PAIN VARIABILITY IN PRETERM INFANTS

By

Amos Hundert

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

At

Dalhousie University
Halifax, Nova Scotia
August 2020

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Background: Infants born preterm are exposed to repeated medically necessary painful procedures during their neonatal intensive care unit (NICU) admission. Particularly in preterm infants, trajectories of pain reactivity and regulation are not well understood.

Objectives: (1) classify individual pain response trajectories over two-minutes following medically indicated heel lances in preterm infants during their NICU admission; (2) compare mean pain scores within each trajectory class to the sample mean; (3) investigate the stability of classes over time within infants; and (4) explore whether pain treatment, sex, gestational age at birth, previous pain exposure, and postnatal age at time of procedure are associated with pain trajectory class.

Methods: This study used existing data collected by the Trial of Repeated Analgesia with Kangaroo Mother Care (TRAKC) study. TRAKC was a single-blind, three-arm, parallel group randomized clinical trial examining the efficacy of kangaroo care and sucrose alone and in combination as methods of infant pain control during repeated procedures in the NICU. Pain was measured using the Premature Infant Pain Profile at 30, 60, 90, and 120 seconds following the heel lance. Group based trajectory modeling was used to classify pain response in this two-minute time period.

Results: 236 infants contributed 610 painful procedures. Median gestational age at birth was 33 weeks. A model with five pain trajectory classes best fit the data. Three of the trajectories were stable over time at different levels of intensity from low-mild to low-moderate pain. One trajectory reflected a linear reduction from high- to low-moderate pain. The final trajectory showed variable pain at moderate-high levels. At all procedures and at all times points, three classes were at least one-point different from the overall sample mean pain score. Overall, 89 (38%) infants were assigned a different class for each available procedure, 126 (53%) maintained the same class for two procedures, and 21 (9%) maintained the same class for all three procedures. No examined infant or treatment characteristic was found to be meaningfully different in a given class.

Conclusions: In this sample of preterm infants receiving pain relieving interventions, most pain response trajectories reflected mild to low-moderate pain that was stable in the two-minutes post heel lance initiation. Pain trajectories were not consistent over multiple procedures within infants, and an overall mean pain score for the sample may misrepresent subgroups of higher and lower pain.

LIST OF ABBREVIATIONS USED

AIC	Akaike information criterion
BIC	Bayesian information criterion
CONSORT	Consolidated Standards of Reporting Trials
EEG	Electroencephalography
GBTM	Group based trajectory modeling
GRoLTS	Guidelines for Reporting on Latent Trajectory Studies
ICC	Intraclass correlation coefficient
IQR	Interquartile range
KC	Kangaroo Care
MBPS	Modified behavior pain scale
MLE	Maximum likelihood estimate
MRI	Magnetic Resonance Imaging
NFCS	Neonatal facial coding system
NICU	Neonatal intensive care unit
PIPP	Premature infant pain profile
PIPP-R	Premature infant pain profile revised
RCT	Randomized controlled trial
REB	Research ethics board
SD	Standard deviation
SEM	Structural equation modeling
SNAP-II	Score for neonatal acute physiology II
TRAKC	Trial of Repeated Analgesia with Kangaroo Mother Care

CHAPTER ONE: INTRODUCTION

Early exposure to procedural pain in infants can result in short and long-term adverse outcomes including altered brain development and pain response later in life.[1] Preterm or ill infants admitted to the neonatal intensive care unit (NICU) have lengths of hospitalization typically around 19 days.[2] While admitted, infants receive multiple painful procedures each day on average.[3] The most frequent procedures conducted as part of routine care include heel lance, suctioning, and venipuncture. Current best practices for pain management during these procedures include pharmacological, sweet-tasting solutions, and non-pharmacological strategies such as skin-to-skin care (also known as Kangaroo Care) and breastfeeding, which are known to effectively reduce pain.[4,5]

Following an acute painful stimulus, the infant pain response can be conceptually divided in to an immediate pain reactivity period, and a subsequent regulation period where indicators of bio-behavioural pain response regulate back to baseline levels. Together these components constitute the pain response trajectory.[6,7] A multitude of contextual factors such as infant age and previous pain exposure contribute to individual differences in pain response trajectories.[8] Yet these factors are not well understood in infants. Existing research has begun to characterize individual variability in pain response trajectories. In healthy full term infants between two and twelve months of age, Pillai-Riddel and colleagues identified distinct classes of infants who followed a similar pain response trajectory after regular immunization procedures.[9] This study of 747 infants found that most were able to regulate pain, with varying time to regulation distinguishing the different classes of infants. At every age, average pain response within at least one class was meaningfully different to the overall sample average. To date, research in preterm infants lacks an adequate sample size to make strong conclusions, though a study of nine infants found a high level of variability in pain response, indicating the need for more research.[10]

As a result, the aim of this research is to classify and understand trajectories of immediate pain response in preterm infants. Further elucidation of infant pain variability both within and across infants over time is important for both clinical and research purposes. Understanding variability in pain trajectories following a painful procedure is important for guiding the appropriate measurement of pain in clinical research studies.

Additionally, if pain trajectories are highly variable between subjects, average pain in the sample may not be an appropriate primary research outcome as it could result in potentially misleading interpretation of the intervention effects. Clinically, understanding how infants regulate pain and whether different contextual factors drive individual responses to pain could help to inform more individualized and effective pain relief during routine procedures.

CHAPTER TWO: LITERATURE REVIEW

This chapter expands on the concepts summarised in Chapter One and is intended to provide the necessary theoretical and empirical background for the thesis. The chapter begins with an overview of exposure to pain in infants and an examination of the short and long-term impacts of pain exposure. Next, a description of the mechanisms of pain processing, followed by defining the components of acute pain response trajectories in infants, is provided. This is followed by an overview of infant pain assessment methods, evidence-based pain management strategies, and factors associated with variability in pain response. Finally, the chapter concludes with a detailed list of study objectives.

Pain in Infants

Ubiquitous Exposure to Pain in the Neonatal Intensive Care Unit

Infants who require intensive care at birth are admitted to the NICU and the majority (65%) of admissions to Canadian NICUs are for preterm birth.[2] Infants born before 37 weeks of gestation are considered to be preterm. Almost all (>96%) infants admitted to the NICU survive to discharge and, of those admitted for more than 24 hours, the mean length of stay is approximately 19 days.[2] During NICU admission, exposure to procedural pain during routine care is frequent. Infants are estimated to receive between seven and 17 painful procedures per day on average during NICU admission.[3] The most frequent painful procedures are heel lance, suctioning, and venipuncture.[3] A heel lance is a pinprick puncture in the heel to obtain blood. The procedure is conducted for a variety of clinical needs including routine screening tests, toxicology, and blood counts.[11,12] While the procedures are common, routine pain management is unfortunately often still lacking despite clear evidence and guidelines for the use of effective pain reduction methods.[5]

Because of the high level of exposure to pain, guidelines from the Canadian Pediatric Society and the American Academy of Pediatrics recommend that a pain management program exist in all relevant institutions caring for infants.[5] Broadly, recommendations include steps to: 1. minimize the number of painful procedures; 2. implement pain control and prevention strategies, including those for minor procedures; and 3. routinely assess pain.

Impact of Pain Exposure

Infants are exposed to repeated painful and stressful stimuli and procedures during their stay in the NICU while undergoing a period of rapid and critical brain development. This exposure has both short- and long-term impacts. A clear understanding of the direct effects of both preterm birth and repeated pain exposure early in life is challenging due to confounding factors such as illness, differing pain management strategies and pain exposure, as well as the need for large, high quality cohort studies with long-term follow-up.[13]

In the short-term, prevention and management of acute pain is desirable in order to reduce distress for the infant and family.[5,14,15] Infants also become hypersensitive to pain with increased short-term exposure.[16] Hospitalized newborn infants who receive repeated heel lances exhibit an increase in pain response during subsequent procedures.[17] Evidence also suggests that a mother's recall of their infants' exposure to frequent painful procedures in the NICU is related to posttraumatic stress symptomatology in the mother at discharge.[18] Such findings which indicate that painful procedures can be traumatic for caregivers substantiate the need to use interventions which reduce distress in both the infant and caregiver.

Long-term, preterm birth and increased exposure to pain is associated with altered brain structure and connectivity in sensory, cognitive, and emotional networks.[8,19,20] Research demonstrates that an increase in procedural pain in preterm infants is associated with reduced white and grey matter in the brain at term-equivalent age.[21] These alterations are linked to later negative changes in motor, sensory, and visual functions.[21] Pain exposure is also associated with modified thalamic growth which results in reduced thalamic volume, a potential disruption of brain development in somatosensory processing regions, and ultimately an impact on functional outcomes.[22] There is evidence that the negative effects of pain exposure on brain development are sex-specific, with females potentially more vulnerable.[23,24] In longitudinal follow-up studies, cortical alterations have been identified in association with reduced cognitive scores at three years of age, and poor behavioural regulation outcomes at seven and eight years of age.[25–27] Exposure to pain can also affect pain processing and potentially increase the risk of developing persistent pain.[20,28] While further research is needed, taken together, the

evidence indicates a strong rationale for the prevention and management of pain early in life given the clear impacts on long-term development.

Mechanisms of Pain

Pain experience is learned over the lifetime and composed of more than the response in nociceptive pathways.[30] It is widely recognized to include sensory, emotional, cognitive, and social components.[29,30] Nociception refers to the ability of a nerve to detect noxious stimuli and transmit the information to the brain.[31] Nociception is propagated through action potentials along the ascending pathway from the site of transduction throughout the sensory nervous system to the spinal cord and then brain.[32] This activates descending pathways that exert inhibitory modulation effects on the transmission of noxious stimuli.[32] Early pain processing occurs during a period of rapid developmental changes, particularly in preterm infants, and differs from pain processing later in life.[33] While preterm infants have fully developed ascending pathways by approximately 30 weeks gestational age, the descending inhibitory controls are not fully developed at birth.[34] As a result, preterm infants have a limited ability to modulate pain resulting in increased sensitivity to stimuli. Additionally, preterm infants have larger receptor fields in dorsal horn cells which begin to decline in size as they age. This may enhance the strength of pain perception and result in a less localized pain response.[32] Due to the lack of inhibitory control, stimuli such as light or heavy touch trigger a response in newborn infants similar to that triggered by pain; this makes the determination of pain response and behaviour in infants difficult.[33] Because of this critical period of development, it is recognized that minimizing exposure to pain and effective pain management is vital to healthy infant development.

Pain Trajectories

For the purpose of this research the infant pain trajectory will be divided in to three temporal periods: baseline, immediate reactivity, and regulation. The baseline period occurs prior to the initiation of a painful stimulus. While baseline is not part of the pain response, changes in behavioural and other pain indicators are frequently measured as a change from baseline levels.[35] The subsequent period, immediate pain reactivity, begins with initiation of the painful stimulus.[6] Immediate reactivity is often considered to last for between 30 seconds and one minute from the stimulus.[7] This immediate reaction

is primarily automatic and reflexive.[36] It is also the period that is typically used as the primary outcome in research studies.[7,37] Following immediate reactivity is a period of regulation where indicators of pain will typically regulate back to baseline levels.[6] The length of the regulation period varies, and will often be considered in terms of time to return to baseline levels of pain indicators. A recent Cochrane review of non-pharmacological pain relieving interventions in infants considered a pain response that was measured during the first 30 seconds following a stimulus to be during the reactivity period, with anything after 30 seconds considered the regulation period.[9] The same review found that of 51 total studies, 20 reported on immediate pain reactivity only, 16 reported on regulation only, and 15 reported on both periods. Given that there is no consistent definition for the infant pain trajectory, studies will often evaluate pain over different lengths of time. There is a need to better understand the trajectories of pain response in infants in order to inform consistent study design and outcome selection. This can lead to improve generalizability across studies and a better understanding of which interventions may be most effective at time points across the pain trajectory.

Methods of Infant Pain Assessment

The assessment of pain is critical to neonatal clinical care and research. Without valid and reliable assessment, clinical management cannot be as effective and research cannot accurately evaluate the effects of pain treatment options.[38] Pain assessment is challenging due to the fact that infants cannot report their own pain. Importantly, verbal report is only one behaviour used to communicate pain. Behaviours such as facial expression and vocalization of distress, as well as physiological responses are also frequently used.[39] Most existing infant pain assessment tools for research and clinical care are multidimensional and those with the greatest evidence for reliability and validity encompass physiological, behavioural, and contextual factors. These items are typically incorporated as a summary score to form a measure of pain intensity.[40] More than 40 infant pain assessment tools have been developed but most have not undergone adequate assessment for reliability and validity.[40] The American Academy of Pediatrics identified five pain scales that they determined to have undergone adequate psychometric evaluation. They include the Neonatal Facial Coding System (NFCS), Premature Infant Pain Profile (PIPP), Neonatal Pain and Sedation Scale, Behavioral Infant Pain Profile,

and Douleur Aiguë du Nouveau-né.[5] Recent guidelines for clinical trial design recommend the PIPP or NFCS as the primary pain intensity outcome measure for infant pain trials.[38] Neurophysiologic methods such as magnetic resonance imaging (MRI) and electroencephalography (EEG[38]) have been used to assess for neurologic pain signatures in infants.[41] But these methods have not undergone adequate evaluation of reliability and validity, and feasibility of application in clinical and research contexts is presently limited.[5]

Pain Management

Pain control for infants includes pharmacological, sweet tasting solutions, and non-pharmacological strategies. Non-pharmacological strategies and sweet tasting solutions are recommended for most mild and moderate procedures while high pharmacological strategies can be used alone or in combination with non-pharmacologic strategies when persistent or severe pain is present such as ventilation, intubation, and during the perioperative period.[5,42]

Based on recent reviews, four non-pharmacological strategies have sufficient evidence to support their efficacy for pain relief in preterm infants: 1. Kangaroo Care; 2. swaddling/tucking; 3. breastfeeding; and 4. non-nutritive sucking.[7,37] Skin-to-skin contact, often referred to as Kangaroo Care, is direct ventral contact between the infant and mother or caregiver, with the caregiver holding the unclothed infant upright on their bare chest, the two covered together with a blanket or sheet.[43,44] Evidence supports Kangaroo Care as one of the most effective interventions for pain relief during common procedures in preterm infants.[7,44] The mechanisms of effectiveness for pain-relieving strategies such as Kangaroo Care are multi-sensorial including tactile, auditory, and olfactory mechanisms. These multiple mechanisms may contribute to a differing pattern of pain reduction and regulation compared to pharmacological strategies acting on a single mechanism.[45–49] Another recommended intervention for pain relief, breastfeeding, is as or more effective than sucrose and other common interventions, while showing evidence for lower heart rate and crying time.[37] Additional sucking interventions such as non-nutritive sucking (often by pacifier) also produce a calming effect in infants.[50] But non-nutritive sucking interventions are likely less effective when compared to Kangaroo Care and breastfeeding, particularly during regulation.[4,7] Swaddling and facilitated tucking

interventions are other common methods of pain control. They limit the infant's boundaries and encourage self-regulating behavior. Adequate evidence supports the efficacy of swaddling and facilitated tucking in preterm infants, however it may not be as effective at regulating pain compared to Kangaroo Care.[50,51] Adjuvant non-pharmacological treatment strategies for routine painful procedures include varieties of holding, rocking, and sucking interventions.[4,52] These interventions are not as effective without a skin-to-skin component and may be used as adjuvant interventions by caregivers.[4] A strength of many of these non-pharmacological interventions is the cost-neutral nature of the predominantly caregiver-led interventions. Challenges to implementation exist around feasibility in terms of availability of the caregiver during procedures, health care providers accommodating the interventions, and access to the infant to conduct the necessary clinical procedure while the intervention is being implemented.[4,5]

The most extensively researched strategies for mild to moderate painful procedures are sweet tasting solutions such as sucrose and glucose.[53,54] Despite some variation, generally an oral dose of 0.1 to 1 mL of 24% sucrose two-minutes prior to a needle is recommended.[55] Concentrations of less than 20% are less effective, while those greater than 50% do not yield increasing benefit.[53] Sucrose may also be combined with non-pharmacological interventions. Despite its known efficacy, the long-term effects and safety of repeated use of sucrose for pain relief in the neonatal period have not been studied well. Furthermore, the mechanism of induced analgesia is not fully known, with some research indicating a risk of reduced neurodevelopmental scores associated with repeated use of sucrose.[56,57] Persistent pain and severe painful procedures are best managed with pharmacologic agents including morphine, fentanyl, and other opioids.[5]

Pain Variability

Current Evidence

There is currently a lack of research characterizing variability in infant pain response. The largest study to date was conducted by Pillai-Riddell and colleagues, they investigated whether healthy full term infants could be separated into distinct classes of pain response trajectories.[9] Pain was assessed over a two-minute period post-needle at two, four, six, and twelve-month immunization appointments in Canadian pediatric

clinics with up to 574 infants at a given time point. Pain was measured using the Modified Behavior Pain Scale (MBPS), which yields scores from 0 to 10.[58] Classes were discerned at each of the four ages, with the number of classes ranging from three to six at the different time points. Classes were differentiated based on initial pain response and patterns of regulation over the two-minutes. Overall, the results indicate significant variability in infant pain response trajectories over the two-minute time period post needle. Additionally, at ages six and twelve months, the mean pain scores for the study sample at the one and two-minute time points after the needle were clinically significantly different (defined as greater than one point on the MBPS) from the means within the classes. This finding highlights that a sample mean may not adequately represent the variability in pain response when used as a primary outcome in clinical research. A sample mean could lead to clinically significant misrepresentations of subgroups within the population of infants. While similar research in preterm infants has not been completed, Cignacco et al. conducted a feasibility study in nine preterm infants.[10] The authors found significant variability within each infant when examining the pain response following five heel lance procedures during the first two weeks of life. Variability was highest during the heel lance procedure, compared to baseline and recovery, however, the sample size limited the precision and generalizability of findings. As a result, existing studies highlight the need for further research to better understand pain variability in infants.

