From THE DEPARTMENT OF NEUROBIOLOGY, CARE SCIENCES AND SOCIETY, DIVISION OF NURSING
Karolinska Institutet, Stockholm, Sweden

DO YOU SEE MY PAIN? ASPECTS OF PAIN ASSESSMENT IN HOSPITALIZED PREVERBAL CHILDREN

Randi Dovland Andersen

Karolinska Institutet

Stockholm 2018

Published by Karolinska Institutet
Printed by Eprint AB 2018
Cover picture: Kjersti Veel Krauss

ISBN 978-91-7831-015-9

DO YOU SEE MY PAIN?
Aspects of pain assessment in hospitalized preverbal children

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Randi Dovland Andersen

Principal Supervisor:
Senior Lecturer Leena Jylli
Karolinska Institutet
Department of Neurobiology, Care Sciences and Society
Division of Nursing

Co-supervisor(s):
Professor Ann Langius-Eklöf
Karolinska Institutet
Department of Neurobiology, Care Sciences and Society
Division of Nursing

Professor Tomm Bernklev
University of Oslo
Institute of Clinical Medicine
Department of Transplantation

Professor Britt Nakstad
University of Oslo
Institute of Clinical Medicine
Division of Medicine and Laboratory Sciences

Opponent:
Professor Karin Enskär
Jönköping University
Department of Nursing
School of Health and Welfare

Examination Board:
Professor Per-Arne Lönnqvist
Karolinska Institutet
Department of Physiology and Pharmacology

Professor Louise von Essen
Uppsala University
Faculty of Medicine
Department of Women's and Children's Health

Associate professor Stefan Nilsson
University of Gothenburg
The Sahlgrenska Academy
Institute of Health and Care Sciences

In loving memory of my brother Anders

"There'll be no more tears in heaven"

Eric Clapton

ABSTRACT

Background and aim: Pain in hospitalized preverbal children is underassessed and undermanaged. According to the Social Communication Model of Pain, pain is both a personal experience and a social construction, influenced not only by the child in pain, but by the observer and the context. Nurses' pain assessment is biased towards underestimation. The use of structured pain scales is strongly advocated, but pain scales have been difficult to implement into clinical practice. To improve clinical pain assessment and reduce unnecessary pain for hospitalized preverbal children, a better understanding of aspects concerning these scales is needed, and nurses' views regarding clinical pain assessment and their understanding and practical use of structured pain scales need to be further explored. The overall aim of this thesis was to contribute to knowledge regarding how to reduce unnecessary pain and suffering in hospitalized preverbal children by exploring aspects that influence nurses' assessment of pain in the clinical setting.

Material and Methods: This PhD thesis consists of four different studies using both qualitative and quantitative methods. In study I the COMFORT behavioral scale was translated into Norwegian using the forward-back-translation method and culturally adapted in 12 cognitive interviews with clinicians who would later be using the scale in clinical practice. The translated scale's responsiveness to change and inter-rater reliability were tested in study II, based on repeated measurements from 45 preverbal children before and after minor outpatient surgery. Study III was a systematic review appraising the evidence underlying the recommendations presented in 14 systematic reviews on the measurement properties of observational pain scales. Study IV was a semi-structured interview study with 22 nurses in Norway and Canada and examined their pain assessment practices based on self-selected clinical examples.

Results: Cognitive interviews identified several problems with the content validity of the Norwegian and original versions of the COMFORT behavioral scale. The responsiveness of the translated version was supported for assessment of sedation, but not for assessment of pain/distress. Scale recommendations given in systematic reviews addressing the measurement properties of observational pain scales had low evidence value and should be interpreted with caution. Observational pain scales were infrequently used in clinical practice and pain scores were not considered pain –specific. Instead; nurses expressed strong preferences for pain assessment based on clinical judgment and individually tailored to the child and the situation. When assessing pain, nurses combined experience-based and child-specific knowledge with one or more specific strategies to interpret observations of and information from the child. Described strategies included identifying a probable cause for pain, eliminating other sources of distress, evaluating behavioral change and/or effect of interventions on behavior, using a personal and contextual approach, and using behavioral pain scores.

Conclusions: A preverbal child's pain will probably be better seen, evaluated and managed if nurses apply a systematic and comprehensive assessment approach that integrates clinical judgement and structured pain scales.

LIST OF SCIENTIFIC PAPERS

The thesis is based on the following publications and manuscripts, referred to with roman numerals in the text:

I. Andersen RD, Jylli L, Ambuel B. Cultural adaptation of patient and observational outcome measures: a methodological example using the COMFORT behavioral rating scale. *International Journal of Nursing Studies* 2014;51(6):934-942.

II. Andersen RD, Bernklev T, Langius-Eklöf A, Nakstad B, Jylli L. The COMFORT behavioural scale provides a useful assessment of sedation, pain and distress in toddlers undergoing minor elective surgery. *Acta Paediatrica* 2015;104(9):904-909.

III. Andersen RD, Langius-Eklöf A, Nakstad B, Bernklev T, Jylli L. The measurement properties of pediatric observational pain scales: a systematic review of reviews. *International Journal of Nursing Studies* 2017;73(8):93-101.

IV. Andersen RD, Nakstad B, Jylli L, Campbell-Yeo M, Anderzen-Carlsson A. The complexities of nurses' pain assessment in hospitalized preverbal children. *Submitted manuscript*.

CONTENTS

LIST OF ABBREVIATIONS

AMSTAR	Assessment of Multiple Systematic Reviews
CI	Confidence Interval
COSMIN	COnsensus-based Standards for the selection of health Measurement INstruments
CTT	Classical Test Theory
FACS	Facial Action Coding System
FLACC	Face, Legs, Activity, Cry, Consolability
ICC	Intraclass Correlation Coefficient
IRT	Item Response Theory
ISPOR	International Society For Pharmacoeconomics and Outcomes Research
Kw	Weighted Cohen's Kappa
MCID	Minimum Clinical Important Difference
NFCS	Neonatal Face Coding System
NICU	Neonatal Intensive Care Unit
NSD	Norwegian Social Sciences Data Services
PICU	Pediatric Intensive Care Unit
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROM	Patient Reported Outcome Measure
PROSPERO	International prospective register of systematic reviews
ROBIS	Risk of Bias in Systematic Reviews
SD	Standard Deviation
VAS	Visual Analog Scale
VASobs	Visual Analog Scale used by an observer

FOREWORD

Eighteen years ago, in January 2000, I started working in the neonatal intensive care unit (NICU). I transferred from an adult orthopedics ward and was used to caring for patients in pain. After some time in the NICU I started to notice that pain was seldom an issue; we rarely used the word pain, but would sometimes talk about "discomfort". I vividly remember taking care of a premature little boy. He was on a ventilator without any analgesia. His face was contorted, he was thrashing, arching his head backwards, breathing against the ventilator, and making the alarms go off repeatedly. Then all of a sudden he became completely still, his face lax, his body limp and the ventilator resumed all breathing for him. A colleague passing by remarked "oh, good, he's finally relaxing" and I remember thinking – is he really?

A university course in pediatric pain followed by extensive reading made me realize that neonatal pain was in fact a huge issue, but one that we seldom addressed at that time. To make a long story short, with support from management, my colleagues and I carried out a 3-year practice-improvement project in two neighboring NICUs, funded by the Norwegian Extra Foundation for Health and Rehabilitation. Through this project, and with the generous support from my current main supervisor Leena Jylli and from the Research Department at Telemark Hospital, I took my first bumbling steps into research and found my two passions that ultimately led me towards starting this PhD-project in 2013, more than 10 years later. I realized that I love research – sorting out the puzzle, making new connections and discoveries – but also the nitty-gritty everyday details necessary for the end result. But not just any type of research. My passion is research that can help relieve and reduce pain in children, most of all those who are unable to speak and advocate for themselves.

1 INTRODUCTION

The starting point for this PhD-project was a strong desire on my part to improve pain management and decrease pain-related suffering in hospitalized children outside the NICU. I had observed first-hand the positive effects of introducing pain measurement scales into our NICU and strongly believed that structured assessment of pain was a necessary foundation to base treatment decisions on and as such the logical first step towards better management of pain. Locally we wanted to extend this practice to children outside the NICU, and a multidisciplinary group had selected the COMFORT behavioral scale for implementation and use in non-verbal children across units. As ours was a relatively small general hospital, it made sense at the time to select a scale that we assumed could be used across different units and in both intubated and spontaneously breathing children. My thesis work was supposed to comprise translation and validation of the scale followed by implementation into clinical practice at our hospital. However, initial findings changed the direction of this work towards issues regarding scale validity in general, nurses' pain assessment practices and the nurse assessor's influence on the assessment of pain.

2 BACKGROUND

2.1 WHAT IS PAIN?

Pain is a warning signal (1, 2). Pain alerts the individual to possible bodily danger and subsequently prompts escape from the dangerous situation, recovery and healing (1). An equally important feature is the ability pain has to grab the attention of others and elicit help (1, 3), demonstrating that pain has social aspects.

2.1.1 Both an individual experience and a social construction

The experience of pain is constructed in the brain based on information from multiple sources, including incoming nociceptive or danger signals, information from the senses (vision, touch, hearing) and other modulating factors such as attention, distraction, expectations, anxiety, stress, the physical and social context, and past experiences (2, 4). Everyone has his or her own individual understanding of and experiences with pain (5, 6).

Williams and Craig recently defined pain as "a distressing experience associated with actual or potential tissue damage with sensory, emotional, cognitive, and social components" (7, p. 2420).

The sensory component includes how much it hurts (pain intensity), what the pain feels like (quality), where it hurts (location), and how long it hurts (duration) (8). The emotional component concerns emotions or feelings associated with the experience of pain as an unpleasant sensation (9). The cognitive component includes all thought processes or intellectual activity related to pain, including beliefs, appraisals, expectations, and meanings attached to pain (10). Cognitive aspects of the experience influence both the emotional and sensory components of pain as well as pain-related behavior. The sensory, emotional and cognitive components all reflect the subjective nature of pain. However, pain is not only a subjective experience constructed in the brain of an individual: it is also a social construction. Pain is always experienced in a social setting and the social features of pain include how others respond to the person experiencing pain (11) and how the behaviors of others and the social environment in turn influence the person in pain (12). This understanding of pain as the outcome of a dynamic and ongoing social transaction between the person experiencing pain and the caregiver is further described in the theoretical framework for this thesis.

2.2 THEORETICAL FRAMEWORK

2.2.1 The Social Communication Model of Pain

Through work spanning more than two decades, Craig and colleagues (12-17) have proposed and subsequently refined this communication model of pain that addresses the processes by which humans of all ages express and perceive pain within a social context. Well-known, earlier pain models, for example the gate control theory of pain (18) and the neuromatrix theory (19), have had a profound impact on our understanding of pain but are limited to biological and psychological intrapersonal processes, or pain as a subjective experience. The

social communication model of pain not only acknowledges pain as a bio-psycho-social phenomenon, it also specifically includes social factors in addition to the intrapersonal mechanisms described in earlier models (12). It is a generic model, not limited to any specific patient group.

2.2.2 Overview of the model

In the Social Communication Model of Pain (Figure 1) there are two actors: the person experiencing pain (in this case a preverbal child) and the assessor. The assessor may be either a primary caregiver or a professional, but in this thesis the assessor is a nurse. Pain is experienced and assessed within a defined physical and social setting. The child is referred to as "he" and the nurse as "she" throughout the thesis to make reading easier.

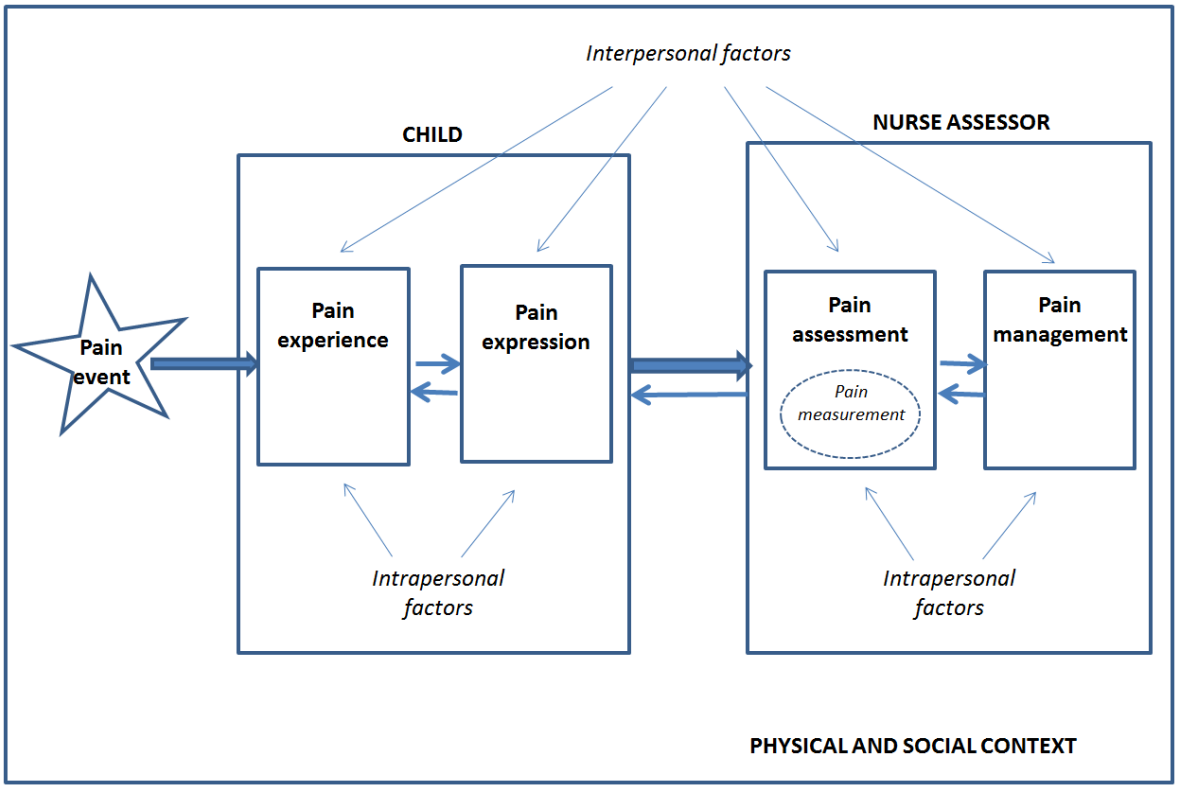

Figure 1. The Social Communication Model of Pain. A conceptual biopsychosocial model depicting the interaction between the child experiencing pain and the nurse assessor. Craig, K. D. The Social Communication Model of Pain. *Canadian Psychology*, Volume 50, pp. 22-32, Copyright © 2009, American Psychology Association. DOI: 10.1037/a0014772. Reproduced and adapted with permission.

In the model an episode of pain is described as a series of interdependent stages starting with a pain event – real or perceived, the child's experience and expression of pain, and the nurse assessor's assessment and management of pain. Each of the stages involves complex and dynamic processes within (intrapersonal) as well as between (interpersonal) the child in pain and the nurse assessor (12, 17, 20). Interpersonal factors also include influences that the social and physical context exerts on the child and the nurse assessor (12).

2.2.3 How the model has been used in this thesis

This thesis specifically explores the assessment stage of the pain process described in the model. Consequently the main actor under study is the nurse assessor. Based on the understanding of pain as a social construction, influenced by both the child experiencing pain and the assessor, an understanding of factors influencing the assessor and her assessment of pain is necessary for a comprehensive understanding of the child's pain. Still, the child and his experience and expression of pain are described to provide an understanding of how the child may influence the assessment. The use of structured pain measurement scales is considered an important aspect of pain assessment and has been added to the figure, although the use of scales in the assessment of pain is not specifically addressed in the model.

Although the Social Communication Model of Pain (12) was only used in the development of study IV, it will be applied as a theoretical lens to understand and interpret findings from all four studies included in this thesis.

2.3 THE PREVERBAL CHILD

The starting point and nexus for all nursing care and nursing research is the patient. This thesis is centered on children between 0 and 3 years of age. This period is characterized by rapid development with the child growing from a helpless, crying newborn baby to a walking, talking 3-year-old (21, 22). These outward changes are accompanied by cognitive and mental developmental changes influencing the child's understanding of and interaction with his or her physical and social environment (23). There are different opinions in the literature in regard to what to call this age group, but for this thesis the term preverbal was chosen.

In the Social Communication Model of Pain (12) there are only two visible actors: the preverbal child and the nurse assessor. A preverbal child is not an independent individual, but emotionally and physically dependent upon his primary caregiver(s). The primary caregiver is most commonly a parent and the term parent is used throughout this thesis. One consequence of this dependence is that pediatric nursing care, including pain assessment, is organized around the child–parent dyad and not the individual child. As such, parents are a major intrapersonal influence on the child's experience and expression of pain (24, 25).

2.3.1 Pain exposure in hospitalized preverbal children

Hospitalized children are frequently exposed to pain from injury, disease or procedures (26). Studies from the last decade reported that between 24 and 72% of hospitalized children experienced moderate to severe pain, defined as a score of 4 or higher on a Numeric Rating Scale or on an observational pain scale ranging from 0 to 10 points (27-31). In general higher pain scores were reported in studies based on interviews with children and/or parents compared to studies based on data extracted from chart reviews that only take into account documented pain scores. Birnie and colleagues (32) found that 62% of hospitalized children had experienced clinically significant pain during the past 24 hours measured as the difference between self- or proxy-reported pain intensity and pain threshold.

2.3.2 Experience of pain

All preverbal children have the functional ability to experience pain, regardless of age. Ascending nociceptive pathways are fully functional at birth and provide the necessary sensory input to the brain together with information from the other senses (4). There is no pain center in the brain. A network of multiple parts of the brain –the neuromatrix - is involved in the construction of pain in the brain (19). As such, the experience of pain is a result of the brain's judgment of incoming and stored information (2).

It is impossible to know exactly how preverbal children experience pain and distress because they cannot describe it verbally. Instead, we use knowledge from other populations, for example older children, together with estimates regarding how developmental stage and maturity may affect the experience. Pain in older children and adults is a synthesis of sensory input, emotions, thoughts and social influences. Their experience of pain can be downregulated by the brain, for example through cognitive understanding of the situation in which pain occurs (4). Limited cognitive abilities (17) and lack of understanding of the pain severely limits preverbal children's capacity for central downregulation of pain (4). Unfortunately, central upregulation of pain, for example caused by negative emotions, emerges earlier in development; the ability to anticipate and fear pain has been seen from around 6-8 months (33). Preverbal children also seem to have difficulties discriminating between the sensory experience of a noxious stimulus that most individuals will experience as painful and what more mature individuals would consider non-painful emotional distress such as fear or anxiety (34, 35). Taken together, this indicates that small children on average may experience an injury as more painful than older children and adults.

2.3.3 Expression of pain

Preverbal children express their pain experience through behavioral and physiological cues. Behavioral cues include facial expressions, early language and/or paralinguistic features, and body movements (17). Although pain expression changes with age (20) and developmental changes are profound over the preverbal period, observed reactions in preverbal children are very similar to those seen in older children and adults in response to similar stimuli (17).

Facial expression encodes information about emotions and pain (13) and is considered the best studied behavioral expression associated with pain (36, 37). Several studies have been performed in adults based on the Facial Action Coding System (FACS), a comprehensive system describing the complete set of facial actions or muscle movements the face is capable of (38) and acute pain is associated with distinct and different subsets of facial actions (39). The Neonatal Face Coding System (NFCS) is an adaptation of FACS to infant pain. Five facial actions (brow lowering, eyes squeezed tightly shut, deepening of nasolabial furrow, open lips and mouth, and taut cupped tongue) in NFCS have been associated with pain in newborns (40, 41) and infants (42).

Crying and paralinguistic features (vocal effects that are not words or phrases, for example grunting and moaning) (12) are considered relatively non-specific signs of distress (43-45),

but they are effective in signaling that something is wrong and attract the attention of caregivers (17). The beginning of a pain-related vocabulary with use of words such as “hurt” “ow” and “ouch” gradually emerges between 1.5 and 3 years of age. To what extent children at this age fully comprehend the meaning of these words is not clear (46). Although children as young as 2 years old can provide some verbal information about their pain (5, 47) and should be believed (47), it is not until around 3 years of age that verbal and cognitive abilities are sufficiently developed to enable more consistent, early self-report (48, 49).

Body movements include both generalized movements, for example flailing limbs in newborns and younger infants, and more specific protective or flight responses (12). An example of a protective response is the guarding of a painful injury, while an example of a flight response is a small child trying to get away from the pain.

Examples of physiological cues associated with pain are changes in heart rate, respiration and blood pressure (4). These cues are part of a fight-flight response and not specific to pain (17).

2.3.4 Factors influencing the experience and expression of pain

Unrelieved pain may interfere with all aspects of life, including physical and social activities, normal development and learning, emotions and sleep (50). It may result in increased anticipatory distress (51, 52), increased pain responses (53-56), and diminished effect of analgesics (57, 58) on subsequent procedures. Unrelieved acute pain increases the risk of chronic pain states (59). Ongoing nociception and unrelieved pain may also influence how pain is expressed. Specifically, it may result in withdrawal or a decrease in observable signs of pain (60). Such individual differences and changes in how pain is expressed make it more difficult for others to detect and assess the pain.

