten-item NFCS is time-consuming for clinicians (Grunau & Craig, 1987; Grunau et al., 1990; Grunau, Oberlander, Holsti, & Whitfield,
1998; Pereira et al., 1999; Peters et al., 2003). In addition, this four-item subset of the NFCS produces reliable and valid scores for pain measurement in infants, and it has also been illustrated to be feasible for use at the bedside (Grunau et al., 1998; Harrison, Johnston, & Loughnan, 2003).

Both the PIPP-R and the NFCS were used as referent standards in this study. Although brow bulge, eye squeeze, and naso-labial furrow have been included in PIPP-R, the indicator “open lips” in the four-item NFCS is also one of the most frequently occurring facial expression during painful procedures in infants, and the scoring standards of the indicators are different. The PIPP-R is focused on the duration of facial indicators, and the four-item NFCS pays attention to whether the facial indicators are present or not (Grunau & Craig, 1987; Stevens et al., 2014). In addition, the PIPP-R is a multidimensional pain measurement, and the four-item NFCS is a unidimensional single-domain measurement. Using both referent standards could inform a full understanding of the validity of SC as a strictly physiological outcome in relation to different types of pain assessment methods.

2.6 Data collection procedure

The SC apparatus and video camera were set up by author J. Hu or S. M., both of whom were trained in the use of the equipment. The equipment was set up at least 10 minutes prior to the painful procedure to ensure the devices were prepared before the data collection and the infant’s behavioral state had time to return to baseline after handling due to the application of the SC electrodes. The infants’ baseline behavioral state, highest heart rate, and lowest oxygen saturation were recorded 15 seconds prior to data collection (i.e., three minutes before a procedure), to collect the baseline data for the PIPP-R (Stevens et al., 2014). The SC and video recording commenced three minutes prior to the painful or non-painful procedure and continued for three minutes after completion of the procedure.

There was no attempt to standardize painful or non-painful procedures, but details of the procedures and any medication change, patient care (e.g., handling or touching), or unexpected painful procedures (e.g., an urgent airway suctioning during one of the observed procedures)
during data collection were recorded. The length of time between completion of data collection during a painful procedure and commencement of a non-painful procedure was at least 10 minutes, or until the infant had stabilized following the procedure. The 10-minute time length was determined based on the previous scoping review, which identified ten minutes as the most commonly used time length for SC stabilization (Hu et al., 2019). In addition, the 10-minute time length was feasible for health care providers to cluster the procedures.

2.7 Video coding

Video coding for the PIPP-R was done using 30 second observation time intervals (Stevens et al., 2014). Previously published time periods for observing and coding facial expressions using the four-item NFCS varied from 2 seconds to 1 minute (Chen et al., 2012; Chimello, Gaspardo, Cugler, Martinez, & Linhares, 2009; Grunau & Craig, 1987). In this study, coding the four-item NFCS via the video recording used the same observation interval as the PIPP-R, of 30 seconds.

The video recordings were divided into the 30-second segments for the three observation phases, and the sequencing of the segments was randomized. All videos were coded by two independent researchers (S.M. and J.C.), who were blinded to the category and phase of the procedure. These two researchers had undergone training on video coding the PIPP-R and the four-item NFCS. Training tools and education included a coding handout, education sessions, and training videos. As this study aimed to obtain the most accurate results of referent standards, any discrepancies in the results of video coding were resolved through a consensus process or by consulting a third senior researcher (D.H.), and thus the interrater reliability scores were not calculated.

2.8 SC Value calculation

The NWps was used as the parameter of SC measurement in this study. The time window for calculating the NWps in this study was 15 seconds, which is demonstrated as an appropriate time frame by previous studies and recommended in the Med-Storm SC product manual (Ledowski, Albus, Stein, & Macdonald, 2011; Med-Storm Innovation AS, 2012). The refreshing time was one second, which means a new value of NWps was obtained every second.
The maximal value of NWps, which was measured in a 30 second time period, was extracted. It was the same 30 seconds observation time interval as the two referent pain scales, before, during, and following completion of the procedure. Using the maximal value of NWps was compatible with the scoring systems of the PIPP-R and four-item NFCS, which capture the maximal intensity of facial expression, highest heart rate, and lowest oxygen saturation (Grunau & Craig, 1987; Stevens et al., 2014).

