Assessing Pain in Children With Profound Cognitive Impairment: The Development of the Checklist Pain Behavior

Het Beoordelen van Pijn bij Kinderen Met een Zeer Ernstige Verstandelijke Handicap: De Ontwikkeling van de Checklist PijnGedrag

ISBN 90-9018062-1

Photograph cover: Naomi Westerhuis, Lorenzo Jonker & Youri van Rees
Printed by the Document Factory, Leiden, The Netherlands
Cover design: Marcel Terstegen
General support: Maarten Wirtz

The study reported in this thesis was performed at the Department of Child and Adolescent Psychiatry and the Department of Pediatric Surgery, Erasmus MC - Sophia Children's Hospital, Rotterdam, the Netherlands.
This research was financially supported by a grant (MW-NWO 940-031-044) of the Netherlands Organization for Scientific Research (NWO, The Hague).

Printing of this thesis was financially supported by:
David Vervat Stichting Rotterdam
Pijnkennis Centrum Rotterdam
Nutricia Nederland B.V.

Assessing Pain in Children With Profound Cognitive Impairment: The Development of the Checklist Pain Behavior

Het Beoordelen van Pijn bij Kinderen Met een Zeer Ernstige Verstandelijke Handicap: De Ontwikkeling van de Checklist PijnGedrag

Proefschrift

ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de
Rector Magnificus

Prof.dr. S.W.J. Lamberts

en volgens besluit van het College voor Promoties

De openbare verdediging zal plaatsvinden op
woensdag 19 mei 2004 om 9.45 uur
door

Chantal Margaretha Terstegen
geboren te Zwolle

Promotiecommissie

Promotoren:	Prof.dr. J.M. Koot Prof.dr. D. Tibboel
Overige leden:	Prof.dr. A.J. van der Heijden Prof.dr. J. Passchier Prof.dr. C. Vlaskamp
Co-promotor:	Dr. J.B. de Boer

Paranimfen: Mariëlla Kasper
Suzanne Koudijs

Voor alle kinderen met onbegrepen pijn
In de hoop dat ze beter te begrijpen zijn

Contents

Chapter 1

Introduction

Introduction

Pain is a subjective phenomenon and therefore complex in its assessment. Self-report is considered to be the gold standard for assessing pain, but several groups of persons lack this ability due to their limited expressive repertoire. Generally, they are excluded from pain studies because (1) they are rather difficult to investigate and (2) a lack of reliable instruments for the assessment of pain exists in these individuals. Only recently, studies became more focused on the group of persons who have difficulties in communicating their experiences to their caregivers. Some studies focused on (post-surgical) pain responses in neonates and infants (Grunau & Craig 1987; Lawrence & Alcock 1993; Horgan & Choonara 1996; Peters 2001; Van Dijk 2001), while others are currently studying pain responses in elderly people who lost their communicative abilities, e.g., due to Alzheimer's disease (Wary & Doloplus 1999; Rainfray et al. 2003; Van Herk et al. 2004).

In this thesis children with profound cognitive impairment (PCI) are studied. The fact that studying pain in this group of children has been neglected for a long time might be explained by the observation that the capability of children with PCI to experience pain was questioned. It has even been suggested that some children with profound cognitive impairment might be insensitive to pain (Biersdorff 1991). Although Oberlander and colleagues (1999a) mentioned that parents and caregivers participating in their study noticed an increased pain tolerance in these children, there is no conclusive evidence to suggest that they are insensitive or indifferent to pain (Stallard et al. 2001). In contrast, there is recent evidence of intact pain sensitivity in at least a substantial proportion of these children (Gilbert-MacLeod et al. 2000; Malviya & Voepel-Lewis 2001). Recently, Breau et al. reported that caregivers with greater knowledge about children with cognitive impairment (CI) held stronger beliefs that these children experience pain less than children without CI (Breau et al. 2003).

Although studies on the level of pain thresholds in children with PCI are very sparse, it is clear that children with PCI do experience pain. However, they are unable to provide clearly interpretable self-report of pain, thus are not able to express themselves verbally in a way we usually understand. For example, they are able to vocalise but do not use words. This group of children process information and communicate distress and pain in a different way than normally developed children do (Gilbert-MacLeod et al. 2000; Breau et al. 2001; Stallard et al. 2001). Their only way to communicate is through non-verbal behavior. Finley and colleagues (1998) were the first to suggest that interpreting behavioral manifestations could be used as a form of self-report. However, due to somatic conditions like contractures, spastic tetraplegia and recurrent epileptic seizures, most of these children are hardly able to express themselves behaviorally. It would be helpful to know which types of expression (e.g., facial, vocal) are most informative, not only because of the limited expressive repertoire of the child, but also as a consequence of the observer's uncertainty about the interpretation of available behavior as a possible sign or expression of pain (Breau et al.2002a). Because of the developmental delays, it seems evident that the children with PCI express pain in accordance

with their state of cognitive and physical development rather than their calendar age. For instance, a social-communicative deficit has been documented in children with cognitive impairment. Social signaling behaviors (e.g., making eye contact, vocalizing, pointing to an object) are less developed and effective in children with cognitive impairment (Berger & Cunningham 1981) and they often display inappropriate facial actions during positive social interactions, e.g., playing a game (Bufkin & Altman 1995). During the expression of emotions in children with cognitive impairment, such as happiness, sadness or pain, facial activity is present, but often attenuated (Maurer & Newbrough 1987; Oberlander et al. 1999b; LaChapelle & Hadjistavropoulos 1999). In addition, Breau et al. (2000) reported that behavior of individuals with cognitive impairment is idiosyncratic and typical pain behavior (e.g., moaning, facial changes) may commonly occur in these children even when they are not in pain.

Validated instruments to assess pain in children with PCI were not available for a long time. Fortunately, in a relatively short period of time, several recent studies aimed at the development of a rating scale for assessing pain in children with PCI (Giusiano & Jimeno 1995;Collignon & Giusiano 2001; Breau et al. 2000, 2002a, 2002b; Hunt & Goldman 2002; Stallard et al. 2002a, 2002b).

Because of the children's characteristics described above, it can be concluded that observational scales, which are available to assist the observer in post-surgical pain assessment in normally developed children also without the ability of adequate self-report (Van Dijk et al. 2000), probably are not appropriate for children with PCI. Therefore, it is obvious that the assessment of pain in this population might profit substantially from empirical evidence on the validity of observable behaviors as indicators of pain. The use of specific assessment scales together with a tailored analgesic regimen will increase the possibility to assess and monitor pain more adequately and minimize the possibility of over- or under-treatment of pain in children with PCI.

Purpose of the study

The purpose of this study is to develop an observational scale to assess and monitor post-surgical pain in children with PCI which can provide assistance to caregivers (e.g., nurses, parents, physicians) in assessing pain and help them to improve the administration of analgesia in children with PCI in a post-surgical setting.

Initially, we included children with a severe and profound cognitive impairment. However, during the investigation it became evident that the level of the child's motor impairments, coincides with the severity of the child's cognitive impairment. Children with severe cognitive impairment were generally better capable of expressing themselves through behavior because their additional motor impairments were less severe compared to children with profound cognitive impairment. Besides, children with severe cognitive impairment (estimated IQ 20-40) are usually capable of using one or more words to express themselves and are relying on

their non-verbal behavior to a smaller degree than children with a profound cognitive impairment (estimated IQ ≤ 20) (Goorhuis & Schaerlaekens 2000). In fact, Hadden & Von Baeyer (2002) reported that in their study the ability to communicate pain verbally, influenced the way these children express their pain. Some behaviors were reported with a significantly greater frequency for children who could verbally communicate and were directly related to the level of motor capacities ('gestures to or touches part of body that hurts' and 'protects or favours part of the body that hurts'). Nonverbal children scored with significantly higher frequency on behaviors such as 'stiff, spastic, tense and rigid'.
As children with a profound cognitive impairment generally seem to express different pain behavior than children with a severe cognitive impairment, we decided to only include the nonverbal, profoundly cognitively impaired children. Although this decision reduced the number of children for which the observational instrument would be appropriate, it will be optimally tuned to this specific group of children.

For adequate pain treatment in children with PCI, three steps have to be taken. First, we need to know how children with PCI express their pain. Exploration of these behaviors can result in a pool of pain indicators which can be used in the development of a pain assessment scale. Second, we need to construct and develop an observational pain scale that is reliable and valid, and which can also be administered by people who are not (very) familiar with the child. These two steps (constructing, developing and validating an observational pain assessment scale) are necessary for a third and last step: the evaluation of pain treatments and analgesic regimen in children with PCI, enabling implementation in daily care.