Behavioral cues may be viewed as either reflexive or purposive, based on whether or not they are under voluntary control (14, 61). Shortly after birth all behavior is reflexive, with cognitive control over behavior emerging slowly through the first years of life. This may be observed as a gradual change towards more deliberate movements serving to protect against or withdraw from injury or in response to or in anticipation of pain (33). Facial expression and paralinguistic features have been categorized as mainly reflexive behaviors, while the use of language is considered a purposive behavior (61). Body movements may be both; generalized movements and withdrawal reflexes are considered reflexive behaviors, while more specific protective or flight responses are considered purposive behaviors (12, 61). With increased cognitive control and maturity both automatic and purposive behaviors may to some extent be consciously modified by the person in pain (12), for example, facial expression of pain can be exaggerated or suppressed (62).

The distinction between reflexive and purposive behaviors may be important because reflexive behaviors are thought to be a more “honest” reflection of the experience (12). Consequently, the interpretation of observed behavioral cues associated with pain may be influenced by to what degree they are considered reflexive or under voluntary control (61).

A newborn child has an inborn ability to signal pain-related and other types of distress, but is otherwise completely helpless and dependent upon caregivers to survive. A parent reacts to the infant's distress, and, over time, specific patterns of attachment develop based on parental sensitivity and how well the parent manages to respond to the infant's distress. The quality of the attachment between the child and the parent shapes how the developing child gradually learns to self-regulate distress and express pain (1, 63).

Children continue to learn about pain by observing and modeling, and by receiving support for pain within the family and from their parents (12, 32). Both reflexive and purposive behaviors are influenced by social factors (12). How the child expresses pain can be reinforced or diminished based on both environmental responses (64) and situational demands (12).

Parental behavior can promote both coping and distress. Distraction (65) and high emotional availability (which means that the parent are good at picking up and responding adequately to their child's distress signals) (66) have been associated with decreased pain behavior, whereas parental reassurance during a painful procedure is associated with increased distress (67, 68). The influence of parental behavior on the child's behavior increases with the age of the child (64). Cultural background and norms are believed to influence both how pain is expressed and subsequently how this behavior is interpreted by others (69). Where the youngest children are concerned, it has been speculated that the social influence of larger social systems like family, hospital setting or culture, is mediated through their influence on the parent rather than acting directly on the infant (64, 70).

2.4 THE NURSE ASSESSOR

The alleviation of suffering is one of the four fundamentals of nursing care (71). Pain must be detected before it can be alleviated, and the assessment of pain is an essential part of nurses' responsibilities (72, 73). Pediatric nurses frequently have to assess pain in preverbal children as they make up a large proportion of pediatric patients. Around half (46.6-59.6%) of all pediatric admissions to North American pediatric hospitals involved children younger than 3 years (27, 28, 30, 31).

2.4.1 Pain assessment

Pain assessment is a systematic and holistic approach to the child's situation (12, 48) taking into account all bio-psycho-social factors that are associated with pain including, but not limited to, sensory aspects (pain intensity, quality, location and duration), vocalization, physiological and behavioral cues, parental assessment and opinions, cause of pain, influencing factors and the overall judgment of the nurse (74, 75).

Within the framework of the Social Communication Model of Pain, pain and its assessment are understood as an ongoing and dynamic transaction within the child–nurse dyad (15). A transaction implies both an interaction where messages (verbal and non-verbal) are exchanged between the child in pain and the nurse, and that the outcome of this exchange

extends the simple sending and receiving of messages. Ideally the transaction should be something like this: The child experience pain and signals distress. To assess pain, the nurse interprets and responds to the distress; the dialogue goes back and forth, and the outcome of the child-nurse interaction is a pain diagnosis and a subsequent treatment decision (76).

Information is lost in the transfer between the child in pain and the observer (13) indicating that the observer's interpretation will always to some degree differ from the child's experience. A substantial body of literature shows a bias towards underestimation of pain in children (13, 77-85), although parental overestimation of pain has also been reported (86). In general, both parents and clinicians underestimate pain, but most of the time parental assessment was more accurate (77). Although these studies have by necessity included older, verbal children it is reasonable to assume that their findings can be extrapolated to preverbal children. Reasons for the persistent underestimation of pain are not clear (32). Intrapersonal factors found to influence perception and assessment of pain include personal factors like gender (87-89) and age (90-92), psychological and cognitive factors (88, 89, 93-96), and prior personal and professional experience with pain (88, 89).

2.4.2 Pain management

Management of pain is outside the scope of this thesis, but is briefly described here to provide a complete overview of the elements in the Social Communication of Pain model. Pain management is usually based on a "3-P" approach where the 3 P's represent the psychological, physical and pharmacological domains of pain management (97). A multimodal approach that includes a combination of strategies from all three domains is considered more effective than single strategies and provides greater pain relief (97, 98). A fourth P for prevention may be added to emphasize that the most powerful pain reducing approach is to avoid inflicting it whenever possible (99) or choosing the least painful approach if several alternatives exist (100).

Different distraction strategies are among the best studied psychological strategies (101) and may be used with infants of 10 months and older (102). Examples of appropriate strategies for preverbal children are bubbles, play, and non-procedural talk (103). Distraction using toys or videos has shown some benefit for preverbal children although the quality of existing evidence is currently low (104). Older preverbal children need age-appropriate preparation prior to a procedure (103), but overall for this age group a majority of the preparation is directed towards the parents to make them feel more secure and help them support their child the best way possible (105). Whenever possible, parents should be with their child in painful and stressful situations (106, 107). Being held in the arms of a parent is the best position for distraction of infants younger than 12 months. Older infants should still be held close, but in a position of the child's choosing (108).

The most common pharmacological interventions for nociceptive pain used across all age groups are opioids (109), non-opioids (non-steroidal anti-inflammatory drugs and paracetamol) (110), and topical anesthetics like EMLA® for procedures that break the skin

(98, 111). Age-related differences in pharmacokinetics and –dynamics influence dose requirements (112). Sweet tasting solutions (glucose or sucrose) (98, 113) are considered effective for procedural pain relief up to 12 months (114), whereas the evidence for their use beyond one year is inconclusive (113).

2.4.3 Factors influencing the assessment and management of pain

An observer's response to another person's pain may be characterized as a dual process including a mixture of reflexive behaviors and behaviors that are under cognitive control, similar to the expression of pain. Automatic reactions, for example from seeing a burn injury, reflect the workings of more fundamental biological systems, while the more complex cognitive processes are associated with consciously trying to interpret observations in light of prior experiences with a judgment regarding the observed situation as the goal. It is assumed that reflexive displays of pain trigger reflexive reactions, while controlled behaviors are more likely to trigger cognitive and controlled responses (12, 115).

Assessment of pain is further influenced by the relationship between the person in pain and the observer (12). Factors related to the person in pain that have been shown to influence the observers' perception of pain in experimental studies, include pain intensity (88, 90), presence or absence of observable cues (89), sources and type of evidence (78, 116), known cause of pain (89), personal characteristics (89), perceived credibility (78, 116) and treatment effect (117). Three of these, including presence or absence of observable cues (118), known cause of pain (118-123), and personal characteristics (118, 123, 124) have also been identified in clinical studies as influencing nurses' assessment of pain.

Craig (12) has suggested that contextual and organizational factors may be important contributors to the persistence of inadequate pain management practices. Assessment of pain in a child may be influenced by the parent (118, 125-127) and the physical and social context (128) in which the assessment takes place. Studies have shown that ward culture impacts pain assessment practices (126, 129) and nurses often attribute deficiencies in pain assessment practices to staffing issues and heavy workloads (15, 130-132).

Management of pain is dependent upon the assessment of pain and factors associated with the assessor/caregiver and the setting (12), but nurses pain assessment practices are not widely studied and clinical studies are sparse (74, 133). Most studies describing nurses' pain assessment practices or how nurses think and what they do when they assess pain in preverbal children were published during the 1990's and early 2000's (118, 120, 123, 134-145). In the past decade, this area of research has drawn even less attention and no published studies have examined nurses' assessment practices in situations where structured pain measurement scales are available.

2.5 PAIN MEASUREMENT

Measurement of pain is one aspect of pain assessment and concerns the use of a structured pain scale to quantify one dimension of pain, most commonly pain intensity (48). The use of a structured pain measurement scale to obtain a numerical pain score is considered a prerequisite for effective treatment of pain (75, 146) and the fundament for a scientific approach to pain (147). In the late 1980's and early 1990's several studies attributed sub-optimal pain management to the lack of objective and appropriate methods for the assessment of pain in children (148-150). As a result, several structured scales for measurement of pain were developed over the next decades. Structured pain scales are based on either self- or proxy-report of pain. Structured self-report of pain intensity may be obtained from verbal individuals using a Faces Scale or a Numeric Rating Scale (48). For non-verbal individuals, a proxy or an observer provides a rating of pain. Structured proxy assessment of pain entails the use of an observational pain scale (5, 151).

2.5.1 Observational pain measurement scales

Observational pain scales are based on structured evaluation of behavioral and/or physiological changes or cues considered to be indicators of pain (4, 14) and are used in situations where children are unable to verbalize their pain experience due to age, illness or severe cognitive or mental impairments. The main premise underlying these scales is that the biobehavioral responses included in the scale are a valid representation of the pain experience (17, 152). One example of an observational scale used to assess pain is the COMFORT behavioral scale.

2.5.2 The COMFORT behavioral scale

The COMFORT behavioral scale (153) is a modified version of the COMFORT scale (154). The original COMFORT scale was developed to assess the efficacy of interventions to reduce distress in intubated children in a pediatric intensive care setting. The authors defined behavioral distress as behaviors resulting from negative affect resulting from pain, fear or anxiety. As such, the concept distress includes pain, but distress is not necessarily painful. The COMFORT scale consists of 8 items – 2 physiological (Blood pressure, Heart rate) and 6 behavioral (Alertness, Calmness/agitation, Respiratory response, Physical movement, Muscle tone, Facial tension). Each of the items includes 5 behaviorally anchored ordinal levels, scored from 1 to 5 points. Item scores are added together to produce a sum score ranging from 8-40 points (154). The lower range of scores indicate sedation, middle range of scores comfort or a normal state and higher scores reflect increasing levels of distress and pain.

In the modified COMFORT behavioral scale (153) used in this thesis, the 2 physiological items have been removed from the scale and a behavioral item "Crying" has been added, extending the use of the scale to spontaneously breathing children. The items "Crying" and "Respiratory response" are mutually exclusive; "Crying" is scored in spontaneously breathing children and "Respiratory response" in intubated children. Thus, the child's behavior is evaluated on 6 different items and the sum score ranges from 6-30 points. In addition a visual

analog scale (VAS) with the anchors "No pain" and "Worst pain" has been added to the scoring form to collect a global or unstructured evaluation of the child's pain from the observer (VASobs). An algorithm had also been developed to assist in treatment decisions depending on the collected COMFORT behavior and VASobs scores (155). The most recent version of the COMFORT behavioral scale may be found at www.comfortassessment.nl.

The construction of the COMFORT/COMFORT behavioral scale differs from most other pain scales. Within each item the behavioral anchors range from a sedated state to a pain/distress state, with a normal state as the neutral middle. Consequently the aggregated COMFORT scores range from sedation to pain/distress with no pain as a neutral middle score. In a most other pain scales, for example FLACC (156), the behavioral anchors and total score range from no pain to worst pain. As a consequence, a COMFORT score is not directly comparable to scores from other pain scales.

A majority of the studies validating the original COMFORT scale and published prior to 2010 were performed in the intended target group for the scale (children between 0 and 8 years) in a North American (USA and Canada) pediatric intensive care (PICU) setting (157-160). Some studies were done on children in Dutch (153, 161), Spanish (162), and Brazilian (163) PICUs, and neonates in an American NICU (164). The COMFORT scale had also frequently been used as an outcome measure in treatment studies (165-172) and to validate other scales or measurement methods (173-178). The validity of the COMFORT behavioral scale had been evaluated in normally developing Dutch children between 0 and 3 years undergoing major surgery (153, 155, 161) and a Swedish translation had been developed and evaluated in children younger than 10 years in the PICU (179). The COMFORT behavioral scale had also been used as an outcome measure in treatment studies (180-184) and to validate other scales or measurement methods (185-188).

2.5.3 Use of pain scales in clinical practice

The development of pain assessment measures did not alleviate the problem of sub-optimal pain management in clinical practice. Several studies have documented that although measures became available, they were not widely used (126, 189-192) or that measurement results were not documented in a systematic manner (121, 126, 193).

In a Norwegian hospital or indeed any non-English-speaking context, one of the main barriers to the implementation and use of structured pain scales was that most existing scales had been developed for use in an English-speaking setting. Unpublished Norwegian versions of published scales had started to emerge in clinical practice. A major problem with unpublished translations is the lack of documentation concerning how the translation was carried out or regarding the measurement properties of the translated version of the scale. As a translation does not automatically inherit the measurement properties of the original scale, rigorous translation and cultural adaptation is necessary to ensure the validity of a scale when used in a new language and setting (194).

2.5.4 Alternatives to the use of observational pain scales

Facial actions are considered the best behavioral indicators of pain (36) and a promising approach is the use of computer software to analyze facial movements associated with pain, for example during painful procedures or postoperatively (195, 196). However, automatic facial analysis is not yet available or feasible for daily use in a clinical setting (196). Although neurophysiological indicators like skin conductance (197), NIRS (near-infrared spectroscopy), EEG (electroencephalography), PET (position emission tomography), and MRI (magnetic resonance imaging) (198), as well as biomarkers like cortisol (199, 200) and heart rate variability (201) and have been suggested as potentially more objective alternatives to clinical observations or the use of observational pain scales, their validity and feasibility vary and currently none may be considered an independent valid indicator of pain (202).

2.6 MEASUREMENT PROPERTIES OF PAIN SCALES

2.6.1 Why measurement properties matter

Measurement quality is dependent upon the scale's measurement properties. The use of structured pain scales with questionable or inadequate measurement properties puts children at unnecessary risk as these scales may both over- and underestimate pain. Underestimation of pain may result in lack of treatment and cause additional suffering for the child. In addition, untreated pain increases the risk of developing chronic pain conditions (203, 204). Overestimation of pain, on the other hand, may expose the child to unnecessary pharmacological pain management with increased risks of negative side effects (109).

2.6.2 Taxonomy

The labeling and definition of measurement properties varies widely in both the methodological literature and in published measurement studies (205, 206). This thesis adheres to the taxonomy put forward by the COSMIN (COnsensus-based Standards for the selection of health Measurement Instruments) group (207) where measurement properties include the domains reliability, validity and responsiveness (Figure 2). The term has a wider application than the more frequently used "psychometric properties" as the term measurement properties is applied for studies that use either a classical test theory (CTT) approach or the more sophisticated item response theory (IRT) (205) approach. Still CTT or psychometric theory is by far the most common approach for validation of observational pain scales. Since none of the studies validating the COMFORT scale/COMFORT behavioral scale have included the use of IRT, the description of measurement properties and how they may be tested, is limited to the use of CTT. To validate observational pain scales or to evaluate the validation profile of a given scale, an understanding of the different measurement properties is necessary.

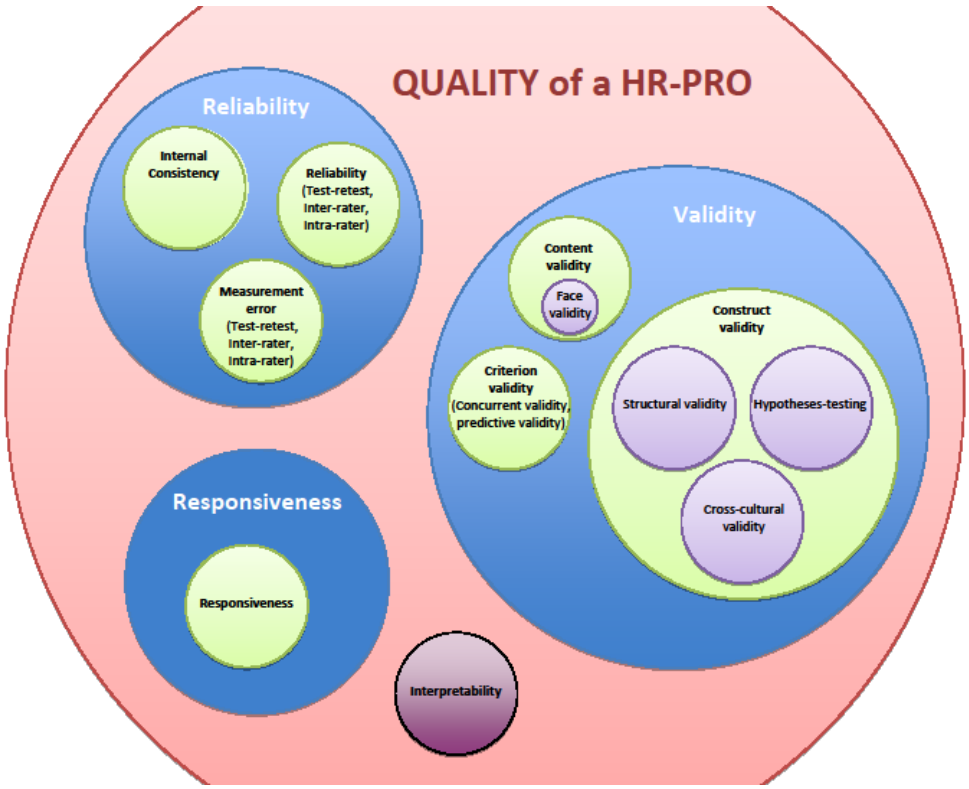

Fig. 2. COSMIN taxonomy of relationships of measurement properties. Abbreviations: COSMIN, Consensus-based Standards for the selection of health Measurement Instruments; HR-PRO, health related-patient reported outcome. Reprinted from Journal of Clinical Epidemiology. Vol. 63, no. 7, Mokkink LB, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes, Pages 737-45, Copyright 2010, with permission from Elsevier.

2.6.3 Reliability

Reliability concerns to what extent a measurement performed with a scale is free from measurement error (207). As such, reliability is a characteristic of scale use in a specific sample and setting, not of the scale itself, and needs to be reassessed each time the scale is applied in a different sample and setting. The reliability of an observational scale is influenced by variations related to the patient, the assessor, the context, and the scale itself. Consequently reliability can be improved by assessor training, restrictions in the assessment situation (for example by assessing all children at a predefined time point after administration of pain medication), and by averaging repeated measurements (205). The taxonomy defines four aspects of reliability: internal consistency, test-retest, inter-rater and intra-rater reliability, of which internal consistency and inter-rater reliability are the most relevant for observational pain scales. Internal consistency is defined as the degree of interrelatedness among the items in the scale (207) or to what extent the items in the scale measure the same construct. Inter-rater reliability concerns to what extent scores agree when a scale is used by two different raters on the same occasion (205).

2.6.4 Validity

Validity concerns to what extent a scale measures the construct it is supposed to measure (207), or to what extent a pain measurment scale measures pain. The COSMIN taxonomy defines three types of validity although only content and construct validity are applicable to observational pain scales. The third, criterion validity, is dependent on the existence of a true gold standard, which does not exist for pain, and is not further discussed.

Content validity is addressed during construction of the measure and has two aspects, face validity and content validity. Both are judgment-based, qualitative evaluations. Face validity concerns whether the scale looks like a good reflection of the construct, while construct validity is an assessment of whether the scale is an adequate representation of the construct in regard to relevance and comprehensiveness. Content validity should be assessed by those who are going to use the scale (205). One specific aspect of content validity concerns questions related to the translation and cultural adaptation of scales. The validity of the translated version of the scale is dependent upon how the translation and cultural adaptation were carried out (194). The process of establishing content validity of a translated scale is limited by the original version as the translation needs to maintain fidelity towards the original version (208).

Construct validity concerns whether the instrument provides expected scores (206). The COSMIN taxonomy divides construct validity into structural validity, hypothesis testing and cross-cultural validity. Structural validity concerns to what extent the scale scores adequately reflect the dimensions of the construct and are assessed by confirmatory factor analysis. Hypothesis testing concerns differences in scores between groups or relationship of scores with scores from other scales measuring similar or dissimilar constructs. Correlation between test scores and scores from other measures are dependent on the validity of the comparator and is an indirect test of the construct. As such, correlation between scores from different scales can only provide circumstantial evidence for validity. Cross-cultural validity is assessed using correlation of scores from a translated version of the scale with scores obtained with the original scale after a standardized translation and qualitative testing of the construct validity of the translated version (205).

2.6.5 Responsiveness

Responsiveness is an aspect of validity (205). While validity concerns the validity of single scores or differences between individuals or groups, responsiveness concerns the validity of change scores or intrapersonal differences. The COSMIN taxonomy defines responsiveness as a scale's ability to detect change over time in the construct (207). Responsiveness is assessed with hypothesis-testing strategies, for example of hypotheses regarding how a pain score will change between before, during and after a procedure within an individual or a group.

2.6.6 Interpretability

Although interpretability is not a measurement property, it is still important as it concerns how to interpret or understand the scores or change scores from a measure (205). If an observational scale provides scores ranging from 0 to 10, what does a score of 5 mean, or a change from 9 to 7? Response shift is an interpretability issue that concerns subtle changes that may occur over time in the observer's understanding of pain, the child's expression of pain and/or how the measurement is performed. These changes may influence measurement scores (205).

2.6.7 Scale validity

The sections above briefly describe the different aspects of validity, reliability and responsiveness and how these can be tested in validation studies in accordance with the COSMIN checklist (209).