2.9 Data analysis

Statistical analyses were performed using the IBM SPSS for Windows Version 25 statistical package (IBM Inc, Chicago, IL, USA). Continuous data such as gestational and postnatal age, pain scores of PIPP-R or the four-item NFCS, and the NWps, were described by using means, standard deviations (SD) and range if normally distributed, and medians and interquartile range (IQR) if non-normally distributed. Categorical variables such as sex, category of procedure, and medications, were summarized using frequencies and percentages. The normal distribution of the data was evaluated using the Kolmogorov-Smirnov Test.

2.9.1 Primary outcomes

There were three primary data analyses in this study to evaluate the validity of SC for acute pain assessment in mechanically ventilated infants. Parametric or nonparametric tests were used, depending on the distribution of SC data. To evaluate the validity of SC for pain assessment in relation to the category of procedure, Paired T-test or Wilcoxon signed ranks test was used to test the hypothesis: Infants being mechanically ventilated have statistically significantly different NWps during painful procedures compared to non-painful procedures. To evaluate the validity of SC in relation to the different phases of procedure, Repeated measures ANOVA (RM-ANOVA) with post hoc test or Friedman test with Dunn's multiple comparison test was used to test two hypotheses: Infants being mechanically ventilated have statistically significantly different NWps during painful procedures compared to before or after painful procedures; There is no difference in NWps in infants being mechanically ventilated between during non-painful procedures and before or after non-painful procedures. To evaluate the validity of SC in relation to referent pain
measurements, Pearson’s correlation or Spearman’s correlation was analyzed to test the hypothesis: The NWps of infants being mechanically ventilated has statistically significant correlations with PIPP-R and the four-item NFCS before, during and after procedures. For inferential statistics, the probability of type I error (\(\alpha\)) was set at 0.05, and two-tailed tests were used. The variables of missing data were deleted pairwise in data analysis.

2.9.2 Secondary outcomes

The relationships between SC and demographic and medical data were also analyzed. Spearman’s correlation was used to evaluate the association of SC with gestational age, postnatal age, current weight, number of days in NICU or PICU, number of days on ventilation, NTISS scores, and time length between painful and non-painful procedures. Kruskal-Wallis test was used to evaluate the association of SC with diagnosis, mechanical ventilation status, and type of painful and non-painful procedures. Mann-Whitney U test was used to evaluate the association of SC with sex, administration of oral sucrose prior to painful procedures, and use of analgesics or sedatives administered within the 6 hours before observation.

If the findings showed the validity of SC in relation to PIPP-R in this study, diagnostic test accuracy (Buntinx & Knottnerus, 2009) was analyzed to provide further validation of SC: how well does SC identify the presence or absence of moderate-to-severe pain in mechanically ventilated infants? The study definition of moderate-to-severe pain was if PIPP-R was greater than or equal to 13 (Stevens et al., 2014; Stevens et al., 1996). Receiver operator curve (ROC) analysis that described a pair of sensitivity and specificity values for every individual cut-off of SC was conducted (Linnet, Bossuyt, Moons, & Reitsma, 2012). The area under the ROC was a global measurement of diagnostic accuracy and was used to estimate the degree of discriminative power of SC for pain assessment in mechanically ventilated infants (Linnet et al., 2012).

The highest Youden’s index (sensitivity+specificity-1) was used to determine the optimal cut-off value of SC (Hajian-Tilaki, 2018). The measurements of diagnostic test accuracy relating to this optimal cut-off value of SC were described, including sensitivity and specificity, positive and negative predictive values, and likelihood ratio for positive test results and negative test.
results (Mallett, Halligan, Thompson, Collins, & Altman, 2012). The area under the ROC, sensitivity, specificity, and positive and negative predictive values were evaluated as excellent (value ≥90%), very good (≥80%), good (≥70%), sufficient (≥60%) and poor (<60%) (Šimundić, 2009). The likelihood ratio for positive test results more than 10 provides a convincing ruling-in discriminative accuracy, and likelihood ratio for negative test results less than 0.1 provides a convincing ruling-out discriminative accuracy (Šimundić, 2009).