The aims of the present study are:
1. To determine the value of salivary cortisol levels used as objective physiological parameter in evaluating pain in children with PCI next to behavioral observations.
2. To test the sensitivity of a large pool of non-verbal candidate indicators for the assessment of surgical pain in children with PCI.
3. To construct and validate an observational scale containing indicators which are sensitive to post-surgical pain in children with PCI.
4. To compare pain behavior displayed by children with PCI with that of normally developed children.
5. To determine the observational scale's applicability and reliability in daily practice for the assessment of everyday, more chronic pain responses in children with PCI.

In this Introduction the main targets and brief outcomes of this thesis are described for each chapter, together with an account of the importance of the studied matters. The last section concerns the outline of this thesis.

Physiological parameter for distress/pain

Pain is principally a subjective experience, and self-report is often described as the gold standard in pain assessment. Because children with profound cognitive impairment signal their pain in other ways than self-report, there remains a degree of uncertainty about the child's pain experience and one must be aware of the knowledge that one can never completely capture the child's experience (Hunt et al. 2003). In the absence of a gold standard for the assessment of pain in children with PCI, it could be informative to use a more objective physiological parameter such as cortisol levels which can be addressed next to behavioral observations when evaluating pain. Rises in cortisol levels may be a useful indicator that a person is in pain (Gunnar 1992). This may be especially true for children with a severe cognitive impairment because they often have very little possibilities to express themselves verbally or purposefully through other behaviors. To be able to recognise and interpret changes in cortisol levels in reaction to stressful events, the availability of reference values for stress- and pain-free children is essential. These reference values have only been published for normally developed, healthy children (Kiess et al. 1995), but it is not known whether these are comparable with the levels typical for children with a severe or profound cognitive impairment and can be used as normative data in studies on these children. Taken the numerous potential painful events in this particular group of children and the lack of a well-validated pain assessment instrument, these baseline data are urgently needed. Since these children tend to experience a number of physical problems and associated stress or pain, their average cortisol levels may be elevated, even if they are in a stable condition. Further, it is unknown whether cortisol secretion in children with a severe or profound cognitive impairment follows a circadian rhythm, as it normally does in healthy individuals (Kiess et al. 1995). This circadian rhythm may be altered due to the unusual sleep-wake pattern in these children, which may be most evident in suppressed early morning cortisol levels.

From item pool to Checklist Pain Behavior

Because of the lack of any validated pain assessment tools developed for this group of children which could serve as a frame of reference, we decided to start with the assessment of post-surgical pain. Despite the administration of analgesia, we were sure that the child was exposed to a directly inflicted pain stimulus (a surgical procedure) and that most of the children, as a consequence, would experience pain. In a later phase of this study, the results will be translated to non-surgical painful situations, namely the application of the observational scale in everyday, more chronic pain situations. We describe the development of the observational pain assessment scale, following three important topics:

Constructing the item pool

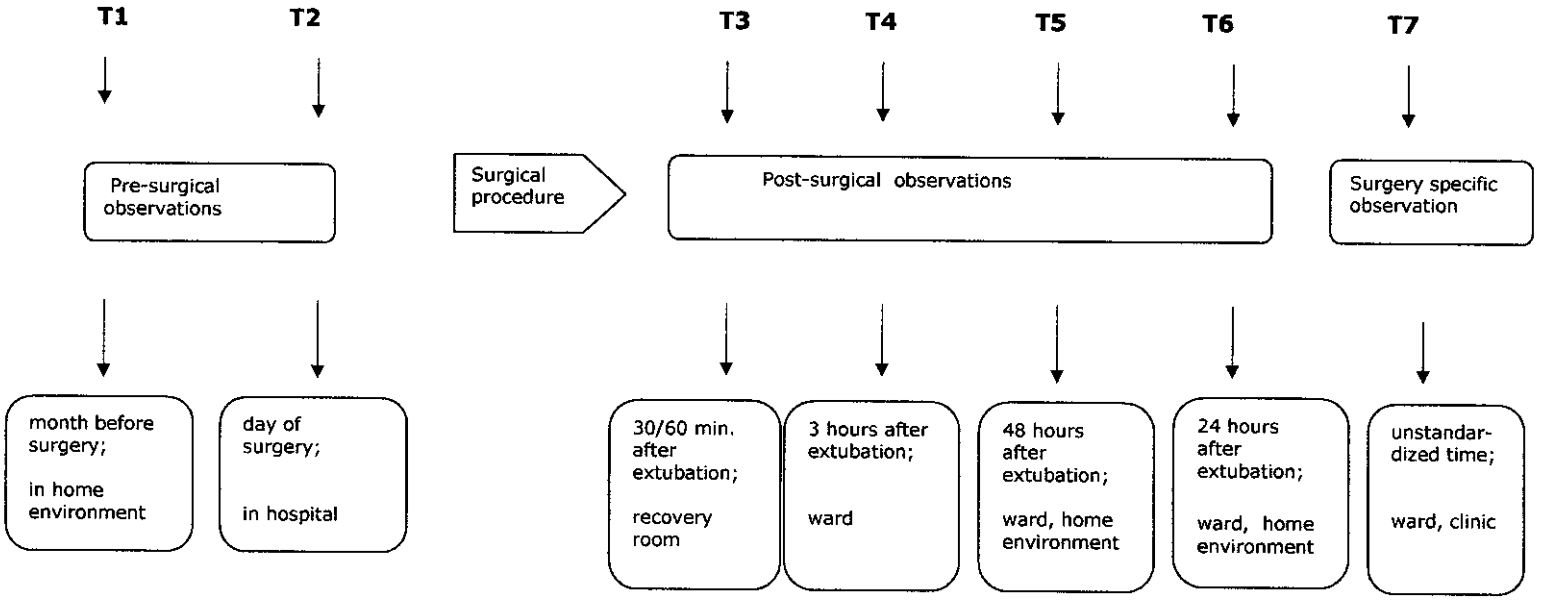
To construct an item pool consisting of behaviors that may be considered to indicate post-surgical pain in children with PCI, a qualitative study was conducted to gain insight into non-verbal pain expressions in children with PCI. First, focused semi-structured interviews were conducted by two researchers with parents (n=5) and caregivers (n=24; 9 nurses,

4 physicians, 5 physical therapists, 3 psychologists, 1 neuro-psychologist, 1 dentist, 1 mouth-hygienist). Participants were sampled on their wide variety of knowledge and experience in caretaking for this group of children. Second, researchers observed behavior of children and adolescents with severe and profound cognitive impairment (n=32) during several possible painful events, e.g., dental treatment, physical therapy and vaccination against influenza. A total number of 169 indicators were observed or reported by the interviewees. Adding 40 pain expressions from existing literature on pain assessment in non-verbal populations (neonates, infants and children with cognitive impairment), resulted in an item pool including a total of 209 indicators (Van Dongen et al. 1999). These indicators were used in a pilot study on 15 children with PCI who underwent various surgical interventions at the Erasmus MC-Sophia Children's Hospital. Expressions not suitable for direct clinical observation of pain were discarded. For example, all expressions relating to 'Activities of daily life' (e.g., changes in sleeping and eating pattern) were removed from the list since the pilot study showed that these indicators were related to a specific context and were mostly not present during standardized observation moments. Further, behaviors that were entered twice under different descriptions, were reduced to one indicator. All remaining indicators were presented in an item pool and consisted of 138 candidate pain behaviors in children with PCI, categorized in six domains: Facial expression, Motor activity, Social/Emotional behavior, Behavior towards painful body part, Vocalisation and Physiological signs.

Selecting the items: pain indicators

Six-minute video recordings were made in order to observe the child's behavior. Initially, we started with observing the child's behavior for 10 minutes. However, the pilot-study showed that in general within six minutes the same behaviors were observed as in 10 minute observations. We concluded that observing longer than six minutes did not contribute anything new to the already observed behaviors in the first six minutes. Further, observation in clinical practice benefits from shorter observational episodes. We decided to set the episodes at 6 minutes and made the following subdivision in order to prevent the observer from missing subtle, attenuated behaviors or signs: 2 minutes of the whole body, 2 minutes of face and trunk and 2 minutes of the face. Seven observational episodes for each recruited child included video recordings of which 2 were pre-surgical: 1 episode at home within one month before surgery when the child was not in pain or distress and 1 episode when the child was already hospitalized but before surgery. Five post-surgical episodes were videotaped: at 30 minutes and at 3 hrs, 24 hrs and 48 hrs after extubation, and one during an pain inducing intervention depending on the type of surgery (Figure 1).