Scale validation is an ongoing process where evidence of the scale's validity is accumulated through scale development and subsequent validation studies. Every validation study adds to the evidence underlying the scale, but the contribution of a given study is dependent upon the methodological quality of the study (study validity) and its risk of bias. Risk of bias and methodological quality are partly overlapping concepts, but risk of bias pertains to threats to a study's internal validity, whereas methodological quality pertains to both internal and external validity (study validity) (210). Scale validity concerns the validity of the scores obtained with the scale, not the scale itself (205). To evaluate scale validity, findings from all validation studies concerning that scale need to be synthesized, taking the number of studies, their methodological quality, and the consistency of results into consideration. Studies also need to have sufficient homogeneity or similarities in the construct being measured, the purpose of the study, and the study population to enable aggregation of study results (205, 211).

The quality of the evidence supporting structured pain scales differs (146, 212). To reduce the risk of over- and underestimation of pain, only scales with good measurement properties should be implemented and used in clinical practice. Systematic reviews evaluating the measurement properties of pain scales often provide recommendations of pain scales. Similar to primary studies, the risk of bias or presence of systematic errors in a systematic review will influence its study validity (213) and subsequently the validity of the recommendations given (214).

2.7 RATIONALE FOR THE THESIS

Hospitalized preverbal children are frequently exposed to pain and this pain is still underassessed and undermanaged. This age group is more vulnerable to the negative effects of pain than older children. Their ability to perform self-care to handle pain and distress is limited, as is their ability to communicate their experience in order to elicit the necessary help from others to manage pain and distress.

Management of pain and distress depends upon the nurse's or caregiver's ability to recognize and correctly interpret the child's expression of pain. The Social Communication Model of Pain describes pain assessment as a transaction between the child and the assessor and as such suggests that aspects related to the nurse performing the assessment need to be further explored.

Assessment of pain in preverbal children is biased towards underestimation. It is a commonly held belief that use of structured pain measurement scales is a prerequisite for better assessment of pain. A better understanding of aspects concerning these scales and their use is needed. Non-English speaking contexts suffer from poor availability of pain scales in the native language, and the development of valid versions of existing pain scales and ways of dealing with methodological issues in the translation and cultural adaptation process are of particular interest. Using a scale with insufficient measurement properties may expose the child to over- or underassessment of pain and unnecessary suffering. Scale selection is difficult; there are many scales to choose from and the evidence underlying existing scale recommendations is usually unavailable.

Nurses' reluctance towards using pain scales has been proposed as a possible reason for the difficulties implementing these scales into clinical practice. In order to improve clinical pain assessment and reduce unnecessary pain for hospitalized preverbal children, nurses' own views and practice regarding clinical pain assessment and their understanding and practical use of structured pain scales need to be further explored.

3 GENERAL AND SPECIFIC AIMS

3.1 GENERAL AIM

The overall aim of this thesis was to contribute to knowledge regarding how to reduce unnecessary pain and suffering in hospitalized preverbal children by exploring aspects influencing nurses' assessment of pain in the clinical setting.

3.2 SPECIFIC AIMS

The specific aims were to:

Explore the use of cognitive interviews in the translation and cultural adaptation of observational measures, exemplified using the COMFORT behavioral scale, and to demonstrate a structured approach to the analysis of data from cognitive interviews (study I)

Test the responsiveness* of the COMFORT behavioral scale in pre- and early verbal children undergoing minor elective surgery and determine inter-rater reliability among participating nurses (study II)

Describe how systematic reviews have evaluated and recommended observational pain scales for use in children aged 0–18 years and appraise the evidence underlying these recommendations (study III)

Explore how nurses assess pain in hospitalized preverbal children (study IV)

**The term "construct validity" was used in the article from study II, but since hypotheses regarding change over time were tested, the correct term in accordance with the COSMIN taxonomy used in this thesis framework is "responsiveness".*

4 MATERIAL AND METHODS

4.1 GENERAL OVERVIEW

This PhD thesis consists of four different studies using both qualitative and quantitative methods (Table 1). Data were collected between 2010 and 2016. Studies I and II were carried out at Telemark Hospital, and study IV in hospitals in the South-Eastern Norway Health Region and in a tertiary children's hospital in Canada. Study III was a systematic review and setting-independent. Since there are no methodological similarities across studies, each study is described separately.

Table 1. Overview of included studies

Study	Design	Sample and setting	Data collection	Data analysis
I	Translation and cultural adaptation of the COMFORT behavioral scale	Clinicians (N=12) including 8 nurses, one nurse assistant and 3 physicians at Telemark Hospital	Individual cognitive interviews	Qualitative; deductive and interactive approach
II	Prospective observational measurement study	Toddlers 12-36 mo. (N=45) at a surgical outpatient unit at Telemark Hospital	Repeated assessments with COMFORT behavioral scale pre- and post-surgery	Quantitative; statistical analyses / CTT
III	Systematic review	Systematic reviews (N=14) on measurement properties of observational pain scales	Data extracted using AMSTAR and ROBIS	Qualitative; narrative synthesis
IV	Interview study	Nurses (N=22) from 5 pediatric units in Norway and Canada	Semi-structured individual interviews	Qualitative; inductive thematic analysis

4.2 STUDY I

Study I was a systematic translation followed by cultural adaptation of the translated scale.

4.2.1 Translation and cognitive interviews

The forward-back-translation method is a commonly used translation approach. It entails translating a document into the target language (forward translation) and back into the original language (back-translation). The back translation is done by an independent translator not familiar with the original document. Several translation guidelines exist and although the forward-back-translation method is advocated in all, their approaches differ. This study was based on the guideline published by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) task force for translation and cultural adaptation (194).

Although the ISPOR guideline recommended cognitive debriefing, no method was described so cognitive interviewing as described by Willis (215) was adapted for cultural adaptation of the COMFORT behavioral scale. Cognitive interviews identify problems with how members of the target group understand the content of written material, process the information given there, and respond to it, with the aim of removing weaknesses and ambiguities from the material (215-217). Cognitive interviewing relies on two main interview techniques: Thinking aloud and Verbal probing (215, 216). When using Thinking aloud, the participant completes a task and simultaneously verbalizes his/her thought process related to the task. During Verbal probing the interviewer administers predefined questions to determine how the participant understands the material (instructions, scale items etc.), how information is recalled from memory, the adequacy of response categories, and how the participant selected his/her answer. The techniques complement each other and are often used together (215).

4.2.2 Translation of the COMFORT behavioral scale

The copyright holders gave permission to translate the COMFORT behavioral scale into Norwegian. Four translators (three certified translators and the PhD student) with Norwegian or English as their native language were engaged in the translation process. First two separate Norwegian translations were developed and subsequently merged into one, based on consensus, before two back-translations into English were developed. All translators translated into their native language and the back translators were blinded to the original scale. The bilingual committee (the PhD student and two clinicians, one Norwegian and one from the US, both with experience from both the American and the Norwegian health care systems) reviewed the back translations against the original and the translated versions of the scale to identify any discrepancies in conceptual meaning and agreed upon a preliminary Norwegian version of the COMFORT behavioral scale. This preliminary version was subsequently tested in cognitive interviews.

4.2.3 Participants and setting (cognitive interviews)

The study was carried out at Telemark Hospital in 2009-2010. Participants were recruited from staff at the pediatric ward, the neonatal intensive care unit (NICU), the pain treatment services, the general intensive care unit and the post-operative recovery unit using purposive followed by snowball sampling (218, 219). Clinicians with experience of caring for children and Norwegian as their mother tongue were eligible for participation. Thirteen clinicians were approached and 12 consented to participate, including two nurses, six specialist nurses, one nursing assistant, one pediatrics resident and two attending pediatricians. Their work experience ranged from three to 36 years. All participants provided written consent to participate.

4.2.4 Data collection and analysis

Data were constructed in individual interviews. After reading through the COMFORT behavioral scale, the participant viewed a 2-3 minute long video vignette showing an infant in distress and used the scale to score the infant's behavior. Two video vignettes from the COMFORT CD-ROM (220) were used alternately. The respondent was encouraged to think aloud while viewing the video vignette and during completion of the scale. Afterwards, predefined probes from a semi-structured interview guide were administered. All interviews took place in a quiet space chosen by the participant, were digitally recorded, and lasted between 30 and 65 minutes. The interview guide was piloted in one interview and found sufficient. Results from this interview were included in the data analysis. All recordings were transcribed verbatim and rechecked against the recording for accuracy by the interviewer. Further analysis was based on the written transcripts.

The aim of the data collection and analysis was to identify problems with the translated scale, and devise and test solutions to identified problems. Interviews were conducted in three rounds with interim data analysis and revisions of the preliminary COMFORT behavioral scale between each round. Each interview round was concluded when one or several problems were sufficiently illuminated (215). The first round compromised seven, the second three and the third two interviews. Data analysis was based on a deductive and interactive approach (221) combined with case and cross-case analysis using data matrices (219). Each interview was considered a case and a preliminary case analysis was performed after each interview and findings used to guide data collection in subsequent interviews. Cross-case analysis was initiated after the first interview round based on predefined organizational categories (221). These included the seven items in the COMFORT behavior scale (Alertness, Calmness/Agitation, Respiratory response, Crying, Physical movement, Muscle tone and Facial tension), the VAS-scale, the use of video vignettes, and the overall use of the scale. Problems identified in the scale were related to one of these categories. Problems identified after each interview round were discussed within the research group and with the scale developers and the preliminary version of the scale adjusted. This adjusted version was then applied in the next interview round. After the third and last interview round no adjustments were necessary.

Upon completion of data collection, the entire data material was reanalyzed independently by two researchers. Findings were then discussed and a final problems matrix that included identified problems and their solutions was developed (219). Matrix content was systematically re-checked against the data and revised when needed. A final Norwegian version of the COMFORT behavioral scale and an English back-translation was developed and approved by the developer of the COMFORT scale (Bruce Ambuel) (Appendix).

4.3 STUDY II

Study II was a prospective, observational measurement study.

4.3.1 Classical test theory (CTT)

This study used CTT to test the responsiveness and inter-rater reliability of the Norwegian translation of the COMFORT behavioral scale. CTT, validity, reliability and responsiveness are previously described and defined in section 2.6. The measurement concepts originally used in study II was based on Polit and Beck's terminology (222) as the COSMIN taxonomy (207) was not yet published at the time the study was designed. As a consequence, the concept "construct validity" is used in the article. The COSMIN taxonomy on the other hand uses the concept "responsiveness" for hypotheses concerning change scores within patients (207) and that term is used here.

4.3.2 Participants and setting

During eight months in 2011, 45 toddlers (69% boys) between 12 and 36 months (mean age 24.8 months) were consecutively included from a surgical outpatient unit for all ages at Telemark Hospital. Sample size was determined in advance to be able to detect a mean change of 2.5 points in the pain/distress range of the COMFORT behavioral scale with 90% power and a significance level of 95%.

Healthy children younger than 3 years (without disease or pre-admission use of medication which significantly influenced motor activity, facial expression, cognition or emotional state) admitted for elective surgery were eligible to participate. During the inclusion period, 22 of 67 (32.8%) eligible preverbal children were not included. The most frequent causes were parental dissent/lack of written consent (n=13) and surgery or post-operative care not completed (n=7). One child was not invited and one had participated earlier and was excluded.

All surgical procedures were minor. The most frequent were ear, nose and throat surgery (paracentesis, adenoidectomy and clipping of short tongue frenulum) (n=27) and urologic surgery (retentio testis, phimosis/balanitis and hydrocele testis) (n=14). Other surgical procedures included umbilical and inguinal hernias, removal of osteosynthesis material and excision of nevus (n=4). Pre-operative sedation (midazolam) was given to all children. All surgeries were performed under general anesthesia and 44 (97.8%) children also received

analgesics during surgery (opioids, paracetamol, corticosteroids and/or local anesthetics). Ten (22.2%) received analgesics post-operatively (opioids and/or paracetamol).

4.3.3 Data collection

COMFORT behavior scale scores were collected at pre-determined time points pre- and post-operatively – see table 2. Additional assessments could be made if and when the nurse believed the child was experiencing pain and participants were encouraged to perform parallel assessments whenever possible to enable evaluation of inter-rater reliability.

Table 2. Data collection. Assessments and interpretation

	Pre-operatively		**Post-operatively**							
Assessment	At admission	After premedication	On arrival at post-op (0min)	30 min	60 min	90 min	120 min	180 min	240 min	At discharge
Interpretation	Baseline (no pain/distress)	Light sedation	Deep sedation*	Highest score=pain/distress						

*If a laryngeal mask was still in place

All staff nurses (n=10) were trained in use of the COMFORT behavioral scale prior to data collection. Training consisted of a two-hour group session followed by 10 bedside training sessions, based on published recommendations (153, 155). Data collection was piloted before study start and findings used to optimize the data collection procedure. Data collection was closely monitored during the study to ensure data completeness. Data on background variables and the surgical procedure/unit stay were collected from the medical journal after discharge.

4.3.4 Data analysis

Central measure and spread was calculated for demographic variables and COMFORT scale scores (mean and standard deviation (SD) for normally distributed continuous variables and median and range for ordinal variables and non-normally distributed continuous variables). Floor–ceiling effects of the COMFORT scores as well as use of each of the five behaviorally anchored categories within each item were evaluated together with the fraction of scores indicating sedation (<14) and need for pain treatment (≥17) (155).
A mean change in COMFORT scores ≥ 2.5 was considered clinically important. Paired samples t tests were used to test three a priori hypotheses concerning (1) the difference between baseline and a pain/distress state (primary hypothesis), (2) baseline and light sedation, and (3) light and deep sedation. $p<0.05$ was considered statistically significant. Inter-rater reliability was calculated with intra class correlation coefficient (ICC) for the total COMFORT score and with weighted Cohen's kappa (Kw) for each scale variable. Strength of agreement was interpreted based on the arbitrary cut-off points proposed by Landis and Koch (223). Confidence intervals (CI) were calculated to assess the precision of both reliability coefficients (224) and mean differences in COMFORT scores. Analyses were

performed using SPSS version 19.0 (IBM Corp, 2010), except for weighted kappa values and corresponding confidence intervals, that were calculated with Stata (StataCorp, 2011).

4.4 STUDY III

Study III was a systematic review of systematic reviews.

4.4.1 Systematic review

A systematic review is characterized by a clearly formulated research question, the use of systematic and explicit methods to identify, select and evaluate relevant studies, and to extract and analyze data from the included studies. Statistical or narrative methods may be used to analyze and summarize extracted data (225). A key feature of systematic reviews is the systematic evaluation of risk of bias in the included studies (213). Generic guidelines for how to conduct systematic reviews have been published and this systematic review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (226) to help ensure its quality.

4.4.2 Identification and selection of studies

A general search strategy was developed using search terms describing the population (children from birth to 18 years of age), type of study (review), and the phenomenon under study (assessment of pain using observational pain measurement scales) and subsequently adapted to each of the five major databases The Cochrane Library, PubMed/MEDLINE, CINAHL, Web of Science, and PsychINFO. All databases were searched without any language or time limitations and the searches last updated in September 2016. Reference lists, articles, book chapters and grey literature were screened for additional studies. Systematic reviews evaluating and reporting on one or more measurement properties of observational pain scales for use in children from birth to 18 years of age were eligible for inclusion. After removal of duplicates, all studies were screened by two reviewers independently in two rounds, first titles and abstracts followed by screening of included full texts. Full texts were retrieved and reviewed if requested by one of the reviewers. After screening of full texts, any disagreements between reviewers were resolved through discussion and 14 review studies were included.

4.4.3 Review methods (data extraction)

Methodological quality of the included reviews (study validity) (213) was evaluated using the Assessment of Multiple Systematic Reviews (AMSTAR) checklist (227) a widely used generic checklist (228). AMSTAR consists of eleven items, with the answering categories "Yes", "No", "Can't answer" and "Not applicable".

Risk of bias (internal validity) (213) was measured using the Risk of Bias in Systematic Reviews (ROBIS) tool (229) available from www.robis-tool.info. ROBIS has a domain-based structure and is completed in three phases: (1) Assessment of relevance; (2) Identifying concerns within each of the four domains "Study eligibility criteria", "Identification and

selection of studies", "Data collection and study appraisal", and "Synthesis and findings"; and (3) Judgment of overall risk of bias in the review (229). ROBIS was developed to assess the risk of bias in review studies in general. For this study, the ROBIS guidance (230) was adapted to reviews on measurement properties using The COSMIN initiative's (www.cosmin.nl) approach to systematic reviews on measurement properties (205).

Two reviewers independently extracted descriptive information about each of the reviews, names of included and recommended observational pain scales, methodological quality, and risk of bias.

Their results were compared using simple percentage inter-rater agreement (231) and found acceptable for both AMSTAR (93.9%) and ROBIS (88.2%) scores. In cases of disagreement, consensus was established through a predefined verification and disagreement procedure.

4.4.4 Data analysis

Due to wide variability in review methods and presentation of results, it was not possible to aggregate findings from the included reviews in a meta-analysis. Data were analyzed narratively and included an overview of evaluated and recommended observational pain scales, in addition to details on the measurement properties of the most frequently recommended scale and how these were evaluated in the reviews recommending this scale.

AMSTAR responses were dichotomized; items scored "Yes" were considered as having high methodological quality while items scored "No" or "Can't answer" were considered as having low methodological quality. No items were scored "Not applicable". The methodological quality was evaluated and reported individually for each item and as a simple summary score (number of items scored "Yes") (227, 232) and mean/median, range and standard deviation (SD) were calculated.

Risk of bias was evaluated and reported for each of the four ROBIS domains and as overall risk of bias in the review (229, 230). The overall risk of bias reflected the evidence of the recommendations given in each of the included reviews.

4.5 STUDY IV

Study IV was a qualitative interview study.

4.5.1 Thematic analysis

Thematic analysis is frequently used in nursing research (233) and for this study we chose Braun and Clarke's inductive approach to thematic analysis (234). Thematic analysis is a basic and yet flexible method for qualitative analysis as is not limited to a specific theoretical or epistemological approach. This versatility makes it important to clarify the theoretical framework for a given analysis in advance. A thematic analysis is used to identify, analyze and report patterns or themes across a data set, for example a set of interviews (234). A

characteristic that made the approach suitable for this study was its emphasis on context (233), which is in accordance with the overarching theoretical framework for this thesis.

4.5.2 Participants and settings

For this study we used purposeful sampling (235). In a two-step sampling strategy we first identified eligible units and then recruited eligible nurses from these units. Units were considered eligible if they regularly cared for preverbal children and confirmed that one or more observational pain scales for use in this age group were available in the unit. Five units representing different hospitals, level of care and type of unit in two different countries were selected to ensure a diverse sample.

Information about the study was disseminated in each of the five units with the help of a designated contact person and the use of multiple written and verbal sources, including information in staff meetings, posters, letters of invitation, internal hospital e-communication, and informal conversations. Eligible nurses (fluent in Norwegian or English and with ≥ 1 year experience caring for preverbal children) volunteered for participation by submitting a short 1-page demographic questionnaire (age group, gender, educational background, and clinical experience).

Twenty-two nurses, 4-5 from each unit, were purposefully selected from the submitted questionnaires (n=45) to ensure maximal variation on these demographic variables and all gave their written consent to participate in an interview. Participating nurses were all female, median age 38 years, and had a most of their work experience with preverbal children.

4.5.3 Data collection

Data were constructed in individual semi-structured interviews led by the PhD student during a 6-month period in 2016. On average the interviews lasted 49 minutes (range 33-75 minutes). They were conducted during the participant's working hours and in a quiet room at the participant's work place. Using the Social Communication Model of Pain (12) as a theoretical starting point, our questions aimed to understand the pain transaction from the nurses' viewpoint, as well as internal and external factors influencing their assessment. The interview guide was piloted in both languages. Pilot data were not included in the analysis, but based on the findings from these interviews (n=3) participants were asked in advance to be prepared to discuss two self-selected examples from their clinical practice concerning pain assessment in a preverbal child.

The starting point for each interview was the self-selected examples. The participant was asked to describe each example and the interviewer asked questions as needed to gain a rich description of the child, the pain situation, the social and physical setting, the nurse's thoughts and actions, and internal and external factors influencing the situation. Only after discussing both examples was the participant asked about use of structured pain assessment scales, both related to the examples she provided and in her everyday practice. The participant was also given the opportunity to expand upon any subject related to assessment

of pain in preverbal children. All interviews were digitally recorded and all recordings but one were of very good quality. The recordings were transcribed verbatim by native-speaking transcribers. Transcripts were subsequently checked against the recordings for accuracy by the interviewer. In addition, we collected descriptive information on the participating units from the unit manager or a designate, together with copies of available pain assessment scales and procedures.

4.5.4 Data analysis

Data were thematically analyzed based on the six phases described by Braun and Clarke (234).

1. Familiarizing oneself with the data. Transcripts were systematically checked against the recordings by the interviewer to ensure accuracy. Further data analysis was based on the written transcripts. Amended transcripts were read and reread several times and initial ideas noted down. The ideas generated during this process guided further data analysis.
2. Generating initial codes. Two authors separately coded the first two interviews in detail before comparing codes and ideas. Ideas generated by each of the authors and from this comparison were used to devise an initial coding structure to guide further coding. New codes were added as needed. Coding was guided by the preliminary aim "to describe nurses' thoughts and actions related to assessment of pain and the use of observational pain scales in children younger than 3 years". The list of codes was reviewed by all co-authors when half of the interviews were coded. Each author read 2 interview transcripts in detail to determine the "fit" of the coding structure.
3. Searching for themes. Codes were sorted into themes and sub-themes, and relationships between different codes and themes explored. Examining the initial codes, we discovered that several were related either to strategies used by nurses to assess pain, or the idea of using of not using pain scales. This insight was used to formulate a more targeted aim and codes that did not fit into the more targeted aim were discarded. At the end of this phase we had a list of themes, sub-themes and data coded to them.
4. Reviewing themes. Themes were refined and a description or map of how they were connected was developed and discussed between authors, before the entire data set was re-read by the PhD student to determine the fit of the map to the data and code additional data to its corresponding theme.
5. Defining and naming themes. Next, themes were further defined and refined and the content within each theme analyzed and written in detail. Co-authors verified the analysis of each theme against the data coded to that theme.
6. Producing the report. Findings and insights from the previous five steps were used to tell the story of how nurses assess pain in preverbal children.
The analysis process was not as linear as described here, but ended up as a series of loops moving back and forth between steps 3, 4 and 5.