2.10 Sample size determination

The sample size was determined to ensure that the three primary data analyses had sufficient statistical power. Fixing the probability of type I error at 5%, the sample size, 44 infants, was necessary to provide 80% power to detect a different NWps during different categories of procedure (i.e., painful and non-painful), 50 infants for identifying a different NWps in the different phases of painful or non-painful procedure (i.e., before, during and after), and 29 infants for examining a relationship between NWps and the referent pain measurements (i.e., the PIPP-R and the four-item NFCS). In this study, the largest sample size (i.e., 50 infants) was used and increased to 55 infants to allow for a 10% loss of data due to technical errors in recording or other attrition reasons.

2.11 Ethical considerations

This study was approved by the affiliated hospital Research Ethics Board (17/85X) and affiliated university Research Ethics Board (A08-17-06). Potentially eligible infants were identified by unit staff in the circle of care. Parents/guardians who were interested in the study were then referred to the researchers. An explanation of the study, along with an information letter and consent form and a research media records release form, was offered by the researchers. Informed consent was voluntary and free from coercion. Parents/guardians were free to withdraw their infants from the study at any time upon their request. If an infant was withdrawn from the study, parents/guardians could choose to allow the research team to have their infant’s data already collected in the study kept for analysis or have the researchers destroy all the data.

3. Results
3.1 Sample characteristics

From October 2017 to November 2018, 130 eligible infants were identified, parents of 92 eligible infants were approached, as most parents were not available before extubating (Figure 1). Finally, 78 consents were obtained, and data for 55 infants were collected, as 23 infants had no painful procedure ordered before extubating or transfer (Figure 1). Three infants were recruited from the PICU, the other 52 infants were from the NICU, and no infant was withdrawn from the study (Figure 1). There were 25 females (45%) and 30 males (55%). The infants had a median gestational age of 35 weeks 5 days (IQR=28 weeks to 38 weeks 1 day), and the median postnatal age at study was 9 days (IQR=3 days to 39 days). The median weight was 2755 grams (IQR=1880 to 3650 grams). The infants had a median of 4 days (IQR=2 to 9 days) in the NICU or PICU. The most common diagnosis was a surgical condition (n=18, 33%) and 17 of these infants underwent surgery before observation during the study period (Table 1). The mean number of days after the most recent surgery at the time of study was 5 (SD=4.62, range=1 to 18). The mean NTISS score was 15.73 (SD=3.27, range=10 to 26). To illustrate, this mean NTISS score represented a typical infant in the study who was mechanically ventilated and administered 1 to 3 antibiotic agents; had monitoring of vital signs, cardiorespiratory, noninvasive oxygen saturation and intake and output; received gavage feeding, intravenous fat emulsion, amino acid solution, insulin, and potassium, and had phototherapy, urinary catheterization and peripheral intravenous lines in the previous 24 hours. The infants had a median of 3 days (IQR=2 to 5 days) on mechanical ventilation at the time of study. Most infants (n=29, 53%) were receiving mechanical ventilation using pressure control ventilation mode (Table 1).

3.2 Medications and procedures

Most infants (n=49, 89%) received 1 to 4 analgesics or sedative medications in the 6 hours prior to observation (Table 2). Heel lance for blood sampling (n=25, 45%) was the most commonly observed painful procedure followed by an arterial puncture for blood sampling (n=11, 20%) and venipuncture for blood sampling (n=8, 15%) (Table 3). Eight infants had more than one attempt of painful procedures, and one of these infants received two types of painful...
procedures: heel lance for blood sampling and venous blood sampling (Table 3). In terms of non-painful procedures, every observed infant had a diaper change, but 21 of the 55 infants (38%) had diaper changed during the same study period as abdominal circumference measurement, temperature measurement, or lung auscultation (Table 3). The median time between painful and non-painful procedures was 11 minutes 20 seconds (IQR=10 minutes 28 seconds to 13 minutes 7 seconds).