Factors Associated with Pain Variability

A pain response depends on more than the intensity of the painful procedure itself. Variability in pain response is also influenced by contextual factors.[8] Known contextual factors include demographic (e.g. sex), health (e.g. severity of illness), and environmental (e.g. noise) factors. Understanding the factors that contribute to pain variability in infants is complex given the challenges of measuring pain and other variables in infants who are unable to self-report.[59] There remains a dearth of research on the topic, while existing research shows conflicting findings.[8,60] One factor contributing to the inconsistency in study results is low agreement between studies using a physiologic versus a behavioural measure of pain.[60,61] Consequently, the following section distinguishes between behavioural and physiologic measures of pain, and contrasts relevant findings.

Postnatal age is the factor most frequently studied in association with pain response.[8] Some studies have shown that premature infants may be more sensitive to pain and other stimulation compared to full-term, likely related to the lack of inhibitory pain control development. Other research contradicts those conclusions, and has found that infants born at earlier gestational ages have a reduced behavioural response to pain.[8] Lower gestational age is associated with reduced facial activity and other movement and this likely contributes to lower levels of observed pain given that most pain scales rely, at least in part, on facial expression to determine pain response.[62–64]

Previous pain exposure is also associated with alterations in subsequent pain response. Again, studies have shown both a negative and positive relationship between pain response and previous exposure. Overall, based on a systematic review, more studies have identified that an increased number of previous pain exposure events is associated with a lower behavioural pain response in preterm infants.[8] Potential reasons for this finding include exhaustion related to increased handling from NICU procedures, and a protective effect of endorphins generated from previous procedures.[8,65] Although the measured association is heterogeneous among studies and dependent on the measure of pain used. The association is also likely confounded by factors such severity of illness, which is associated with the number of procedures and length of NICU stay.[61]

Demographic factors such as sex and weight have also been examined in relation to pain response. Most studies indicate no significant associations, although some evidence indicates that males have a higher physiologic pain response compared to females.[8,61,66] Other research found female infants display greater facial indicators of pain.[67] Immediate contextual factors are also likely to contribute to variability in pain response while also contributing to the heterogeneity in research findings in the field. Environmental factors such as light, choice of pain relieving intervention, maternal characteristics, noise and even music can have an impact on infant physiological and behavioural responses to pain.[4,10,68,69] The evidence clearly shows the importance of contextual factors and the need for further research to better understand the complex interaction of these factors with individual infant responses to pain.

Objectives

      This study was a secondary analysis using data collected from the Trial of Repeated Analgesia with Kangaroo Mother Care (TRAKC) study. TRAKC was a single-blind, three-arm, parallel group randomized controlled trial (RCT) examining the efficacy of Kangaroo Care and 24% sucrose alone and in combination as methods of pain control during painful procedures for infants in the NICU.[44,47]Given the gap in knowledge related to pain response and variability in preterm infants, the objectives of this thesis were to:

1. Classify individual pain response trajectories over two-minutes following medically-indicated heel lances in preterm infants during the NICU admission;

2. Identify whether the overall mean pain score for the sample is different from the mean score within each identified pain trajectory class;

3. Determine if pain trajectory classes are stable over time within infants;

4. Explore whether pain treatment group (sucrose, Kangaroo Care with sucrose, Kangaroo Care with placebo), sex, gestational age at birth, previous pain exposure, and postnatal age at time of procedure are associated with pain trajectory class.

CHAPTER THREE: METHODS

Chapter Three contains a description of the methodological approaches used in this thesis. It begins with an overview of the TRAKC study sample, procedures, and interventions. This is followed by a description of study data collection and measures. Analytical methods used are then presented sequentially by order of study objective. First, the pain trajectory modeling methods are discussed along with planned sensitivity and exploratory analyses, followed by the descriptive analysis methods applied to objectives two through four.

Trial of Repeated Analgesia with Kangaroo Mother Care Study Overview

The TRAKC study is a single-blind, three-arm, parallel group RCT examining the efficacy Kangaroo Care and sucrose (alone and in combination) as methods of pain control during painful procedures for neonates in the NICU.[47] Study recruitment was completed between July 2012 and March 2016 at a tertiary-level Canadian hospital (IWK Health Centre, Halifax, Nova Scotia). Participants were randomized to one of three groups: 1. Kangaroo Care with oral placebo (sterile water); 2. Kangaroo Care with 24% oral sucrose; and 3. 24% oral sucrose. Following randomization, pain was assessed during three medically necessary painful heel lance procedures with participants receiving the study pain management strategy as allocated. At 32, 36, and 40 weeks corrected gestational age, neurodevelopmental assessments were conducted, and salivary cortisol samples were collected from mothers. The primary findings have since been published.[10] Subsequent sections will outline the study sample and procedures relevant to the objectives of this secondary analysis of the collected trial data.

Sample

Inclusion and Exclusion Criteria

Infants were eligible for inclusion in the TRAKC study if:

1. they were born at less than or equal to 36 6/7 weeks of gestational age (preterm) according to a dating ultrasound performed in the first trimester, and;

2. they could be enrolled within 7 days of birth, and;

3. they were clinically stable enough to receive maternal Kangaroo Care, as determined by the neonatal attending team, and;

4. their mother provided consent and was available for Kangaroo Care for the duration of the hospital admission.

Infants were excluded if any of the following criteria were met:

1. diagnosed with a major congenital anomaly, or;

2. required surgery.

For the purposes of this secondary analysis, only infants with data available for at least one of the three possible painful procedures was included.

Procedure

General

A research nurse not involved in clinical care identified eligible infants and obtained written informed consent from each mother. Following randomization, the research nurse described the allocated Kangaroo Care intervention to the mother (sucrose allocation was concealed) and coordinated video recording during three medically necessary painful procedures. A maximum of three painful procedures per infant were video recorded; infants received many more medically necessary painful procedures that were not recorded. Hospital staff were instructed to follow the allocated study protocol intervention for all painful procedures undergone by the infant, with any deviation from the protocol recorded in the infant's chart.

Randomization and Allocation Concealment

Randomization was conducted over a secure web based system accessible by the study research nurse only. The randomization allocation sequence used permuted blocks and was stratified by gestational age at birth (less than 32 weeks and 32 weeks or greater). Parents and health care providers could not be blinded to Kangaroo Care, but they were kept blind to group assignment in terms of sucrose or placebo. Research staff who coded pain scores were was also blinded to group allocation. Syringes (24% sucrose, sterile water) were prepared by the study pharmacy and labelled with a code known to the pharmacy and research nurse only.

Interventions

The three interventions groups were: 1. Kangaroo Care and oral placebo (sterile water); 2. Kangaroo Care and oral 24% sucrose solution; and 3. 24% oral sucrose solution alone. Infants receiving Kangaroo Care were placed wearing a diaper upright on their

mother's bare chest at least 15 minutes prior to the painful procedure. The goal was to maximize contact between the mother and infant. Given that sucrose and placebo intervention allocation was blinded, the interventions were administered with the same procedures. Sucrose syringes contained 24% oral sucrose solution, while placebo syringes contained sterile water. Two-minutes prior to the procedure the infant received one-quarter of the NICU recommended volume (0.4 – 1.0 mL depending on infant weight) of solution to the tongue by syringe with the remaining volume given as needed in increments during the procedure. Those infants receiving sucrose alone were placed in the supine position 30 minutes prior to the procedure and remained in the cot or incubator for the duration of the procedure.

Data Collection

Every attempt was made to time the three painful procedures for which pain response measurements were made with the first collected immediately post-recruitment at approximately 32 weeks, the second at 36 weeks, and the third prior to discharge. Given that all painful procedures were required to be part of routine care, particularly for younger infants, this timing was not always feasible. At minimum, each recorded procedure was required to be at least 24 hours from the preceding recorded procedure. All study procedures were video recorded from before the initiation of the allocated intervention, to no earlier than two-minutes after the procedure. Those infants receiving Kangaroo Care were first recorded for one minute in the incubator. Physiologic heart rate and oxygen saturation recordings were captured in real time during the procedures. The same recording equipment was used for all videos and infants.

Measures

Demographics, Disease, and Admission Characteristics

Participant demographic and disease characteristics were obtained from parent report and patient charts by the research nurse. Severity of medical risk at birth was assessed with the Score for Neonatal Acute Physiology (SNAP)-II.[70] The daily number of non-tissue damaging and tissue damaging procedures was obtained from the medical chart. A daily diary provided to the mother was used to record self-reported number of hours spent providing Kangaroo Care to their infant.

Pain

Premature Infant Pain Profile

The PIPP is a rigorously evaluated indicator of pain intensity in infants that is regularly used in both research and clinical care.[71] The PIPP uses a multidimensional approach incorporating behavioural, physiological, and contextual factors.[35] It is composed of three facial indicators: brow bulge, eye squeeze, and nasolabial furrow; two physiological indicators: heart rate and oxygen saturation; and two contextual indicators: gestational age and behavioural state.[35] The seven indicators are each scored on a four-point scale from zero to three, indicating increased change from baseline values. Scores are summed for a total score ranging from zero (no pain) to 21 (highest pain). See Appendix A for detailed PIPP scoring instructions. In clinical care, PIPP scores are conducted in real time by the clinician. When used for research, the PIPP is typically scored retroactively using stored data and video recordings of the infant.[72] This ensures reliability and reproducibility of PIPP scores. PIPP scores are typically assessed over a 30-second epoch. When used in research multiple epochs are often evaluated back to back, such as four epochs for a total of two-minutes. Scores are also calculated during different phases immediately following the procedure, and at rest once the infant is in the recovery period. Physiological indicators are calculated based on a change from baseline values, therefore collection of resting baseline values prior to the procedure is important. More recently, the PIPP-Revised (PIPP-R) was developed based on user feedback, with the goal of improving ease of clinical use and validity.[73,74] The changes include minor scoring and instructional revisions. Pain scored using both the PIPP and PIPP-R are highly correlated.[73,74] Given that the PIPP-R was not validated at the time the TRAKC study was developed, the PIPP was used.

Trained coders used the close up video recordings of the infant's face to conduct the PIPP scoring for the TRAKC study. All coders were trained with the same training videos until they were reliable with an intraclass correlation coefficient (ICC) no less than 0.95, indicating excellent reliability.[75] Subsequent inter- and intra-rater reliability was evaluated throughout the study on a subset of procedures with all coders maintaining greater than 0.92 ICC for the duration of the study.[44] Coding was conducted in four phases: baseline one, baseline two, heel lance, and recovery. Figure 1 provides an

overview of the coding procedures. The first baseline phase (2x30 seconds) was started at the beginning of the monitoring while the infant was in the incubator, followed by a second baseline (2x30 seconds) one minute prior to initiation of the heel lance procedure. The heel lance phase (4x30 seconds) began immediately following the heel lance initiation. The final phase, recovery (minimum of 4x30 seconds), started once the nurse completed all blood sample collection procedures and ended once the infant's heart rate returned to baseline values. The research nurse inserted coloured cards in the videos during recording as phase markers. Coding was conducted one painful procedure at a time, with the same coder assessing each phase in sequence. Coding was conducted following the same method for the three video recorded painful procedures in each participant. PIPP scores for this study were calculated based on changes observed from the first baseline period.
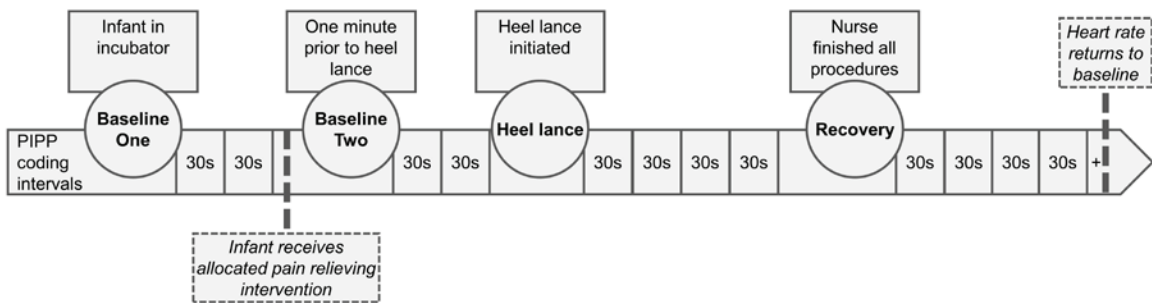
Figure 1. Premature Infant Pain Profile (PIPP) coding flow diagram.

Ethics

The standard of care with respect to procedural pain management in the IWK NICU when the TRAKC study was initiated was the administration of 24% sucrose prior to tissue breaking procedures. In addition, adjuvant non-pharmacological interventions such as Kangaroo Care were implemented as necessary based on provider discretion. As a result, all infants in the study, regardless of treatment arm, received the standard of care.[5] The study was approved by the IWK Research Ethics Board (REB 1009503).

Analyses

Objective One

To classify individual pain response trajectories over 2-minutes following medically indicated heel lances in preterm infants during three assessment time points throughout the NICU admission

Group-based trajectory models (GBTM) were used to classify individual pain responses over the two-minute period post heel lance in the preterm infants using the data from the TRAKC study.[76] GBTM is a modeling approach used to identify distinct groups of individuals who follow a similar trajectory in an outcome over time. In health research, GBTM is used in the field of pediatric development, as well as for examining trajectories of conditions such as depressive symptoms and cardiovascular health over a multi-year time period.[77–79]

The assumption of the GBTM approach is that the population contains distinct groups of individuals (classes) with similar trajectories.[80] The heterogeneity is captured with a latent categorical variable representing the classes. Latent variables are not observed directly (unmeasured) but inferred through mathematical modelling from other observed (measured) variables. Of course, the assumption that the population is composed of categorical classes, while useful for describing and understanding the population of study, is generally incorrect.[76] The assumption of distinct classes is made for the purposes of modeling and the likely continuous distribution of potential classes is grouped into clinically meaningful classes that are useful in understanding the question of interest. The GBTM approach estimates the probability of class membership for all individuals, with individuals being placed in the class for which they have the highest probability. As a result, a class is composed of a group of individuals who generally follow a similar trajectory. The important distinction is that individuals do not belong to a class, but are assigned a class.[76]

GBTM is one of several related methods to develop trajectories over time based on structural equation modeling (SEM).[81] GBTM may also be referred to as latent class growth analysis in some research, and terminology is inconsistent. A study evaluating real and simulated data sets reviewed three different SEM approaches to trajectory modeling and identified GBTM as generally performing best in terms of identifying linear

trajectories.[79,81] Using the GBTM approach, there is no within class variation so that all participants within a class are assumed to have the same growth trajectory shape.[80,81] Other methods allow within class variation, where individuals within a class can differ in growth trajectory. As a result, GBTM is less likely to classify individuals with different growth trajectories in the same class. This can result in GBTM generating more unique classes in the optimal model.[79]

One of the most critical choices, as well as one of the major challenges in GBTM is determining the appropriate number of classes.[82] Modeling is typically done by first fitting a single class (k) and then increasing the number by one ($k + 1$), iteratively testing the models with increasing number of classes against the last until the best fit is identified. A criterion based on the maximum likelihood estimate (MLE) such as Bayesian information criterion (BIC) or Akaike information criterion (AIC) is used to determine best fit and resulting number of classes in the population.[76,83] Simulation studies have demonstrated that the BIC outperforms AIC, with the BIC optimally identifying the number of classes in more models.[84] Such rule-based methods used in isolation may not obtain trajectories that are clinically and theoretically meaningful, interpretable, and address the research question of interest. Recommendations include applying additional judgement-based rules to improve the meaningfulness of results.[80,82] For example, trajectories that account for minor variations may be collapsed to improve interpretability. Regardless of how the final model is selected, transparency in the method used is critical to ensuring reproducibility.[82]

In this study, GBTM analyses were conducted using the Stata plugin by Jones and Nagin.[85] The Stata plugin is widely used with over 240 citations in the literature as of March 2020. The outcome of interest was infant pain intensity, measured using the PIPP from the three painful procedures. The PIPP was captured at four time points, 30, 60, 90, and 120 seconds from the heel lance. Data from the three painful procedures per individual infant were included in the model. The model used a censored normal distribution with a maximum possible PIPP value of 21 (highest possible scale score). Parameter estimates from the model were used to calculate the probability of class membership for each individual in each class. Classes fit from the final model were

plotted over time (30, 60, 90, 120-seconds) with individual raw scores included as individual lines on the plot.