Collecting data from multiple sites in two different countries and the application of investigator triangulation in coding and analytic decisions increased the trustworthiness of the

findings (236). A detailed description of the sample and settings was provided to assist the reader in determining the transferability of the findings. NVivo qualitative data analysis Software (QSR International Pty Ltd. Version 10, 2012) for Windows was used to code interviews and facilitate the analysis process.

4.6 ETHICAL CONSIDERATIONS

Special considerations are needed when vulnerable children are included in research (237). In this project, children were only included in study II. This was an observational study with minimal risks for the child, with the exception that touching the child to assess muscle tone during the immediate recovery might disturb the child. In these situations the nurses were instructed to rely on visual observations only. Having their child participate in a pain assessment study might result in increased parental stress and worry. To diminish the parental burden, parents were given information about the study and decided upon participation prior to admission. They also received ongoing support and information from the attending nurses during the child's hospital stay. The study was approved by the Regional Ethics Committee of Southern Norway (2010/1268) and registered in Clinical Trials (NCT01181687). Clinicians participated in studies I and IV. This type of research is outside the remit of the Norwegian Health Research Act of 2008 and not reviewed by the Norwegian Regional Committees for Medical and Health Research Ethics. Instead study I was approved by the head of department at Telemark Hospital and by the Norwegian Social Sciences Data Services (NSD) (21863). NSD is the Data Protection Official for Research at Telemark Hospital. Study IV was approved by NSD (46336), the Data Protection Official for Research at Oslo University Hospital (OUS) (20162813), the Research Ethics Board at IWK Health Center, Halifax, Canada (1021714) and by the heads of all participating units. The risks for interview participants were considered minimal, although there was a small risk that the interview might illuminate practices and attitudes the participant was not comfortable with acknowledging or sharing. To address this risk, the participant was given time for debriefing after the interview if needed. In all three studies (I, II, and IV) participants were given both written and oral information about the study, including how participation was voluntary and that the participant was free to withdraw at any time. Data from all three studies were collected and maintained in accordance with national and local laws and regulations. Digital recordings were directly and securely transferred to an encrypted and password protected area on Telemark Hospital's server. Copies of recordings from Canadian interviews as well as transcripts of these interviews were exchanged using IWK's secure data transfer service. As a systematic literature review, study III did not require any approval, but was registered in the International Prospective Register of Systematic Reviews (PROSPERO) (CRD42016035264).

5 MAIN FINDINGS

Findings are presented thematically under the two main headings “Pain measurement scales” and “Assessment of pain”.

5.1 PAIN MEASUREMENT SCALES

5.1.1 Measurement properties of the COMFORT behavioral scale

Content validity of the Norwegian translation of the COMFORT behavioral scale was improved by cognitive interviews with future users of the scale. Several problems concerning both the translated and the original versions of the scale were identified and addressed, either by a change in the translation or as an aspect to be addressed in future user training. In contrast, no problems were revealed during the standardized and commonly advocated forward-back-translation procedure (study I).

Overall inter-rater reliability in the sample and setting being studied was high (ICC=0.96; 95% CI 0.93-0.98), with the limitation that half of the scores used to calculate inter-rater reliability were between 6 and 8 points on a scale ranging from 6 to 30 points. Although the inter-rater reliability for “Muscle tone” and “Facial tension” was substantial (Kw>0.75), it was somewhat lower than for the other items in the scale (study II). This correlated with our find during the cognitive interviews that assessments varied more between raters on these items than on any of the others (study I). Feedback from nurse assessors participating in study II pointed towards the COMFORT behavioral scale not being the best choice of scale for non-verbal children outside the PICU setting. It worked well enough as long as the child was in bed, but was considered unsuitable for ambulant children (personal communication).

Interview participants, of whom a majority were nurses, understood several words and phrases in the COMFORT behavioral scale differently from what the scale developers intended. The participants also had difficulties with descriptive text added to some of the categories in the COMFORT behavioral scale. Although participants expressed appreciation for the descriptive texts when reading through the scale, they made actual use of the scale more difficult because the observed behavior frequently did not fit any of the descriptions. All in all, more problems were identified in the part of the interview based on actual use of the scale than when the participant was just thinking aloud while reading through the scale (study I).

The responsiveness of the Norwegian version of the COMFORT behavioral scale was supported for assessment of sedation, but not for assessment of pain/distress in children 12-36 months old after outpatient surgery. Each participant (n=45) was assessed between 3 and 9 times (median 7). More than half of the recorded scores (175/307; 56.9%) were in the sleep/sedation range of the scale (between 6 and 14 points) and the results showed clinically important differences (>2.5 points) between a normal, pain-free state and light sedation and between light and deep sedation respectively. Consequently, the Norwegian translation appeared useful to assess sedation before and after minor surgery in non-ventilated toddlers

(12-36 months). However, the primary hypothesis was not confirmed. Although the difference between a normal, pain-free state and a pain/distress state was statistically significant in this sample, it was not big enough to be clinically important. Still, a wide range of pain/distress scores were recorded, all behavioral anchors in the scale were used at least once and recorded COMFORT scores covered all values between 6 and 28 points (scale range 6-30 points) (study II).

5.1.2 Scale validity and recommendations for pain scales

Although the systematic reviews included in study III (n=12) showed little consensus on which of the multitude of published scales (n=65) to recommend, the scales most frequently recommended were FLACC (156)/rFLACC (238) in 7 of 12 reviews and COMFORT (154)/COMFORT behavioral scale (153) in 5 of 12 reviews. However, these and other recommendations given regarding pain scales for use in clinical practice had low evidence value and should be interpreted with caution. Overall the included reviews demonstrated low methodological quality with a mean AMSTAR score of 3.3 of 11 points (SD 2.24). Although risk of bias varied within and across reviews, none demonstrated an overall low risk of bias and 10 exhibited high risk of bias. Weaknesses were identified in all steps of the review process, including study identification and selection, data extraction and study appraisal, and synthesis and interpretation of findings. Primary studies were included regardless of study type; the methodological quality of the included studies was infrequently assessed, and data analysis was mainly limited to a reiteration of values and findings from included studies without any attempt to synthesize or aggregate these findings. How findings from the included primary studies contributed to the evaluation of the pain scales under review was not described. Data from the reviews could not be aggregated per pain scale as authors used different rating systems, different criteria for inclusion of studies, and did not synthesize the findings from the primary studies included (study III).

5.2 ASSESSMENT OF PAIN

Assessment of pain was in general considered more difficult in preverbal children expressing pain through behavior than in older and verbal children (study IV).

5.2.1 Use of pain measurement scales

Structured pain scales were infrequently used in clinical practice. Neither pain scores nor behavioral responses associated with pain were considered pain-specific. The primary reason for non-use in Canadian units was lack of awareness and access. Reasons for non-use in Norwegian units included both personal and organizational factors, such as lack of competence in using the scale, forgetfulness, scale not used by colleagues, and/or implementation of scale not facilitated or supported by the management. Organizational factors such as having the pain scale integrated into the patient records and scale scores being routinely evaluated by physicians had a positive influence on scale use. Pain scales were never used in isolation but integrated as an aspect in a composite assessment of pain (see 5.2.2). The Norwegian nurses perceived pain scores as helpful in communication with

colleagues about the child's pain management, to justify the need for pain management and to evaluate treatment effect. One nurse used pain scores to verify her clinical assessment. Canadian nurses had little or no experience with structured observational pain scales and identified fewer benefits from their use (study IV). Behaviorally anchored FLACC scores ranging from 0 to 10 points were considered equivalent to self-reported scores on a Numeric Rating Scale (unpublished data).

5.2.2 An individualized and complex process

Instead of using pain scales, nurses preferred to assess pain using their clinical judgment. Pain assessment was described as a complex, individualized and non-linear process. Although similar strategies were used across assessors and situations, it was not possible to identify a generic pain assessment process. The nurse selected and combined strategies from a wide repertoire, based on the child's situation and expression of distress. These strategies were organized under the three themes "Gaining knowledge about the individual child", "Having experience-based knowledge", and "Combining knowledge with observations". Strategies from all three themes were included in all assessments (study IV).

The assessment was always grounded in previous knowledge and experiences as well as being individually tailored to the situation and child.

Well, I mean every child is different. It depends on what's wrong with them, and why they're here, and what they've had done. And you kind of see a broad picture that, okay, well, kids who normally have this done, they have…you know, usually they tolerate this much pain or they don't need this much pain medication. But everybody has a different pain threshold, and everybody reacts differently. So even – even little ones. (R22)

Knowledge about the individual child was gained by establishing and cultivating good relationships with both child and parents (study IV). Clinicians, of whom a majority were nurses, perceived assessment as more difficult and less reliable when they did not know the child and/or were not able to touch or interact with the child (studies I and IV). Lack of child knowledge made the nurse more dependent on colleagues' or the parents' evaluation of the situation. Parents were in general the main information provider regarding the child's normal behavior and were considered good at identifying subtle changes in the child's behavior. Further inclusion of parents in the assessment of pain varied considerably (study IV).

Nurses considered previous experience a prerequisite for correct assessment of pain (study IV) and lack of experience-based knowledge was associated with variability across participants (studies I and IV). Previous experience facilitated detection of pain and sometimes resulted in an expectation of pain even before seeing the child. However, expectations could also negatively influence assessment and result in underassessment of pain (study IV).

Experience-based and child-specific knowledge was combined with information from current observations of the child. Different strategies contributed information used to reach a decision

on whether or not the child was experiencing pain. These strategies included identifying a probable cause for pain, eliminating other sources of distress, using behavioral pain scores, evaluating behavioral change and/or effect of interventions on behavior, and using a personal and contextual approach (study IV).

It was considered much easier to assess pain in children when a probable cause of pain could be identified than in cases where no diagnosis of procedure could explain observations of behaviors indicative of pain. Behavioral cues were seldom considered pain-specific and nurses systematically tried to eliminate other possible causes for the observed behavior such as hunger or anxiety. Pain was considered a diagnosis based on exclusion; the nurse concluded that the child was probably experiencing pain only if the child continued to display pain-related behavior after an exhaustive elimination process. A few nurses included scores from a structured pain scale in their assessment. The effect of analgesics on the child's behavior was used to verify a pain hypothesis and changes in behavioral and sometimes physiological cues over time were used to evaluate changes in pain intensity. Although known generic pain cues were used to assess pain, nurses emphasized individual and child-specific signs and changes or deviation from what was considered the child's normal behavior. The child's previous experiences were also taken into account, together with influences from the physical and social environment (study IV).

6 DISCUSSION

6.1 SUMMARY OF RESULTS

Cognitive interviews identified several problems with the content validity of the Norwegian and original versions of the COMFORT behavioral scale. The responsiveness of the translated version was supported for assessment of sedation but not for assessment of pain/distress. Scale recommendations given in systematic reviews addressing the measurement properties of these scales had low evidence value and should be interpreted with caution. Observational pain scales were infrequently used in clinical practice and neither pain scores nor behavioral cues associated with pain were considered pain –specific. Instead, nurses expressed strong preferences for pain assessment based on clinical judgment and individually tailored to the child and the situation. When assessing pain, nurses combined experience-based and child-specific knowledge with one or more specific strategies to interpret observations of and information from the child. Described strategies included identifying a probable cause for pain, eliminating other sources of distress, evaluating behavioral change and/or effect of interventions on behavior, using a personal and contextual approach, and using behavioral pain scores.

The discussion of study results that follows is structured according to the Social Communication Model of Pain (12) and addresses the pain assessment step in the model as well as intra- and interpersonal factors influencing the assessment of pain (Figure 1).

6.2 PAIN ASSESSMENT

6.2.1 Preference for assessment based on clinical judgment

Nurses first and foremost found pain assessment in preverbal children challenging and complex. They preferred to assess pain based on their clinical judgment (study IV), and this preference is also well described in the literature (125, 128, 239-241). Clinical judgment of pain entailed combining experience-based and child-specific knowledge with information from observations of the child and information from applying different, well-defined strategies such as identifying a probable cause for pain, eliminating other sources of distress, using behavioral pain scores, evaluating behavioral change and/or effect of interventions on behavior, and using a personal and contextual approach (study IV). This description of clinical judgment concurs with the definition of pain assessment as a systematic and holistic approach (12, 48). The assessment and the application of the different assessment strategies (for example, elimination of other sources of distress) were also dependent upon interaction between the child and the nurse. With a pain diagnosis as an outcome, the described assessment practice complied with the concept of pain as a social and dynamic transaction (12).

The strategies used by the nurses in study IV in the assessment of pain had to some extent been described earlier, including identifying a probable cause for pain (118-122, 242), eliminating other sources of distress (120, 135, 145), evaluating behavioral change and/or

effect of interventions on behavior (118-121, 128, 145, 243, 244), and using a personal and contextual approach (118, 120, 125, 132, 142, 242, 245). However, no previous study described the full set of approaches including the use of structured pain scales as part of a clinical judgment. Only a handful of clinical studies have addressed nurses' assessment of pain; two doctoral dissertations have described clinical assessment of pain in preverbal children (120, 135). The dissertations are over a decade old and pain scales were not part of the assessments. A few studies have described assessment as part of nurses' decision-making process in regard to medication (118, 121, 123, 134). Al in all, there has been a noticeable lack of research on nurses' pain assessment in the last decade, except for a few studies on their perception of how pain is expressed (244), opinions regarding barriers and facilitators to effective pain management (132), nurses' experience of being comfortable or not when caring for children in pain (128), and a questionnaire study on pain assessment practices (125).

6.2.2 Structured pain scales seldom used

Few nurses described using pain measurement scales in the assessment of pain and the common belief was that the use of pain scales could not replace clinical judgment. Scale availability and positive attitudes towards them were not enough to instigate actual use of these scales (study IV). The limited use of structured pain scales was in accordance with previous studies (27, 121, 125, 128). A systematic review of nurses' pain management practices concluded that children's behavior seemed to influence nurses' pain assessment more than structured pain scales (241). The belief that pain scales could not replace clinical judgment concurred with nurses' views in a previous study (244) and with the definition of pain measurement as one aspect of pain assessment (48) and not a substitute. The findings further reflect literature reports of the persisting difficulties implementing pain scales into clinical practice (74, 246) and the persistent gap between what is advocated in mainstream research, guidelines and policies, and what is done in practice to assess and manage pain in children (133).

6.3 INTRAPERSONAL FACTORS INFLUENCING PAIN ASSESSMENT

6.3.1 Beliefs regarding pain expression

Behavioral cues were not considered specific to pain (study IV). This may explain the number of strategies nurses used to interpret observations as part of their clinical judgment. The strong emphasis on knowing the child's normal and nurses' inclusion of deviations from the child's normal as an aspect of pain assessment, spoke towards a strong recognition of behavioral expressions of pain not being consistent within or across individuals.

Nurses believed that older preverbal children sometimes exaggerated their pain (study IV). This reflected an understanding of pain behavior being under voluntary control and not always an "honest" reflection of the child's experience (12). Williams (1) has suggested that the purpose of pain behavior may be to elicit help from others to manage escape, recovery and healing from pain. If the child does not receive help, this may result in a stronger attempt

towards getting the necessary attention from his surroundings by exaggerating the pain behavior. Exaggerations were sometimes met with distrust (study IV); instead of recognizing it as a result of untreated pain from lack of help, some nurses considered it a way to elicit sympathy or gain benefits.

Children who showed no or diminished signs of pain in response to a situation known to be painful were not discussed in the interviews which may be a concern. The body always strives toward homeostasis or balance and the conservation of energy (60, 247). Ongoing and untreated pain may result in a change in the outward signs of pain; they may be different, toned down or missing completely (60) to limit the expenditure of the body's energy resources. Studies have shown infants that presumably do not react at all to procedures known to be painful (248, 249). Bonavita and De Simone (60) have suggested that escapable pain results in a fight or flight response while inescapable pain results in more passive coping strategies, for example reduced responsiveness. Think of two verbal children, one with chronic pain from rheumatoid arthritis and one with post-operative pain. They may report the same pain intensity, but their behavioral expressions will most certainly not be identical. The process resulting in suppression of pain expression is probably shorter in the youngest or otherwise most fragile children, who have more limited energy resources than older and healthier individuals.

6.3.2 Biased towards underestimation

Nurses found that lack of experience made pain assessment difficult, while having experience-based knowledge made them feel more secure in their ability to assess pain. General knowledge about children was believed to improve the ability to assess pain in children (study IV). This contradicts the research associating longer experience with greater degree of underassessment (77, 78, 82), but nurses are probably not aware of biases influencing their judgment. It has been hypothesized that the association between increased underestimation of pain and increased professional experience may be explained by previous exposure to high levels of pain in others (vicarious pain) (78) and subsequent desensitization (89), maybe to protect the professional observer from unwanted or negative consequences associated with repeated exposure to vicarious pain (72). This hypothesis is supported by both experimental and clinical studies. Prkachin and colleagues (79) showed that prior exposure to facial expression of high pain intensity made participants less willing to acknowledge pain in others. Using brain imaging techniques, Coll and colleagues (250) demonstrated that short and intense exposure to vicarious pain altered the cerebral responses to subsequent episodes of vicarious pain, indicating a decrease in the perceived importance of this pain. A qualitative study described how nurses' guarded themselves emotionally against children's pain and the authors concluded that this defense mechanism may lead to underassessment (251).

6.4 INTERPERSONAL FACTORS INFLUENCING PAIN ASSESSMENT

6.4.1 Relationship between the child and the nurse

Knowing the child was considered an important fundament for pain assessment (studies I and IV). This view was also reflected in nurses' actions where assessments were individually tailored to the individual child (study IV). Clinicians in study I found the use of a pain scale more difficult when they were not able to touch or interact with the child. When assessment is based on a film, there is no opportunity for interaction. Early studies on assessment of pain in preverbal children relied on filmed or written vignettes to elicit cues nurses used to assess pain (136-139, 141, 142), factors influencing their assessment (140, 143, 144) or to describe different modes of assessment (145). A major limitation with these studies is that they are decontextualized and impersonal, which probably makes assessment and assessment decisions different from how they would be in clinical practice. This approach is further based on the assumption that pain can be "read" from a child's expression or behavioral cues, and does not correspond with the understanding of pain as a social transaction.

Nurses had different approaches towards communication and interaction with the child. Some nurses seemed to rely mainly on observations of the child and information from parents; others included verbal and non-verbal communication with the child, either directly or indirectly through the parents (study IV). Coyne (252) recently argued that the child's perspective and agency in pediatric nursing care needs to be strengthened in compliance with the rights of the child (253). This entails a conceptual shift from nurse-parent to nurse-child collaboration where care is planned around the child's own perspective and preferences while still provided in the context of the family and community.

6.4.2 Parents

The children themselves were too young to provide the necessary information and parents were considered the main source of knowledge about the child (study IV). Studies have shown that nurses frequently use parents to provide information about the child (119-121, 127, 242, 245). This concurs with the increased emphasis placed on parents' role in their children's pain management during the past 20 years from virtually non-existent (122, 254) to being viewed as an important aspect of care (118, 125). However, few nurses involved the parents more closely in the assessment process (study IV), which was in accordance with previous studies (121, 128, 255) and indicates that this view is not necessarily transformed into practice (125).

6.4.3 Scale aspects

Structured pain scales played a small part in nurses' pain assessment and many did not use them at all (study IV). This finding may in part be explained by aspects with the scales themselves.

The large number of published observational scales (study III) makes the selection of scales difficult. Many published scales have undergone very limited validity testing, clearly

demonstrating that this field has been more focused on developing new scales than on validating those that already exist. This may be interpreted as attempts to cure perceived problems with existing scales by developing new ones. Still, an overview of the items in existing behavioral scales shows considerable overlap (256), indicating that a new scale seldom means a new and different approach. Although new scales are needed for some populations and settings, there is no need to develop new scales for the acute procedural pain. Instead, focus should be directed towards further validation of existing scales (146).

Recommendations given in systematic reviews addressing the measurement properties of observational pain scales had low evidence value and should be interpreted with caution (study III) indicating that even when the validity of a given scale has been evaluated in several studies, determining scale validity is difficult. The aggregation of results from primary validation studies was hampered by the use of different taxonomies, resulting in similar concepts having different interpretations or different concepts the same interpretation (study III). To exemplify, hypothesis testing regarding change scores in study II was defined as construct validity testing in accordance with Polit and Beck (222), but as responsiveness testing according to the COSMIN taxonomy (207). The main reason for the development of the consensual COSMIN taxonomy was the prevailing lack of consensus regarding how to define the different measurement concepts (205, 206). This lack of consensus is probably one of the reasons why reviews included in study III merely reiterated the results from the primary studies. Although some presented a full or partial taxonomy, few extracted data from the included primary studies in accordance with a common taxonomy, making synthesis of findings and valid comparisons across studies impossible.