Figure 1. Time and place of observational episodes



On the basis of the pre-surgical episodes the common behavior of the child can be determined. The several standardized post-surgical moments provide more information about the course of the child's pain over time.

All videotaped observational episodes were scored using the 138 behaviors included in the item pool, which were rated on a 5-point scale, with 0=never shown, 1=sometimes shown, 2=regularly shown, 3=often shown, 4=always shown. A scale like this with differentiation in answering categories is essential in children with PCI. Due to the severity of the cognitive, motor and often sensory impairments, most children with PCI have a very limited expressive repertoire. If there are behaviors the child uses to express pain, but are also shown when the child is not in any pain or distress, this behavior will not be recognized as pain indicator when using a less nuanced dichotomous answering category (0=behavior absent, 1= behavior present). For instance, a certain behavior is 'sometimes shown' when the child is not in pain, but 'often shown' when the child indeed is experiencing pain. In case of a dichotomous answering scale, this behavior will be scored as 'behavior present' both in painful and non-painful situations and thus will not be recognized as pain indicator, while, for this particular child, showing customary behavior with increased frequency is a signal of pain.

Further, all video recordings of these episodes were observed and scored with the Visual Analogue Scale (VAS) additional to the 138 behaviors included in the item pool. The VAS is a horizontal continuous 100 mm line with the anchors 'no pain' (= score 0) on the left side and 'the worst pain' (= score 10) on the right side (Huskisson 1974; McGrath & Johnson 1985). The VAS was applied for two reasons, (1) to estimate the convergent validity and (2) to obtain a criterion for the presence of pain. A VAS $\geq$ 4 is considered to indicate the presence of pain (Buchholz et al. 1998; Van Dijk et al. 2002).

The initial item pool of 138 was further reduced to 30 indicators, using any significant difference in occurrence across 7 pre- and post-surgical measurements as a criterion to keep the indicator included. The most important quality of a pain assessment tool is the ability to discriminate between episodes with likely presence versus absence of pain (Breau et al. 2002b). Scores on 7 of the 30 indicators did not significantly differ before and after surgery and were eliminated from the item pool. The remaining 23 indicators showed sensitivity to the pre- versus post-surgical situations, most probably related to the absence versus presence of pain.

Constructing the observational scale

In order to construct the observational scale from the 23 indicators listed in the pool of pain indicators, internal structure was studied and the indicators were summarized into meaningful and reliable clusters. These formed the basis for specific and overall observational subscales to be included in the observational scale, the Checklist Pain Behavior (CPG). Four subscales were clearly distinguishable: Facial expression (9 items), Social behavior/Mood (4 items), Vocalisation (4 items) and Physiological signs (4 items). However, because an uni-

dimensional solution was also satisfactory, these subscales can be combined into a total score without loss of information.

As the scores on the CPG appeared to be highly skewed and also because observation in clinical practice benefits from fewer response categories, we tested the goodness of fit of the models (one and four dimension solution) using four response categories by combining the 'regularly shown' and 'often shown' responses into one 'often shown' category. Since the models based on four and five response categories had an equal fit to the data, we maintained this four-point scale in the further analyses.

Finally, when data reduction was completed and the construction of the Checklist Pain Behavior was set up in a final version, special attention was given to the CPG's qualities on the following three topics.

First, the CPG is tested on its potential to *discriminate between continuous versus acute pain* by comparing CPG scores obtained in the post-surgical episodes with those obtained during an acute painful surgery specific procedure.

Second, a pain assessment scale should not only detect pain but should also be able to *monitor the presence of pain over time*. For a pain instrument to be useful in a surgical setting, it has to be reliable and sensitive to differences in pre- and post-surgical signs of pain, and sensitive to changes in pain expression during the post-surgical period when adequate pain management is critical. This topic has not been addressed in any investigation so far. Because we have data from 6 post-surgical observational episodes at our disposal, we were able to study the sensitivity of the CPG to changes in pain behavior over time by comparing the CPG ratings across subsequent post-surgical episodes.

Third, the sensitivity of the CPG ratings to pain intensity was studied by comparing scores obtained from children admitted to surgical procedures that were assumed to *differ in the degree of painfulness (low, moderate or high)* as rated by medical professionals. Since no gold standard for pain assessment in these children exists, it is important to know whether ratings of children's pain are related to independent indicators of the painfulness of the surgical procedure.

Specificity of CPG pain behaviors

Children with cognitive impairment express their pain in a different manner than normally developed children do (Gilbert-MacLeod et al. 2000; Breau et al. 2001). However, in which way they differ is a question that still remains to be answered.

In order to investigate to what extent behaviors indicating pain in children with PCI are unique for this population, differences in pain behavior between children with and without PCI should be investigated in a standardized post-surgical setting using the CPG. Comparing pain behavior between children with and without PCI undergoing comparable surgical procedures with potentially high pain intensity, while receiving the same standardized analgesia, did not receive much attention so far. We compared observations of pain behavior in two groups of children, with and without PCI, but with a similar mental age, undergoing similar procedures.

Applicability in practice: adequacy for the assessment of everyday pain responses

Everyday pain in daily practice may differ from post-surgical pain (Breau et al. 2002b). The CPG was used in daily practice to test its reliability, validity and feasibility when used for observations of everyday, often more chronic pain, related to physiological impairments and diseases.

It is unknown whether nurses involved in the caretaking of these children in different settings can be trained to score the child's behavior reliably at the bed-side using the CPG. In addition, we do not know if scoring with the CPG is influenced by knowledge of the child's common behavior. Therefore we investigated if there are differences regarding reliability of scoring with the CPG between nurses who are familiar with the observed child (and with children with PCI in general) and nurses who are not (or less) familiar with the child.

The CPG was completed by nurses working on a neurology ward of the Erasmus MC-Sophia Children's Hospital and by nurses in a residential healthcare facility for children with PCI. In the hospital we decided to work with nurses on a neurology ward because the aim was to observe everyday pain and on this ward most children with PCI were admitted for other reasons than surgery.

Outline of this thesis

In *chapter 2*, information is provided about baseline values of circadian salivary cortisol levels for children with PCI. To be able to recognise and interpret changes in cortisol levels in reaction to painful events, the availability of reference values for stress- and pain-free children is essential. Cortisol is potentially an important physiological parameter in pain related research projects in this particular group of children and can be addressed next to behavioral observations.

Chapter 3 describes the development of the observational measure, from item pool to Checklist Pain Behavior. Although the specific expression of pain may be very individual, there appears to be a set of shared generic indicators expressed by children with PCI. The identification of this item pool of observable pain behaviors enables the development of an observational pain assessment scale for this particular group of children.

Chapter 4 reports the psychometric properties of the Checklist Pain Behavior (CPG). We investigated whether all 23 indicators, which were listed in the pool of pain indicators in the preceding study (chapter 3), could be summarized into meaningful and reliable clusters for subscales of the CPG. Further, reliability - internal consistency, interrater agreement and interscale correlations for the resulting subscales and overall scale - was determined, and construct and convergent validity of the CPG were assessed.

Chapter 5 compares the differences in post-surgical pain responses between children with and without PCI. Sixteen children with PCI and 12 normally developed preverbal children were

Chapter 6 reports on a study that examined the use of the CPG in daily practice and investigates whether this observational measure could detect every day, more chronic, pain in children with PCI in two settings that differed in degree of acquaintance with the child, i.e., a children's hospital and a residential healthcare facility.

Finally, in *chapter 7*, the main findings and conclusions of this thesis are discussed. Directives for future research and recommendations for use in daily practice are presented.

Chapter 2

Salivary cortisol in children with severe and profound cognitive impairment

Chantal Terstegen
Hans M. Koot
Suzanne M. Koudijs
Josien B. De Boer
Els M. Weijers
Frank H. de Jong
Dick Tibboel

Summary in:
Developmental Medicine and Child Neurology, 2003, 45, 139-140

Abstract

Cortisol levels are potential indicators of distress, such as pain, in children with cognitive impairment, but reference values of circadian salivary cortisol levels have only been published for normally developed children. The aim of this study was to provide these reference values for children with cognitive impairment and assess differences with normally developed children. Using a cross-sectional design, saliva samples were collected at 7.30 AM, 12.30 PM and 17.30 PM in 49 children (1-18 years) with severe cognitive impairment and in 44 normally developed children matched by age. Cortisol levels were determined by radio immuno assay. Cortisol levels and their circadian variation in children with cognitive impairment were not significantly different from those of normally developed children, although in more than one third of the impaired children a circadian rhythm was lacking. Our results indicate that reference values of salivary cortisol for normally developed children can also be used for children with severe cognitive impairment. Cortisol may be an important parameter in pain related research projects in this particular group of children.