The BIC was used as a criterion for model selection, starting with a single class model and increasing by one class incrementally, with a lower BIC indicating better model fit. Selection of the final model also used the following criteria as recommended by Nagin: 1. a strong correspondence between the estimated probability of group membership and the proportion assigned to that group based on posterior probability; 2. an average posterior probability among all procedures assigned to a class greater than 0.7 for each class; and 3. a reasonable sample within each class so that no class represents only a small subset of the sample.[80] Reporting of results and the model selection process were guided by the Guidelines for Reporting on Latent Trajectory Studies (GRoLTS).[82] The guidelines include a checklist of 16 items for transparent and well-conducted studies using GBTM. GRoLTS items include reporting the number of fitted models, sample size and missing data procedures, and how the final model was selected. See Appendix B for the full list of GRoLTS items.

The PIPP score contained missing epoch times within a procedure (e.g., PIPP at 60 seconds) as well as fully missing procedures. For instances of missing epochs within a procedure, the procedure remained in the model data with the missing time point. Procedures that were fully missing with no available data were not included. With the GBTM approach, the maximum likelihood estimates yield parameter estimates that are unbiased if data are missing at random.[76] Accordingly, patterns of missing data were descriptively explored and reported. All results present the procedure and or infant sample sizes at each time point and a sensitivity analysis was conducted including only those procedures with complete data at all four time points within a procedure. Another sensitivity analysis was conducted including only procedure one data to examine the impact of including up to three procedures from each infant in the trajectory models.

Sequential measurement during the same procedure was not considered in the initial development of the PIPP, but it is often used in research.[7] Gestational age and baseline state components of the PIPP score are assigned constant values within a procedure. As such, pain scores will always remain above zero over multiple measurements for infants assigned a score greater than zero on either of the two stable

components. As well, a PIPP score indicating the presence of pain can be recorded with no corresponding observed behavioural or physiologic indicators of pain. As a result, an exploratory analysis was conducted to investigate the impact of the stable score components. Gestational age and baseline state scores were subtracted from the total PIPP score and GBTM models were fitted following the same procedures as the primary analysis. The modified PIPP score contains a possible range in values from zero to 15.

A second exploratory analysis was conducted to examine the PIPP scores captured during the heel lance recovery phase defined previously in Figure 1. The amount of time between the initial 120 seconds post heel lance and the start of the recovery period was based on the time to the end of the nurse conducting all procedures for the heel lance, which varied. All other relevant time points in the study were assessed at consistent 30 second intervals. The GBTM method used in this study treats the data as time-structured, with infants assessed at exactly the same time intervals.[86] Due to the variations in the time between the heel lance and recovery phase from one procedure to another, separate trajectories were fitted for the recovery period alone which included the PIPP measured at four 30-second intervals.[82]

Objective Two

To identify whether the overall mean pain score for the sample is different from the mean score within each identified pain trajectory class

To determine whether the mean PIPP score within each trajectory class differs from the overall sample mean, the smallest change in pain that could be perceived as meaningful was selected as recommended by the Optimal Strategies for Reporting Pain in Clinical Trials.[87] Based on previous research and investigator judgement, the minimally clinically significant difference of one point on the PIPP was selected.[55,88] A one-point change may be considered clinically meaningful, as any reduction in pain is desirable.[55] There is no clear consensus on a clinically meaningful reduction in PIPP score. Both a one and two-point reduction have been used, along with consideration for whether the PIPP score is reduced from what is considered a higher to lower pain category.[88] Generally, a score of less than six indicates minimal pain, a score between six and twelve indicates moderate pain, and a score of 13 or more indicates severe pain.[71] This was considered when interpreting findings. Infants were assigned to the class for which they

have the highest probability at each procedure. Mean PIPP scores in the identified classes were descriptively compared to the full sample mean at each procedure to identify if clinically significant differences were present.

Objective Three

To determine if pain trajectory classes are stable over time within infants

To address whether pain trajectory classes are stable over time within infants, a descriptive approach was taken. Infants were assigned to the trajectory class for which they have the highest probability of being in at each procedure. Descriptive statistics were used to report the frequency of infants who would be placed in the same trajectories over repeated procedures, based on the trajectories fit at each procedure point.

Objective Four

To explore whether pain treatment group (sucrose, Kangaroo Care with sucrose, Kangaroo Care with placebo), sex, gestational age at birth, previous pain exposure, and postnatal age at time of procedure are associated with pain trajectory class

To determine if TRAKC pain treatments, gestational age at birth, postnatal age at time of procedure, and previous pain exposure as associated with pain trajectory class the variables were descriptively explored and reported for each class at each procedure. Frequencies were examined for clinically meaningful differences between classes which were consistent across procedure times. The variables examined were selected based on the existing literature identifying their association between these factors and variability with infant pain intensity.

The a priori analysis plan was to conduct multinomial logistic regression modelling to determine the association between the characteristics of interest and pain trajectory classes. This methodology was not pursued due to the small number of infants in some of the pain trajectory classes identified. The sample size of infants within classes was prohibitively small to run a convergent model once stratified by procedure and independent variables.

21

CHAPTER FOUR: RESULTS

The chapter starts with an outline of the study participant characteristics. The results identifying the best fitting model of pain trajectory classes are then presented, followed by an overview of the selected final trajectory classes. This is followed by an examination of the impact of missing data and the results of the exploratory analyses. Subsequently the pain scores within each class are compared and contrasted to the overall sample, stability of trajectory classes over time are evaluated, and factors potentially associated with trajectory classes are examined. All tables and figures for the results section are presented following the text, at the end of this chapter.

Sample Characteristics

Overall 1276 infants were screened for inclusion in the TRAKC study. Of those, 714 were ineligible, 207 declined to participate, 113 were not approached, and 242 infants were randomized in the TRAKC study. A full Consolidated Standards of Reporting Trials (CONSORT) flow diagram is published along with the primary study results.[44,89] Six infants had no painful procedure data, resulting in 236 infants included in the sample for the current study. Median gestational age at birth was 32.9 weeks. See Table 1 for additional maternal and birth characteristics and Table 2 for infant characteristics. Characteristics were generally similar across the three TRAKC intervention groups.

Objective One

Assessment of Group Based Trajectory Model Fit

The BIC improved with the addition of classes from a single class model up to six, based on initial quadratic polynomial specifications for all models and classes. The maximum number of classes fitted was a six class solution. One of the classes within the six class solution had only 17 (2.8%) procedures assigned to it and based on this low sample size within the class, no further classes were added to the models. See Figure 2 for trajectories fit from all initial models and Table 3 for model development criteria.

Following the criteria defined a priori for model selection, a model with five trajectory classes was identified as the best fit and combinations of polynomial function forms were tested. The polynomial degree was increased incrementally from intercept (degree zero; i.e., a mean straight line) for each trajectory class. Increasing the

polynomial specification was done to best account for slopes and curvatures of the trajectory beyond a mean straight line. See Appendix C, Table 1 for the evaluation of four fit options. The final specification selected was an intercept, linear, intercept, linear, and cubic polynomial for classes one through five respectively. The model demonstrated good assignment accuracy based on the mean posterior probabilities, with the average posterior probability ranging from 0.85 to 0.96 for each class. The estimated and observed proportion of class membership corresponded strongly (within one percentage point). See Table 4 for specifications of the final model fit.

Pain Response Trajectories

Primary Results

Of the five distinct trajectory classes identified, class one (12% of procedures, n = 22), class two (41% of procedures, n = 251), and class three (33% of procedures, n = 205) all reflected stable trajectories over the two-minute post heel lance period. Stable was defined as a trajectory where the mean PIPP score over time did not deviate by more than one point. From 30 to 120 seconds, the mean value for each class did not deviate by more than one point on the PIPP score for class one, two, and three. Class four (10% of procedures, n = 59) demonstrated a linear reduction in pain from 30 seconds (mean PIPP score, 10.6) to 120 seconds (mean PIPP score, 7.3) with a difference of 3.3 points. Class five (4% of procedures) included the smallest number of procedures (n=22) and the highest pain score at each time point. The trajectory was stable with the mean pain score ranging from 11.3 to 11.9 except for a spike in pain at 60 seconds to 13.1. The mean at each time point is plotted along with the estimated trajectory lines in Figure 3. As well, Figure 4 includes the raw PIPP scores for each individual procedure within the class it was assigned. Overlap is observable between the upper and lower bounds of procedures assigned to similar classes.

Missing Data and Sensitivity Analyses

Overall the sample contained 610 painful procedures, including 229 (95% of 242 randomized infants), 211 (87%), and 170 (70%) in procedures one through three respectively. No painful procedure with a PIPP score available was excluded from the primary analyses. One hundred and thirty-three infants (55%) had completed data for all four time points (30, 90, 90, 120 seconds) within all three procedures. Of the 610

procedures, 40 (7%) contained one or more missing time point within the procedure, the remaining 570 (93%) contained complete procedure data. See Appendix D Table 1 for an outline of all missing data across procedures. Demographic and other characteristics of the missing compared to non-missing samples at each procedure are also available in Appendix D Table 2.

A sensitivity analysis was conducted on procedures with complete PIPP scoring only to examine the impact of missing data. Complete procedures included data for all four time points within the procedure (N=570). Figure 5 displays the resulting trajectories. All classes showed minimal change to the mean PIPP scores at each time point compared to the primary results (Figure 2), with none changing by more than 0.5 of a point. The proportion of procedures within each class also remained consistent compared to the primary results. The largest change was two percentage points in class three (33 to 35%). See Appendix C, Table 2 for data on the sensitivity model fit.

In order to include each infant only once in the data, the final model was fitted to data from procedure one (N=229 infants). Results shown in Figure 6 found that generally the classes followed the same trajectory patterns, with no changes in terms of stability or slope. Some change in the proportion of procedures assigned to each class was observed. Class four had approximately a one-point lesser reduction in mean PIPP score from 30 to 120 seconds and class five demonstrated an approximately one point higher peak PIPP score at 60 seconds. See Appendix C, Table 2 for data on model fit.

Exploratory Trajectory Modeling

Trajectories modeled using the modified PIPP scores without gestational age and baseline state also resulted in selection of a five class solution. See Appendix C, Table 2 for model fit information. Figure 7 displays the resulting five classes. Caution should be used in interpretation of the modified PIPP score values, as they do not correspond to a validated pain score.[35] The trajectories of the resulting exploratory classes demonstrated similar results to the primary analysis with three stable classes representing the majority of procedures (89%), a fourth class demonstrating a reduction in the modified PIPP score from 30s to 120 seconds, and a fifth class with higher and variable pain. Table 5 shows the agreement of procedures between the primary trajectory classes and the modified classes. Results show that the majority of procedures remained within the same class,

with the exception of modified class three, where 60% of procedures had been assigned to class two in the primary analysis. Overall, the majority of any movement was to a class closest and most similar in terms of mean score and trajectory (e.g. from class one to class two).

The median time from initiation of the heel lance to start of recovery period was 3.6 minutes (IQR 2.7, 5.2). Again, a five class solution was solution was selected as the optimal fit for the trajectories modelled using the heel lance recovery phase data. See Figure 7 for results and Appendix C, Table 3 for fit data on the selected final model. All five trajectories were relatively stable, class four (20%) demonstrated the greatest downward slope over time, but the change in mean PIPP score from 30 seconds (6.7) to 120 seconds (5.9) was less than one point (0.8). Class three was the largest, representing 47% of procedures. A cross-tabulation of the primary reactivity and recovery trajectories is shown in Table 6. Generally, procedures were not highly stable in terms of maintaining the same class number from primary reactivity classes to those during recovery. Class three was the most stable with 50% of procedures maintaining the same class. Movement was most often a distance of one class (e.g. from class one to class two).

Objective Two

Pain Scores Within Trajectory Classes

The overall sample mean PIPP pain intensity scores at each time point for each procedure were compared to those across the five classes. At procedure one and two, at all times points all classes except class three were at least one-point different the sample mean. At procedure three, the sample mean more closely matched class two, and was within one point on the PIPP score at three of the four time points. Results for all procedures are shown in Table 7, with those scores within one-point of the sample mean highlighted in grey. Generally, class five was the furthest from the sample mean, with values ranging from 11.0 to 14.5, compared to the overall sample which ranged from a mean of 5.0 to 6.2 across the procedures and time points.

Objective Three

Stability of Infants Within Classes Across Procedures

The stability of individuals within classes over multiple procedures was examined descriptively. Overall, 126 infants (53%) maintained the same class for two procedures, and 21 (9%) maintained the same class for all three procedures. The remaining 89 (38%) infants were assigned a different class for all procedures. Only classes two and three contained infants who maintained the same class for all procedures, these classes were the largest two classes by procedure count, and both demonstrated stable pain over time. In the classes with higher pain levels, nine infants maintained class four for two procedures (representing 31% of class four procedures) and two infants maintained class five (representing 18% of class five procedures). A cross-tabulation between the procedures and classes is presented in Table 8.

Objective Four

Factors Associated with Pain Trajectory Class

The association of class with various study and infant characteristics was examined descriptively. See Table 9 for results by class, procedure, and characteristic. In terms of TRAKC study treatment group, there was no consistent trend observed for any class across all three procedures where a greater proportion of infants were observed to be in a given treatment group for that class. The proportion of male and female infants in classes one through four was generally similar to the sample overall (56% male). In class five 16 (73%) of the procedures were male infants. There were no meaningful differences in gestational age at birth across the five classes. Likewise, for gestational age at procedure the median across the five classes was within one week for each procedure, showing no differences. There was no trend observed between the classes over the three procedures for age at procedure. No differences were observed in the number of previous pain related procedures between classes. At procedure two in class five the median value was 46 painful procedures, an increase of more than 10 compared to any other class. But only six procedures were assigned to this class and time point, and a similar high value was not observed at procedure one or three.

Table 1. Maternal demographic characteristics, N=236 mothers.

Characteristic	Study group			
	KC & Placebo	KC & Sucrose	Sucrose	Total
	n=79	n=78	n=79	N=236
Maternal age in years, median (IQR)	31 (28, 35)	31 (26, 34)	31 (26, 34)	31 (26.5, 34)
Caesarean delivery				
No	37 (47.8%)	40 (51.3%)	44 (55.7%)	121 (51.3%)
Yes	42 (52.2%)	38 (48.7%)	35 (44.3%)	115 (48.7%)
Prior Kangaroo Care experience				
No	56 (70.9%)	51 (65.4%)	56 (71.8%)	163 (69.4%)
Yes	23 (29.1%)	27 (34.6%)	22 (28.2%)	72 (30.6%)
Missing	0	0	1	1
Maternal ethnicity, white				
No	4 (5.1%)	7 (9.0%)	11 (13.9%)	22 (9.3%)
Yes	75 (94.9%)	71 (91.0%)	68 (86.1%)	214 (90.7%)
Two-parent family				
No	0 (0.0%)	3 (3.8%)	1 (1.3%)	4 (1.7%)
Yes	79 (100.0%)	75 (96.2%)	78 (98.7%)	232 (98.3%)
Maternal college or university education				
No	14 (18.2%)	12 (16.0%)	24 (31.2%)	50 (21.8%)
Yes	63 (81.8%)	63 (84.0%)	53 (68.8%)	179 (78.2%)
Missing	2	3	2	7

Table 2. Neonatal demographics, N=236 infants.