A majority of the review studies did not present criteria for how test results should be interpreted either. Only one defined criteria for reliability coefficients, but not for validation tests (study III). Just as hypotheses should be defined prior to the study, it is important to define in advance cut-off scores for tests. The use of the arbitrary framework proposed by Landis and Koch (223) to evaluate reliability coefficients in study II is an example of the use of predefined cut-off scores to interpret findings in a primary study.

Published reviews tended to include all validation studies, regardless of study type. Some also included treatment studies where the scale had been used as an outcome measure or studies where the scale under review had been used as a gold standard in the validation of another scale or measurement approach (study III). Different types of validation studies contribute differently to the overall validity of the scale, and from a methodological point of view, the extensive use of correlation of scores from different scales in validation studies does little to bring the field forward. These studies are very easy to conduct, which probably explains much of their popularity and frequent use. Their validity is weaker than that of well executed studies testing construct validity or responsiveness based on predefined hypotheses, as the results are dependent on the validity of another scale or another representation of the construct and are not evaluated based on comparisons with the construct itself (205). Moreover, authors have seldom defined a cut-off score for a sufficient correlation in advance.

These correlation studies mainly serve to maintain a circular logic that seems to permeate this field. From the outset, structured pain scales were developed as a more objective alternative to clinical judgment. Still, the first scales were frequently validated against a VASobs score (see for example references 156, 157, and 257-260). In reality this meant that the scale was validated against clinicians' judgment which is what they wanted to get away from by developing the scale in the first place. As more observational/behavioral scales were published, they were frequently validated against each other (see for example references 156, and 260-263). In addition, physiological methods suggested as more objective approaches to pain assessment, like skin conductance, have been validated against behavioral scales (see for example references 264-266).

6.4.4 Contextual and organizational factors

Canadian units did not have structured pain scales available. Instead, what they referred to as a scale was a locally compiled list of 11 behavioral and verbal cues associated with pain. This list was used to document the presence or absence of pain in the nursing records. As such, their practice was in compliance with hospital regulations as nurses routinely assessed, documented and managed pain (study IV). Still, the presumed benefits of a structured pain scale, for example the ability to grade pain, were still missing. Documentation of pain scores is frequently considered a quality indicator and treatment algorithms have been developed where pain scores dictate the treatment, often with unwarranted and negative consequences (267). Studies have described how nurses develop work-arounds to comply with guidelines and practice standards (268) presumably to avoid time-consuming practices that do not make sense to them. An observational study showed that nurses frequently recorded and documented pain scores without a scale being used (126). In a retrospective chart study 97% of recorded scores on the Children's and Infants' Postoperative Pain Scale (CHIPPS) (269) on the first postoperative day were either 0 or 4. The authors speculated that nurses did not actually use the CHIPPS scale, but instead determined whether an intervention was needed or not as pain scores of 4 or higher required an intervention according to local standards (270).

6.4.5 Nurses' understanding and use of pain scales

Pain scores were not considered specific to pain. This may be another reason for pain scales not being used as nurses believed that the use of pain scales could not replace clinical judgment (study IV). The use of structured scales limits assessment to a standardized set of items in a structured scale. One advantage of this generic approach is that it focuses the assessment on specific behaviors or cues associated with pain. A disadvantage is that individual aspects are not taken into account. In their assessments, nurses in study IV routinely relied on what they knew about the individual child and deviations from the child's normal behavior in addition to more generic behavioral cues associated with pain.

Behavioral pain scores were misinterpreted as being equivalent to self-report scores of pain (study IV). This view on behavioral scales and checklists used in patients unable to self-report pain are common (271). However, a pain scale based on observation of behavior

provides a pain behavior score reflecting the number of observable pain cues present (272) and there is no scientific data supporting the assumption of equivalence. On the contrary, studies have shown that behavioral indicators do not correlate well with self-reported pain intensity (273, 274).

Nurses who used structured scales described an assessment practice characterized by a multifaceted approach, very similar to their colleagues, with the difference that pain scales were one of several strategies they applied and pain scores were integrated into their clinical judgment (study IV). Interestingly, this mirrors a change in our approach to pain assessment previously suggested in the literature (5, 275) where a set of assessment strategies, of which one is a structured pain scale, are combined in a composite assessment of pain. Study IV is the first to describe the actual use of this approach in clinical practice.

Guidelines from the American Society of Pain Management Nurses (5) recommend pain assessment based on a hierarchical approach where a set of assessment strategies are applied in a predetermined order with self-report using a validated scale as the first step whenever possible, followed by a search for probable cause, evaluation of child behavior, parental report of pain and changes in behavior, performance of an analgesic trial, use of systematic assessment, use of behavioral scales, minimal emphasis on physiological indicators, and reassessment and documentation. Twycross and colleagues (275) have suggested a "bundled approach" where pain scores, behavioral cues, information from parents, and information regarding the child's condition (diagnosis) are systematically integrated in a comprehensive clinical judgment. The strategies applied by the nurses in study IV partly overlap with suggested strategies in both approaches. However, the nurses' overall approach to pain assessment was not hierarchical, but more closely resembled a bundled approach as selection and sequencing of strategies varied across assessors and situations in response to child and contextual factors.

6.5 METHODOLOGICAL CONSIDERATIONS

The main strength across all four studies was the systematic approach to study design, data collection and data analysis. The use of different research methods and approaches resulted in findings that illuminate pain assessment from different angles; findings regarding the measurement properties of observational pain scales helped in understanding findings regarding nurses' pain assessment practices and vice versa. Strengths and weaknesses are thematically discussed in the next sections.

6.5.1 Selection of the COMFORT behavioral scale

The COMFORT behavioral scale was originally selected for translation and later use at Telemark Hospital based on scale content or face validity and the number of validation studies published. The modifications making the scale usable in non-ventilated children were taken as an indication that it would be useable also outside the pediatric intensive care setting. It was taken into account whether validation studies had been done or not, and if they had, the authors' conclusions were borne in mind, but the methodological quality of these studies was

not considered. This approach was in line with the most prevalent way of thinking about validation studies and the validity of pain scales, as illustrated by a well-known set of criteria frequently used to determine scale validity (276). According to these criteria, all that is needed to define a scale as well-established is that it has been presented in at least two peer-reviewed articles by two different researchers or groups, that the scale and a scoring manual are available, and that statistics indicating good validity and reliability has been presented in at least one article. Note that what constitutes "good reliability and validity" is not further defined. Results from the first two studies (I and II) indicated problems with this approach and study III demonstrated the necessity of taking study validity into consideration when evaluating scale validity.

6.5.2 Piloting of data collection

The data collection procedure used in studies I, II and IV was piloted before study start. In study I the pilot test went well and the pilot interview was included in the data. In studies II and IV the results from the pilot tests uncovered problems that were subsequently addressed before data collection started. Although the changes made in study II ensured high quality data, limitations in the study sample and setting negatively impacted the results and study generalizability. In study IV on the other hand, modifications to the interview guide as a result of the pilot study contributed to rich, in-depth data.

The original plan for study II was to include children undergoing both outpatient and inpatient surgery. The pilot study showed that data collection did not work for pediatric surgical inpatients due to the complexity of the care environment. Because nurses from several units (pediatric ward, pre-op, post-op or intensive care) were responsible for different parts of the data collection for each patient, assessments were forgotten, forms were misplaced, etc., even though all nursing personnel had undergone systematic training, were informed about the study and a study folder was included in the papers that accompanied the child. This problem might have been solved by having a study nurse follow each included patient through the process, but the study did not have that sort of funding. Instead inclusion was limited to children admitted to the outpatient unit. This organization was much simpler and the same handful of nurses performed all assessments. The drawback was that only children scheduled for minor surgery were included. They all received good pain treatment peri-operatively, the post-operative stay was relatively short, and consequently few displayed behavioral cues indicative of pain or distress postoperatively. However, the collected scores covered almost the whole range of the COMFORT behavioral scale (6-28 of 6-30 points) indicating that the pain/distress hypothesis may be verified in a different sample and setting. Moreover, the scores used for reliability testing demonstrated a distinct floor effect, with a majority of the scores in the lowest range of the scale (<8 points), meaning that the inter-rater reliability for scores in the pain/distress rage of the scale could not be determined in this study.

In study IV the use of self-selected examples based on concrete patient situations contributed to rich data. In this, the pilot study was instrumental. Two pilot interviews were performed on

the Norwegian interview guide. They were difficult for both the interviewer and the participant and resulted in thin data. When participants afterwards were asked what could have been different, both suggested being asked to think about examples beforehand, as they spent a lot of energy trying to remember suitable examples. Gimbler-Berglund and colleagues (118) encountered the same problem in their interview study on factors influencing nurses' pain management in children.

Pilot data were not included in studies II and IV because of low quality; this was caused by missing data in the study II pilot and thin data in the first two pilot interviews for study IV. Moreover, the Norwegian pilot interviews for study IV resulted in significant changes in the interview guide and a different type of interview. The Canadian pilot interview for study IV was also excluded because the participant was a nurse with previous experience from the NICU, and this was outside the setting defined for this study.

6.5.3 Eliciting users' perspectives

In study I cognitive interviews were used to elicit users' perspectives on the Norwegian translation of the COMFORT behavioral scale and its use, as advocated by the translation guideline (194). The inclusion of the users' perspectives in translation and validation of measures is in accordance with the recommendations from the COSMIN group (205). Although cognitive interviews are considered a monolingual approach that only addresses issues with the translated version of the scale (277), the interviews were also able to identify some weaknesses and ambiguities in the original scale, indicating problems with overall content validity of the scale. This finding has been replicated in a separate study translating and culturally adapting a different scale (278). Together these studies indicate that problems in the original versions of the scales went unnoticed during scale development and subsequent validation. Unfortunately such weaknesses are not easily corrected later (206). It seems reasonable to assume that these problems would have had a better chance of being detected during scale development if scale users had been involved systematically, yet the user perspective is frequently missing from scale development. Although scale items are often generated based on a literature review, expert opinion and/or interviews with clinicians, patients or clients, it is generally experts who ultimately review and approve the compiled scale. How users understand and use the scale is not commonly reported, indicating that scale developers seldom take users' understanding of scale items into account. Cognitive interviews seem like a valuable approach in the development phase to strengthen the construct validity of the scale under construction.

6.5.4 Cultural influences

Study I demonstrated how cultural factors influenced nurses' understanding of words and content in a pain scale. In study IV cultural factors influenced data collection as it was carried out in both Norway and Canada. Cross-cultural interviewing increases the possibility for misunderstandings (235), but several steps were taken to minimize this risk, including a single bilingual interviewer with nursing background and in-depth knowledge of the field, a

bilingual interview guide pretested in both cultural settings, familiarization with the settings prior to the interviews, asking participants for clarification if needed, use of professional native transcribers, and a cross-cultural research group. Further, the social and health care systems in the two countries are fairly similar, which facilitated a common understanding. One major and unexpected limitation resulted from different understanding of the concept "pain scale". The intention was to include units that had structured and published pain scales available. To the research group the word scale implied a published, structured pain scale, for example the FLACC (156) or COMFORT behavioral scale (153), but this was not articulated. It turned out that the Canadian units used the term for a locally compiled list of behaviors associated with pain (i.e. what is often referred to as a non-validated scale in the literature), and the nurses did not know about the structured pain scales that were recommended for preverbal children in the hospital's pain management guideline. As a consequence, the sample included nurses working in units with and without access to structured pain assessment scales. Questions related to use of scales were only introduced after participants had presented and discussed their examples in detail, which means that this problem was not discovered until the first Canadian interviews had been carried out and the rest of them already planned.

6.5.5 Statistical significance vs. clinical importance

In study II a statistically significant difference was found for all three hypotheses, but the result from the pain/distress hypothesis was not considered clinically important as the group difference between baseline and the highest postoperative score was less than 2.5 points. This underscores the importance of not evaluating results solely based on statistical significance; although such significance can easily be achieved with a large enough sample, the difference may still be too small to have any clinical implications. An arbitrary clinically important difference was used in this study as this had not been previously determined for the COMFORT behavioral scale. The rationale was that a difference of 1 point would not make any difference in how clinicians would manage a situation, but a difference over 2 points probably would. A recent systematic review included 37 studies with data from 8479 patients and found that the minimum clinically important difference (MCID) in acute pain varied widely across studies (279). Their conclusion was that MCID needs to be determined case by case, and as such supports the approach chosen in study II.

6.5.6 Selection of analysis strategy for qualitative studies

A deductive approach to data analysis was devised and applied in study I. Items and other elements in the COMFORT behavioral scale made up a set of predefined analysis categories. The structured, matrix-based approach (219) gave a good overview of the data and eased the process of interpreting relevance and finding solutions to each problem. A structured approach (280) previously used to evaluate cognitive interviews as a research method (281) was initially tried, but did not address the relevance of identified problems or how to handle them. Traditional analysis of cognitive interviews using an unstructured and largely undocumented approach (215, 217, 280) had low reliability (217) and was not an alternative.

Data from study IV were analyzed using inductive thematic analysis (234). The rich narrative material was well suited for thematic analysis. Another alternative was content analysis (282) as both are basic, flexible, easy for a novice researcher to learn and use, and not limited to a specific theoretical or epistemological approach. However, an interest in both latent and manifest interview content, a preference for analysis of content in relation to its context, as well as a lack of interest in quantification, all favored thematic analysis (233).

6.5.7 Selection of tools to evaluate systematic reviews

A challenge in study III was to identify and apply methods for data extraction and evaluation of study quality and to devise an analysis strategy. Methodological quality was initially assessed using the well-established AMSTAR checklist (227, 232, 283) along with questions adapted from the COSMIN checklist for methodological quality of studies on the measurement properties of patient reported outcome measures (PROMs) (209).

ROBIS, a domain-based tool developed to assess risk of bias in systematic reviews (229) in accordance with the Cochrane group's current approach to the assessment of study quality (210), was not published until the analysis was already underway. After learning about the new tool, the analysis approach in study III was changed to include both ROBIS and AMSTAR. As the ROBIS tool was newly developed, its validity was not very well established; therefore the more well-established AMSTAR checklist was retained. AMSTAR and ROBIS assess the partly overlapping concepts methodological quality and risk of bias, respectively (227, 229). The overlap was confirmed by the similar pattern in results; reviews that demonstrated low risk of bias measured with ROBIS tended to exhibit high methodological quality measured with AMSTAR, and vice versa.

ROBIS is a generic tool for evaluation of systematic reviews regardless of their aim and scope. As such a limitation was the lack of specific guidance related to reviews on measurement properties. This was sought addressed by adapting ROBIS to this type of reviews using the COSMIN initiative's approach (205) and the COSMIN checklist for selection of health status measurement instruments (209). The COSMIN group recently published a guideline (284) for how to conduct systematic reviews of PROMs. This guideline builds on the guidance that was used in study III and would probably have been an even better alternative if it had been available when the study was done.

6.5.8 Generalizability/transferability

The generalizability/transferability varied between studies. It was limited for the results of study I as this study was performed on one translation of one scale at one hospital. However, the generalizability of the methodological approach described in this study has been demonstrated (278). The generalizability of study II was limited by the study design and by the sample and setting. Findings are probably generalizable to Norwegian children 12-36 months in the first hours after minor outpatient surgery. How the scale will work outside this age group and setting is undetermined. Study III used a systematic and comprehensive method to ensure inclusion of all available studies. Study selection, data extraction and

analysis had high internal validity. Consequently the results of this systematic review are considered generalizable. The similarities in findings across different countries, units and nurses, as well as the use of investigator triangulation support the transferability of the findings in study IV. However, transferability needs to be determined by the reader not the researcher (236).

7 CLINICAL IMPLICATIONS

In caring for preverbal children pediatric nurses are frequently faced with the unspoken question “Do you see my pain?” and the challenge of assessing pain in this age group. Findings in this thesis do not support the use of structured pain scales as a substitute for a complex clinical judgment by skilled observers. Neither do they support abandoning the use of structured pain scales. Instead findings point towards integration of scale scores into a comprehensive assessment of pain as a better approach. However, it is important to recognize that both clinical judgment and all observational pain scales have limitations. Recommendations in published reviews regarding pain scales for use in clinical practice have low evidence value and should be interpreted with caution. Until new and better recommendations have been published the following aspects should be taken into consideration if pain scales are selected for use in clinical practice:

- Developed and/or validated for the patient group and setting in which it will be used
- Available in the target language with evidence of a systematic translation and cultural adaptation process
- Validation profile; number and type of validation studies. Look separately for evidence of content validity, construct validity, responsiveness and inter-rater reliability. Use the revised COSMIN checklist to determine risk of bias in the primary studies and the COSMIN guideline to systematically draw conclusions regarding measurement properties
- How easy the scale is to use (feasibility) and the probability for use if it is made available in your unit (clinical utility)
- The resources needed to educate and train clinicians in proper scale use, to implement the scale and to support continued use

8 CONCLUSIONS

The Norwegian version of the COMFORT behavioral scale can be used for assessment of sedation, while its properties regarding assessment of pain remain undetermined. Problems identified with the content and construct validity of the COMFORT behavioral scale point towards validity issues with observational pain scales in general. The validity of observational pain scales could not be determined from published reviews as their recommendations had low evidence value and should be interpreted with caution.

Although systematic use of structured pain scales is strongly advocated as a prerequisite for pain relief in suffering children, nurses did not consider these scales specific to pain. A majority of nurses did not use structured pain scales despite expressing a positive view towards these scales and having access to them. Instead they preferred to use a comprehensive clinical judgment that was individually tailored to the child and the situation to see preverbal children's pain. Nurses who used structured scales integrated scale scores as an aspect in this judgment.

A preverbal child's pain will probably be better seen, evaluated and managed if nurses apply a systematic and comprehensive assessment approach that integrates clinical judgement and structured pain scales. This kind of approach may facilitate the implementation and use of such scales into clinical practice, as it acknowledges the nurse's clinical competence and knowledge of the individual child by taking into account factors currently not included in these scales, such as cause of pain, elimination of other causes of distress, and response to pain-relieving interventions.

9 FUTURE PERSPECTIVES

Evaluation of the evidence base for individual observational pain scales was outside the scope of this thesis. A clearer picture of strengths and deficiencies in existing scales is needed. Instead of reproducing variations of the same approach by publishing new scales with a similar approach to pain measurement, more and targeted research is needed to improve the measurement properties of existing pain scales for use in preverbal children, for example through critical revisions of scale items as suggested by Chang and colleagues (36), cognitive interviewing to elicit scale users' understanding of scale content and use as described in study I, and targeted measurement studies to evaluate specific measurement properties where sufficient evidence is lacking. To identify these gaps, better systematic reviews of the measurement properties of these scales are needed. With the publication of the COSMIN methodology for this type of systematic review (284) a necessary tool for this type of evaluation is finally available. For the Norwegian translation of the COMFORT behavioral scale specifically, new validation studies in better suited samples and settings are needed to evaluate whether the responsiveness of the scale as a measure of pain/distress is supported and whether participants' understanding of the measure changes if the measure is applied in a clinical setting.

The comprehensive clinical judgment and the specific strategies used by nurses in the assessment of pain should be evaluated in larger quantitative studies to determine their generalizability. How nurses integrate a structured scale into a comprehensive pain assessment, and involve parents in pain assessment also needs to be further explored. Future work examining the validity, feasibility, acceptability and effectiveness of a comprehensive pain assessment process that embeds a structured and valid pain scale and encompasses additional child, parent, situational and contextual factors is warranted, including evaluation of its effect on pain assessment practices, pain treatment and outcomes for the child.

10 ACKNOWLEDGEMENTS

Research often feels like lonely work and more so when the majority of this project has been carried out outside the university and in a hospital setting where few others have been involved in research on a daily basis. I am extremely grateful for those who have supported me and those who have connected me to the larger research community throughout this journey.

The project and this thesis would not have been possible without my main supervisor Leena Jylli. Thank you for taking a chance on me more than 15 years ago and for supporting me through projects and educations ever since. Our ongoing dialogue has been a major influence on my understanding of pediatric pain assessment. Thank you for being calm and patient when I am not.

I am also grateful for my three co-supervisors Ann Langius-Eklöf, Tomm Bernklev and Britt Nakstad. Thank you, Ann, for bringing your extensive experience in nursing research and instrument development into the project. Tomm: thank you for providing hands-on support during the first years. I learned a lot and you made the journey a lot less lonely. I want to thank you Britt for your enthusiasm and willingness to engage in this project when I wanted to include a physician perspective.

Two of the included studies have had co-authors outside the research group. Professor emeritus Bruce Ambuel, developer of the COMFORT scale, sadly died earlier this year. I remember him for his considerable engagement in study I. Study IV would have been very different without the support from Marsha Campbell-Yeo and Agneta Anderzen-Carlsson. Thank you Marsha for inviting me to visit you research group – the MOM-LINC Lab - at the Centre for Pediatric Pain Research, for making the Canadian data collection possible and for providing your insight and perspective to the analysis. Agneta: Thank you for supervising me on this study, for inspiring me and for generously sharing your considerable knowledge and experience in qualitative research.

Maria Gradin has been my mentor throughout this project. Thank you for being the best possible mentor, colleague and friend. We have discussed research and life, and shared both laughter and tears.

There would be no research without participants. I am very grateful to the children, parents and clinicians who have participated and made these studies possible.

I am thankful for the high-quality masters' and doctoral education Karolinska Institutet and the Department for Neurobiology, Care Sciences and Society have provided. I have benefitted greatly from knowledgeable and enthusiastic professors and lecturers and from discussion and collaboration with fellow students.