Introduction

Cortisol levels are well known as a measure of the activity of the hypothalamic-pituitary-adrenocortical (HPA) system. This system induces the production of cortisol in reaction to various forms of stress, and is believed to play an essential role in stress-regulation in humans (Kirschbaum & Hellhammer 1989). The production of cortisol is regulated by the pituitary hormone adreno-cortico-tropic hormone (ACTH), the secretion of which is in turn stimulated by corticotropin-releasing hormone (CRH) from the hypothalamus.
Rises in cortisol level are usually detectable shortly after distressing events but also correlate with active and successful attempts to cope with possible distress. This emphasises, that cortisol has a primary function in the organism's biologic *adaptation* to stress (Gunnar 1992).

The production of cortisol in humans exhibits a well-defined circadian rhythm. This circadian rhythm of cortisol release is under the influence of the sleep-wake activity of a person. Under basal conditions, levels of cortisol are the highest during the early morning, just prior to waking. After awakening, the cortisol level starts to decrease and reaches a minimum after midnight. If a new sleep-wake rhythm is adopted, this secretion pattern changes over several days to adjust to the new pattern (Gunnar 1992; Born et al. 1999). Gunnar and colleagues have provided evidence that in healthy children the circadian rhythm in cortisol should certainly be present at 15 months of age (Gunnar et al. 1996).

There are two important reasons why diurnal cortisol levels in severely cognitively impaired children may be disturbed. First, children with a severe cognitive impairment are often coping with many physical problems. They often suffer chronically from neurological or motor problems such as epilepsy and spasticity (Fanurik et al. 1999), gastrointestinal symptoms, e.g., gastroesophageal reflux, oesophagitis, abdominal diseases, constipation (Del Giudice et

al. 1999) and orthopaedic problems like scoliosis, hip luxations and contractures (Nolan et al. 2000).

Second, although empirical evidence is lacking, clinical impressions indicate that children with a severe or profound cognitive impairment tend to show disturbed night-time sleep behaviour, awakening several times at night or early in the morning. They also tend to take more naps during the day than normally developed children and sometimes even switch their day-night rhythm. Preliminary evidence shows that, at least in normally developed infants, peak cortisol values do not occur in children with significantly reduced duration of night sleep (Larson et al. 1998). Rises in cortisol levels may be a useful indicator that a person is in pain (Gunnar 1992). This may be especially true for children with a severe cognitive impairment because they often have very little possibilities to express themselves verbally. To be able to recognise and interpret changes in cortisol levels in reaction to stressful events, the availability of reference values for stress- and pain-free children is essential. These reference values have only been published for normally developed, healthy children (Kiess et al. 1995), but it is not known whether these are comparable to the levels typical of children with a severe cognitive impairment and can be used as normative data in studies on these children. Taken the numerous potential painful events in this particular group of children and the almost complete absence of a well validated pain assessment instrument, especially in this field of medicine, these baseline data are urgently needed.

This study aimed to answer two questions. The first question was whether cortisol levels of children with a severe cognitive impairment are comparable to those of normally developed children. Since these children tend to experience a number of physical problems and associated stress or pain, their average cortisol levels may be expected to be elevated even if they are in a stable condition. The second aim was to evaluate whether cortisol secretion in children with a severe cognitive impairment follows a circadian rhythm, as it normally does in healthy individuals (Gunnar 1992). This circadian rhythm may be expected to be altered due to the unusual sleep-wake pattern in these children, which may be most evident in suppressed early morning cortisol levels.

Procedure

The medical ethical committee of the academic hospital approved this study and informed consent was obtained from parents of participating children.

Participants

A cross-sectional study was undertaken to measure salivary cortisol levels in 49 cognitively impaired children (27 boys, 22 girls; ages 1-18 years, mean age 8) and 44 normally developed children (18 boys, 26 girls; ages 1-18 years, mean age 9.1) in order to obtain

information about possible differences in mean cortisol baseline levels. The children with cognitive impairment who participated in this study were enrolled in a study on pain assessment in 55 cognitively impaired children supported by the Dutch Research Council, of whom 50 yielded sufficient data. To set up a group of normally developed children, a sample of 54 children was drawn from their birth register, of whom 44 yielded sufficient data.

None of the cognitively impaired children was documented to have a cognitive level higher than 12 months. They may all be regarded severely or profoundly cognitively impaired according to the AAMR definition (Luckasson 1992). The origins of the included children's cognitive impairment were diverse (Table 1).

Table 1. Origin of the children's cognitive impairment (n=49)

Origin cognitive impairment		N*	% of total
Congenital / metabolic anomalies:			
- Syndrome		10	20%
Infections		2	4%
Major structural cerebral abnormalities		6	12%
	Total	**18**	**36%**
Cognitive impairment during birth:			
Perinatal asphyxia		11	22%
Posthypoxic encephalopathy		7	14%
	Total	**18**	**36%**
Cognitive impairment after birth: (between 2 months - 2 years old)			
Encephalopathy		3	6.2%
Myocardial infarct		1	4.8%
Meningo- or pneumococcal sepsis		3	6.2%
	Total	**7**	**16.4%**
Origin remained unknown		6	12%

*more than 1 origin per child is possible

Although all children with cognitive impairment were physically in a stable condition at the day of cortisol sampling, chronic physical diseases were frequently documented (Table 2). All but one of these children underwent one or more surgical procedures (mean=4, range=1-7) under general anaesthesia during their lifetime. None of them had surgery within the preceding month of the surgical procedure they were included for. The group of normally developed children did not have any physical disabilities.

Table 2. Chronic physical disorders in children with cognitive impairment (n=49)

Type of disorder	N*	%
Epilepsy	40	80%
Malnutrition	26	52%
Poor vision or blindness	24	48%
Cortical blindness	6	12%
Gastroesophageal reflux	24	48%
Constipation	22	44%
Contractures	22	44%
Bronchial and pulmonary infections	18	36%
Spastic quadriplegia	16	32%
Hearing loss or deafness	14	28%
ENT(ear, nose, throat) infections	11	22%
Asthma	6	12%
Cardiac disease	6	12%
Scoliosis	5	10%
Hyper- and hypothermia	4	8%
Anaemia	3	6%
Urinary tract infection	2	4%
Eczema	2	4%
Adipositas	2	4%

*more than 1 disorder per child is possible

Twenty-four children with cognitive impairment were permanently living in an institution for residential care and 31 children were living at home and visited a day care centre. All participating institutionalised children were living in the same institution at the day of sampling. Eight of the children in the institution suffered from behavioural disorders and five of them were living in a separate group for that reason. All normally developed children were living at home with their parent(s).

For both groups of children included in this study, possible confounders on the measured cortisol levels were documented. The use of medication was evaluated for both groups because of their influence on cortisol production, especially analgesic and anti-convulsive medication (Putignano et al. 1998; Coolens et al. 1987). Of all participating children, 41.9% (n=39, all cognitively impaired) suffered from mild to severe epilepsy for which they were all treated with a wide variety of anti-convulsant drugs. No stressful events or overt epileptic features were present within 2 hours before sampling. None of the children used analgesics on a regular basis nor used them on the day of sampling.

Because cortisol production increases in response to food intake and in persons with adipositas (Kiess et al. 1995), the child's body weight was reported. As children with severe cognitive impairment often suffer from growth deficiencies, their bodyweight was compared with their length instead of their age to determine over- or underweight. Two cognitively impaired children suffered from severe overweight, while 26 cognitively impaired children were coping with malnutrition due to swallowing disorders, regurgitation and/or vomiting.

For the normally developed children and children with cognitive impairment, the activity (sleep/wake pattern) (Born et al. 1999; Kiess et al. 1995), use of milk products which may

yield false high cortisol results because cortisol or cortisol-like substances seem to be present in milk (Magnano et al. 1989; Kirschbaum & Hellhammer 1994), and possible stressful events prior to saliva sampling were reported by parents, personal caretakers or the researcher on the day of saliva sampling. Furthermore the parents or caretakers were asked for their perception of the child's (chronic) pain or distress and the way the child expressed this pain/distress.

To be included in this study, children needed to be living in their present situation for at least one month, to make sure they did not recently go through any major environmental changes. Children with abnormal renal/adrenal function or abnormal liver function were excluded.

Measurements and analysis

Saliva samples were taken at three times within one day to assess the average levels and the circadian rhythm of cortisol secretion. Because Kiess et al (Luckasson 1992) reported variation in mean cortisol levels between age groups, we grouped the participants according to age as follows (respectively cognitively impaired/normally developed): 1-4 yr (n=7/5), 5-7yr (n=15/12), 8-18yr (n=28/27).