Characteristic	Study group			
	KC & Placebo	**KC & sucrose**	**Sucrose**	**Total**
	n=79	n=78	n=79	N=236
Gestational age at birth (weeks), median (IQR)	33.1 (31.4, 34.6)	33.1 (31.6, 34.1)	32.4 (31.3, 33.6)	32.9 (31.4, 34.1)
Birth weight (g), median (IQR)	1780 (1560, 2230)	1874 (1480, 2140)	1860 (1510, 2126)	1845 (1510, 2170)
Infant sex, n (%)				
Male	41 (51.9%)	49 (62.8%)	42 (53.2%)	132 (55.9%)
Female	38 (48.1%)	29 (37.2%)	37 (46.8%)	104 (44.1%)
Twin birth, n (%)				
No	55 (69.6%)	51 (65.4%)	48 (60.8%)	154 (65.3%)
Yes	24 (30.4%)	27 (34.6%)	31 (39.2%)	82 (34.7%)
SNAP-II total score, median (IQR)	0 (0, 9)	0 (0, 9)	0 (0, 9)	0 (0, 9)

Table 3. Initial group-based trajectory model development: comparison of models fit with a one through six class solution. All preliminary models based on quadratic polynomial shape, N=610 procedures.

Model and class	Bayesian information criterion		Mean posterior probability	Group membership		Sample size			
	Value	Diff.		Est. %	Sample %	Total N=610	P1 n=229	P2 n=211	P3 n=170
One class model									
One	-5685		1.00	100	100	610	229	211	170
Two class model									
One	-5296	389	0.96	72.3	72.8	444	168	151	125
Two			0.91	27.6	27.2	166	61	60	45
Three class model									
One	-5093	203	0.90	37.1	37.4	228	78	78	72
Two			0.92	51.7	52.0	317	127	109	81
Three			0.95	11.2	10.6	65	24	24	17
Four class model									
One	-5016	77	0.89	16.3	15.7	96	35	28	33
Two			0.91	51.7	51.1	318	121	112	85
Three			0.89	26.6	26.9	164	59	58	47
Four			0.96	5.4	5.3	32	14	13	5
Five class model									
One	-4955	61	0.90	13.9	13.7	83	27	25	31
Two			0.88	28.5	27.4	167	69	55	43
Three			0.89	46.7	48.4	295	110	108	77
Four			0.89	7.7	6.4	39	10	14	15
Five			0.93	4.2	4.3	26	13	9	4
Six class model									
One	-4882	73	0.84	11.0	11.8	72	22	22	28
Two			0.88	39.8	40.7	248	98	86	64
Three			0.91	34.8	33.0	201	76	70	55
Four			0.91	4.3	4.3	26	11	10	5
Five			0.89	7.5	7.5	46	13	18	15
Six			0.93	2.6	2.8	17	9	5	3

Table 4. Final immediate pain reactivity trajectory model fit, N=610 procedures.

Model and class	Type[a]	Bayesian information criterion		Mean posterior probability	Group membership		Sample size			
		Value	Diff.		Est. %	Sample %	Total N=610	P1 n=229	P2 n=211	P3 n=170
One	I	-4942	13	0.85	11.3	12.0	73	23	22	28
Two	L			0.90	42.2	41.1	251	97	90	64
Three	I			0.86	32.9	33.6	205	81	71	53
Four	L			0.90	9.9	9.7	59	16	22	21
Five	C			0.96	3.7	3.6	22	12	6	4

[a]Type of polynomial fitted: I: intercept, L: linear, Q: quadratic, C: cubic.

Table 5. Cross-tabulation between the primary classes and the exploratory classes based on the modified Premature Infant Pain Profile score excluding baseline state and gestational age, N=610 procedures.

Class	Modified class					
	One	Two	Three	Four	Five	Total
One	**45 (62%)**	28 (38%)	0	0	0	73
Two	56 (22%)	**184 (73%)**	10 (4%)	1 (1%)	0	251
Three	0	124 (60%)	**71 (35%)**	9 (4%)	1 (1%)	205
Four	0	0	20 (34%)	**36 (61%)**	3 (5%)	59
Five	0	0	0	6 (27%)	**16 (73%)**	22
Total	101 (16%)	336 (55%)	101 (16%)	52 (9%)	20 (3%)	610

Table 6. Cross-tabulation between the primary reactivity classes and the exploratory recovery classes, N=606 procedures.

Class	Recovery class					
	One	Two	Three	Four	Five	Total
One	**17 (24%)**	49 (68%)	6 (8%)	0	0	72
Two	2 (1%)	**83 (33%)**	146 (58%)	17 (7%)	2 (1%)	250
Three	0	16 (8%)	**101 (50%)**	78 (38%)	8 (4%)	203
Four	2 (3%)	6 (10%)	25 (42%)	**17 (29%)**	9 (15%)	59
Five	0	5 (23%)	5 (23%)	9 (41%)	**3 (13%)**	22
Total	21 (3%)	159 (26%)	283 (47%)	121 (20%)	22 (4%)	**606**

Table 7. Mean Premature Infant Pain Profile scores within each trajectory class and for the overall sample at each time point, N=236 infants.

Class	Time from heel lance initiation							
	30 seconds		60 seconds		90 seconds		120 seconds	
	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)
Procedure one								
One	23	2.3 (1.1)	23	2.2 (0.9)	23	2.6 (1.0)	21	2.3 (0.8)
Two	96	4.9 (1.3)	96	4.5 (0.8)	97	4.5 (0.8)	97	4.5 (0.8)
Three	79	6.6 (1.4)	81	7.1 (1.5)	81	6.6 (1.5)	79	6.6 (1.5)
Four	16	10.1 (1.8)	16	9.5 (1.9)	16	8.3 (1.9)	16	7.7 (1.6)
Five	11	11.5 (3.5)	12	13.9 (1.9)	11	12.1 (2.1)	11	11 (2.6)
Total	**225**	**5.9 (2.7)**	**228**	**6 (2.9)**	**228**	**5.7 (2.4)**	**224**	**5.6 (2.3)**
Procedure two								
One	22	2.3 (0.9)	22	2.2 (1)	22	2.3 (0.9)	21	2.4 (1.4)
Two	89	5 (1.4)	88	4.4 (0.9)	88	4.4 (0.9)	83	4.4 (0.9)
Three	71	6.7 (1.5)	69	6.6 (1.1)	70	6.9 (1.8)	68	6.9 (1.7)
Four	22	11.4 (2.6)	21	9.4 (2.1)	20	8.6 (1.5)	19	7.5 (2.1)
Five	6	11.8 (3.1)	6	13.3 (3.7)	6	11.7 (2)	5	13.2 (1.8)
Total	**210**	**6.2 (2.9)**	**206**	**5.7 (2.6)**	**206**	**5.6 (2.5)**	**196**	**5.6 (2.5)**
Procedure three								
One	28	2.4 (1.3)	28	2 (1)	27	1.9 (0.9)	26	1.9 (0.9)
Two	63	5.2 (1.5)	63	4.4 (0.9)	63	4.3 (1)	61	4.2 (1.1)
Three	53	6.9 (1.5)	53	6.8 (1.6)	51	6.6 (1.5)	51	6.4 (1.4)
Four	20	10.4 (1.7)	21	9.4 (1.8)	20	7.8 (1.5)	17	7 (0.9)
Five	4	11.3 (2.2)	4	11 (2.2)	4	11.8 (1.7)	4	11.3 (2.9)
Total	**168**	**6 (2.8)**	**169**	**5.5 (2.7)**	**165**	**5.2 (2.5)**	**159**	**5 (2.3)**

Table 8. Cross-tabulation between procedure and pain trajectory class, N=610 procedures.

Procedure	Class					
	One	Two	Three	Four	Five	Total
One	59 (20%)	94 (32%)	83 (28%)	41 (14%)	18 (6%)	295
Two	14 (6%)	112 (44%)	104 (41%)	18 (7%)	4 (2%)	252
Three	0	45 (71%)	18 (29%)	0	0	63
Total	73 (12%)	251 (41%)	205 (34%)	59 (10%)	22 (4%)	610

Table 9. Infant and clinical care characteristics by procedure and pain trajectory class, N=236 infants.