Thank you to the Norwegian Nurses Association and Telemark Hospital, both the Department of Research and the Department of Child and Adolescent Health Services, for

funding this project. Over the years I have also benefitted greatly from participation in the weekly research lunches and bi-annually research seminars arranged by the Research Department. The value of having regular meeting points with other researchers when you mostly work alone cannot be overestimated. Special thanks to researchers and fellow PhD students for knowledge-sharing and lively discussions during these lunches and seminars.

I specifically want to mention and give thanks to previous Head of Research Jan Fredrik Pape who got me started on research, now retired hospital librarian Mirjam Håndlykken, who took me under her wing at the old hospital library, Martin Veel Svendsen for statistical advice, and Janet Holmén for language editing. Kari Fornes, thank you for helping me get through difficult times. Professor Emeritus Geir Hoff, thank you for being my inspiration and living proof that it is possible to be a world-class researcher even at a small county hospital.

The Canadian-based Pain in Child Health (PICH) research training initiative has been a central and invaluable part of my PhD training. Through participation in monthly webinars and yearly conferences and workshops I have had the pleasure of learning from and interacting with fellow students from all over the world as well as world-class scientists in the pediatric pain community.

Professor Mats Eriksson is responsible for introducing me to exciting research opportunities and collaborations outside my own PhD project. Thank you! I am grateful for our collaboration in the Nordic Pain in EARly Life (PEARL) research network and I want to extend my thanks to all my other Nordic research colleagues in the network as well.

This project would not have been even remotely possible without the solid support from my leader at Telemark Hospital, Head of Department Gunnar Gausel. Thank you for your unwavering support and encouragement. I also want to thank current and previous colleagues in the Department for Child and Adolescent Health Services and in the neonatal intensive care unit.

Last, but not least, friends and family for helping me get outside my own head occasionally. Kristina for long conversations over endless cups of tea and Marit for great and inspiring exercise and engaging conversations. My parents, Åse and Jørg, for believing in me and for your love and support. My sister Ingrid, her husband Hans Tore and my nephew Hans Emil and niece Vilde Andrea for always being there. My late brother Anders: You live on in our hearts and I consider my work towards relieving children’s pain a tribute to your life.

My beautiful and courageous daughter Ragnhild. You bring joy and love to my life and to the world. I am grateful every day for being your mom.

My husband Runar. My anchor and safe haven. I love you!

11 SAMMENDRAG (NORWEGIAN SUMMARY)

Bakgrunn og mål: Smerter hos preverbale barn innlagt på sykehus blir undervurdert og underbehandlet. The Social Communication Model of Pain beskriver smerter som både en personlig opplevelse og en sosial prosess. Smerter påvirkes ikke bare av barnet som har vondt, men av den som observerer smertene og av omgivelsene. Sykepleiere tenderer mot å underestimere smerter hos andre. Bruk av strukturerte smertevurderingsskalaer blir anbefalt, men disse skalaene har likevel vært vanskelige å få innført i klinisk praksis. En bedre forståelse av disse skalaene og av sykepleieres oppfatninger om og praktisk bruk av dem er nødvendig for å kunne forbedre smertevurdering i klinisk praksis og redusere unødvendig smerter for preverbale barn innlagt i sykehus. Det overordnede målet for denne avhandlingen var derfor å bidra til kunnskap om hvordan unødvendige smerter og lidelse hos preverbale barn innlagt i sykehus gjennom en utforsking av aspekter som påvirker sykepleieres vurdering av smerter i klinisk praksis.

Metode: Avhandlingen består av fire studier gjennomført med både kvalitativ og kvantitativ metode. I studie I ble COMFORT atferdsskala oversatt til norsk ved hjelp av en systematisk fram- og tilbakeoversettelse og deretter kulturelt tilpasset i 12 kognitive intervjuer med klinikere og framtidige brukere av skalaen. Studie II testet den oversatte skalaens følsomhet for endring og samsvar i scorer mellom ulike brukere basert på gjentatte målinger fra 45 barn før og etter mindre dagkirurgiske inngrep. Studie III var en systematisk oversikt som gransket det vitenskapelige underlaget for anbefalinger om smertevurderingsskalaer gitt i 14 systematiske oversiktsartikler. Studie IV utforsket 22 sykepleieres smertevurderingspraksis i norske og kanadiske avdelinger gjennom delvis strukturerte intervjuer basert på sykepleiernes eksempler fra egen praksis.

Resultater: Kognitive intervjuer identifiserte flere problemer både med den originale og den oversatte versjonen av COMFORT atferdsskala. Skalaens følsomhet for endring er uavklart for smerte og stress. De anbefalingene som blir gitt i systematiske oversiktsartikler basert på skalaenes måleegenskaper hadde et svakt vitenskapelig grunnlag og var ikke til å stole på. Strukturerte vurderingsskalaer ble sjelden brukt og hverken smerteskårer eller atferdstegn assosiert med smerter ble ansett som spesifikke for smerter. I stedet foretrakk sykepleiere smertevurdering basert på klinisk skjønn og individuelt tilpasset til barnet og situasjonen. Smertevurderingen var en kombinasjon av erfaringsbasert kunnskap og kunnskap om det individuelle barnet med spesifikke strategier for å tolke observasjoner og informasjon fra barnet. Disse strategiene inkluderte å identifisere en sannsynlig årsak til smerter, eliminere andre årsaker til observert atferd, bruk av smerteskårer, vurdering av atferdsendringer og/eller effekt av intervensjoner på atferd, samt å bruke en individuell og situasjonstilpasset tilnærming.

Konklusjoner: Preverbale barns smerter vil sannsynligvis bli bedre sett, evaluert og håndtert dersom sykepleiere integrerer klinisk skjønn og strukturerte smertevurderingsskalaer i en systematisk og helhetlig vurdering av smerter.

12 REFERENCES

1. Williams AC. Facial expression of pain: an evolutionary account. Behav Brain Sci. 2002;25(4):439-55; discussion 55-88.

2. Butler DS, Moseley GL. Explain Pain. 2nd ed. Adelaide: Noigroup Publications; 2013.

3. Pillai Riddell RR, Stevens BJ, Cohen LL, Flora DB, Greenberg S. Predicting maternal and behavioral measures of infant pain: the relative contribution of maternal factors. Pain. 2007;133(1-3):138-49.

4. Fitzgerald M. Central nociceptive pathways and descending modulation. In: McGrath PJ, Stevens BJ, Walker SM, Zempsky WT, editors. Oxford textbook of paediatric pain. Oxford: Oxford University Press; 2014. p. 74-81.

5. Herr K, Coyne PJ, McCaffery M, Manworren R, Merkel S. Pain assessment in the patient unable to self-report: position statement with clinical practice recommendations. Pain Manag Nurs. 2011;12(4):230-50.

6. Classification of Chronic Pain [Internet]. 2nd ed. (revised). Seattle: IASP Press.; 1994. Part III: Pain terms, a current list with definitions and notes on usage. [Updated 2012; cited 2018 Mar 10]. Available from: https://www.iasp-pain.org/files/Content/ContentFolders/Publications2/ClassificationofChronicPain/Part_III-PainTerms.pdf

7. Williams AC, Craig KD. Updating the definition of pain. Pain. 2016;157(11):2420-3.

8. Gracely RH. Pain in humans, sensory-discriminative aspects. In: Gebhart GF, Schmidt RF, editors. Encyclopedia of pain. 2nd ed. Berlin Heidelberg: Springer-Verlag; 2013. p. 2646-50.

9. Affective component (aspect, dimension) of pain. In: Gebhart GF, Schmidt RF, editors. Encyclopedia of pain. 2nd ed. Berlin Heidelberg: Springer-Verlag; 2013.

10. Turk DC. Cognitive-behavioral perspective of pain. In: Gebhart GF, Schmidt RF, editors. Encyclopedia of pain. 2nd ed. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013. p. 697-700.

11. Goubert L, Vervoort T, Craig KD. Empathy and pain. In: Gebhart GF, Schmidt RF, editors. Encyclopedia of pain. 2nd ed. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013. p. 1128-34.

12. Craig KD. The Social Communication Model of Pain. Can Psychol. 2009;50(1):22-32.

13. Prkachin KM, Craig KD. Expressing pain: The communication and interpretation of facial pain signals. J Nonverbal Behav. 1995;19(4):191-205.

14. Hadjistavropoulos T, Craig KD. A theoretical framework for understanding self-report and observational measures of pain: a communications model. Behav Res Ther. 2002;40(5):551-70.

15. Schiavenato M, Craig KD. Pain assessment as a social transaction: beyond the "gold standard". Clin J Pain. 2010;26(8):667-76.

16. Craig KD. Social communication model of pain. Pain. 2015;156(7):1198-9.

17. Craig KD, Korol CT, Pillai RR. Challenges of judging pain in vulnerable infants. Clin Perinatol. 2002;29(3):445-57.

18. Melzack R, Wall PD. Pain mechanisms: a new theory. Science. 1965;150(3699):971-9.

19. Melzack R. From the gate to the neuromatrix. Pain. 1999;Suppl 6:S121-6.

20. Craig KD, Lilley CM, Gilbert CA. Social barriers to optimal pain management in infants and children. Clin J Pain. 1996;12(3):232-42.

21. Grossman LS. Normal child development. In: Elzouki AY, Harfi HA, Nazer HM, Stapleton FB, Oh W, Whitley RJ, editors. Textbook of clinical pediatrics. Berlin, Heidelberg: Springer; 2012. p. 571-82.

22. Gilmore Karen MD, Meersand Pamela PD. Normal child and adolescent development. In: Hales RE, Yudofsky SC, Weiss Roberts L, editors. The American psychiatric publishing textbook of psychiatry. 6th ed. Arlington, VA: American Psychiatric Publishing; 2014. p. 139-74.

23. Hetherington EM, Parke RD, Gauvain M, Locke VO. Child psychology. A contemporary vievpoint. 6th ed. New York: McGraw-Hill Education; 2006.

24. Brown EA, De Young A, Kimble R, Kenardy J. Review of a parent's influence on pediatric procedural distress and recovery. Clin Child Fam Psychol Rev. 2018. Epub ahead of print

25. Racine NM, Pillai Riddell RR, Khan M, Calic M, Taddio A, Tablon P. Systematic review: predisposing, precipitating, perpetuating, and present factors predicting anticipatory distress to painful medical procedures in children. J Pediatr Psychol. 2016;41(2):159-81.

26. Stevens BJ, Abbott LK, Yamada J, Harrison D, Stinson J, Taddio A, et al. Epidemiology and management of painful procedures in children in Canadian hospitals. CMAJ. 2011;183(7):E403-10.

27. Stevens BJ, Harrison D, Rashotte J, Yamada J, Abbott LK, Coburn G, et al. Pain assessment and intensity in hospitalized children in Canada. J Pain. 2012;13(9):857-65.

28. Friedrichsdorf SJ, Postier A, Eull D, Weidner C, Foster L, Gilbert M, et al. Pain outcomes in a US children's hospital: a prospective cross-sectional survey. Hosp Pediatr. 2015;5(1):18-26.

29. Walther-Larsen S, Pedersen MT, Friis SM, Aagaard GB, Romsing J, Jeppesen EM, et al. Pain prevalence in hospitalized children: a prospective cross-sectional survey in four Danish university hospitals. Acta Anaesthesiol Scand. 2017;61(3):328-37.

30. Shomaker K, Dutton S, Mark M. Pain prevalence and treatment patterns in a US children's hospital. Hosp Pediatr. 2015;5(7):363-70.

31. Taylor EM, Boyer K, Campbell FA. Pain in hospitalized children: a prospective cross-sectional survey of pain prevalence, intensity, assessment and management in a Canadian pediatric teaching hospital. Pain Res Manag. 2008;13(1):25-32.

32. Birnie KA, Chambers CT, Fernandez CV, Forgeron PA, Latimer MA, McGrath PJ, et al. Hospitalized children continue to report undertreated and preventable pain. Pain Res Manag. 2014;19(4):198-204.

33. Craig KD, McMahon RJ, Morison JD, Zaskow C. Developmental changes in infant pain expression during immunization injections. Soc Sci Med. 1984;19(12):1331-7.

34. Goodenough B, Thomas W, Champion GD, Perrott D, Taplin JE, von Baeyer CL, et al. Unravelling age effects and sex differences in needle pain: ratings of sensory intensity and unpleasantness of venipuncture pain by children and their parents. Pain. 1999;80(1-2):179-90.

35. Wennstrom B, Bergh I. Bodily and verbal expressions of postoperative symptoms in 3- to 6-year-old boys. J Pediatr Nurs. 2008;23(1):65-76.

36. Chang J, Versloot J, Fashler SR, McCrystal KN, Craig KD. Pain assessment in children: validity of facial expression items in observational pain scales. Clin J Pain. 2015;31(3):189-97.

37. Stevens B, McGrath P, Gibbins S, Beyene J, Breau L, Camfield C, et al. Determining behavioural and physiological responses to pain in infants at risk for neurological impairment. Pain. 2007;127(1-2):94-102.

38. Ekman P, Oster H. Facial Expressions of Emotion. Annu Rev Psychol. 1979;30:527-54.

39. Craig KD, Prkachin KM, Grunau RE. The facial expression of pain. In: Turk DC, Melzack R, editors. Handbook of pain assessment. 3rd ed. New York: The Guilford Press; 2011. p. 117-33.

40. Grunau RV, Craig KD. Pain expression in neonates: facial action and cry. Pain. 1987;28(3):395-410.

41. Grunau RV, Johnston CC, Craig KD. Neonatal facial and cry responses to invasive and non-invasive procedures. Pain. 1990;42(3):295-305.

42. Johnston CC, Stevens B, Craig KD, Grunau RV. Developmental changes in pain expression in premature, full-term, two- and four-month-old infants. Pain. 1993;52(2):201-8.

43. Craig KD. The facial display of pain. In: Finley GA, McGrath PJ, editors. Measurement of pain in infants and children. Seattle: IASP Press; 1998. p. 103-22.

44. Fuller BF. Acoustic discrimination of three types of infant cries. Nurs Res. 1991;40(3):156-60.

45. Hatfield LA, Ely EA. Measurement of acute pain in infants: a review of behavioral and physiological variables. Biol Res Nurs. 2015;17(1):100-11.

46. Stanford EA, Chambers CT, Craig KD. A normative analysis of the development of pain-related vocabulary in children. Pain. 2005;114(1-2):278-84.

47. American Academy of Pediatrics. The assessment and management of acute pain in infants, children, and adolescents. Pediatrics. 2001;108(3):793-7.

48. Von Baeyer CL. Self-report: the primary source in assessment after infancy. In: McGrath PJ, Stevens BJ, Walker SM, Zempsky WT, editors. Oxford textbook of paediatric pain. Oxford: Oxford University Press; 2014. p. 370-8.

49. Chan JY, von Baeyer CL. Cognitive developmental influences on the ability of preschool-aged children to self-report their pain intensity. Pain. 2016;157(5):997-1001.

50. Amtmann D, Cook KF, Jensen MP, Chen WH, Choi S, Revicki D, et al. Development of a PROMIS item bank to measure pain interference. Pain. 2010;150(1):173-82.

51. Taddio A, Shah V, Gilbert-MacLeod C, Katz J. Conditioning and hyperalgesia in newborns exposed to repeated heel lances. JAMA. 2002;288(7):857-61.

52. Racine NM, Pillai Riddell RR, Flora DB, Taddio A, Garfield H, Greenberg S. Predicting preschool pain-related anticipatory distress: the relative contribution of longitudinal and concurrent factors. Pain. 2016;157(9):1918-32.

53. Hermann C, Hohmeister J, Demirakca S, Zohsel K, Flor H. Long-term alteration of pain sensitivity in school-aged children with early pain experiences. Pain. 2006;125(3):278-85.

54. Taddio A, Katz J. The effects of early pain experience in neonates on pain responses in infancy and childhood. Paediatr Drugs. 2005;7(4):245-57.

55. Fitzgerald M, Walker SM. Infant pain management: a developmental neurobiological approach. Nat Clin Pract Neurol. 2009;5(1):35-50.

56. Beggs S. Long-Term Consequences of Neonatal Injury. Can J Psychiatry. 2015;60(4):176-80.

57. Weisman SJ, Bernstein B, Schechter NL. Consequences of inadequate analgesia during painful procedures in children. Arch Pediatr Adolesc Med. 1998;152(2):147-9.

58. Crock C, Olsson C, Phillips R, Chalkiadis G, Sawyer S, Ashley D, et al. General anaesthesia or conscious sedation for painful procedures in childhood cancer: the family's perspective. Arch Dis Child. 2003;88(3):253-7.

59. Rabbitts JA, Fisher E, Rosenbloom BN, Palermo TM. Prevalence and Predictors of Chronic Postsurgical Pain in Children: A Systematic Review and Meta-Analysis. J Pain. 2017;18(6):605-14.

60. Bonavita V, De Simone R. Pain as an evolutionary necessity. Neurol Sci. 2011;32(Suppl 1):S61-6.

61. McCrystal KN, Craig KD, Versloot J, Fashler SR, Jones DN. Perceiving pain in others: validation of a dual processing model. Pain. 2011;152(5):1083-9.

62. Poole GD, Craig KD. Judgments of genuine, suppressed, and faked facial expressions of pain. J Pers Soc Psychol. 1992;63(5):797-805.

63. Waxman J, Martin J, Pillai Riddell R. Understanding infant pain responding within a relational context. In: Buonocore G, Bellieni CV, editors. Neonatal pain suffering, pain, and risk of brain damage in the fetus and newborn. 2nd ed. Cham: Springer International Publishing; 2017. p. 89-104.

64. Pillai Riddell R, Racine NM, Craig KD, Campbell L. Psychological theories and biopsychosocial models in paediatric pain. In: McGrath PJ, Stevens BJ, Walker SM, Zempsky WT, editors. Oxford textbook of paediatric pain. Oxford: Oxford University Press; 2014. p. 85-94.

65. Uman LS, Chambers CT, McGrath PJ, Kisely S. A systematic review of randomized controlled trials examining psychological interventions for needle-related procedural pain and distress in children and adolescents: an abbreviated Cochrane review. J Pediatr Psychol. 2008;33(8):842-54.

66. Atkinson NH, Gennis H, Racine NM, Pillai Riddell R. Caregiver emotional availability, caregiver soothing behaviors, and infant pain during immunization. J Pediatr Psychol. 2015;40(10):1105-14.

67. McMurtry CM, Chambers CT, McGrath PJ, Asp E. When "don't worry" communicates fear: children's perceptions of parental reassurance and distraction during a painful medical procedure. Pain. 2010;150(1):52-8.

68. McMurtry CM, McGrath PJ, Chambers CT. Reassurance can hurt: parental behavior and painful medical procedures. J Pediatr. 2006;148(4):560-1.

69. Craig KD, Pillai Riddell R. Social influences, culture and ethnicity. In: Finley GA, McGrath PJ, editors. Pediatric pain: biological and social context. Seattle: IASP Press; 2003. p. 159-82.

70. O'Neill MC, Pillai Riddell R, Garfield H, Greenberg S. Does caregiver behavior mediate the relationship between cultural individualism and infant pain at 12 months of age? J Pain. 2016;17(12):1273-80.

71. The ICN code of ethics for nurses. Geneva, Switzerland: International Council of Nurses; Revised 2012.

72. Campbell-Yeo M, Latimer M, Johnston C. The empathetic response in nurses who treat pain: concept analysis. J Adv Nurs. 2008;61(6):711-9.

73. Cox F, Cannons K, Lewis S. RCN Pain knowledge and skills framework for the nursing team. London: Royal College of Nursing; 2015.

74. Franck LS, Bruce E. Putting pain assessment into practice: why is it so painful? Pain Res Manag. 2009;14(1):13-20.

75. Gaffney A, McGrath PJ, Dick B. Measuring pain in children: developmental and instrument issues. In: Schechter NL, Berde CB, Yaster M, editors. Pain in infants, children and adolescents. 2nd ed. Philadelphia: Lippincott Williams & Wilkins; 2003. p. 128-41.

76. Hadjistavropoulos T, Craig KD, Duck S, Cano A, Goubert L, Jackson PL, et al. A biopsychosocial formulation of pain communication. Psychol Bull. 2011;137(6):910-39.

77. Boerner KE, Chambers CT, Craig KD, Pillai Riddell RR, Parker JA. Caregiver accuracy in detecting deception in facial expressions of pain in children. Pain. 2013;154(4):525-33.

78. Prkachin KM, Solomon PE, Ross J. Underestimation of pain by health-care providers: towards a model of the process of inferring pain in others. Can J Nurs Res. 2007;39(2):88-106.

79. Prkachin KM, Rocha EM. High levels of vicarious exposure bias pain judgments. J Pain. 2010;11(9):904-9.

80. Chambers CT, Reid GJ, Craig KD, McGrath PJ, Finley GA. Agreement between child and parent reports of pain. Clin J Pain. 1998;14(4):336-42.

81. Kamper SJ, Dissing KB, Hestbaek L. Whose pain is it anyway? Comparability of pain reports from children and their parents. Chiropr Man Therap. 2016;24:24.

82. Zhou H, Roberts P, Horgan L. Association between self-report pain ratings of child and parent, child and nurse and parent and nurse dyads: meta-analysis. J Adv Nurs. 2008;63(4):334-42.

83. Kelly AM, Powell CV, Williams A. Parent visual analogue scale ratings of children's pain do not reliably reflect pain reported by child. Pediatr Emerg Care. 2002;18(3):159-62.

84. Goubert L, Vervoort T, Cano A, Crombez G. Catastrophizing about their children's pain is related to higher parent-child congruency in pain ratings: an experimental investigation. Eur J Pain. 2009;13(2):196-201.