3.3 Primary outcomes

Three infants had missing facial score data before, during, and after painful and non-painful procedures. For these three infants, facial expressions were unable to be scored from the video footage due to the face being partially covered by tapes or tubes (Figure 1). Due to artifacts in the SC recording, two infants had missing data of NWps before, during and after painful and non-painful procedures, two infants had missing data only before, during and after painful procedures, and three infants had missing data only before, during and after non-painful procedures (Figure 1). The median NWps was 0.27 (IQR=0.2 to 0.4) during painful procedures and 0 (IQR=0 to 0.09) during non-painful procedures. The NWps was statistically significantly higher during painful procedures compared to non-painful procedures (n=48, Z=5.71, P<0.001, effect size=0.82). The median NWps was 0 (IQR=0 to 0.07) both before and after painful procedures, 0 (IQR=0 to 0.02) before non-painful procedures, and 0 (IQR=0 to 0.07) after non-painful procedures. There was a statistically significant different NWps in the different procedure phase for painful (n=51, Chi-Square=69.90, P<0.001, effect size=0.69) and non-painful procedure (n=50, Chi-Square=19.93, P<0.001, effect size=0.20). Dunn's multiple comparison test, using Bonferroni correction, showed infants had statistically significantly different NWps according to the phase of painful procedures; with during painful procedures being higher than at baseline (n=51, Z=5.81, P<0.001, effect size=0.83) and in the recovery phase (n=51, Z=5.86, P<0.001, effect size=0.82). However, there was also statistically significantly higher NWps during non-painful procedures compared to baseline (n=50, Z 2.95, P=0.003, effect size=0.42) and recovery (n=50, Z 2.81, P=0.005, effect size=0.4). Before, during
and after painful or non-painful procedures, NWps showed statistically significant positive correlations with PIPP-R (Spearman’s rho=0.4-0.62) and the four-item NFCS (Spearman’s rho=0.31-0.67) separately (Table 4).

3.4 Secondary outcomes

No statistically significant relationships between SC and demographic or medical data were found. The area under ROC was 0.979 (95% confidence interval=0.962 to 0.996) reaching an excellent level (Figure 2). The cut-off point of 0.235 of NWps was identified having the highest Youden’s index (0.88) with excellent sensitivity (92.31%), specificity (95.42%) and negative predictive value (99.21%) but only sufficient positive predictive value (66.67%) to recognize infants who scored moderate-to-severe pain on the PIPP-R. It also resulted in a 20.15 likelihood ratio for positive test results and a 0.08 likelihood ratio for negative test results, providing convincing ruling-in discriminative accuracy and ruling-out discriminative accuracy.

4. Discussion

4.1. Main findings

To our best knowledge, 34 previous studies focusing on the use of SC in infants have been published, including 28 studies synthesized in the recent scoping review and six studies published after the scoping review (Hu et al., 2019; Maillard, Garnier, Saliba, & Favrais, 2019; Meesters et al., 2019; Oji-Mmuo et al., 2019; Passariello et al., 2019; Pettersson, Olsson, Ohlin, & Eriksson, 2019; Roue et al., 2018). This present study is the first to evaluate different sources of validity evidence of SC for pain assessment in the same population of infants during the same time period. Specifically, SC showed good validity for pain assessment in mechanically ventilated infants in relation to the category of procedure (i.e., painful and non-painful), the phase of painful and non-painful procedures (i.e., before, during and after) and referent pain measurements (i.e., PIPP-R and four-item NFCS).

The most common validation approach in previous studies was analyzing the SC changes before, during and after painful procedures, with the aim of demonstrating whether SC responded immediately to the painful stimuli and recovered when the painful stimuli ceased.
Results in this study showed a statistically significant increase in SC during the painful procedures and a statistically significant decrease after painful procedures. The finding was in agreement with 15 of the previous 16 studies evaluating the response of SC to painful procedure in infants (Hu et al., 2019; Passariello et al., 2019) and showed good validity in relation to the three different phases of painful procedure with respect to pain assessment in mechanically ventilated infants.

In terms of non-painful procedures, there were also significant differences between different phases. The phenomena of SC responding to both painful and non-painful procedures accords with the biological mechanism of SC, which is based on how the sympathetic nervous system responds to general stress rather than pain specifically. However, the effect sizes of the SC changes before, during, and after non-painful procedures were small. Additionally, the median values of NWps before, during, and after non-painful procedures were all zero, although the data distributions and IQRs were different. The IQRs of NWps during non-painful procedures were also smaller than the SC cut-off point of moderate-to-severe pain identified in this study. These results provided evidence that the SC responses to non-painful procedures were minimal and not clinically significant. This evidence could help to illustrate that the SC values during non-painful procedures indicated the presence of a general stress response rather than pain.