Procedure	Characteristic	Assigned class				
		One n=73	Two n=251	Three n=205	Four N=59	Five n=22
One	Sample size, n (%)	23 (10.0)	97 (42.4)	81 (35.4)	16 (7.0)	12 (5.2)
Two		22 (10.4)	90 (42.7)	71 (33.6)	22 (10.4)	6 (2.8)
Three		28 (16.5)	64 (37.6)	53 (31.2)	21 (12.4)	4 (2.4)
	Study group, n (%)					
One	KC and placebo	9 (11.7)	30 (39.0)	26 (33.8)	6 (7.8)	6 (7.8)
	KC and sucrose	7 (9.5)	33 (44.6)	26 (35.1)	4 (5.4)	4 (5.4)
	Sucrose	7 (9.0)	34 (43.6)	29 (37.2)	6 (7.7)	2 (2.6)
Two	KC and placebo	8 (11.4)	32 (45.7)	17 (24.3)	10 (14.3)	3 (4.3)
	KC and sucrose	5 (7.6)	28 (42.4)	27 (40.9)	4 (6.1)	2 (3.0)
	Sucrose	9 (12.0)	30 (40.0)	27 (36.0)	8 (10.7)	1 (1.3)
Three	KC and placebo	9 (16.1)	19 (33.9)	19 (33.9)	7 (12.5)	2 (3.6)
	KC and sucrose	10 (17.2)	25 (43.1)	16 (27.6)	6 (10.3)	1 (1.7)
	Sucrose	9 (16.1)	20 (35.7)	18 (32.1)	8 (14.3)	1 (1.8)
	Sex, n (%)					
One	Male	13 (10.2)	55 (43.0)	43 (33.6)	8 (6.3)	9 (7.0)
	Female	10 (9.9)	42 (41.6)	38 (37.6)	8 (7.9)	3 (3.0)
Two	Male	10 (8.3)	56 (46.7)	40 (33.3)	11 (9.2)	3 (2.5)
	Female	12 (13.2)	34 (37.4)	31 (34.1)	11 (12.1)	3 (3.3)
Three	Male	17 (17.7)	30 (31.3)	35 (36.5)	10 (10.4)	4 (4.2)
	Female	11 (14.9)	34 (45.9)	18 (24.3)	11 (14.9)	0 (0)
One	Gestational age (weeks) at birth, median (IQR)	33.1 (32.0, 35.0)	32.7 (31.3, 34.0)	32.7 (31.4, 34.1)	33.3 (32.1, 34.2)	32.9 (30.7, 35.3)
Two		32.6 (31.6, 34.6)	33.1 (31.7, 34.1)	32.4 (30.3, 33.7)	32.9 (30.6, 34.1)	32.6 (29.3, 33.7)
Three		33.0 (31.7, 34.1)	32.6 (31.5, 33.8)	32.4 (30.3, 33.7)	32.7 (30.3, 33.4)	32.6 (30.7, 33.4)
One	Gestational age (weeks) at procedure, median (IQR)	33.4 (32.5, 35.5)	33.3 (32.1, 34.4)	33.2 (32.1, 34.6)	34.0 (32.6, 34.5)	33.3 (31.4, 35.6)
Two		33.5 (32.4, 35.2)	34.0 (32.5, 35.0)	33.2 (31.5, 34.6)	33.3 (32.3, 35.1)	33.2 (31.3, 34.5)
Three		34.3 (33.4, 36.0)	34.1 (33.3, 35.1)	33.5 (32.6, 35.1)	34.4 (33.0, 35.0)	34.3 (32.4, 34.9)
One	Age (days) at procedure, median (IQR)	4.0 (3.0, 6.0)	5.0 (4.0, 7.0)	4.0 (4.0, 6.0)	5.0 (3.0, 6.0)	4.0 (3.0, 6.5)
Two		6.0 (5.0, 7.0)	7.0 (5.0, 10.0)	7.0 (5.0, 10.0)	6.0 (4.0, 12.0)	6.0 (5.0, 9.0)
Three		13.0 (6.0, 14.0)	13.0 (7.0, 14.0)	14.0 (8.0, 14.0)	14.0 (9.0, 15.0)	14.0 (11.0, 15.0)
One	Previous pain related procedures, median (IQR)	19.0 (12.0, 32.0)	20.0 (12.0, 26.0)	19.0 (14.0, 26.0)	18.5 (16.0, 26.5)	19.0 (12.5, 33.0)
Two		25.0 (22.0, 30.0)	27.0 (19.0, 41.0)	29.0 (22.0, 37.0)	25.0 (18.0, 36.0)	46.0 (25.0, 65.0)
Three		25.0 (20.5, 37.0)	24.0 (19.5, 35.0)	31.0 (22.0, 37.0)	29.0 (17.0, 38.0)	29.0 (26.5, 48.0)
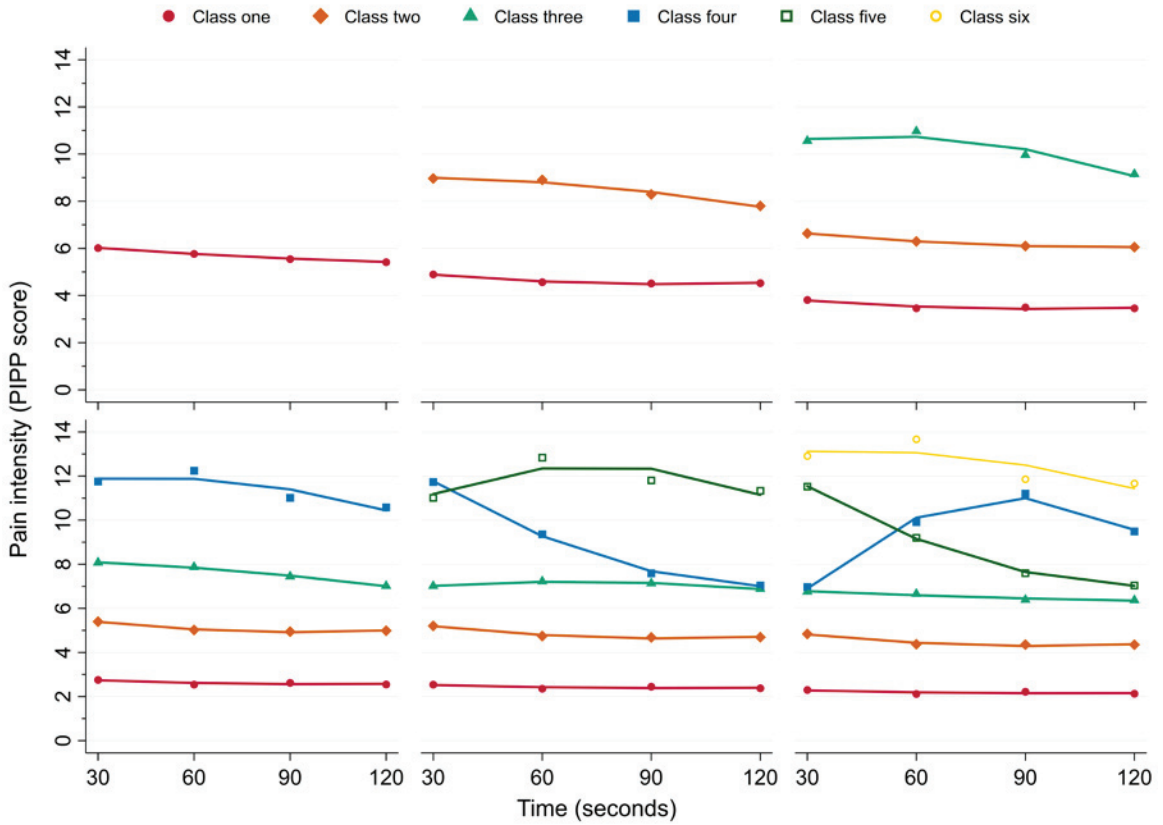
Figure 2. Initial group-based trajectory model development: Fit of models with a one through to six classes model of pain response trajectories over two-minutes post heel lance initiation in preterm infants. Estimated trajectory values (line), observed class means (points), and percent of procedures assigned to each class, N=610 procedures.
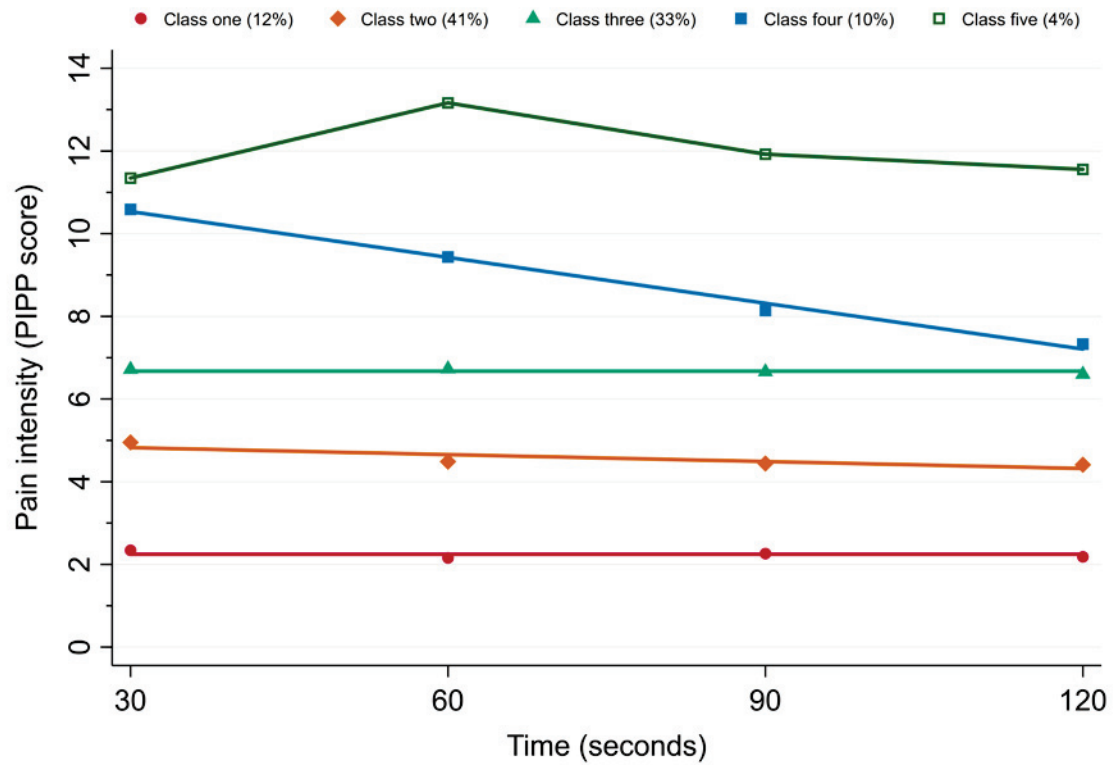
Figure 3. Pain response trajectory classes over two-minutes post heel lance initiation in preterm infants. Estimated trajectory values (line), observed class means (points), and percent of procedures assigned to each class, N=610 procedures.
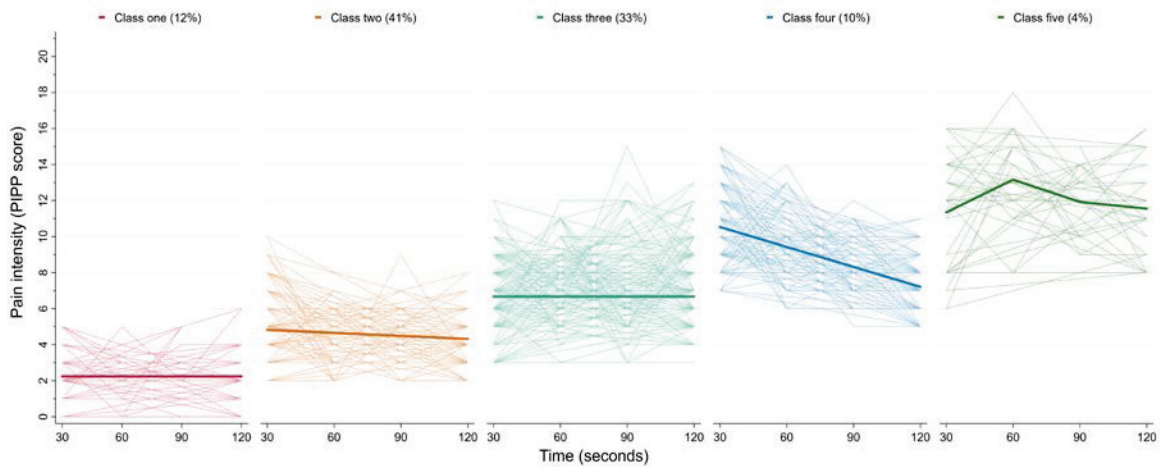
Figure 4. Pain response trajectory classes over two-minutes post heel lance initiation in preterm infants. Estimated trajectory values (thick line), individual raw values (thin lines), and percent of procedures assigned to each class, N=610 procedures.
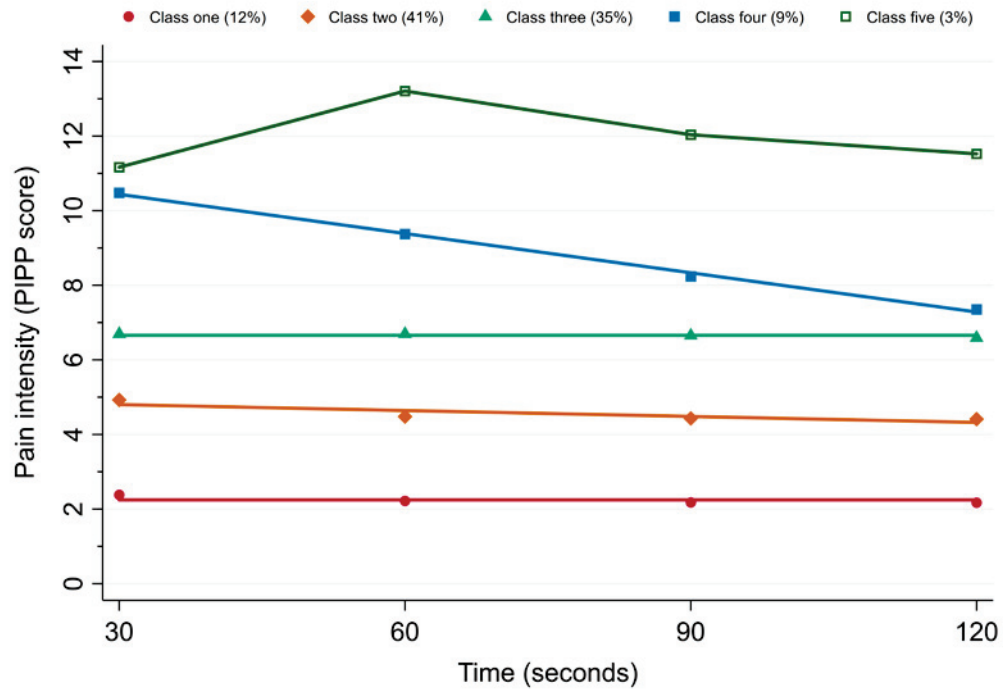
Figure 5. Sensitivity analysis with complete procedures only, pain response trajectory classes over two-minutes post heel lance initiation in preterm infants. Estimated trajectory values (line), observed class means (points), and percent of procedures assigned to each class, N=570 procedures
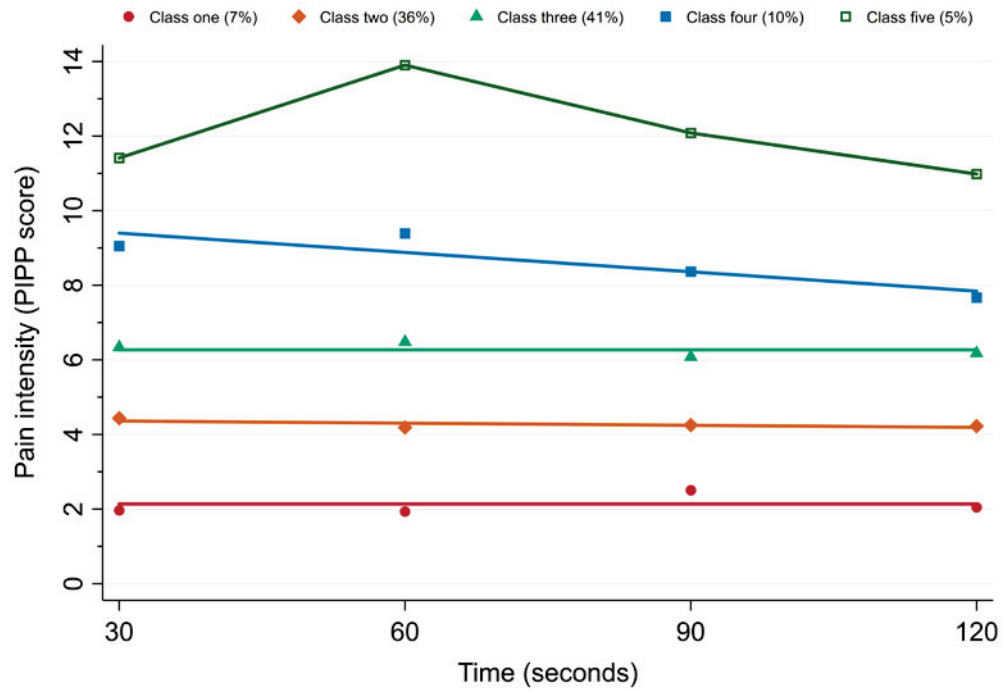
Figure 6. Sensitivity analysis using data from procedure one only, pain response trajectory classes over two-minutes post heel lance initiation in preterm infants. Estimated trajectory values (line), observed class means (points), and percent of procedures assigned to each class, N=229 infants.
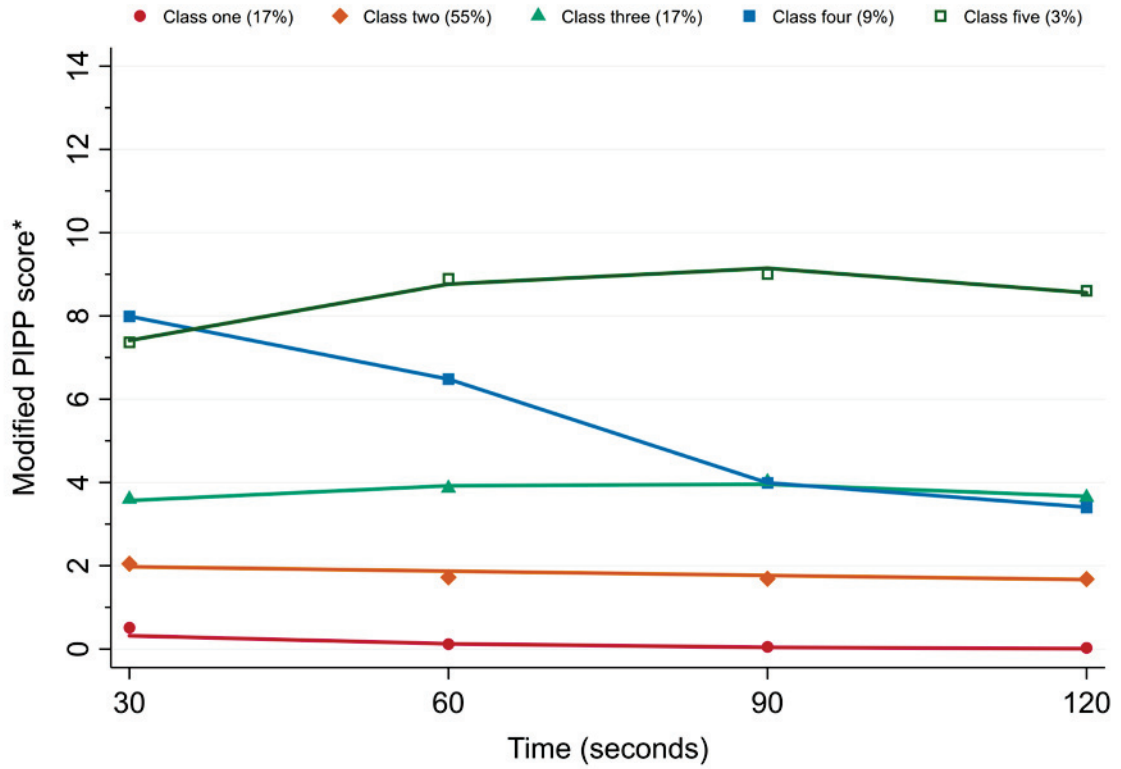
Figure 7. Exploratory analysis using a modified Premature Infant Pain Profile (PIPP) score for pain response trajectory classes over two-minutes post heel lance initiation in preterm infants. Estimated trajectory values (line), observed class means (points), and percent of procedures assigned to each class, N=610 procedures. *Modified PIPP excludes baseline state and gestational age components from final score, possible values range from 0 to 15.
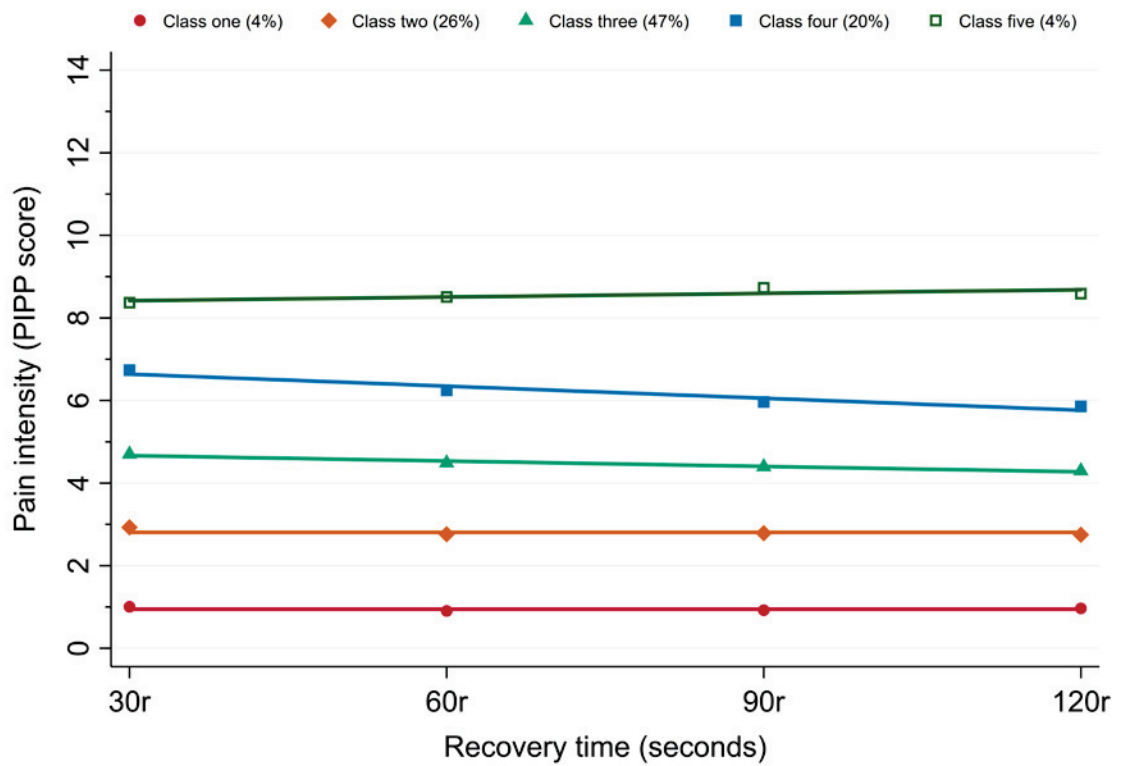
Figure 8. Exploratory analysis of pain response trajectory classes over two-minutes following heel lance recovery in preterm infants. Estimated trajectory values (line), observed class means (points), and percent of procedures assigned to each class, N=617 procedures.

CHAPTER FIVE: DISCUSSION

This chapter begins by interpreting the findings of each objective in a clinical context and in the context of existing research before weighing the strength and generalizability of the findings. This is followed by the strengths and limitations of the study and a discussion of any implications for clinical practice and future research prior to making concluding statements.