85. St-Laurent-Gagnon T, Bernard-Bonnin AC, Villeneuve E. Pain evaluation in preschool children and by their parents. Acta Paediatr. 1999;88(4):422-7.

86. Chambers CT, Giesbrecht K, Craig KD, Bennett SM, Huntsman E. A comparison of faces scales for the measurement of pediatric pain: children's and parents' ratings. Pain. 1999;83(1):25-35.

87. Robinson ME, Wise EA. Gender bias in the observation of experimental pain. Pain. 2003;104(1-2):259-64.

88. Ruben MA, Hall JA. "I know your pain": proximal and distal predictors of pain detection accuracy. Pers Soc Psychol Bull. 2013;39(10):1346-58.

89. Coll MP, Gregoire M, Latimer M, Eugene F, Jackson PL. Perception of pain in others: implication for caregivers. Pain Manag. 2011;1(3):257-65.

90. Deyo KS, Prkachin KM, Mercer SR. Development of sensitivity to facial expression of pain. Pain. 2004;107(1-2):16-21.

91. Gregoire M, Bruneau-Bherer R, Morasse K, Eugene F, Jackson PL. The perception and estimation of others' pain according to children. Pain Res Manag. 2016;2016:9097542.

92. Lautenbacher S, Hofer W, Kunz M. Age Differences in decoding pain from the facial expression of healthy individuals and patients with dementia. Pain Med. 2016;17(4):685-91.

93. Goubert L, Craig KD, Vervoort T, Morley S, Sullivan MJ, de CWAC, et al. Facing others in pain: the effects of empathy. Pain. 2005;118(3):285-8.

94. Sullivan MJ, Martel MO, Tripp DA, Savard A, Crombez G. Catastrophic thinking and heightened perception of pain in others. Pain. 2006;123(1-2):37-44.

95. Rash JA, Prkachin KM, Campbell TS. Observer trait anxiety is associated with response bias to patient facial pain expression independent of pain catastrophizing. Pain Res Manag. 2015;20(1):39-45.

96. Green AD, Tripp DA, Sullivan MJ, Davidson M. The relationship between empathy and estimates of observed pain. Pain Med. 2009;10(2):381-92.

97. Taddio A, Appleton M, Bortolussi R, Chambers C, Dubey V, Halperin S, et al. Reducing the pain of childhood vaccination: an evidence-based clinical practice guideline (summary). Cmaj. 2010;182(18):1989-95.

98. Shah V, Taddio A, McMurtry CM, Halperin SA, Noel M, Pillai Riddell R, et al. Pharmacological and combined interventions to reduce vaccine injection pain in children and adults: systematic review and meta-analysis. Clin J Pain. 2015;31(10 Suppl):S38-63.

99. Harrison D, Bueno M, Reszel J. Prevention and management of pain and stress in the neonate. Midwifery Digest. 2015;25(2):253.

100. Taddio A, Chambers CT, Halperin SA, Ipp M, Lockett D, Rieder MJ, et al. Inadequate pain management during routine childhood immunizations: the nerve of it. Clin Ther. 2009;31 Suppl 2:S152-67.

101. Agoston AM, Sieberg CB. Nonpharmacologic treatment of pain. Semin Pediatr Neurol. 2016;23(3):220-3.

102. Kuttner L. A Child in Pain: What health professionals can do to help. Carmarthen, Wales: Crown House Publishing Ltd; 2010.

103. LeBlanc CK, Chambers CT. Child life interventions in pediatric pain. In: McGrath PJ, Stevens BJ, Walker SM, Zempsky WT, editors. Oxford textbook of paediatric pain. Oxford: Oxford University Press; 2014. p. 543-52.

104. Pillai Riddell R, Taddio A, McMurtry CM, Chambers C, Shah V, Noel M, et al. Psychological interventions for vaccine injections in young children 0 to 3 years: systematic review of randomized controlled trials and quasi-randomized controlled trials. Clin J Pain. 2015;31(10 Suppl):S64-71.

105. Sieberg CB, Manganella J. Family beliefs and interventions in pediatric pain management. Child Adolesc Psychiatr Clin N Am. 2015;24(3):631-45.

106. Pillai Riddell R, Taddio A, McMurtry CM, Shah V, Noel M, Chambers CT, et al. Process interventions for vaccine injections: systematic review of randomized controlled trials and quasi-randomized controlled trials. Clin J Pain. 2015;31(10 Suppl):S99-108.

107. Brasher C, Gafsous B, Dugue S, Thiollier A, Kinderf J, Nivoche Y, et al. Postoperative pain management in children and infants: an update. Paediatr Drugs. 2014;16(2):129-40.

108. Taddio A, Shah V, McMurtry CM, MacDonald NE, Ipp M, Riddell RP, et al. Procedural and physical interventions for vaccine injections: systematic review of randomized controlled trials and quasi-randomized controlled trials. Clin J Pain. 2015;31(10 Suppl):S20-37.

109. Strassels SA. Opioids in clinical practice. In: McGrath PJ, Stevens BJ, Walker SM, Zempsky WT, editors. Oxford textbook of paediatric pain. Oxford: Oxford University Press; 2014. p. 457-73.

110. Anderson BJ. The non-steroidal anti-inflammatory drugs and acetaminophen. In: McGrath PJ, Stevens BJ, Walker SM, Zempsky WT, editors. Oxford textbook of paediatric pain. Oxford: Oxford University Press; 2014. p. 436-48.

111. Zempsky WT. Topical anaesthetics and analgesics. In: McGrath PJ, Stevens BJ, Walker SM, Zempsky WT, editors. Oxford textbook of paediatric pain. Oxford: Oxford University Press; 2014. p. 486-94.

112. Hathway GJ. Developmental pharmacology of opioids. In: McGrath PJ, Stevens BJ, Walker SM, Zempsky WT, editors. Oxford textbook of paediatric pain. Oxford: Oxford University Press; 2014. p. 449-56.

113. Harrison D, Anseloni VCZ, Yamada J, Bueno M. Sucrose and sweet taste. In: McGrath PJ, Stevens BJ, Walker SM, Zempsky WT, editors. Oxford textbook of paediatric pain. Oxford: Oxford University Press; 2014. p. 508-16.

114. Harrison D, Stevens B, Bueno M, Yamada J, Adams-Webber T, Beyene J, et al. Efficacy of sweet solutions for analgesia in infants between 1 and 12 months of age: a systematic review. Arch Dis Child. 2010;95(6):406-13.

115. Craig KD, Versloot J, Goubert L, Vervoort T, Crombez G. Perceiving pain in others: automatic and controlled mechanisms. J Pain. 2010;11(2):101-8.

116. Kappesser J, Williams AC, Prkachin KM. Testing two accounts of pain underestimation. Pain. 2006;124(1-2):109-16.

117. Lamm C, Batson CD, Decety J. The neural substrate of human empathy: effects of perspective-taking and cognitive appraisal. J Cogn Neurosci. 2007;19(1):42-58.

118. Gimbler-Berglund I, Ljusegren G, Enskar K. Factors influencing pain management in children. Paediatr Nurs. 2008;20(10):21-4.

119. LaFond CM, Van Hulle Vincent C, Corte C, Hershberger PE, Johnson A, Park CG, et al. PICU nurses' pain assessments and intervention choices for virtual human and written vignettes. J Pediatr Nurs. 2015;30(4):580-90.

120. Noviello SR. A grounded theory study of pain management behaviors of nurses who care for preverbal children [dissertation on the internet]. Augusta: Medical College of Georgia; 2006 [cited 2017 Sep 15]. Available from: http://proxy.kib.ki.se/docview/304954685?accountid=11794

121. Smyth W, Toombes J, Usher K. Children's postoperative pro re nata (PRN) analgesia: nurses' administration practices. Contemp Nurse. 2011;37(2):160-72.

122. Woodgate R, Kristjanson LJ. A young child's pain: how parents and nurses 'take care'. Int J Nurs Stud. 1996;33(3):271-84.

123. Hamers JP, Abu-Saad HH, Halfens RJ, Schumacher JN. Factors influencing nurses' pain assessment and interventions in children. J Adv Nurs. 1994;20(5):853-60.

124. Ely B. Pediatric nurses' pain management practice: barriers to change. Pediatr Nurs. 2001;27(5):473-80.

125. Zisk-Rony RY, Lev J, Haviv H. Nurses' report of in-hospital pediatric pain assessment: examining challenges and perspectives. Pain Manag Nurs. 2015;16(2):112-20.

126. Twycross A. Children's nurses' post-operative pain management practices: an observational study. Int J Nurs Stud. 2007;44(6):869-81.

127. Twycross A, Finley GA, Latimer M. Pediatric nurses' postoperative pain management practices: an observational study. J Spec Pediatr Nurs. 2013;18(3):189-201.

128. Ljusegren G, Johansson I, Gimbler Berglund I, Enskar K. Nurses' experiences of caring for children in pain. Child Care Health Dev. 2012;38(4):464-70.

129. Lauzon Clabo LM. An ethnography of pain assessment and the role of social context on two postoperative units. J Adv Nurs. 2008;61(5):531-9.

130. Shoqirat N. 'We are nurses, they are doctors': barriers to nurses' roles in pain management following surgery in Jordan. Int J Nurs Pract. 2015;21(2):200-6.

131. Namnabati M, Abazari P, Talakoub S. Identification of perceived barriers of pain management in Iranian children: a qualitative study. Int J Nurs Pract. 2012;18(3):221-5.

132. Twycross A. Nurses' views about the barriers and facilitators to effective management of pediatric pain. Pain Manag Nurs. 2013;14(4):e164-72.

133. Voepel-Lewis T, Piscotty RJ, Jr., Annis A, Kalisch B. Empirical review supporting the application of the "pain assessment as a social transaction" model in pediatrics. J Pain Symptom Manage. 2012;44(3):446-57.

134. Ross RS, Bush JP, Crummette BD. Factors affecting nurses' decisions to administer PRN analgesic medication to children after surgery: an analog investigation. J Pediatr Psychol. 1991;16(2):151-67.

135. Joslyn BJ. How pediatric nurses assess infant pain [dissertation on the internet]. New York City: Columbia University Teachers College; 2005 [cited 2017 Sep 15]. Available from: http://proxy.kib.ki.se/docview/305013652?accountid=11794

136. Fuller BF, Conner DA. The effect of pain on infant behaviors. Clin Nurs Res. 1995;4(3):253-73.

137. Fuller BF, Conner DA. Distribution of cues across assessed levels of infant pain. Clin Nurs Res. 1996;5(2):167-84.

138. Fuller B, Thomson M, Conner DA, Scanlan J. Relationship of cues to assessed infant pain level. Clin Nurs Res. 1996;5(1):43-66.

139. Fuller BF. The process of infant pain assessment. Appl Nurs Res. 1998;11(2):62-8.

140. Fuller BF, Conner DA. The influence of length of pediatric nursing experience on key cues used to assess infant pain. J Pediatr Nurs. 1997;12(3):155-68.

141. Hudson-Barr DC, Duffey MA, Holditch-Davis D, Funk S, Frauman A. Pediatric nurses' use of behaviors to make medication administration decisions in infants recovering from surgery. Res Nurs Health. 1998;21(1):3-13.

142. Ramelet AS. Assessment of pain and agitation in critically ill infants. Aust Crit Care. 1999;12(3):92-6.

143. Hamers JP, Abu-Saad HH, van den Hout MA, Halfens RJ, Kester AD. The influence of children's vocal expressions, age, medical diagnosis and information obtained from parents on nurses' pain assessments and decisions regarding interventions. Pain. 1996;65(1):53-61.

144. Hamers JP, van den Hout MA, Halfens RJ, Abu-Saad HH, Heijltjes AE. Differences in pain assessment and decisions regarding the administration of analgesics between novices, intermediates and experts in pediatric nursing. Int J Nurs Stud. 1997;34(5):325-34.

145. Seymour E, Fuller BF, Pedersen-Gallegos L, Schwaninger JE. Modes of thought, feeling, and action in infant pain assessment by pediatric nurses. J Pediatr Nurs. 1997;12(1):32-50.

146. Stevens BJ, Pillai Riddell RR, Oberlander TE, Gibbins S. Assessment of pain in neonates and infants. In: Anand KJS, Stevens BJ, McGrath PJ, editors. Pain in neonates and infants. 3rd ed. Amsterdam: Elsevier; 2007. p. 67-90.

147. Finley GA, McGrath PJ. Introduction: The roles of measurement in pain management and research. In: Finley GA, McGrath PJ, editors. Measurement of pain in infants and children. Seattle: IASP Press; 1998. p. 1-4.

148. Franck LS. A national survey of the assessment and treatment of pain and agitation in the neonatal intensive care unit. J Obstet Gynecol Neonatal Nurs. 1987;16(6):387-93.

149. Pigeon HM, McGrath PJ, Lawrence J, MacMurray SB. Nurses' perceptions of pain in the neonatal intensive care unit. J Pain Symptom Manage. 1989;4(4):179-83.

150. Nagy S. A comparison of the effects of patients' pain on nurses working in burns and neonatal intensive care units. J Adv Nurs. 1998;27(2):335-40.

151. Drendel AL, Kelly BT, Ali S. Pain assessment for children: overcoming challenges and optimizing care. Pediatr Emerg Care. 2011;27(8):773-81.

152. Anand KJ, Craig KD. New perspectives on the definition of pain. Pain. 1996;67(1):3-6.

153. Van Dijk M, de Boer JB, Koot HM, Tibboel D, Passchier J, Duivenvoorden HJ. The reliability and validity of the COMFORT scale as a postoperative pain instrument in 0 to 3-year-old infants. Pain. 2000;84(2-3):367-77.

154. Ambuel B, Hamlett KW, Marx CM, Blumer JL. Assessing distress in pediatric intensive care environments: the COMFORT scale. J Pediatr Psychol. 1992;17(1):95-109.

155. Van Dijk M, Peters JW, van Deventer P, Tibboel D. The COMFORT behavior scale: a tool for assessing pain and sedation in infants. Am J Nurs. 2005;105(1):33-6.

156. Merkel SI, Voepel-Lewis T, Shayevitz JR, Malviya S. The FLACC: a behavioral scale for scoring postoperative pain in young children. Pediatr Nurs. 1997;23(3):293-7.

157. Marx CM, Smith PG, Lowrie LH, Hamlett KW, Ambuel B, Yamashita TS, et al. Optimal sedation of mechanically ventilated pediatric critical care patients. Crit Care Med. 1994;22(1):163-70.

158. Carnevale FA, Razack S. An item analysis of the COMFORT scale in a pediatric intensive care unit. Pediatr Crit Care Med. 2002;3(2):177-80.

159. Grap MJ, Pickler RH, Munro C. Observation of behavior in sedated, mechanically ventilated children. Pediatr Nurs. 2006;32(3):216-20.

160. Bear LA, Ward-Smith P. Interrater reliability of the COMFORT Scale. Pediatr Nurs. 2006;32(5):427-34.

161. Ista E, Van Dijk M, Tibboel D, de Hoog M. Assessment of sedation levels in pediatric intensive care patients can be improved by using the COMFORT "behavior" scale. Pediatr Crit Care Med. 2005;6(1):58-63.

162. Lamas A, Lopez-Herce J, Sancho L, Mencia S, Carrillo A, Santiago MJ, et al. Assessing sedation in critically ill children by bispectral index, auditory-evoked potentials and clinical scales. Intensive Care Med. 2008;34(11):2092-9.

163. Brunow de Carvalho W, Lucas da Silva PS, Paulo CS, Fonseca MM, Belli LA. Comparison between the Comfort and Hartwig sedation scales in pediatric patients undergoing mechanical lung ventilation. Sao Paulo Med J. 1999;117(5):192-6.

164. Blauer T, Gerstmann D. A simultaneous comparison of three neonatal pain scales during common NICU procedures. Clin J Pain. 1998;14(1):39-47.

165. Jin HS, Yum MS, Kim SL, Shin HY, Lee EH, Ha EJ, et al. The efficacy of the COMFORT scale in assessing optimal sedation in critically ill children requiring mechanical ventilation. J Korean Med Sci. 2007;22(4):693-7.

166. Allegaert K, Daniels H, Naulaers G, Tibboel D, Devlieger H. Pharmacodynamics of chloral hydrate in former preterm infants. Eur J Pediatr. 2005;164(7):403-7.

167. de Wildt SN, de Hoog M, Vinks AA, Joosten KF, van Dijk M, van den Anker JN. Pharmacodynamics of midazolam in pediatric intensive care patients. Ther Drug Monit. 2005;27(1):98-102.

168. Murray DM, Thorne GC, Rigby-Jones AE, Tonucci D, Grimes S, Tooley MA, et al. Electroencephalograph variables, drug concentrations and sedation scores in children emerging from propofol infusion anaesthesia. Paediatr Anaesth. 2004;14(2):143-51.

169. Jaarsma AS, Knoester H, van Rooyen F, Bos AP. Biphasic positive airway pressure ventilation (PeV+) in children. Crit Care. 2001;5(3):174-7.

170. Bouwmeester NJ, Anand KJ, Van Dijk M, Hop WC, Boomsma F, Tibboel D. Hormonal and metabolic stress responses after major surgery in children aged 0-3 years: a double-blind, randomized trial comparing the effects of continuous versus intermittent morphine. Br J Anaesth. 2001;87(3):390-9.

171. Campbell ML, Bizek KS, Thill M. Patient responses during rapid terminal weaning from mechanical ventilation: a prospective study. Crit Care Med. 1999;27(1):73-7.

172. el-Khatib MF, Chatburn RL, Potts DL, Blumer JL, Smith PG. Mechanical ventilators optimized for pediatric use decrease work of breathing and oxygen consumption during pressure-support ventilation. Crit Care Med. 1994;22(12):1942-8.

173. Lamas A, Lopez-Herce J, Sancho L, Mencia S, Carrillo A, Santiago MJ, et al. Assessment of the level of sedation in children after cardiac surgery. Ann Thorac Surg. 2009;88(1):144-50.

174. Froom SR, Malan CA, Mecklenburgh JS, Price M, Chawathe MS, Hall JE, et al. Bispectral Index asymmetry and COMFORT score in paediatric intensive care patients. Br J Anaesth. 2008;100(5):690-6.

175. Twite MD, Zuk J, Gralla J, Friesen RH. Correlation of the bispectral index monitor with the COMFORT scale in the pediatric intensive care unit. Pediatr Crit Care Med. 2005;6(6):648-53; quiz 54.

176. Grap MJ, Borchers CT, Munro CL, Elswick RK, Jr., Sessler CN. Actigraphy in the critically ill: correlation with activity, agitation, and sedation. Am J Crit Care. 2005;14(1):52-60.

177. Courtman SP, Wardurgh A, Petros AJ. Comparison of the bispectral index monitor with the Comfort score in assessing level of sedation of critically ill children. Intensive Care Med. 2003;29(12):2239-46.

178. Crain N, Slonim A, Pollack MM. Assessing sedation in the pediatric intensive care unit by using BIS and the COMFORT scale. Pediatr Crit Care Med. 2002;3(1):11-4.

179. Johansson M, Kokinsky E. The COMFORT behavioural scale and the modified FLACC scale in paediatric intensive care. Nurs Crit Care. 2009;14(3):122-30.

180. Peeters MY, Prins SA, Knibbe CA, DeJongh J, van Schaik RH, van Dijk M, et al. Propofol pharmacokinetics and pharmacodynamics for depth of sedation in nonventilated infants after major craniofacial surgery. Anesthesiology. 2006;104(3):466-74.

181. Prins SA, Peeters MY, Houmes RJ, van Dijk M, Knibbe CA, Danhof M, et al. Propofol 6% as sedative in children under 2 years of age following major craniofacial surgery. Br J Anaesth. 2005;94(5):630-5.

182. Prins SA, Van Dijk M, Van Leeuwen P, Searle S, Anderson BJ, Tibboel D, et al. Pharmacokinetics and analgesic effects of intravenous propacetamol vs rectal paracetamol in children after major craniofacial surgery. Paediatr Anaesth. 2008;18(7):582-92.

183. Bouwmeester NJ, Hop WC, van Dijk M, Anand KJ, van den Anker JN, Tibboel D. Postoperative pain in the neonate: age-related differences in morphine requirements and metabolism. Intensive Care Med. 2003;29(11):2009-15.

184. Bouwmeester NJ, van den Anker JN, Hop WC, Anand KJ, Tibboel D. Age- and therapy-related effects on morphine requirements and plasma concentrations of morphine and its metabolites in postoperative infants. Br J Anaesth. 2003;90(5):642-52.

185. Schasfoort FC, Formanoy MAG, Bussmann JBJ, Peters JWB, Tibboel D, Stam HJ. Objective and continuous measurement of peripheral motor indicators of pain in hospitalized infants: a feasibility study. Pain. 2008;137(2):323-31.

186. De Jonghe B, Cook D, Griffith L, Appere-de-Vecchi C, Guyatt G, Theron V, et al. Adaptation to the Intensive Care Environment (ATICE): development and validation of a new sedation assessment instrument. Crit Care Med. 2003;31(9):2344-54.

187. Peters JW, Koot HM, Grunau RE, de Boer J, van Druenen MJ, Tibboel D, et al. Neonatal Facial Coding System for assessing postoperative pain in infants: item reduction is valid and feasible. Clin J Pain. 2003;19(6):353-63.

188. Hartrick CT, Kovan JP. Pain assessment following general anesthesia using the Toddler Preschooler Postoperative Pain Scale: a comparative study. J Clin Anesth. 2002;14(6):411-5.