Results also showed a statistically significantly higher NWps with a large effect size in infants undergoing painful procedures compared to non-painful procedures. This finding was in agreement with one previous study comparing SC during different categories of procedures (painful vs. non-painful) in healthy full-term infants (Eriksson, Storm, Fremming, & Schollin, 2008). This finding demonstrated that SC could distinguish painful from non-painful events in mechanically ventilated infants.

Correlations of SC with PIPP-R and the four-item NFCS indicated that the SC values had good validity for pain assessment in mechanically ventilated infants in relation to referent pain measurements. However, the magnitudes of the correlations were not so high to suggest that SC and the referent pain measurements are evaluating the same construct. Additionally, most of
these correlations were weaker than the correlations between PIPP-R and the four-item NFCS. These findings could be expected as SC and these two referent measurements are not evaluating the same dimension of pain responses in mechanically ventilated infants. SC is based on the sympathetic nervous system’s response to stimuli; facial expressions in PIPP-R and the four-item NFCS are behavioral reactions, specifically, the reflexive reactions to nerve impulse transmitted to the brainstem; and the two physiological indicators in PIPP-R, changes in heart rate or oxygen saturation are induced by both the sympathetic and parasympathetic nervous system (Cowen, Stasiowska, Laycock, & Bantel, 2015; Fitzgerald, 2015).

The low likelihood ratio for negative test results and the excellent specificity and negative predictive value indicated that SC was able to detect an infant without moderate-to-severe pain based on the negative SC result (i.e., NWps<0.235). The high likelihood ratio for positive test results and excellent sensitivity indicated that SC was able to detect an infant with moderate-to-severe pain based on the positive SC result. However, the positive predictive value did not reach an excellent level. The imperfect positive predictive value might illustrate that one-third of infants, who had positive SC results, did not have PIPP-R-scores indicative of moderate-to-severe pain, and further indicated that each infant who was identified to have pain by SC actually had one-third probability of false results. On the other side, the imperfect positive predictive value might be due to the higher validity of SC for pain assessment than the PIPP-R. SC did respond to pain in these infants, who did not show sufficient behavioral responses to pain. More sources of SC validity (e.g. in relation to pain treatment, other pain measurements) need to be evaluated to provide more evidence and stronger validity of SC for assessing acute pain in mechanically ventilated infants.

Only one previous SC study conducted ROC in infants and compared SC with NFCS in 34 infants aging from 6 to 12 months for assessing acute postoperative pain in the post-anesthesia care unit (Dalal et al., 2013). Although the Med-Storm SC device was also used in that study, the time window for obtaining NWps and observation time interval were set differently, at 5 seconds and 15 minutes respectively. Compared to the present study, the cut-off point of NWps for
discriminating moderate-to-severe pain (i.e., NFCS ≥ 4 out of 9 points) in Dalal et al., (2013) study was higher (i.e., 0.571), and the parameters of diagnostic test accuracy were lower. Many reasons may contribute to the differences, for example, the different referent pain measurements, the different populations, pain-related contexts, and SC calculation methods.

4.2 Implications for clinical practice and future research

This study showed that SC is a promising approach to assessing acute pain in mechanically ventilated infants. However, the values of SC must be interpreted with caution, due to the imperfect correlations with the referent pain measurements and low positive predictive value. As long as infants are able to respond behaviorally to pain, it may be useful to use SC with other pain scales, for example, PIPP-R and the four-item NFCS, as these scales provided a different dimension of pain in infants from SC and may exclude the false positive cases compared to when only using SC (Wideman et al., 2019). Additionally, pain assessment using SC needs to be considered along with the status of the individual infant, personal expression, and pain-related context (Craig, 2015).

The validity of SC for assessing ongoing (e.g., postoperative pain) or distress of no known cause in infants was not assessed in this study and needs to be comprehensively evaluated in the specific context before its use. Discriminating between pain and non-painful distress is of great clinical utility. Although the validity of SC in relation to known non-painful procedures (e.g., diaper change) was evaluated in our study, SC measurements during periods of distress out of the context of the known painful procedure and non-painful handling did not occur. There certainly remain important knowledge gaps about assessment of ongoing pain and distress of no known cause for sick hospitalized infants. Thus, it is recommended to test the validity of SC for pain assessment in relation to the intensity of ongoing pain or non-painful distress in infants in future studies.