Primary Results

Sample

The demographic and clinical characteristics of the infants in the TRAKC study sample was reflective of typical preterm sample of Canadian NICU infants and other studies in this population.[2,43] Infants were predominantly moderate to late preterm, low birthweight (between 1500 and 2500 g), and more likely to be male. Maternal and family characteristics were homogenous. The majority of the sample was white, two parent families with college or university education.

Classification of Pain Response Trajectories

Five distinct trajectories best described patterns of pain response observed in the two-minutes following heel lance initiation in this sample of preterm infants in the NICU. Approximately 85% of the sample procedures aligned with three trajectory classes which indicated stable pain over the two-minutes. Varying levels of severity differentiated the classes which ranged from low-minimal to high-moderate levels of pain on the PIPP scale.[71] Two classes (class one and two), representing more than half of all procedures, contained stable pain maintained below a mean PIPP score of six, indicating minimal pain. In class three pain was also stable at a mean of approximately 6.5, indicating low moderate pain. Class four showed a linear and clinically significant reduction in pain from 30 to 120 seconds with mean pain falling from a high moderate to a low moderate category. Class four was the only class which demonstrated regulation of pain, defined in terms of a meaningful (one point) reduction in pain over the 120 seconds post heel lance initiation. The pain levels observed within classes one through to four overall indicate well managed pain and efficacious pain treatment during the heel lance.[55] Class five represented the smallest number of procedures and the highest pain scores. Mean pain was high-moderate at 30 seconds, low severe at 60 seconds, and then reduced back to the

high-moderate range from 90 to 120 seconds. The pain scores observed in class five, which remained at a high-moderate level after 120 seconds indicate that for some procedures within the class, the treatment provided could be considered to have failed to adequately manage pain intensity.[55]

Mean Pain Scores Within Trajectory Classes

There were clinically significant differences (>1 point on the PIPP) between the overall sample mean pain score and the mean within all classes at each procedure. Generally, the sample mean reflected closely the mean of one class, while all other classes were distinct from 30 to 120 seconds. This suggests subgroups within the sample which may not be reflected by reporting only the sample mean in clinical research studies.

Stability of Infants Within Classes Across Procedures

Findings indicated that most infants were not stable within a given class over multiple procedures. None of the five classes had a larger observed proportion of infants who were stable over multiple procedures, indicating that stability was not found to be related to pain severity. The lack of stability within infants may indicate that environmental or other contextual factors are contributing to the variability within infants.

Factors Associated with Pain Trajectory Class

A descriptive examination of potential infant and treatment characteristics associated with trajectory class did not identify any variables strongly associated with a given class. Despite the multisensory nature of Kangaroo Care combined with sucrose, there was not a greater proportion of infants receiving Kangaroo Care in the classes with stable mild pain, or class four which demonstrated a reduction in pain over the 120 seconds. There were also no clear patterns when examining the association between gestational age at birth or age at procedure and the trajectory classes. There was some indication that male sex may be more frequent in the high-moderate pain class (class 5), however the strength of this finding is low given the descriptive nature of the analysis and the small sample size within class five. Generally previous research has not identified a strong association between sex and pain response in premature infants.[8] The finding is supported by a study which found male infants, while overall having similar pain, demonstrated increased heart rate following a painful procedure and a corresponding

larger change in heart rate from baseline to post-procedure.[66] But research on sex differences in pain is not conclusive and males and females may communicate pain in different ways, resulting in varying research findings.[90] The number of previous pain related procedures was also not clearly associated with a given class. A similar number of previous pain related procedures was observed across classes and procedures, despite the broad range in previous pain exposure found in the sample. Overall, caution should be applied when interpreting these findings as the analysis was descriptive in nature and limited by the sample size within classes at a given procedure time point. The conflicting nature of the existing literature related to factors associated with pain variability further limits potential interpreations.[8]

Missing Data and Exploratory Findings

An examination of missing data found minimal impact of incorporating the procedures with missing epoch times within the trajectory models. Furthermore, similar trajectories were identified when fitted using data from only one procedure, supporting the use of multiple procedures from each infant within the models for the primary results.

An exploratory analysis on the PIPP scores calculated from when the nurse finished all clinical procedures and had applied a bandage to the heel lance location found that pain was highly stable over this recovery period, with all classes generally showing stable pain. Three classes representing 80% of the sample were in the minimal PIPP score range while class four and five showed low and mid-moderate mean pain respectively. None of the classes demonstrated further regulation in pain during this period. While infants did not stay in the same numbered class (e.g. stay within class four) from the primary heel lance classes to the recovery period classes, they did generally stay within a class that fell within the same range of PIPP scores (e.g. minimal pain). Overall during recovery pain was highly stable and the classes reflected lower levels of pain when compared to the initial heel lance trajectories.

An exploratory analysis examining the sensitivity of the PIPP score to change over the 120 second period was done to better understand the impact of the constant values in the PIPP score contributed by gestational age and baseline state. Removing these factors from the PIPP score resulted in trajectories that reflected patterns similar to those found in the primary analysis. Although the modified PIPP scores cannot directly be

interpreted as pain intensity, the trends reflected similar patterns over the 120 seconds. These findings indicate that the high level of stable low pain trajectories identified in this research cannot be attributed to the constant components of the PIPP score alone.

Results in the Context of Other Research

The high proportion of mild pain trajectories observed in this sample support the primary efficacy results reported for the TRAKC study which found that Kangaroo Care remains efficacious over time and for repeated painful procedures and is comparable to 24% oral sucrose.[44] But there may not be added benefit of combining the interventions, based on the primary results and no clear association between trajectory class and TRAKC intervention group.

While no directly comparable research has been conducted in preterm infants, trajectories identified in healthy term infants yielded different findings.[9] In term infants from two to 12 months of age receiving routine immunization, between three and six trajectories were identified. In contrast to this study, the trajectories in term infants showed increased levels of pain regulation with the majority of the sample at each age group showing a clinically meaningful (One point on the MBPS) regulation in pain by the two-minute mark. The study found that regulation increased with age. Several factors potentially contribute to the difference in findings. The immunization needle is a more acute painful stimulus compared to the more sustained stimulus in the heel stick and no consistent pain relieving intervention was provided. This is evident in the higher initial pain identified in the term sample, with most trajectories showing an initial pain rating ranging between 8 and 10 (out of a highest possible score of 10). Given the higher initial pain, more regulation is expected. As well, there are differences expected due to the different age of the infants at the procedure, as well as the differences in age at birth. It is possible that prematurity increases the time to regulation due to reduced inhibitory pain control.[91] Yet due to the many other differences between studies, it is difficult to determine the amount of these differences which is due to prematurity alone. Further directly comparable research is needed in term and preterm infants at similar time points to more clearly elucidate differences and similarities. The two studies did identify a similar trend with overall sample mean pain scores being clinically significantly different

compared to those within trajectory classes, lending support to the strength of this finding.

Previous research has recognised the need to investigate within infant pain variability.[9,44] Existing studies have identified high levels of variability in preterm infant pain response over multiple procedures.[10,92] Our findings add to this body of evidence. Overall less than 10% of infants maintained the same trajectory class for all three procedures, with infants often assigned to a different class for each procedure. This finding, combined with the lack of clear association between hypothesized demographic and clinical factors and pain trajectory indicates that additional influences such as environmental factors which were not measured in the current research may play a key role in infant pain variability. Previous research has demonstrated a relationship between noise and light reduction and physiological reactivity in newborn infants.[69,93,94] Unfortunately it is rarely measured in clinical studies examining the efficacy of pain relieving interventions.

The GBTM approach used produced five distinct trajectories. Other methods which allow for within class variation may have produced fewer or differing trajectories.[95] This is an important consideration when comparing this research to other trajectory modeling studies which may use different methodology. Notably, the trajectories developed in term infants used an approach allowing for within class variation and based model selection on both the BIC and AIC values.

Strengths and Limitations

To our knowledge this was the first study to classify trajectories of pain over time in preterm infants. This study is based on data from an RCT with rigorous controls and consistent procedures between clinicians using standardized and documented pain care. Additionally, pain assessments were done using video tapes of the procedures by trained, reliable coders. We used a transparent approach to trajectory development incorporating both clinical and statistical criteria in model development and reported details to ensure a reproducible modeling approach.

The study had limits in terms of generalizability. The results reflect those of a single sample from one NICU. Data reflects infants who were undergoing procedures where pain was generally well managed. A PIPP score of thirteen or more (severe pain)

was only observed in approximately ten percent of procedures recorded.[44] As a result, the pain trajectories identified in this study are most generalizable to lower pain exposure and may not reflect the variability in trajectories identified in a population exposed to more severe or unmanaged pain. A similar result can be expected in other Canadian NICUs where pain management is improving and few procedures that involve high pain are indicated in most preterm infants.[15,52,96]

Additionally, future research which does not provide adequate pain relief during procedures is not ethical and should not be conducted.[15,97] Despite this, numerous RCTs have been conducted in recent years without adequate pain control, where mean PIPP scores are clinically significantly higher in the no treatment group.[3,97–99] Likely in these samples with unmanaged pain, or in samples where more severe painful procedures are conducted, different pain response trajectories could be identified. In order to better understand the range of possible pain trajectories in preterm infants, research which (ethically) examines trajectories of pain response in samples of infants with a greater variance in pain levels, including more severe pain, is needed.

The primary infant pain outcome used in this research was the PIPP. The PIPP is widely considered to be the reference standard for pain intensity in infants.[5,71] The PIPP was more recently updated to the PIPP-R, which modified the scoring related to the constant score components (gestational age, behavioural state). While this study used the PIPP, our exploratory analysis removing the constant factors did not result in altering the identified trajectories. The PIPP is also a multidimensional measure of pain. Conflicting findings related to the relationship of contextual factors and pain are often linked to whether a physiologic or behavioural measure of pain is used.[8] It is possible different trajectories would have been identified using a different measurement of pain. Based on the greater variability observed in physiologic measures, compared to the relative stability in preterm infant facial and other behavioural indicators, it is possible that physiologic measures alone may yield trajectories with greater variability and change over time compared to behavioural ones.[100]

The study did not collect data on the environmental context during pain exposure, such as noise and light, which are theorized to impact pain. As such, we were unable to

determine whether these factors contribute to variability in pain response over time. The study was also limited by sample size.

A desired aim of the research was to analyze the factors hypothesized to be associated with pain trajectory classes in a regression model. Despite a sample of more than 200 infants, it was not feasible given the sample within a given class for a given variable. For example, there were no female infants within class five at procedure three. As a result, all factors were examined at a descriptively. This limited the strength of conclusions which could be made related to this research aim. The descriptive results are nevertheless valuable in terms of indicating areas of focus for further analytical research in larger samples.

Implications for Research and Clinical Practice

Given that this was the first study to identify pain trajectories in preterm infants, more research is needed to extend and replicate the findings. The findings suggest that the mean pain score in the sample does not adequately reflect the variability in pain trajectories across the sample. Clinical research studies examining the efficacy of pain relieving interventions in infants should examine results within subgroups of pain scores to identify if meaningful differences in intervention effectiveness are present. It is possible that treatment effects for the sample may not be representative of the subgroups within the trajectory classes.

For both clinical care and research the identified trajectories in preterm infants demonstrated that with in a sample with well managed pain, pain trajectories will be highly stable at 30, 60, 90 and 120 seconds. As such, defining the reactivity period as the first 30 seconds following initiation of a painful stimulus such as heel lances may not be appropriate.[7] In order to measure regulation and return to baseline values, studies should continue to measure the outcome several minutes past initiation of the procedure. Similarly, clinicians should ensure infants are monitored and interventions for pain relief extend beyond this period.

Infants were not stable within a pain trajectory class across procedures. As a result, clinically it is important to note that an infant's pain response at a given procedure may not indicate a similar response at subsequent procedures despite the provision of similar pain relieving intervention. This study did not identify contextual factors

associated with different trajectories, as such it does not provide a rationale for considering different pain relieving interventions within different subgroups of preterm infants.

The findings demonstrating low and moderate pain across TRAKC study interventions further support the use of Kangaroo Care as an alternative to 24% oral sucrose during minor painful procedures in preterm infants. Critically, there was still a subset of procedures for which the provided pain treatment did not adequately manage pain over the observed trajectory. Accordingly, research which aims to understand the factors contributing to greater pain in this subgroup is critical to better understanding how to manage and treat pain for all infants.

Conclusion and Future Directions

The results of this study support the feasibility of developing distinct classes of pain response trajectories in preterm infants. The pain trajectories identified in this research demonstrate that in a sample of well managed pain, trajectories are highly stable over the two-minutes post heel lance. There is a need for further research in larger samples to replicate these findings and extend on the results by exploring the effect of contextual factors from an analytical approach. Further research would benefit from the measurement of environmental factors such as noise and light at the time of painful procedures. As well, given the stable nature of the identified trajectories over 120 seconds, increasing the monitoring time post painful procedure would be beneficial in order to extend the trajectories to allow for additional time to observe return to baseline pain values. The field would benefit from additional research which examines a variety of pain measurement including both physiologic and behavioural measures to better understand their influence on trajectories. The study results highlight the difficulty of measuring pain in preterm infants. Continued research will advance the field and lead to improved pain care and management during painful procedures.

REFERENCES

1.   Ranger M, Grunau RE. Early repetitive pain in preterm infants in relation to the developing brain. *Pain Manag*. 2014;4(1):57-67. doi:10.2217/pmt.13.61

2.   Lee SK, McMillan DD, Ohlsson A, et al. Variations in practice and outcomes in the Canadian NICU network: 1996-1997. *Pediatrics*. 2000;106(5):1070-1079. doi:10.1542/peds.106.5.1070

3.   Cruz MD, Fernandes AM, Oliveira CR. Epidemiology of painful procedures performed in neonates: A systematic review of observational studies. *Eur J Pain Lond Engl*. 2016;20(4):489-498. doi:10.1002/ejp.757

4.   McNair C, Campbell-Yeo M, Johnston C, Taddio A. Nonpharmacologic Management of Pain During Common Needle Puncture Procedures in Infants: Current Research Evidence and Practical Considerations: An Update. *Clin Perinatol*. 2019;46(4):709-730. doi:10.1016/j.clp.2019.08.006

5.   American Academy of Pediatrics, and Fetus and Newborn Committee. Prevention and Management of Procedural Pain in the Neonate: An Update. *Pediatrics*. Published online February 1, 2016:peds.2015-4271. doi:10.1542/peds.2015-4271

6.   Oberlander TF, Grunau RE, Whitfield MF, Fitzgerald C, Pitfield S, Saul JP. Biobehavioral Pain Responses in Former Extremely Low Birth Weight Infants at Four Months' Corrected Age. *Pediatrics*. 2000;105(1):e6-e6. doi:10.1542/peds.105.1.e6

7.   Riddell RRP, Racine NM, Turcotte K, et al. Non-pharmacological management of infant and young child procedural pain. *Cochrane Database Syst Rev*. 2011;(10). doi:10.1002/14651858.CD006275.pub2

8.   Sellam G, Cignaccol EL, Craigl KD, Engbergl S. Review: Contextual factors influencing pain response to heelstick procedures in preterm infants: What do we know? A systematic review. *Eur J Pain*. 2011;15(7):661.e1-661.e15. doi:10.1016/j.ejpain.2011.01.002

9.   Pillai Riddell R, Flora DB, Stevens SA, et al. Variability in infant acute pain responding meaningfully obscured by averaging pain responses. *Pain*. 2013;154:714-721. doi:http://dx.doi.org/10.1016/j.pain.2013.01.015

10. Cignacco E, Denhaerynck K, Nelle M, Bührer C, Engberg S. Variability in pain response to a non-pharmacological intervention across repeated routine pain exposure in preterm infants: a feasibility study. *Acta Paediatr*. 2009;98(5):842-846. doi:10.1111/j.1651-2227.2008.01203.x

11. Barker DP, Rutter N. Exposure to invasive procedures in neonatal intensive care unit admissions. *Arch Dis Child - Fetal Neonatal Ed*. 1995;72(1):F47-F48. doi:10.1136/fn.72.1.F47

12. Shah VS, Ohlsson A. Venepuncture versus heel lance for blood sampling in term neonates. *Cochrane Database Syst Rev*. 2011;(10). doi:10.1002/14651858.CD001452.pub4

13. Walker SM, Melbourne A, O'Reilly H, et al. Somatosensory function and pain in extremely preterm young adults from the UK EPICure cohort: sex-dependent differences and impact of neonatal surgery. *Br J Anaesth*. 2018;121(3):623-635. doi:10.1016/j.bja.2018.03.035