For each age group mean cortisol values were calculated for morning, afternoon and evening measurements. Saliva samples from the children living in an institution were obtained by the researchers using a standardised procedure. For children living at home, parents were asked to take saliva samples using the same procedure on basis of written instructions. Saliva flow was stimulated by applying 0.3 ml of a solution of citric acid crystals in water (15g/L). Citric acid might decrease the pH of the saliva sample, and a low sample pH may cause a false high cortisol level in immuno assays (Kirschbaum 1989). However, the increase of cortisol levels due to the use of citric acid should not be exaggerated. In a pilot study, which was performed among 10 random members of the research group, we found a non-significant mean increase of 0.5 ± 0.9 nmol/L for saliva that was stimulated by citric acid in comparison to pure unstimulated saliva of the same person. Given this limited effect, we decided to use stimulation by citric acid in this study. Stimulated saliva deserves preference to resting saliva, because a larger volume of saliva can be obtained, the pH gradient between plasma and saliva is smaller and fewer saliva samples are too viscous to allow proper analysis than when unstimulated saliva is used (Gorodischer & Koren 1992).

The saliva was collected by swabbing the child's mouth with a cotton role. The moistened end of the cotton was then put in a needleless syringe and expressed into a collecting tube. A minimum of 400 µL of saliva was needed for proper analysis. Collected saliva was stored at -20°C. A radio immuno assay method, Coat-A-Count (Diagnostic Products Corporation, Los Angeles, CA), was used to measure cortisol in the saliva samples in duplicate. Interessay variation was below 14%, whereas intra-essay variations were between 5 and 7%, with the exception of samples with cortisol levels below 2 nmol/L, where it was 16%. During data collection and analysis, six children with cognitive impairment were excluded from the study.

In 3 cases, the cortisol samples were too viscous, contained too little saliva or were contaminated by blood and no proper analysis in the laboratory could be performed on these samples. In 3 cases the use of medication (anti-inflammatory medication containing corticosteroids) caused extreme elevations of the salivary cortisol level (>3 SD). Ten normally developed children dropped out because seven parents did not give their informed consent to participate and three parents never returned the saliva samples.

The Mann-Whitney test ($p<.05$) was used to compare mean cortisol levels in saliva from 49 children with a cognitive impairment with those in samples from 44 normally developed children, matched by age. Fisher's exact test was performed to investigate whether the difference in a lack of circadian rhythm in both groups was statistically significant. In order to test if children with cognitive impairment with and without a circadian rhythm in cortisol differed in clinical, physiological or sociographic parameters, a Mann-Whitney test was performed.

Results

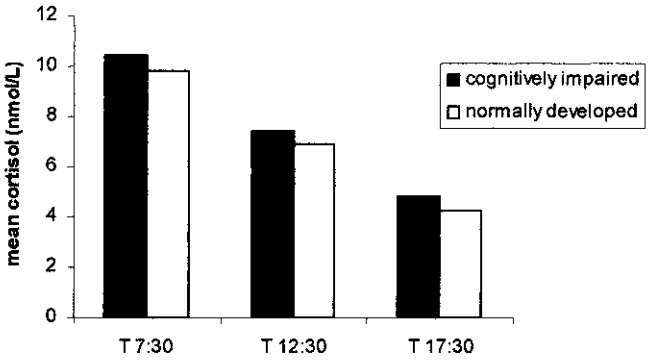
No differences were found in volume of the saliva samples collected by researchers and parents, indicating that parents are perfectly capable of taking saliva samples from their children using written instructions. Despite the presence of numerous sources of stress in cognitively impaired children no significant differences in mean cortisol levels between the children with cognitive impairment and the normally developed children were found (Table 3). Figure 1 shows that, across all measurements, the mean cortisol levels for the cognitively impaired children were on average 0.5 nmol/L higher, and showed more variation than in the group of normally developed children.

Within the group of normally developed children no deviate cortisol levels were found (i.e., ≥ 2 SD) and no possible confounders on cortisol levels were reported.

Table 3. Salivary cortisol levels (median (CI 95%), nmol/L) in children with CI (A) and in normally developed children (B) sampled at on one day

	Age group 1-4 A: N=7 B: N=5	Age group 5-7 A: N=15 B: N=12	Age group 8-18 A: N=27 B: N=27	Total all ages A: N=49 B: N=44
07.30h	A: 7.5 (35.1, -12.5)	8.0 (18.3, -.9)	9.5 (32.3, -8.1)	8.95 (29.1, -6.7)
	B: 10.3 (12.3, 8.8)	9.6 (18.2, 2.8)	8.4 (24.1, -2.3)	9.5 (21.9, -.4)
12.30h	A: 4.5 (11.9, -.3)	3.6 (15.7, -4.7)	4.6 (13.3, -1.5)	4.4 (13, -2.4)
	B: 6.25 (10.6, 2.2)	3.7 (10.1, -.8)	4.1 (11.9, -1.9)	4.3 (11.3, -1.2)
17.30h	A: 3.1 (13.1, -3.4)	2.4 (7.2, -1.4)	2.7 (12.5, -4.5)	2.7 (11.2, -3.6)
	B: 4.5 (12.9, -2.1)	3.4 (6.1, .06)	1.9 (6.1, -1.2)	2.4 (7.1, -1.3)

Figure 1. Cortisol levels (mean ± 1 SD) in saliva collected from children with CI and from normally developed children at various times of the day



The cortisol levels of the 10 cognitively impaired children who were not treated with anti-convulsants, were substantial higher but did not significantly differ from the 39 impaired children who did receive this kind of medication (Table 4). No significant difference in mean cortisol levels was found between children with cognitive impairment suffering from severe overweight, malnourished children and those who had a normal weight for their age.

Table 4. Salivary cortisol levels (median (CI 95%), nmol/L) in children with CI with epilepsy (n=39) and without epilepsy (n=10) at three time points on one day

	Epilepsy absent	Epilepsy present
07.30h	5.4 (17.6, -11.7)	9.0 (23.3, -5.3)
12.30h	3.1 (7.5, -1.3)	5.0 (12.2, -2.2)
17.30h	1.4 (6.2, -3.5)	2.8 (10.6, -5.0)

In 11 children with cognitive impairment fairly low levels were found in the morning (< 1 SD than in the group with normally developed children). Four of these 11 children had been awake most of the night before the day of saliva sampling and four children were awake very early that morning and therefore may already have passed their morning cortisol peak (Gunnar 1992; Born et al. 1999). Ten of these 11 children also showed lower levels on the afternoon and evening measurement. For 16 measurements in eight children with cognitive impairment cortisol levels were elevated (> 1 SD) compared to the mean levels of normally developed children. Seven of these high levels were found in the group of children that suffered from behavioural disorders.

On average, the cortisol levels of the cognitively impaired children showed a circadian rhythm similar to that of normally developed children as presented in Figure 1. However, when looking at intra-individual variation of cortisol levels, in 20 of the 49 cognitively

impaired children the expected circadian rhythm was lacking, i.e. their cortisol production showed no decreasing pattern in time from early morning to evening measurement or cortisol levels were low in the morning and peaked in the afternoon or evening (Figure 2). Fourteen of the 44 normally developed children also showed no circadian rhythm, in all cases because their 12.30h levels peaked (Figure 3). However, the lack of circadian rhythm in both groups did not significantly differ (X^2= .68, α .05).

Figure 2. Individual cortisol levels without circadian rhythm in 20 children with cognitive impairment

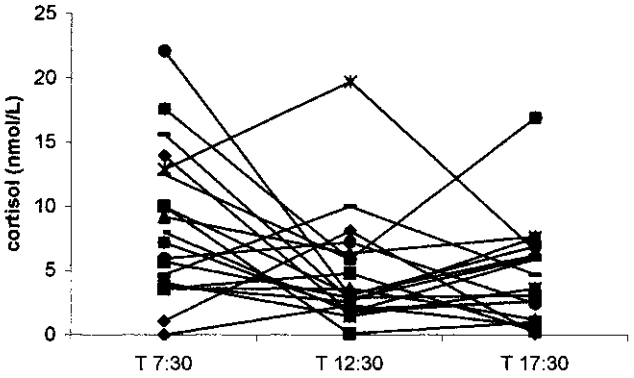
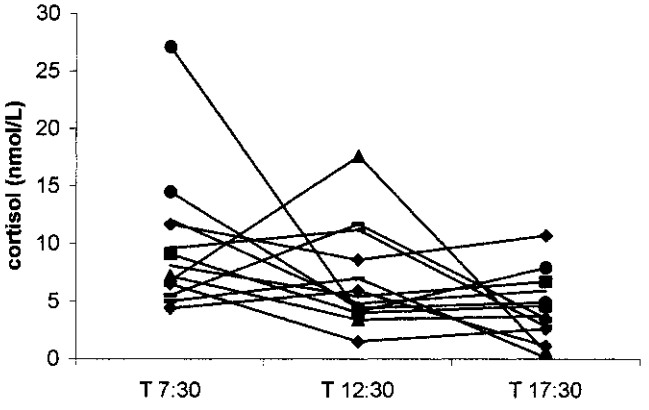


Figure 3. Individual cortisol levels without circadian rhythm in 14 normally developed children



The two groups (presence and absence of a circadian rhythm) within the cognitively impaired group have been studied on characteristics like physical or behavioral problems, origin of cognitive impairment, severity of the impairment and sleep-wake patterns (Table 5).