This study did not test the validity of SC for pain assessment in relation to pain treatment. This response is also important as a valid pain assessment method should be able to detect a change in pain due to the implemented pain treatments. In addition, the existence of SC artifacts
might be a practical issue of using SC in clinical settings. Thus, an interdisciplinary collaborative approach between technologists and pain scientists is warranted to advance the technology of measuring SC.

### 4.3 Strengths and limitations

Due to the absence of a gold standard for pain assessment in infants, this study compared SC during painful and non-painful procedures, in the three different phases of procedure and also compared SC with other pain measurements that are commonly used in the clinical and research setting. These comparisons provided comprehensive validity of SC for pain assessment in mechanically ventilated infants. Additionally, this study provided detailed information on SC value calculation, including the time window for obtaining NWps (i.e., 15 seconds), refreshing time (i.e., 1 second), observation time interval (i.e., 30 seconds) and SC sampling method (i.e., maximal value). This transparent reporting reduced bias and enables the replication of the study. Last, this study used ROC analysis to determine the diagnostic test accuracy of SC based on the validity of SC in relation to PIPP-R. This analysis provided further information on the discriminative performance of SC to recognize moderate-to-severe pain in mechanically ventilated infants.

Despite the study strengths, there were also limitations to the study. First, this study was an observational study, and there was no attempt to standardize painful or non-painful procedures and pain treatments. Also, it is not possible to control all the confounding factors using observational research design in a clinical study, such as the variability of each infant recovering from painful procedures and the potential impact of painful procedures on infants’ responses to non-painful procedures. However, the dependent group statistics were used in this study to ensure that the pair of data was from the same patient during the same period of time and in the same context. This design can reduce the effect of potential variability between infants on SC validation. Second, due to the absence of a gold standard for pain assessment in infants, only surrogate approaches could be used to evaluate the validity evidence of SC for pain assessment in this study. Third, the separation of these 30 seconds observation periods was arbitrarily set a
5. Conclusion

In this study, SC showed good validity in relation to category of procedure, phase of procedure and referent pain measurements. SC is a potential approach to assessing pain in critically ill infants requiring mechanical ventilation. However, further research testing the validity of SC in relation to pain treatments and advancing the technology of measuring SC is needed before it can be recommended for routine clinical use.

Acknowledgments

Thanks are due to the nurses and physicians in the NICU and PICU at the Children’s Hospital of Eastern Ontario (CHEO) located in Ottawa, Canada and all parents who agreed to participate. The authors also acknowledge the contributions of Franco Momoli (scientist at Clinical Epidemiology Program, Ottawa Hospital Research Institute and CHEO Research Institute), who provided consultation to statistics and Juliana Choueiry (undergraduate nursing student at University of Ottawa), who coded the videos.

Author contributions

This study is part of the first author’s (J.Hu) PhD dissertation. Specifically, the author led and was engaged in all aspects of the study, including the conceptualization and design of the studies, literature searching and screening, patient recruitment, data collection, data analysis, interpretation of findings and manuscript preparation. The study was conceived in collaboration with the supervisor (D.Harrison) and committee members (J. Harrold & J.Squires). They provided conceptual and methodological support throughout this scholarly work, participated in the analysis and interpretation of data, critically reviewed and edited each manuscript and the dissertation, approved the final version of the work to be published, and had agreement on the integrity and accuracy of the work. Co-author (S.Modanloo.) participated in patient screening, data collection, and video coding. She also provided feedback on analyzing and interpreting the data.
data, contributed intellectual content, and critically reviewed and approved the final version of the two manuscripts.

References


This article is protected by copyright. All rights reserved


Harrison, D. (2001). The Efficacy Of 25% Oral Sucrose In The Reduction Of Pain Associated With Heel Lancing In The Hospitalised Infant; A Randomised Controlled Trial. The
University of Melbourne.