14. Anand KJS. Clinical Importance of Pain and Stress in Preterm Neonates. *Neonatology*. 1998;73(1):1-9. doi:10.1159/000013953

15. Campbell-Yeo M. 'First, do no harm' – the use of analgesia or placebo as control for babies in painful clinical trials. *Acta Paediatr*. 2016;105(2):119-120. doi:10.1111/apa.13255

16. Taddio A, Katz J, Ilersich AL, Koren G. Effect of neonatal circumcision on pain response during subsequent routine vaccination. *The Lancet*. 1997;349(9052):599-603. doi:10.1016/S0140-6736(96)10316-0

17. Taddio A, Shah V, Gilbert-MacLeod C, Katz J. Conditioning and Hyperalgesia in Newborns Exposed to Repeated Heel Lances. *JAMA*. 2002;288(7):857-861. doi:10.1001/jama.288.7.857

18. Vinall J, Noel M, Disher T, Caddell K, Campbell-Yeo M. Memories of Infant Pain in the Neonatal Intensive Care Unit Influence Posttraumatic Stress Symptoms in Mothers of Infants Born Preterm. doi:info:doi/10.1097/AJP.0000000000000620

19. Vinall J, Grunau RE. Impact of repeated procedural pain-related stress in infants born very preterm. *Pediatr Res*. 2014;75(5):584-587. doi:10.1038/pr.2014.16

20. Walker SM. Early life pain—effects in the adult. *Curr Opin Physiol*. 2019;11:16-24. doi:10.1016/j.cophys.2019.04.011

21. Brummelte S, Grunau RE, Chau V, et al. Procedural pain and brain development in premature newborns. *Ann Neurol*. 2012;71(3):385-396. doi:10.1002/ana.22267

22. Duerden EG, Grunau RE, Guo T, et al. Early Procedural Pain Is Associated with Regionally-Specific Alterations in Thalamic Development in Preterm Neonates. *J Neurosci*. 2018;38(4):878-886. doi:10.1523/JNEUROSCI.0867-17.2017

23. Schneider J, Duerden EG, Guo T, et al. Procedural pain and oral glucose in preterm neonates: brain development and sex-specific effects. *PAIN*. 2018;159(3):515–525. doi:10.1097/j.pain.0000000000001123

24. Walker SM, O'Reilly H, Beckmann J, Marlow N. Conditioned pain modulation identifies altered sensitivity in extremely preterm young adult males and females. *Br J Anaesth*. 2018;121(3):636-646. doi:10.1016/j.bja.2018.05.066

25. Chau CMY, Ranger M, Bichin M, et al. Hippocampus, Amygdala, and Thalamus Volumes in Very Preterm Children at 8 Years: Neonatal Pain and Genetic Variation. *Front Behav Neurosci*. 2019;13. doi:10.3389/fnbeh.2019.00051

26. Brummelte S, Chau CMY, Cepeda IL, et al. Cortisol levels in former preterm children at school age are predicted by neonatal procedural pain-related stress. *Psychoneuroendocrinology*. 2015;51:151-163. doi:10.1016/j.psyneuen.2014.09.018

27. Ranger M, Chau CMY, Garg A, et al. Neonatal pain-related stress predicts cortical thickness at age 7 years in children born very preterm. *PloS One*. 2013;8(10):e76702. doi:10.1371/journal.pone.0076702

28. Walker SM, Franck LS, Fitzgerald M, Myles J, Stocks J, Marlow N. Long-term impact of neonatal intensive care and surgery on somatosensory perception in children born extremely preterm. *Pain*. 2009;141(1-2):79-87. doi:10.1016/j.pain.2008.10.012

29. Williams AC de C, Craig KD. Updating the definition of pain: *PAIN*. 2016;157(11):2420-2423. doi:10.1097/j.pain.0000000000000613

30. Raja SN, Carr DB, Cohen M, et al. The revised International Association for the Study of Pain definition of pain: concepts, challenges, and compromises. *Pain*. Published online May 23, 2020. doi:10.1097/j.pain.0000000000001939

31. Fitzgerald M. The development of nociceptive circuits. *Nat Rev Neurosci*. 2005;6(7):507-520. doi:10.1038/nrn1701

32. Hatfield LA. Neonatal pain: What's age got to do with it? *Surg Neurol Int*. 2014;5(Suppl 13):S479-S489. doi:10.4103/2152-7806.144630

33. Fitzgerald M. What do we really know about newborn infant pain? *Exp Physiol*. Published online November 18, 2015:1451-1457. doi:10.1113/EP085134@10.1002/(ISSN)1469-445X(CAT)VirtualIssues(VI)womenthemedpreview

34. Verriotis M, Chang P, Fitzgerald M, Fabrizi L. The development of the nociceptive brain. *Neuroscience*. 2016;338:207-219. doi:10.1016/j.neuroscience.2016.07.026

35. Stevens B, Johnston C, Petryshen P, Taddio A. Premature Infant Pain Profile: development and initial validation. *Clin J Pain*. 1996;12(1):13-22.

36. Ahola Kohut S, Pillai Riddell R, Flora DB, Oster H. A longitudinal analysis of the development of infant facial expressions in response to acute pain: Immediate and regulatory expressions. *PAIN®*. 2012;153(12):2458-2465. doi:10.1016/j.pain.2012.09.005

37. Shah PS, Herbozo C, Aliwalas LL, Shah VS. Breastfeeding or breast milk for procedural pain in neonates. *Cochrane Database Syst Rev*. 2012;12:CD004950. doi:10.1002/14651858.CD004950.pub3

38. Walco GA, Kopecky EA, Weisman SJ, et al. Clinical trial designs and models for analgesic medications for acute pain in neonates, infants, toddlers, children, and adolescents: ACTTION recommendations. *PAIN*. 2018;159(2):193. doi:10.1097/j.pain.0000000000001104

39. Craig KD. The facial expression of pain Better than a thousand words? *APS J*. 1992;1(3):153-162. doi:10.1016/1058-9139(92)90001-S

40. Duhn LJ, Medves JM. A systematic integrative review of infant pain assessment tools. *Adv Neonatal Care Off J Natl Assoc Neonatal Nurses*. 2004;4(3):126-140.

41. Jones L, Laudiano-Dray MP, Whitehead K, et al. EEG, behavioural and physiological recordings following a painful procedure in human neonates. *Sci Data*. 2018;5. doi:10.1038/sdata.2018.248

42. Axelin A, Salanterä S, Kirjavainen J, Lehtonen L. Oral Glucose and Parental Holding Preferable to Opioid in Pain Management in Preterm Infants. *Clin J Pain*. 2009;25(2):138–145. doi:10.1097/AJP.0b013e318181ad81

43. Johnston C, Campbell-Yeo M, Disher T, et al. Skin-to-skin care for procedural pain in neonates. *Cochrane Database Syst Rev*. 2017;(2). doi:10.1002/14651858.CD008435.pub3

44. Campbell-Yeo M, Johnston CC, Benoit B, et al. Sustained efficacy of kangaroo care for repeated painful procedures over neonatal intensive care unit hospitalization: a single-blind randomized controlled trial. *Pain*. 2019;160(11):2580-2588. doi:10.1097/j.pain.0000000000001646

45. Goubet N, Rattaz C, Pierrat V, Bullinger A, Lequien P. Olfactory experience mediates response to pain in preterm newborns. *Dev Psychobiol*. 2003;42(2):171-180. doi:10.1002/dev.10085

46. Matthiesen A-S, Ransjö-Arvidson A-B, Nissen E, Uvnäs-Moberg K. Postpartum Maternal Oxytocin Release by Newborns: Effects of Infant Hand Massage and Sucking. *Birth*. 2001;28(1):13-19. doi:10.1046/j.1523-536x.2001.00013.x

47. Campbell-Yeo M, Johnston C, Benoit B, et al. Trial of repeated analgesia with Kangaroo Mother Care (TRAKC Trial). *BMC Pediatr*. 2013;13:182. doi:10.1186/1471-2431-13-182

48. Melzack R, Wall PD. Pain mechanisms: a new theory. *Science*. 1965;150(3699):971–979.

49. Kostandy R, Cong X, Abouelfettoh A, Bronson C, Stankus A, Ludington SM. Effect of Kangaroo Care (skin contact) on crying response to pain in preterm neonates. *Pain Manag Nurs Off J Am Soc Pain Manag Nurses*. 2008;9(2):55-65. doi:10.1016/j.pmn.2007.11.004

50. Riddell RP, Racine N, Turcotte K, et al. Nonpharmacological management of procedural pain in infants and young children: An abridged Cochrane review. *Pain Res Manag J Can Pain Soc*. 2011;16(5):321-330.

51. Sleuwen BE van, Engelberts AC, Boere-Boonekamp MM, Kuis W, Schulpen TWJ, L'Hoir MP. Swaddling: A Systematic Review. *Pediatrics*. 2007;120(4):e1097-e1106. doi:10.1542/peds.2006-2083

52. Stevens BJ, Harrison D, Rashotte J, et al. Pain assessment and intensity in hospitalized children in Canada. *J Pain Off J Am Pain Soc*. 2012;13(9):857-865. doi:10.1016/j.jpain.2012.05.010

53. Stevens B, Yamada J, Ohlsson A, Haliburton S, Shorkey A. Sucrose for analgesia in newborn infants undergoing painful procedures. *Cochrane Database Syst Rev*. 2016;(7). doi:10.1002/14651858.CD001069.pub5

54. Bueno M, Yamada J, Harrison D, et al. A systematic review and meta-analyses of nonsucrose sweet solutions for pain relief in neonates. *Pain Res Manag J Can Pain Soc*. 2013;18(3):153-161.

55. Stevens B, Yamada J, Campbell-Yeo M, et al. The minimally effective dose of sucrose for procedural pain relief in neonates: a randomized controlled trial. *BMC Pediatr*. 2018;18(1):85. doi:10.1186/s12887-018-1026-x

56. Johnston CC, Filion F, Snider L, et al. How Much Sucrose Is Too Much Sucrose? *Pediatrics*. 2007;119(1):226-226. doi:10.1542/peds.2006-3001

57. Harrison D, Beggs S, Stevens B. Sucrose for Procedural Pain Management in Infants. *Pediatrics*. 2012;130(5):918-925. doi:10.1542/peds.2011-3848

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

59. Walco GA. Needle Pain in Children: Contextual Factors. *Pediatrics*. 2008;122(Supplement 3):S125-S129. doi:10.1542/peds.2008-1055D

60. Hatfield LA, Ely EA. Measurement of Acute Pain in Infants: A Review of Behavioral and Physiological Variables. *Biol Res Nurs*. 2015;17(1):100-111. doi:10.1177/1099800414531448

61. Sellam G, Engberg S, Denhaerynck K, Craig KD, Cignacco EL. Contextual factors associated with pain response of preterm infants to heel-stick procedures. *Eur J Pain*. 2013;17(2):255-263. doi:10.1002/j.1532-2149.2012.00182.x

62. Slater R, Cantarella A, Yoxen J, et al. Latency to facial expression change following noxious stimulation in infants is dependent on postmenstrual age. *Pain*. 2009;146(1-2):177-182. doi:10.1016/j.pain.2009.07.022

63. Schenk K, Stoffel L, Bürgin R, et al. The influence of gestational age in the psychometric testing of the Bernese Pain Scale for Neonates. *BMC Pediatr*. 2019;19(1):20. doi:10.1186/s12887-018-1380-8

64. Gibbins S, Stevens B, McGrath PJ, et al. Comparison of Pain Responses in Infants of Different Gestational Ages. *Neonatol Basel*. 2007;93(1):10-18.

65. Johnston CC, Stevens BJ, Franck LS, Jack A, Stremler R, Platt R. Factors Explaining Lack of Response to Heel Stick in Preterm Newborns. *J Obstet Gynecol Neonatal Nurs*. 1999;28(6):587-594. doi:10.1111/j.1552-6909.1999.tb02167.x

66. Valeri BO, Gaspardo CM, Martinez FE, Linhares MBM. Pain reactivity in preterm neonates: examining the sex differences. *Eur J Pain*. 2014;18(10):1431-1439. doi:10.1002/ejp.508

67. Guinsburg R, Peres C de A, Branco de Almeida MF, et al. Differences in pain expression between male and female newborn infants. *Pain*. 2000;85(1):127-133. doi:10.1016/S0304-3959(99)00258-4

68. Butt ML, Kisilevsky BS. Music modulates behaviour of premature infants following heel lance. *Can J Nurs Res Rev Can Rech En Sci Infirm*. 2000;31(4):17-39.

69. Slevin M, Farrington N, Duffy G, Daly L, Murphy JFA. Altering the NICU and measuring infants' responses. *Acta Paediatr*. 2000;89(5):577-581. doi:10.1111/j.1651-2227.2000.tb00342.x

70. Richardson DK, Corcoran JD, Escobar GJ, Lee SK. SNAP-II and SNAPPE-II: Simplified newborn illness severity and mortality risk scores. *J Pediatr*. 2001;138(1):92-100. doi:10.1067/mpd.2001.109608

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

72. Hundert AS, Campbell-Yeo M, Brook HR, Wozney LM, O'Connor K. Development and Usability Evaluation of a Desktop Software Application for Pain Assessment in Infants. *Can J Pain*. 2018;2(1):302-314. doi:10.1080/24740527.2018.1540261

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

74. Stevens BJ, Gibbins S, Yamada J, et al. The premature infant pain profile-revised (PIPP-R): initial validation and feasibility. *Clin J Pain*. 2014;30(3):238-243. doi:10.1097/AJP.0b013e3182906aed

75. Fleiss JL, Cohen J. The Equivalence of Weighted Kappa and the Intraclass Correlation Coefficient as Measures of Reliability. *Educ Psychol Meas*. 1973;33(3):613-619. doi:10.1177/001316447303300309

76. Nagin DS, Odgers CL. Group-Based Trajectory Modeling in Clinical Research. *Annu Rev Clin Psychol*. 2010;6(1):109-138. doi:10.1146/annurev.clinpsy.121208.131413

77. Mirza SS, Wolters FJ, Swanson SA, et al. 10-year trajectories of depressive symptoms and risk of dementia: a population-based study. *Lancet Psychiatry*. 2016;3(7):628-635. doi:10.1016/S2215-0366(16)00097-3

78. Hao Guang, Wang Xiaoling, Treiber Frank A., Harshfield Gregory, Kapuku Gaston, Su Shaoyong. Blood Pressure Trajectories From Childhood to Young Adulthood Associated With Cardiovascular Risk. *Hypertension*. 2017;69(3):435-442. doi:10.1161/HYPERTENSIONAHA.116.08312

79. Twisk J, Hoekstra T. Classifying developmental trajectories over time should be done with great caution: a comparison between methods. *J Clin Epidemiol*. 2012;65(10):1078-1087. doi:10.1016/j.jclinepi.2012.04.010

80. Nagin DS, Nagin T and HJHIUP of PP and SDS, NAGIN D. *Group-Based Modeling of Development*. Harvard University Press; 2005.

81. Feldman DE, de Civita M, Dobkin PL, Malleson P, Meshefedjian G, Duffy CM. Perceived adherence to prescribed treatment in juvenile idiopathic arthritis over a one-year period. *Arthritis Rheum*. 2007;57(2):226-233. doi:10.1002/art.22534

82. Schoot R van de, Sijbrandij M, Winter SD, Depaoli S, Vermunt JK. The GRoLTS-Checklist: Guidelines for Reporting on Latent Trajectory Studies. *Struct Equ Model Multidiscip J*. 2017;24(3):451-467. doi:10.1080/10705511.2016.1247646

83. Vrieze SI. Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychol Methods*. 2012;17(2):228-243. doi:10.1037/a0027127

84. Nylund KL, Asparouhov T, Muthén BO. Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. *Struct Equ Model Multidiscip J*. 2007;14(4):535-569. doi:10.1080/10705510701575396

85. Jones BL, Nagin DS. A Note on a Stata Plugin for Estimating Group-based Trajectory Models. *Sociol Methods Res*. 2013;42(4):608-613. doi:10.1177/0049124113503141

86. Coulombe P, Selig JP, Delaney HD. Ignoring individual differences in times of assessment in growth curve modeling. *Int J Behav Dev*. 2016;40(1):76-86. doi:10.1177/0165025415577684

87. Busse JW, Bartlett SJ, Dougados M, et al. Optimal Strategies for Reporting Pain in Clinical Trials and Systematic Reviews: Recommendations from an OMERACT 12 Workshop. *J Rheumatol*. 2015;42(10):1962-1970. doi:10.3899/jrheum.141440

88. Stevens B, Gibbins S. Clinical utility and clinical significance in the assessment and management of pain in vulnerable infants. *Clin Perinatol*. 2002;29(3):459-468. doi:10.1016/S0095-5108(02)00016-7

89. Altman DG, Schulz KF, Moher D, et al. The Revised CONSORT Statement for Reporting Randomized Trials: Explanation and Elaboration. *Ann Intern Med*. 2001;134(8):663-694. doi:10.7326/0003-4819-134-8-200104170-00012

90. Keogh E. Gender differences in the nonverbal communication of pain: A new direction for sex, gender, and pain research? *PAIN*. 2014;155(10):1927–1931. doi:10.1016/j.pain.2014.06.024

91. Walker SM. Neonatal pain. *Paediatr Anaesth*. 2014;24(1):39-48. doi:10.1111/pan.12293

92. Franck LS, Ridout D, Howard R, Peters J, Honour JW. A comparison of pain measures in newborn infants after cardiac surgery. *Pain*. 2011;152(8):1758-1765. doi:10.1016/j.pain.2011.03.017

93. Johnson AN. Practice applications of research. Neonatal response to control of noise inside the incubator. Hayes JS, ed. *Pediatr Nurs*. 2001;27(6):600-605.