They appeared to be significantly different regarding the prevalence of epilepsy and the severity of the cognitive impairment.

Table 5. Characteristics of children with cognitive impairment with (n=29) and without (n=20) a circadian rhythm in cortisol (sum; %)

Child characteristics	Presence circadian rhythm	Absence circadian rhythm
Age	2-14 years old, mean 7.1	2-16 years old, mean 7.4
Gender		
Boys	16 (55.2)	13 (65)
Girls	13 (44.8)	7 (35)
Origin cognitive impairment		
Congenital	8 (27.6)	8 (40)
During birth	8 (27.6)	4 (20)
After birth	10 (34.5)	7 (35)
Unknown	3 (10.3)	1 (5)
Severity of cognitive impairment		
Profound	28 (96.6)	15 (75)
Severe	1 (3.4)	5 (25)
Sensory impairments (blindness or deafness)		
Absent	10 (34.5)	9 (45)
Present	19 (65.5)	11 (55)
Behavioral problems (eating problems, hyperactivity, autistic, automutilation)		
Absent	20 (69)	13 (65)
Present	9 (31)	7 (35)
Physically handicapped (confined to a wheelchair)		
Absent	1 (3.4)	4 (20)
Present	28 (96.6)	16 (80)
Spastic quadriplegia		
Absent	13 (44.8)	11 (55)
Present	16 (55.2)	8 (45)
Epilepsy		
Absent	3 (10.3)	7 (35)
Present	26 (89.7)	13 (65)

Discussion

These are the first data on cortisol levels reported for children with a severe or profound cognitive impairment. On average, cortisol levels appeared to be very similar for children with cognitive impairment compared to normally developed children. Furthermore, the majority of

these children showed a circadian rhythm in cortisol levels comparable to that reported for the general population.

An explanation for the lack of significant differences between mean cortisol levels found in children with cognitive impairment and normally developed children, might lay in the fact that cortisol secretion is regulated through a negative feedback mechanism. Rises of cortisol levels will be answered by a decrease in cortisol secretion and cortisol will remain at a set level. The activity of the HPA axis will adapt to chronic exposure to stress, such as physical problems, and to long-lasting disturbances in the sleep-wake pattern (Kirschbaum & Hellhammer 1998; Pignatelli et al. 1998). Therefore, rises of cortisol levels are probably only detectable shortly after acute, stressful stimuli and not in answer to chronic forms of distress. There are no convincing data available on chronic stress effects on cortisol in saliva (Kirschbaum 1989).

Although on average a circadian rhythm in cortisol levels was found for the cognitively impaired children, we did not find a circadian rhythm in 20 children. However, in the group of normally developed healthy children, a lack of circadian rhythm was also demonstrable (n=14).

In several studies (Kiess et al. 1995; Gunnar 1992) the lack of a circadian rhythm of salivary cortisol levels was reported during the first months of life (up to 1 year old). Although most studies report about interindividual variations of cortisol levels (Kiess et al. 1995), Gunnar et al. (1996) and Kirschbaum and Hellhammer (1989) reported about individual differences and intra-individual variations of cortisol levels at different times of the day. However, nothing has been published about intra-individual circadian variation e.g., the prevalence of a lack of circadian rhythm in these children.

Since all children were 3 years or older, we expected a circadian rhythm in all children, unless this was seriously disturbed by sources of stress or an altered sleep-wake pattern. Gunnar and Vasquez (2001) mentioned that chaotic sleep-wake patterns might be a factor contributing to the disturbed daytime pattern of cortisol levels noted for some populations of children. However, parents of eight children with cognitive impairment reported an altered sleep-wake pattern (they were awake most of the night or awoke very early that morning before saliva sampling), though in all but one of these eight children the cortisol secretion showed a normal circadian rhythm. In a review paper, Heim et al. (2000) concluded that low morning cortisol levels and a dampening of the diurnal rhythm in daytime cortisol is more often found in individuals with stress-related physical disorders (e.g., chronic pelvic pain). However, they based their conclusion on studies with normally developed adults. We were not able to ascertain whether the groups with presence versus absence of a circadian rhythm in this study differed regarding stress-related disorders, because these children were not able to report the presence or severity of stress they might experience due to their physical disorder(s). The normally developed children who were able to communicate about this topic did not report any form of distress at the day of sampling, including those with a lack of

circadian rhythm in their cortisol levels. They also slept well the night before sampling and they all awoke approximately at the same time that morning.

In sum, this study provided no indications for altered salivary cortisol levels in children with a severe or profound cognitive impairment compared to those in normally developed children of the same age.

Although the cognitively impaired children included in this study suffered chronically from many physical problems which are likely to cause distress in these children, this was not reflected in higher cortisol levels.

These data suggest that mean salivary cortisol levels for normally developed children, are also representative for children with a severe or profound cognitive impairment. We therefore suggest that when in future (pain) research salivary cortisol levels will be measured, reference values of normally developed healthy children can also be used for children with a severe or profound cognitive impairment. However, it should be emphasised that for the assessment of cortisol reactivity to acute stimuli, individual baseline data are needed, since a sizeable portion of the children do not show the expected circadian curve in cortisol levels.

Chapter 3

Measuring pain in children with profound cognitive impairment: pain response to surgical procedures

Chantal Terstegen
Hans M. Koot
Josien B. de Boer
Dick Tibboel

Pain, 2003, 103, 187-198

Abstract

This study investigated post-surgical pain in children with profound cognitive impairment (PCI), searching for a core set of cues these children use to express their pain. Fifty-two children were observed while they were admitted to the Sophia Children's Hospital for surgery, twice before and five times after surgery. All observations were scored with the item-pool consisting of 134 possible pain indicators, using a five point scale ranging from 0 (never shown) to 4 (always shown). Second, we used the VAS to give a general impression of the severity of the children's pain during the episodes they were observed. Several analyses provided evidence that 23 observable behaviors are sensitive to post-surgical pain in children with PCI, regardless of the pain intensity of the surgical procedures they underwent. The finding that all indicators, except for one, were scored significantly higher on episodes with VAS ratings ≥ 4, indicates the sensitivity of these indicators concerning absence versus presence of clinically meaningful levels of pain. This study reveals the potential clinical utility of a core set of indicators which can be used to assess post-surgical pain in children with PCI.

Introduction

Children with profound cognitive impairment (PCI) probably experience many episodes of pain due to a variety of painful somatic conditions such as gastro-oesophagael reflux, contractures and epilepsy. These children have a higher risk of experiencing pain while their communication difficulty increases the likelihood that pain will remain unrecognised due to wrongly interpreted observations of behavior and therefore remain untreated (Anand et al. 2000; Stallard et al. 2001). The recognition and assessment of these probable pain experiences is hindered by both the limited expressive repertoire of the child due to a number of cognitive, verbal and motor limitations and uncertainty about the meaning of available behavior as a sign or expression of pain. Because of their developmental delays, it seems evident that these children experience and express pain in accordance with their state of cognitive and physical development rather than their calendar age. For instance, a social-communicative deficit has been documented in children with cognitive impairment. Social signaling behaviors (e.g., making eye contact, vocalizing, pointing to an object) are less developed and effective in children with cognitive impairment (Berger & Cunningham 1981) and they often display inappropriate facial actions during positive social interactions, e.g., playing a game (Bufkin & Altman 1995). During the expression of emotions in children with cognitive impairment, such as happiness, sadness or pain, facial activity is present, but often attenuated (Maurer & Newbrough 1987; Oberlander et al. 1999b; LaChapelle et al. 1999). Given these findings about their usual behavior, it is conceivable that these children exhibit an altered pain response and therefore validated scales which are available to assist the observer in post-surgical pain assessment in normally developed children (Van Dijk et al. 2000, 2002) may not be appropriate for children with PCI.