Linnet, K., Bossuyt, P. M., Moons, K. G., & Reitsma, J. B. (2012). Quantifying the accuracy of a

This article is protected by copyright. All rights reserved


This article is protected by copyright. All rights reserved


Legends for illustrations and tables

Figure 1. Patient recruitment flow diagram

Figure 2. The ROC of NWps to discriminate the moderate-to-severe pain (PIPP-R≥13)

Table 1. Sample characteristics (n=55)

Table 2. Analgesic/Sedation medications (n=55)*

Table 3. Painful and non-painful procedures (n=55)

Table 4. The validity of SC in relation to referent pain measurements
<table>
<thead>
<tr>
<th>Variable</th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diagnostics</strong></td>
<td></td>
</tr>
<tr>
<td>Surgical condition</td>
<td>18 (33)</td>
</tr>
<tr>
<td>Respiratory disease</td>
<td>15 (27)</td>
</tr>
<tr>
<td>Cardiac disease</td>
<td>11 (20)</td>
</tr>
<tr>
<td>Sepsis</td>
<td>6 (11)</td>
</tr>
<tr>
<td>General medical disease</td>
<td>3 (5)</td>
</tr>
<tr>
<td>Non-surgical congenital anomalies</td>
<td>2 (4)</td>
</tr>
<tr>
<td><strong>Type of surgery</strong></td>
<td></td>
</tr>
<tr>
<td>Major abdominal surgery</td>
<td>8 (46)</td>
</tr>
<tr>
<td>Cardiac surgery</td>
<td>3 (18)</td>
</tr>
<tr>
<td>Major chest surgery</td>
<td>2 (12)</td>
</tr>
<tr>
<td>Percutaneous drain insertion</td>
<td>2 (12)</td>
</tr>
<tr>
<td>Neurosurgical surgery</td>
<td>1 (6)</td>
</tr>
<tr>
<td>Endoscopy</td>
<td>1 (6)</td>
</tr>
<tr>
<td><strong>Ventilation mode</strong></td>
<td></td>
</tr>
<tr>
<td>Pressure Control Ventilation</td>
<td>29 (53)</td>
</tr>
<tr>
<td>Volume Control Ventilation</td>
<td>20 (36)</td>
</tr>
<tr>
<td>High Frequency Oscillation</td>
<td>6 (11)</td>
</tr>
</tbody>
</table>
Table 2. Analgesic/Sedation medications (n=55)\textsuperscript{a}

<table>
<thead>
<tr>
<th>Medication</th>
<th>Delivery method</th>
<th>Number of doses</th>
<th>Median (IQR) of dosing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetaminophen</td>
<td>Oral</td>
<td>1</td>
<td>10 mg/kg</td>
</tr>
<tr>
<td>Ibuprofen</td>
<td>Oral</td>
<td>1</td>
<td>10 mg/kg</td>
</tr>
<tr>
<td>Fentanyl</td>
<td>Intravenous bolus</td>
<td>17</td>
<td>2 mcg/kg (1.1-2)</td>
</tr>
<tr>
<td></td>
<td>Intravenous infusion</td>
<td>24</td>
<td>2 mcg/kg/h (0.9-3)</td>
</tr>
<tr>
<td>Morphine</td>
<td>Intravenous bolus</td>
<td>10</td>
<td>60 mcg/kg (50-100)</td>
</tr>
<tr>
<td></td>
<td>Intravenous infusion</td>
<td>11</td>
<td>30 mcg/kg/h (20-60)</td>
</tr>
<tr>
<td>Sucrose</td>
<td>Oral</td>
<td>22</td>
<td>0.2 ml</td>
</tr>
<tr>
<td>Ketamine</td>
<td>Intravenous bolus</td>
<td>3</td>
<td>500 mcg/kg (100-1000)</td>
</tr>
<tr>
<td></td>
<td>Intravenous infusion</td>
<td>1</td>
<td>500 mcg/kg/h</td>
</tr>
<tr>
<td>Midazolam</td>
<td>Intravenous bolus</td>
<td>2</td>
<td>50 mcg/kg (50-50)</td>
</tr>
<tr>
<td></td>
<td>Intravenous infusion</td>
<td>2</td>
<td>12.5 mcg/kg/h (10-15)</td>
</tr>
</tbody>
</table>

Note.