94. Trapanotto M, Benini F, Farina M, Gobber D, Magnavita V, Zacchello F. Behavioural and physiological reactivity to noise in the newborn. *J Paediatr Child Health*. 2004;40(5-6):275-281. doi:10.1111/j.1440-1754.2004.00363.x

95. Erosheva EA, Matsueda RL, Telesca D. Breaking Bad: Two Decades of Life-Course Data Analysis in Criminology, Developmental Psychology, and Beyond. *Annu Rev Stat Its Appl*. 2014;1(1):301-332. doi:10.1146/annurev-statistics-022513-115701

96. Orovec A, Disher T, Caddell K, Campbell-Yeo M. Assessment and Management of Procedural Pain During the Entire Neonatal Intensive Care Unit Hospitalization. *Pain Manag Nurs Off J Am Soc Pain Manag Nurses*. 2019;20(5):503-511. doi:10.1016/j.pmn.2018.11.061

97. Bellieni CV, Johnston CC. Analgesia, nil or placebo to babies, in trials that test new analgesic treatments for procedural pain. *Acta Paediatr Oslo Nor 1992*. Published online September 20, 2015. doi:10.1111/apa.13210

98. Nimbalkar S, Sinojia A, Dongara A. Reduction of Neonatal Pain Following Administration of 25% Lingual Dextrose: A Randomized Control Trial. *J Trop Pediatr*. 2013;59(3):223-225. doi:10.1093/tropej/fms072

99. Sundaram B, Shrivastava S, Pandian JS, Singh VP. Facilitated tucking on pain in pre-term newborns during neonatal intensive care: a single blinded randomized controlled cross-over pilot trial. *J Pediatr Rehabil Med*. 2013;6(1):19-27. doi:10.3233/PRM-130233

100. O'Neill MC, Kohut SA, Riddell RP, Oster H. Age-related differences in the acute pain facial expression during infancy. *Eur J Pain*. 2019;23(9):1596-1607. doi:10.1002/ejp.1436

APPENDIX A. Premature Infant Pain Profile Scoring

Table 1. Scoring method for the PIPP.

Process and item	Values and scoring			
Obtained from chart				
Gestational age in weeks	≥ 36	32-35	28-31	<28
Baseline observation, 15 seconds				
Behavioural state	Active awake	Quiet awake	Active sleep	Quiet sleep
Maximum heart rate value				
Minimum oxygen saturation value				
Observation, 30 seconds				
Increase in heart rate beats per minute from measured baseline maximum	0-4	5-14	15-24	≥ 25
Decrease in oxygen saturation from measured baseline minimum	0-2.4	2.5-4.9	5.0-7.4	≥ 7.5
Brow bulge, % time	0-9	10-39	40-69	≥70
Eye squeeze, % time	0-9	10-39	40-69	≥70
Nasolabial furrow, % time	0-9	10-39	40-69	≥70
Scoring*				
Item score	0	1	2	3

*Score is summed across all items for a total score from 0 to 21, with higher values indicating increased pain.

PIPP scoring developed by: Stevens B, Johnston C, Petryshen P, Taddio A. Premature Infant Pain Profile: development and initial validation. Clin J Pain. 1996;12(1):13-22

APPENDIX B. Guidelines for Reporting on Latent Trajectory Studies

Table 1. Final List of Items of the Guidelines for Reporting on Latent Trajectory Studies (GRoLTS) Checklist: Guidelines for Reporting on Latent Trajectory Studies.

#	Checklist Item
1	Is the metric of time used in the statistical model reported?
2	Is information presented about the mean and variance of time within a wave?
3a	Is the missing data mechanism reported?
3b	Is a description provided of what variables are related to attrition/missing data?
3c	Is a description provided of how missing data in the analyses were dealt with?
4	Is information about the distribution of the observed variables included?
5	Is the software mentioned?
6a	Are alternative specifications of within-class heterogeneity considered (e.g., LGCA vs. LGMM) and clearly documented? If not, was sufficient justification provided as to eliminate certain specifications from consideration?
6b	Are alternative specifications of the between-class differences in variance covariance matrix structure considered and clearly documented? If not, was sufficient justification provided as to eliminate certain specifications from consideration?
7	Are alternative shape/functional forms of the trajectories described?
8	If covariates have been used, can analyses still be replicated?
9	Is information reported about the number of random start values and final iterations included?
10	Are the model comparison (and selection) tools described from a statistical perspective?
11	Are the total number of fitted models reported, including a one-class solution?
12	Are the number of cases per class reported for each model (absolute sample size, or proportion)?
13	If classification of cases in a trajectory is the goal, is entropy reported?
14a	Is a plot included with the estimated mean trajectories of the final solution?
14b	Are plots included with the estimated mean trajectories for each model?
14c	Is a plot included of the combination of estimated means of the final model and the observed individual trajectories split out for each latent class?
15	Are characteristics of the final class solution numerically described (i.e., means, SD/SE, n, CI, etc.)?
16	Are the syntax files available (either in the appendix, supplementary materials, or from the authors)?

Modified from: Schoot R van de, Sijbrandij M, Winter SD, Depaoli S, Vermunt JK. The GRoLTS-Checklist: Guidelines for Reporting on Latent Trajectory Studies. Struct Equ Model Multidiscip J. 2017;24(3):451-467. doi:10.1080/10705511.2016.1247646

APPENDIX C. Group Based Trajectory Model Fitting

Table 1. GBTM final model development: comparison of five class models fit using different polynomial shapes.

Model and class	Type[a]	Bayesian information criterion		Mean posterior probability	Group membership		Sample size			
		Value	Diff.		Est. %	Sample %	Total N=610	P1 n=229	P2 n=211	P3 n=170
Five class model: initial										
One	Q	-4955		0.90	13.9	13.7	83	27	25	31
Two	Q			0.88	28.5	27.4	167	69	55	43
Three	Q			0.89	46.7	48.4	295	110	108	77
Four	Q			0.89	7.7	6.4	39	10	14	15
Five	Q			0.93	4.2	4.3	26	13	9	4
Five class model: alternative one										
One	I	-4946	9	0.94	10.1	8.0	49	12	14	23
Two	I			0.85	40.9	44.6	272	108	95	69
Three	I			0.89	34.8	33.6	205	79	74	52
Four	L			0.90	10.5	10.3	63	18	22	23
Five	C			0.98	3.6	3.4	21	12	6	3
Five class model: alternative two										
One	I	-4946	9	0.94	10.1	8.0	49	12	14	23
Two	I			0.85	40.9	44.6	272	108	95	69
Three	I			0.89	34.8	33.6	205	79	74	52
Four	L			0.90	10.5	10.3	63	18	22	23
Five	Q			0.98	3.6	3.4	21	12	6	3
Five class model: final selected										
One	I	-4942	13	0.85	11.3	12.0	73	23	22	28
Two	L			0.90	42.2	41.1	251	97	90	64
Three	I			0.86	32.9	33.6	205	81	71	53
Four	L			0.90	9.9	9.7	59	16	22	21
Five	C			0.96	3.7	3.6	22	12	6	4

[a]Type of polynomial fitted: I: intercept, L: linear, Q: quadratic, C: cubic.

Table 2. Final fitted GBTM models for all exploratory and sensitivity analyses conducted.

Class	Type[a]	BIC[b]	Mean posterior probability	Group membership		Sample size			
				Est. %	Sample %	Total	P1	P2	P3
Exploratory: recovery period						N=613	n=229	n=212	n=172
One	I	-3808	0.99	3.8	3.6	22	5	8	9
Two	I		0.94	25.1	26.0	159	50	54	55
Three	L		0.96	47.3	46.7	286	122	95	69
Four	L		0.93	20.0	20.2	124	48	41	35
Five	L		0.99	3.8	3.6	22	4	14	4
Exploratory: modified PIPP scores						N=610	n=229	n=211	n=170
One	L	-4458	0.93	16.3	16.6	101	38	32	31
Two	L		0.91	53.2	55.1	336	133	122	81
Three	Q		0.79	18.4	16.6	101	34	33	34
Four	C		0.92	8.6	8.5	52	13	19	20
Five	Q		0.97	3.5	3.3	20	11	5	4
Sensitivity: complete case procedures						N=570	n=220	n=194	n=156
One	I	-4700	0.84	11.2	11.9	68	21	21	26
Two	L		0.90	42.1	41.0	234	95	80	59
Three	I		0.87	34.0	34.6	197	77	69	51
Four	L		0.91	9.3	9.0	51	16	19	16
Five	C		0.93	3.4	3.5	20	11	5	4
Sensitivity: Procedure one						N=229	N=229		
One	I	-1865	0.96	7.0	5.2	12	12		
Two	L		0.85	36.3	38.9	89	89		
Three	I		0.88	41.1	41.5	95	95		
Four	L		0.94	10.3	9.1	21	21		
Five	C		1.00	5.3	5.2	12	12		

[a]Type of polynomial fitted: I: intercept, L: linear, Q: quadratic, C: cubic; [b]Bayesian information criterion

APPENDIX D. Missing Data

Table 1. Patterns and frequency of completed and missing data for PIPP scores.

n	%	Procedure one				Procedure two				Procedure three			
		Time from heel lance				Time from heel lance				Time from heel lance			
		30	60	90	120	30	60	90	120	30	60	90	120
		Complete (1) or missing (0)				Complete (1) or missing (0)				Complete (1) or missing (0)			
133	55.0%	1	1	1	1	1	1	1	1	1	1	1	1
37	15.3%	1	1	1	1	1	1	1	1	0	0	0	0
19	7.9%	1	1	1	1	0	0	0	0	0	0	0	0
6	2.5%	1	1	1	1	1	1	1	0	1	1	1	1
6	2.5%	0	0	0	0	0	0	0	0	0	0	0	0
5	2.1%	1	1	1	1	1	1	1	1	1	1	1	0
5	2.1%	0	0	0	0	1	1	1	1	1	1	1	1
3	1.2%	1	1	1	1	1	1	1	1	1	1	0	0
3	1.2%	1	1	1	1	1	1	1	0	0	0	0	0
3	1.2%	1	1	1	0	0	0	0	0	0	0	0	0
2	0.8%	1	1	1	1	1	1	1	1	0	1	1	1
2	0.8%	1	1	1	1	1	0	0	0	1	1	1	1
2	0.8%	1	1	1	1	0	0	0	0	1	1	1	1
2	0.8%	0	1	1	1	1	1	1	1	1	1	1	1
2	0.8%	0	0	0	0	1	1	1	1	0	0	0	0
1	0.4%	1	1	1	1	1	1	1	1	1	0	0	1
1	0.4%	1	1	1	1	1	1	1	0	1	1	0	0
1	0.4%	1	1	1	1	1	1	0	0	1	1	1	1
1	0.4%	1	1	1	1	1	0	1	1	1	1	1	1
1	0.4%	1	1	1	1	1	0	0	0	1	1	1	0
1	0.4%	1	1	1	1	1	0	0	0	0	0	0	0
1	0.4%	1	1	1	1	0	1	1	1	1	1	1	1
1	0.4%	1	1	1	1	0	0	0	0	1	1	1	0
1	0.4%	1	1	0	0	1	1	1	1	1	1	1	1
1	0.4%	1	0	1	1	1	1	1	1	1	1	1	1
1	0.4%	0	1	1	1	1	1	1	1	0	0	0	0
1	0.4%	0	1	1	0	1	1	1	1	1	1	1	1

Table 2. Infant, maternal, and study characteristics of those with missing and completed PIPP data at each procedure.

Characteristic	Procedure one		Procedure two		Procedure three	
	Missing n=13	Available n=229	Missing n=31	Available n=211	Missing n=72	Available n=170
Study group, n (%)						
KC and placebo	4 (30.8)	77 (33.6)	11 (35.5)	70 (33.2)	25 (34.7)	56 (32.9)
KC and sucrose	6 (46.2)	74 (32.3)	14 (45.2)	66 (31.3)	22 (30.6)	58 (34.1)
Sucrose	3 (23.1)	78 (34.1)	6 (19.4)	75 (35.5)	25 (34.7)	56 (32.9)
Gestational age at birth (weeks), median (IQR)	32 (31, 34.1)	32.9 (31.4, 34.1)	33.9 (31.1, 34.9)	32.7 (31.4 (34)	33.6 (31.6, 34.9)	32.6 (31.3, 33.7)
Birth weight (g), median (IQR)	1700 (1450, 2140)	1850 (1510, 2170)	1680 (1385, 2265)	1850 (1510, 2170)	1745 (1465, 2285)	1855 (1510, 2140)
Infant sex, n (%)						
Male	8 (61.5)	128 (55.9)	16 (51.6)	120 (56.9)	40 (55.6)	96 (56.5)
Female	5 (38.5)	101 (44.1)	15 (48.4)	91 (43.1)	32 (44.4)	74 (43.5)
Twin birth, n (%)						
No	8 (61.5)	150 (65.5)	25 (80.6)	133 (63.0)	50 (69.4)	108 (63.5)
Yes	5 (38.5)	79 (34.5)	6 (19.4)	78 (37.0)	22 (30.6)	62 (36.5)
SNAP-II, median (IQR)	0 (0, 9)	0 (0, 9)	0 (0, 0)	0 (0, 9)	0 (0, 2.5)	0 (0, 9)
Maternal age (years), median (IQR)	29 (26, 36)	31 (26, 34)	29 (22, 34)	31 (26, 34)	30 (27, 34)	31 (26, 34)
Type of delivery						
Spontaneous vaginal	6 (46.2)	104 (45.4)	11 (35.5)	99 (46.9)	24 (33.3)	86 (50.6)
Induced vaginal	0	5 (2.2)	1 (3.2)	4 (1.9)	4 (5.6)	1 (0.6)
Caesarean	7 (53.8)	112 (48.9)	18 (58.1)	101 (47.9)	43 (59.7)	76 (44.7)
Forceps	0	8 (3.5)	1 (3.2)	7 (3.3)	1 (1.4)	7 (4.1)
Prior KC experience, n (%)						
No	10 (76.9)	158 (69.3)	22 (73.3)	146 (69.2)	51 (74.6)	115 (67.6)
Yes	3 (23.1)	70 (30.7)	8 (26.7)	65 (30.8)	18 (25.4)	55 (32.4)
Missing	0	1	1	0	1	1
Ethnicity, white, n (%)						
No	0	22 (9.6)	0	22 (10.4)	3 (4.2)	19 (11.2)
Yes	13 (100)	207 (90.4)	31 (100)	189 (89.6)	69 (95.8)	151 (88.8)
Two-parent family, n (%)						
No	1 (7.7)	4 (1.7)	2 (6.5)	3 (1.4)	2 (2.8)	3 (1.8)
Yes	12 (92.3)	225 (98.3)	29 (93.5)	208 (98.6)	70 (97.2)	167 (98.2)
Maternal college/university education, n (%)						
No	2 (18.2)	50 (22.5)	7 (26.9)	45 (21.7)	14 (21.5)	38 (22.6)
Yes	9 (81.8)	172 (77.5)	19 (73.1)	162 (78.3)	51 (78.5)	130 (77.4)
Missing	2	7	5	4	7	7