189. Lago P, Guadagni A, Merazzi D, Ancora G, Bellieni CV, Cavazza A. Pain management in the neonatal intensive care unit: a national survey in Italy. Paediatr Anaesth. 2005;15(11):925-31.

190. Gibbins S, Maddalena P, Moulsdale W, Garrard F, jan Mohamed T, Nichols A, et al. Pain assessment and pharmacologic management for infants with NEC: a retrospective chart audit. Neonatal Netw. 2006;25(5):339-45.

191. Akuma AO, Jordan S. Pain management in neonates: a survey of nurses and doctors. J Adv Nurs. 2012;68(6):1288-301.

192. de Freitas GR, de Castro CG, Jr., Castro SM, Heineck I. Degree of knowledge of health care professionals about pain management and use of opioids in pediatrics. Pain Med. 2014;15(5):807-19.

193. Reyes S. Nursing assessment of infant pain. J Perinat Neonatal Nurs. 2003;17(4):291-303.

194. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR task force for translation and cultural adaptation. Value Health. 2005;8(2):94-104.

195. Heiderich TM, Leslie AT, Guinsburg R. Neonatal procedural pain can be assessed by computer software that has good sensitivity and specificity to detect facial movements. Acta Paediatr. 2015;104(2):e63-9.

196. Sikka K, Ahmed AA, Diaz D, Goodwin MS, Craig KD, Bartlett MS, et al. Automated assessment of children's postoperative pain using computer vision. Pediatrics. 2015;136(1):e124-31.

197. Hullett B, Chambers N, Preuss J, Zamudio I, Lange J, Pascoe E, et al. Monitoring electrical skin conductance: a tool for the assessment of postoperative pain in children? Anesthesiology. 2009;111(3):513-7.

198. Benoit B, Martin-Misener R, Newman A, Latimer M, Campbell-Yeo M. Neurophysiological assessment of acute pain in infants: a scoping review of research methods. Acta Paediatr. 2017;106(7):1053-66.

199. Heden L, von Essen L, Ljungman G. The relationship between fear and pain levels during needle procedures in children from the parents' perspective. Eur J Pain. 2016;20(2):223-30.

200. Morelius E, He HG, Shorey S. Salivary cortisol reactivity in preterm infants in neonatal intensive care: an integrative review. Int J Environ Res Public Health. 2016;13(3).

201. Oberlander T, Saul JP. Methodological considerations for the use of heart rate variability as a measure of pain reactivity in vulnerable infants. Clin Perinatol. 2002;29(3):427-43.

202. Lee GY, Stevens BJ. Neonatal and infant pain assessment. In: McGrath PJ, Stevens BJ, Walker SM, Zempsky WT, editors. Oxford textbook of paediatric pain. Oxford: Oxford University Press; 2014. p. 353-69.

203. Connelly M, Fulmer RD, Prohaska J, Anson L, Dryer L, Thomas V, et al. Predictors of postoperative pain trajectories in adolescent idiopathic scoliosis. Spine. 2014;39(3):E174-81.

204. Pagé MG, Huguet A, Katz J. Prevention of the development and maintenance of paediatric chronic pain and disability. In: McGrath PJ, Stevens BJ, Walker SM, Zempsky WT, editors. Oxford textbook of paediatric pain. Oxford: Oxford University Press; 2014. p. 39-49.

205. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine. A practical guide. Cambridge: Cambridge University Press; 2011.

206. Nunnally JC, Bernstein IR. Psychometric theory. 3rd ed. New York: McGraw-Hill; 1994.

207. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010;63(7):737-45.

208. Sperber AD. Translation and validation of study instruments for cross-cultural research. Gastroenterology. 2004;126(1 Suppl 1):S124-S8.

209. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res. 2010;19(4):539-49.

210. Reitsma JB, Rutjes AWS, Whiting P, Vlassov VV, Leeflang MMG, Deeks JJ. Chapter 9: Assessing methodological quality. In: Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy [Internet]. Version 1.0.0. The Cochrane Collaboration; 2009. [cited 2017 Sep 15]. Available from: http://srdta.cochrane.org/.

211. Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. Qual Life Res. 2017.

212. Bennett M, Carter B, Dooley F, Goddard J, Gordon J, Hartley J, et al. The recognition and assessment of acute pain in children. Update of full guideline [internet]. London: Royal College of Nursing; 2009. Report No.: 003542. [cited 2017 Sep 15]. Available from: https://www.rcn.org.uk/professional-development/publications/pub-003542

213. Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions [internet]. Version 5.1.0 [updated March 2011; cited 2017 Sep 15]. Oxford, UK: The Cochrane Collaboration; 2011. Available from http://handbook.cochrane.org

214. Shea BJ, Bouter LM, Peterson J, Boers M, Andersson N, Ortiz Z, et al. External validation of a measurement tool to assess systematic reviews (AMSTAR). PloS one. 2007;2(12):e1350.

215. Willis GB. Cognitive interviewing: A tool for improving questionnaire design. Thousand Oaks, CA: Sage Publications; 2005.

216. Collins D. Pretesting survey instruments: an overview of cognitive methods. Qual Life Res. 2003;12(3):229-38.

217. Drennan J. Cognitive interviewing: Verbal data in the design and pretesting of questionnaires. J Adv Nurs. 2003;42(1):57-63.

218. Bogdan RC, Biklen SK. Qualitative research for education. An introduction to theories and methods. 4th ed. Boston: Allyn and Bacon; 2003.

219. Miles MB, Huberman AM. Qualitative data analysis: An expanded sourcebook. 2nd ed. Thousand Oaks, CA: Sage Publications; 1994.

220. Van Dijk M. COMFORT behavioral scale [CD-rom]. Rotterdam: Medisch Multimedia; 2004. [cited 2017 Sep 15]. Available from: https://www.comfortassessment.nl/web/index.php/about/contact-us/

221. Maxwell JA. Qualitative research design: An interactive approach. 2nd ed. Thousand Oaks, CA: Sage Publications; 2005.

222. Polit DE, Beck CT. Nursing research. Generating and assessing evidence for nursing practice. Philadelphia: Lippincott, Williams & Wilkins; 2008.

223. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-74.

224. Gwet KL. Handbook of inter-rater reliability. The definitive guide to measuring the extent of agreement among raters. 3rd ed. Gaithersburg, MD: Advanced Analytics; 2012.

225. Systematic Review. In: Glossary [Internet]. Cochrane Community; 2018 [cited 2018 Mar 10]. Available from: http://community.cochrane.org/glossary

226. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ. 2009;339:b2700.

227. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol. 2007;7:10.

228. Faggion CM, Jr. Critical appraisal of AMSTAR: challenges, limitations, and potential solutions from the perspective of an assessor. BMC Med Res Methodol. 2015;15:63.

229. Whiting P, Savovic J, Higgins JP, Caldwell DM, Reeves BC, Shea B, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol. 2016;69:225-34.

230. Whiting P, Savovic J, Higgins J, Caldwell D, Reeves B, Shea B, et al. ROBIS: Tool to assess risk of bias in systematic reviews. Guidance on how to use ROBIS. Bristol: University of Bristol; 2015 [cited 2016 Jul 22]. Available from: http://www.bristol.ac.uk/media-library/sites/social-community-medicine/robis/robisguidancedocument.pdf

231. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med. 2012;22(3):276-82.

232. Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epidemiol. 2009;62(10):1013-20.

233. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. Nurs Health Sci. 2013;15(3):398-405.

234. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol. 2006;3(2):77-101.

235. Patton MQ. Qualitative research & evaluation methods: integrating theory and practice. 4th ed. Thousand Oaks, CA: SAGE Publications; 2015.

236. Lincoln YS, Guba EG. Naturalistic inquiry. Newbury Park, CA: Sage Publications; 1985.

237. Walco GA, Kelley MC. The ethics of pain control in infants and children. In: McGrath PJ, Stevens BJ, Walker SM, Zempsky WT, editors. Oxford textbook of paediatric pain. Oxford: Oxford University Press; 2014. p. 661-8.

238. Malviya S, Voepel-Lewis T, Burke C, Merkel S, Tait AR. The revised FLACC observational pain tool: improved reliability and validity for pain assessment in children with cognitive impairment. Paediatr Anaesth. 2006;16(3):258-65.

239. Byrd PJ, Gonzales I, Parsons V. Exploring barriers to pain management in newborn intensive care units: a pilot survey of NICU nurses. Adv Neonatal Care. 2009;9(6):299-306.

240. Pölkki T, Korhonen A, Laukkala H, Saarela T, Vehviläinen-Julkunen K, Pietil A. Nurses' attitudes and perceptions of pain assessment in neonatal intensive care. Scand J Caring Sci. 2010;24(1):49-55.

241. Twycross A, Forgeron P, Williams A. Paediatric nurses' postoperative pain management practices in hospital based non-critical care settings: a narrative review. Int J Nurs Stud. 2015;52(4):836-63.

242. Twycross AM, Williams AM, Finley GA. Surgeons' aims and pain assessment strategies when managing paediatric post-operative pain: A qualitative study. J Child Health Care. 2015;19(4):513-23.

243. Twycross A. Managing pain during the first year of life. Infant. 2006;2(1):10-4.

244. Mattsson JY, Forsner M, Arman M. Uncovering pain in critically ill non-verbal children: nurses' clinical experiences in the paediatric intensive care unit. J Child Health Care. 2011;15(3):187-98.

245. Van Hulle Vincent C. Nurses' perceptions of children's pain: a pilot study of cognitive representations. J Pain Symptom Manage. 2007;33(3):290-301.

246. Ista E, van Dijk M, van Achterberg T. Do implementation strategies increase adherence to pain assessment in hospitals? A systematic review. Int J Nurs Stud. 2013;50(4):552-68.

247. Craig AD. A new view of pain as a homeostatic emotion. Trends Neurosci. 2003;26(6):303-7.

248. Slater R, Cantarella A, Franck L, Meek J, Fitzgerald M. How well do clinical pain assessment tools reflect pain in infants? PLoS Med. 2008;5(6):e129.

249. Johnston CC, Sherrard A, Stevens B, Franck L, Stremler R, Jack A. Do cry features reflect pain intensity in preterm neonates? A preliminary study. Biol Neonate. 1999;76(2):120-4.

250. Coll MP, Gregoire M, Prkachin KM, Jackson PL. Repeated exposure to vicarious pain alters electrocortical processing of pain expressions. Exp Brain Res. 2016;234(9):2677-86.

251. Byrne A, Morton J, Salmon P. Defending against patients' pain: a qualitative analysis of nurses' responses to children's postoperative pain. J Psychosom Res. 2001;50(2):69-76.

252. Coyne I, Hallstrom I, Soderback M. Reframing the focus from a family-centred to a child-centred care approach for children's healthcare. J Child Health Care. 2016.

253. Convention on the Rights of the Child (General Assembly resolution 1989:44/25) [internet]. New York: United Nations [cited 2018 Jan 12]. Available from: http://www.ohchr.org/EN/ProfessionalInterest/Pages/CRC.aspx

254. Jacob E, Puntillo KA. A survey of nursing practice in the assessment and management of pain in children. Pediatr Nurs. 1999;25(3):278-86.

255. Axelin A, Anderzen-Carlsson A, Eriksson M, Polkki T, Korhonen A, Franck LS. Neonatal intensive care nurses' perceptions of parental participation in infant pain management: a comparative focus group study. J Perinat Neonatal Nurs. 2015;29(4):363-74.

256. Blount RL, Loiselle KA. Behavioural assessment of pediatric pain. Pain Res Manag. 2009;14(1):47-52.

257. Tarbell SE, Cohen IT, Marsh JL. The Toddler-Preschooler Postoperative Pain Scale: an observational scale for measuring postoperative pain in children aged 1-5. Preliminary report. Pain. 1992;50(3):273-80.

258. Buchholz M, Karl HW, Pomietto M, Lynn A. Pain scores in infants: a modified infant pain scale versus visual analogue. J Pain Symptom Manage. 1998;15(2):117-24.

259. Fournier-Charrière E, Tourniaire B, Carbajal R, Cimerman P, Lassauge F, Ricard C, et al. EVENDOL, a new behavioral pain scale for children ages 0 to 7 years in the emergency department: design and validation. Pain. 2012;153(8):1573-82.

260. Ramelet AS, Rees N, McDonald S, Bulsara M, bu-Saad HH. Development and preliminary psychometric testing of the Multidimensional Assessment of Pain Scale: MAPS. Paediatr Anaesth. 2007;17(4):333-40.

261. Hudson-Barr D, Capper-Michel B, Lambert S, Palermo TM, Morbeto K, Lombardo S. Validation of the Pain Assessment in Neonates (PAIN) scale with the Neonatal Infant Pain Scale (NIPS). Neonatal Netw. 2002;21(6):15-21.

262. Hummel P, Lawlor-Klean P, Weiss MG. Validity and reliability of the N-PASS assessment tool with acute pain. J Perinatol. 2010;30(7):474-8.

263. Milesi C, Cambonie G, Jacquot A, Barbotte E, Mesnage R, Masson F, et al. Validation of a neonatal pain scale adapted to the new practices in caring for preterm newborns. Arch Dis Child Fetal Neonatal Ed. 2010;95(4):F263-6.

264. Eriksson M, Storm H, Fremming A, Schollin J. Skin conductance compared to a combined behavioural and physiological pain measure in newborn infants. Acta Paediatr. 2008;97(1):27-30.

265. Gjerstad AC, Wagner K, Henrichsen T, Storm H. Skin conductance versus the modified COMFORT sedation score as a measure of discomfort in artificially ventilated children. Pediatrics. 2008;122(4):e848-e53.

266. Tristao RM, Garcia NV, de Jesus JA, Tomaz C. COMFORT behaviour scale and skin conductance activity: what are they really measuring? Acta Paediatr. 2013;102(9):e402-6.

267. Lucas CE, Vlahos AL, Ledgerwood AM. Kindness kills: the negative impact of pain as the fifth vital sign. J Am Coll Surg. 2007;205(1):101-7.

268. von Baeyer CL, Pasero C. What nurses' work-arounds tell us about pain assessment. Int J Nurs Stud. 2017;67:A1-A2.

269. Buttner W, Finke W, Hilleke M, Reckert S, Vsianska L, Brambrink A. [Development of an observational scale for assessment of postoperative pain in infants]. Anasthesiol Intensivmed Notfallmed Schmerzther. 1998;33(6):353-61.

270. Avian A, Messerer B, Wunsch G, Weinberg A, Kiesling AS, Berghold A. Postoperative paediatric pain prevalence: A retrospective analysis in a university teaching hospital. Int J Nurs Stud. 2016;62:36-43.

271. von Baeyer CL, Spagrud LJ. Systematic review of observational (behavioral) measures of pain for children and adolescents aged 3 to 18 years. Pain. 2007;127(1-2):140-50.

272. Pasero C, McCaffery M. No self-report means no pain-intensity rating. Am J Nurs 2005;105(10):50-3.

273. Willis MHW, Merkel SI, Voepel-Lewis T, Malviya S. FLACC Behavioral Pain Assessment Scale: a comparison with the child's self-report. Pediatr Nurs. 2003;29(3):195-8.

274. Nilsson S, Finnstrom B, Kokinsky E. The FLACC behavioral scale for procedural pain assessment in children aged 5-16 years. Paediatr Anaesth. 2008;18(8):767-74.

275. Twycross A, Voepel-Lewis T, Vincent C, Franck LS, von Baeyer CL. A debate on the proposition that self-report is the gold standard in assessment of pediatric pain intensity. Clin J Pain. 2015;31(8):707-12.

276. Cohen LL, Lemanek K, Blount RL, Dahlquist LM, Lim CS, Palermo TM, et al. Evidence-based assessment of pediatric pain. J Pediatr Psychol. 2008;33(9):939-55.

277. Maneesriwongul W, Dixon JK. Instrument translation process: a methods review. J Adv Nurs. 2004;48(2):175-86.

278. Olsson E, Anderzen-Carlsson A, Atladóttir SM, Axelin A, Campbell-Yeo M, Eriksson M, et al. Cultural adaptation and harmonization of four Nordic translations of the revised Premature Infant Pain Profile. 2018; Submitted manuscript.

279. Olsen MF, Bjerre E, Hansen MD, Hilden J, Landler NE, Tendal B, et al. Pain relief that matters to patients: systematic review of empirical studies assessing the minimum clinically important difference in acute pain. BMC medicine. 2017;15(1):35.

280. Conrad F, Blair J, editors. From impressions to data: Increasing the objectivity of cognitive interviews. Proceedings of the Section on Survey Research Methods [internet]. 156th Annual Meeting of the American Statistical Association. Aug 4-8, 1996; Chicago, Illinois:1996. [cited 2017 Sep 15] Available from: https://www.bls.gov/osmr/abstract/st/st960080.htm

281. Collins D. Analysing and interpreting cognitive interview data: A qualitative approach. Conference proceedings, Quest conference. April 24-26, 2007; Ottawa, Ontario, Canada: Statistics Canada, 2007.

282. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today. 2004;24(2):105-12.

283. Pieper D, Buechter RB, Li L, Prediger B, Eikermann M. Systematic review found AMSTAR, but not R(evised)-AMSTAR, to have good measurement properties. J Clin Epidemiol. 2015;68(5):574-83.

284. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res. 2018.

COMFORT atferdsskala ©

Navnelapp

Dato _____

Klokkeslett _____

Observatør _____

		Kryss av
Våkenhet	Sover dypt	□ 1
	Sover lett	□ 2
	Døsig	□ 3
	Våken og oppmerksom	□ 4
	Overoppmerksom	□ 5
Ro / engstelse	Rolig	□ 1
	Noe engstelig	□ 2
	Engstelig	□ 3
	Svært engstelig	□ 4
	Panikkslagen	□ 5
Respiratorisk respons (gjelder bare intuberte barn)	Ingen spontan respirasjon	□ 1
	Spontan og respiratorstyrt respirasjon	□ 2
	Rastløshet eller motstand mot respiratoren	□ 3
	Puster aktivt mot respiratoren eller hoster jevnlig	□ 4
	Kjemper mot respiratoren	□ 5
Gråt (gjelder bare barn med egenrespirasjon)	Rolig respirasjon, ingen gråtelyder	□ 1
	Sporadisk hiksting og jamring	□ 2
	Klynking (monoton lyd)	□ 3
	Gråt	□ 4
	Skrik eller hyl	□ 5
Motorisk aktivitet	Ingen bevegelse	□ 1
	Sporadiske (tre eller færre) små bevegelser	□ 2
	Hyppige (mer enn tre) små bevegelser	□ 3
	Kraftige bevegelser, begrenset til ekstremiteter	□ 4
	Kraftige bevegelser, inkludert kropp og hode	□ 5
Muskeltonus	Slappe muskler; ingen muskeltonus	□ 1
	Redusert muskeltonus; mindre motstand enn normalt	□ 2
	Normal muskeltonus	□ 3
	Økt muskeltonus. Knyttede hender og tær	□ 4
	Ekstrem muskelstivhet	□ 5
Spenning i ansiktet	Helt slappe ansiktsmuskler	□ 1
	Normal tonus i ansiktet	□ 2
	Tydelig spenning i enkelte ansiktsmuskler (ikke vedvarende)	□ 3
	Tydelig spenning i alle ansiktsmusklene (vedvarende)	□ 4
	Fordreide ansiktsmuskler og grimaserer	□ 5

Poeng totalt ☐

Opplysninger om medisinering _____

Opplysninger om barnets tilstand _____

Type vurdering _____

(før eller etter medisinering eller standard vurdering)

Gjennomsnittlig arterielt blodtrykk og hjertefrekvens er ikke inkludert i denne versjonen av COMFORT-skalaen.

COMFORT behavior © scale - revised

Patient sticker

Date _____

Time _____

Observer _____

		Please place a mark
Alertness	Deeply asleep	□ 1
	Lightly asleep	□ 2
	Drowsy	□ 3
	Awake and alert	□ 4
	Hyper-alert	□ 5
Calmness/Anxiety	Calm	□ 1
	Slightly anxious	□ 2
	Anxious	□ 3
	Very anxious	□ 4
	Panicky	□ 5
Respiratory response (score only in mechanically ventilated children)	No spontaneous respiration	□ 1
	Spontaneous and ventilator respiration	□ 2
	Restlessness or resistance to ventilator	□ 3
	Actively breathes against ventilator or coughs regularly	□ 4
	Fights ventilator	□ 5
Crying (score only in spontaneously breathing children)	Quiet breathing, no crying sounds	□ 1
	Occasional sobbing or moaning	□ 2
	Whining (monotonous sound)	□ 3
	Crying	□ 4
	Screaming or shrieking	□ 5
Physical movement	No movement	□ 1
	Occasional, slight movements	□ 2
	Frequent, slight movements	□ 3
	Vigorous movements limited to extremities	□ 4
	Vigorous movements including torso and head	□ 5
Muscle tone	Muscles lax; no muscle tone	□ 1
	Reduced muscle tone	□ 2
	Normal muscle tone	□ 3
	Increased muscle tone. Clenched fingers and toes	□ 4
	Extreme muscle rigidity	□ 5
Facial tension	Facial muscles totally lax	□ 1
	Normal facial tone	□ 2
	Tension evident in some facial muscles	□ 3
	Tension evident throughout facial muscles	□ 4
	Facial muscles contorted and grimacing	□ 5

Total score ☐

Details medication _____

Details child's condition _____

Type of assessment _____

(Before or after medication or standard assessment)

Mean arterial blood pressure and heart rate are not included in this version of the COMFORT scale