\textsuperscript{a} Medications administered within the 6 hours before observation
Table 3. Painful and non-painful procedures (n=55)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type of painful procedures</strong></td>
<td></td>
</tr>
<tr>
<td>Heel lance for blood sampling</td>
<td>25 (45)</td>
</tr>
<tr>
<td>Arterial puncture for blood sampling</td>
<td>11 (20)</td>
</tr>
<tr>
<td>Venipuncture for blood sampling</td>
<td>8 (15)</td>
</tr>
<tr>
<td>Intravenous insertion</td>
<td>6 (11)</td>
</tr>
<tr>
<td>Nasogastric tube insertion</td>
<td>4 (7)</td>
</tr>
<tr>
<td>Heel lance for blood sampling &amp; Venipuncture for blood sampling</td>
<td>1 (2)</td>
</tr>
<tr>
<td><strong>Number of attempts of painful procedures</strong></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>47 (85)</td>
</tr>
<tr>
<td>2</td>
<td>6 (11)</td>
</tr>
<tr>
<td>3</td>
<td>1 (2)</td>
</tr>
<tr>
<td>4</td>
<td>1 (2)</td>
</tr>
<tr>
<td><strong>Type of non-painful procedures</strong></td>
<td></td>
</tr>
<tr>
<td>Diaper change</td>
<td>34 (62)</td>
</tr>
<tr>
<td>Diaper change, abdominal circumference measurement, temperature measurement, lung auscultation</td>
<td>13 (24)</td>
</tr>
<tr>
<td>Diaper change, temperature measurement, lung auscultation</td>
<td>8 (14)</td>
</tr>
</tbody>
</table>
Table 4. The validity of SC in relation to referent pain measurements

<table>
<thead>
<tr>
<th>Different phase and procedure</th>
<th>Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before painful procedures</td>
<td>PIPP-R (0, 0-6)</td>
</tr>
<tr>
<td>SC</td>
<td>$r^b=0.50$ (n=48, p=0.000)</td>
</tr>
<tr>
<td></td>
<td>$r(\text{PIPP-R* NFCS})=0.41$ (n=52, p=0.002)</td>
</tr>
<tr>
<td>During painful procedures</td>
<td>PIPP-R (13, 7.25-16)</td>
</tr>
<tr>
<td>SC</td>
<td>$r=0.62$ (n=48, p&lt;0.001)</td>
</tr>
<tr>
<td></td>
<td>$r(\text{PIPP-R* NFCS})=0.87$ (n=52, p&lt;0.001)</td>
</tr>
<tr>
<td>After painful procedures</td>
<td>PIPP-R (5.5, 0-7)</td>
</tr>
<tr>
<td>SC</td>
<td>$r=0.46$ (n=48, p=0.001)</td>
</tr>
<tr>
<td></td>
<td>$r(\text{PIPP-R* NFCS})=0.56$ (n=52, p&lt;0.001)</td>
</tr>
<tr>
<td>Before non-painful procedures</td>
<td>PIPP-R (0, 0-0)</td>
</tr>
<tr>
<td>SC</td>
<td>$r=0.40$ (n=48, p&lt;0.001)</td>
</tr>
<tr>
<td></td>
<td>$r(\text{PIPP-R* NFCS})=0.86$ (n=53, p&lt;0.001)</td>
</tr>
<tr>
<td>During non-painful procedures</td>
<td>PIPP-R (0, 0-6)</td>
</tr>
<tr>
<td>SC</td>
<td>$r=0.54$ (n=48, p&lt;0.001)</td>
</tr>
<tr>
<td></td>
<td>$r(\text{PIPP-R* NFCS})=0.67$ (n=53, p&lt;0.001)</td>
</tr>
<tr>
<td>After non-painful procedures</td>
<td>PIPP-R (0, 0-3.75)</td>
</tr>
<tr>
<td>SC</td>
<td>$r=0.42$ (n=48, p&lt;0.001)</td>
</tr>
<tr>
<td></td>
<td>$r(\text{PIPP-R* NFCS})=0.69$ (n=53, p&lt;0.001)</td>
</tr>
</tbody>
</table>

Note:

\(^a\) median, interquartile range

\(^b\) $r$=Spearman's rho
Author/s:
Hu, J; Harrold, J; Squires, J E; Modanloo, S; Harrison, D

Title:
The validity of skin conductance for assessing acute pain in mechanically ventilated infants: A cross-sectional observational study

Date:
2021-10

Citation:

Persistent Link:
http://hdl.handle.net/11343/298676