Rather than assuming that these children are relatively pain indifferent or insensitive, as suggested by Biersdorff (1991, 1994), it is possible and probably more likely that these children process information and communicate distress in a different manner than normally developed children (Gilbert-MacLeod et al. 2000; Breau et al. 2001). Although Oberlander and colleagues mentioned that parents and caregivers participating in their study noticed an increased pain tolerance in these children, there is no conclusive evidence to suggest that they are insensitive or indifferent to pain (Oberlander et al. 1999a; Stallard et al. 2001). Actually there is recent evidence of intact pain sensitivity in at least a substantial proportion of these children (Gilbert-MacLeod et al. 2000; Malviya & Voepel-Lewis 2001). Breau reported that behavior of individuals with cognitive impairment is idiosyncratic and typical pain behavior (e.g., moaning, facial changes) may commonly occur in these children even when they are not in pain (Breau et al. 2000). Considering the above, it is not surprising that some investigators suggested the possibility that medical personnel is unaware of or unable to accurately interpret the individual expressions of pain in children with PCI (Maurer & Newbrough 1987; Biersdorff 1991; Malviya & Voepel-Lewis 2001). It is obvious that the assessment of pain in this population might profit substantially from empirical evidence on the validity of observable behaviors as indicators of pain.

Few investigations have studied the impact of cognitive impairment on children's pain responses. Three studies subscribed the importance of special attention for this group of children (Collignon & Giusiano 2001; Hunt 2001; Breau et al. 2002a), of which only one specifically studied post-surgical pain (Breau et al. 2002a). All three studies identified possible pain cues in children with PCI showing behaviors from multiple dimensions: vocalisation (e.g., cry, scream, moan), facial expression, movement (both increased and decreased), change in muscle tone (increased and decreased), guarding or protection of painful area, changes in social interaction and changes in eating and sleeping. Physiological changes such as changes in skin color, shivering and sweating were only described by Breau (2001, 2002a).

The aim of this study was to identify indicators of post-surgical pain in children with PCI which are observable in clinical practice and can also be administered by persons unfamiliar with the child. To this end, first, a large pool of possible indicators for pain was derived by a variety of methods and then were tested on their sensitivity to pain. We selected indicators from this pool that reflected clear pre- and post-surgical differences, most probably related to the presence versus absence of pain. Second, we studied the sensitivity of the selected indicators to changes in pain at four time points, from 30 minutes till 48 hours after surgery. Third, in order to evaluate acute pain lasting several days such as post-surgical pain (Gauvain-Piquard et al. 1987) as well as sharp short-lasting pain, we tested the indicators on sensitivity to two types of post-surgical pain: pain measurements over time (four episodes) and directly induced, surgery specific pain. Fourth, we studied if the indicators can be used to assess post-surgical pain caused by several types of surgical procedures varying in pain intensity. Fifth, and finally, we tested the relation between these indicators and (partially independent) pain scores on a visual analogue scale.

Methods

Participants

The children who participated in this study were enrolled in a research project on pain assessment in children with PCI supported by the Dutch Research Council. This cross-sectional study was undertaken to assess post- surgical pain in 52 children (27 boys, 25 girls; ages 3-19 years, mean age 8.0), who were admitted to a level III children's hospital for a surgical procedure. Power analyses ($\alpha < 0.05$) indicated that the power to detect large differences in means or proportions was ≥ 0.99, for medium differences ≥ 0.85 and for small differences ≥ 0.25. To be included in this study, children needed to be living in their present situation for at least one month, to make sure they did not recently go through any major environmental changes. Children with abnormal renal/adrenal function or abnormal liver function were excluded.

Although there is no clear definition available of children with PCI, they can be described as children with a cognitive development which is estimated to be below a calendar age of 2 years old (equivalent IQ 0-20). Since no diagnostic tools are available which can precisely determine a child's IQ level below 40, we also assessed language development using 'minimal speech standards' (Goorhuis & Schaerlaekens 2000). Standards for PCI include the ability of making sounds but not expressing oneself through words. None of the children's abilities exceeded these standards, indicating that the participants were unlikely to express pain verbally. The origins of the included children's cognitive impairments were diverse.

Table 1. Origin of the children's cognitive impairment (n=52)

Origin of cognitive impairment	N	% of total N
Congenital/ metabolic anomalies:		
- Syndrome:	12	23.1
- - *chromosomal disorder*	*2*	
- - *identified metabolic disorder*	*6*	
- - *without identified cause*	*4*	
- Major structural cerebral abnormalities	12	23.1
- Infections	1	1.9
- Metachromatic leukodystrofy	3	5.8
- Intra-uterine asphyxia	1	1.9
Total	**29**	**55.8**
Cognitive impairment during birth:		
- Perinatal asphyxia	6	11.5
- Posthypoxic encephalopathy	2	3.8
Total	**8**	**15.3**
Cognitive impairment after birth: (between 2 months and 4 years old)		
- Encephalopathy	2	3.8
- Meningo- or pneumococcal sepsis	4	7.7
- Near-drowning-syndrome	1	1.9
Total	**7**	**13.4**
Origin remained unknown		
Total	**8**	**15.3**

Physiological impairments and chronic physical diseases were frequently documented (Table 2). All children underwent one or more surgical procedures under general anaesthesia during their lifetime up till the present study (mean=4, range=1-7).

Table 2. Chronic physical disorders in children with cognitive impairment

Type of disorder	N	%
Epilepsy	41	78.8
Fed by a nasogastric or gastrostomy tube	37	71.2
Poor vision or blindness	27	52.8
Gastroesophageal reflux	24	46.2
Bronchial and pulmonary infections	23	44.2
Spastic quadriplegia	22	42.3
Constipation	17	32.7
Malnutrition	16	30.8
Scoliosis	12	23.1
ENT(ear, nose, throat) infections	11	22.0
Hearing loss or deafness	9	17.5
Cardiac disease	8	15.4
Anaemia	6	11.5
Urinary tract infection	6	11.5
Inadequate temperature regulation: Hyper- and hypothermia	4	7.7
Asthma, COPD	4	7.7
Eczema	4	7.7

*Note: more than 1 disorder per child possible (mean = 4 disorders per child, range 1-6).

Measures

Item pool

The original item pool consisted of a list of behaviors that were considered to indicate post-surgical pain in children with PCI. A qualitative study was conducted to gain insight into non-verbal pain expressions in children with PCI. First, focussed semi-structured interviews were conducted (n=29) with parents and caregivers (all working at 3 different institutions for residential care). Second, researchers observed behavior of children and adolescents with PCI during several possibly painful events (n= 32). This resulted in 169 possible pain indicators. Adding pain expressions mentioned in the literature resulted in an item pool including a total of 209 indicators [van Dongen et al, submitted]. These indicators were used in a pilot study on 15 children with PCI who underwent various surgical interventions. Expressions that were not suitable for clinical observation of pain were discarded. For example, all expressions relating to 'activities of daily life' were removed from the list since the pilot-study showed that these indicators were related to a specific context and were mostly not present during standardized observation moments. All remaining indicators were included in the initial item pool, making a total of 138 indicators that can be organized into 6 behavioral categories: facial expression, motor behavior, social behavior/mood, attitude towards sore body part, vocalisation and physiological (Table 3).

Table 3. Initial item pool which was tested on pain sensitivity (138 indicators) with Kappa values

Facial expression		**65.** Moves foot (feet)	.82
1. Tense face***	.72	**66.** Moves leg(s) (excl. stretching, trembling)	.86
2. Facial restlessness, tics***	.52	**67.** Moves trunk, back (excl. stretching)	.71
3. Grimace***	.68	**68.** Restless head movements	.68
4. Yawning	.90	**69.** Turning head away*	.54
5. Trembling lips*	.58	**70.** Moves head (excl. Stretching neck)	.91
6. Pout*	.61	**Social behavior/Mood**	
7. Pursed lips	.50	**71.** Restlessness, being unquiet**	.64
8. Lips pressed together	.49	**72.** Active, lively	.81
9. Lip biting	.68	**73.** Irritable	.51
10. Lip curling	.54	**74.** Panics, panic attack***	.65
11. Corners of mouth downward***	.63	**75.** Agitated	.48
12. Warped mouth	.49	**76.** Uncooperative	.51
13. Mouth wide open ('O' shape)	.51	**77.** Sleepy, drowsy	.56
14. Tense tongue	.60	**78.** Dozing, be half asleep	.61
15. Protruded tongue	.71	**79.** Listless, apathetic	.63
16. Jaws clamped together	.48	**80.** Quiet	.78
17. Suddenly eyes wide open	.51	**81.** No interaction	.69
18. Wide-eyed	.62	**82.** Inconsolable	.64
19. Eyes almost closed**	.61	**83.** Refuses physical contact	.71
20. Eyes squeezed***	.80	**84.** Accepting comfort*	.68
21. Continuous eye blinking	.71	**85.** Searching for comfort	.70
22. Rolling eyes	.53	**86.** Resistant, rebellious***	.79
23. Looking sad, almost in tears**	.55	**87.** Anger	.49
24. Frightened, fearful look***	.63	**88.** Hitting, pinching, scratching, biting	.71
25. Dejected, serious look	.39	**89.** Destructive behavior (to material)	.69
26. Angry look	.58	**90.** Head banging	.61
27. Pull up nose	.60	**91.** Bites oneself	.71
28. Deeper naso-labial furrow**	.56	**92.** Hits oneself	.63
29. Moving nostrils	.68	**93.** Pulls hair	.81
30. Chin pulled up	.71	**94.** Hyperactive behavior	.59
31. Trembling chin*	.78	**95.** Stereotypic, repetitive behavior	.78
32. Raises eyebrows	.63	**Attitude towards sore body part**	
33. Frown eyebrows	.78	**96.** Protects sore body part***	.65
Motor behavior		**97.** Turns away from sore body part**	.68
34. Motor restlessnes	.59	**98.** Points at sore body part	-
35. Striking movements	.66	**99.** Withdraws sore body part	.70
36. Involuntary movements	.72	**100.** Holding on to sore body part	.80
37. Shaking movements	.78	**101.** Rubs sore body part	.71
38. Poverty of motion	.81	**102.** Hits sore body part	-
39. Motionless	.88	**103.** Scratches sore body part	.61
40. Huddling oneself	.70	**104.** Pulls sore body part	-
41. Changing position without help	.81	**105.** Pulls clothes around sore body part	-
42. Rock to and fro	.82	**106.** Takes clothes, shoes off	-
43. Waving hands	.68	**Vocalisation**	
44. Other stereotypical movements	.64	**107.** Cries hard, loudly**	.74
45. Tensed up***	.79	**108.** Cries softly***	.50
46. Stiff, rigid	.60	**109.** Moaning, groaning***	.65
47. Limp	.62	**110.** Screaming, yelling	.68
48. Trembling arm, hand	.70	**111.** Guttural, throaty sounds	.51
49. Trembling leg, foot	.72	**112.** Babbling	.53
50. Stretched trunk and back	.62	**113.** Penetrating sounds of restlessness***	.54
51. Stretched arm	.69	**114.** Stereotypic sounds	.61
52. Stretched fingers	.69	**115.** Growl, snarl	.52
53. Stretched leg	.54	**116.** Angry sounds	-
54. Stretched toes	.49	**117.** Grinds teeth	.71
55. Stretched neck, head backwards*	.69	**118.** Blowing	.61
56. Puts fingers in mouth	.78	**119.** Sucking	.52
57. Puts fingers in ear(s)	.88	**120.** Chewing, smacking lips	.69
58. Clenched fists	.74	**121.** Coughing	.90
59. Rubbing one's face	.80	**122.** Paradoxical laughter	.49
60. Raises shoulders	.61	**123.** Being quiet, not babbling	.54
61. Moves hand	.78	**Physiological signs**	
62. Moves arm(s) (excl. stretching, trembling)	.86	**124.** Breath holding, faltering respiration**	.87
63. Bending toes	.48	**125.** Sharp, brief respiration	.63
64. Knees pulled up**	.80	**126.** Fast respiration	.61

Table 3. (continued)

127. Gasping	.56
128. Shocking respiration***	.77
129. Snoring respiration	.61
130. Perspiration	.51
131. Shivering	.49
132. Looks ashen	.49
133. Looks pale	.48
134. Looks red, turns red***	.66
135. Cyanotic (blue lips)	.41
136. Blotched body	.61
137. Tears***	.93
138. Ruminating	-

Note:30 indicators are selected to be sensitive to pain using levels of significance (Kruskal-Wallis test): *p< .05; **p<.01; *** p<.001 Kappa: Interrater reliability per indicator, never scored behaviors are indicated with: ' - '

Subsequently, the item pool was tested on its sensitivity to post-surgical pain. Seven episodes for each recruited child included video recordings of the whole body (2 min.), face and trunk (2 min.) and face (2 min.) yielding a total of 333 recorded episodes (52 children x 7 episodes = 364 minus 31 missing episodes).

All recorded episodes were scored by 2 researchers on the 138 indicators, using a five point scale ranging from 0 (never shown) to 4 (always shown). The first 30 video recordings were independently scored by two researchers to assess reliability. Average kappa across the 138 indicators was 0.67 (range 0.39-0.91), indicating substantial agreement (Landis & Koch 1977). Table 3 shows kappa's for each separate indicator. Both pre- and post-surgical episodes were scored using all 138 indicators. Scoring disagreements were resolved by discussion.

Visual Analogue Scale (VAS)

Parents and researchers gave an overall impression of the severity of children's pain during the observed episodes on the Visual Analogue Scale (VAS), a 10 cm line with the anchors 0 (no pain) and 10 (worst pain). Parents rated 211 of the episodes and researchers 233. VAS ratings ≥ 4 were considered clear indications of pain (Buchholtz et al. 1998; Van Dijk et al. 2002). A kappa of 0.51 between parents' and researchers' ratings indicated moderate agreement (Landis and Koch 1977).

Procedure

The study was approved by the Medical Ethical Committee of the Erasmus University Academic Hospital. Parents were asked for informed consent when their child was scheduled for elective surgery in the hospital. Information on medical diagnoses associated with their cognitive impairment level was obtained from their medical record. Information about speech behavior of the children was gathered by (1) sending a questionnaire to their parents, (2) using written descriptions of the child's verbal capacities in medical records and (3) through

interpretation of the video recordings by the investigators.

Six-minute video recordings were made in order to observe the child's behavior at 7 standardized episodes (Figure 1, p.17), of which 2 pre-surgical: 1 episode at home within one month before surgery when the child was not in pain or distress (T1) and 1 episode when the child was already hospitalized but before surgery (T2). Five post-surgical episodes were videotaped: at 30 minutes (T3) and 3 (T4), 24 (T5) and 48 hours (T6) after extubation, and one during an intervention specific moment inducing pain and depending on the type of surgery (T7). After each episode pain intensity was scored on the VAS by the researchers and the parents (when available). Video recordings of pre-surgical episodes were only included if the child was not acutely ill or in distress and VAS scores as rated by both researchers and parents were 0.

Painfulness of procedures

In this study participants underwent different types of surgery. These were rated on painfulness (low (1), moderate (2) or high (3)) by 11 medical specialists all working fulltime in the children's hospital with an average number of 18.0 years of experience with the included procedures (1 surgeon, 2 ICU intensivists , 1 gastro-enterologist, 1 orthopedic surgeon, 1 ENT specialist, 1 ophthalmologist, 1 urologist, 1 plastic surgeon and 2 anesthesiologists) (Table 4). The surgical procedure was assigned to one of the three categories of painfulness based on the average ratings from the medical specialists, with priority given to the rating of the executive specialist (e.g., the rating of the orthopedic surgeon counts double regarding 'adductor tenotomy' in comparison with ratings from the other specialists whose mean ratings count only once). All intra-operative anaesthetic and analgesic medication was standardized for each type of surgery and was documented in anaesthetic and surgical records. Post-surgical medications and their doses were also registered (Table 4).

Table 4. Type of surgery and corresponding pain intensity

Type of surgery	N	Pain intensity	Standard post-surgical analgesic regimes
Gastro intestinal tract:	***27(51.9%)***		
Gastroscopy, ph-monitoring	7	1	
Gastroscopy, Percutaneous Endoscopic Gastrostomy (PEG)	5	1	Acetominophen 90-100 mg./kg. per day
Gastroscopy, gastrostomy	2	2	
Gastroscopy, PEG-placement	7	2	
Nissenfundoplication laparoscopic	3	3	Morfine 0.1 mg./kg. loading dose, then 0.01- 0.04 mg./kg./hr. &
Nissenfundoplication aparoscopic, oesophagoscopy, PEG-placement	2	3	Diclophenac 1-2 mg./kg. 3dd (up till 1 day after stopping morfine)
Navel rupture correction	1	2	Diclophenac 1-2 mg./kg. 3 times a day & Morfine 0.1 mg./kg. loading dose, then 0.01- 0.04 mg./kg./hr.
ENT-surgery:	***9 (17.3%)***		
Adenotonsillectomy (ATE)	5	3	Acetominophen 90-100 mg./kg. per day
Middle ear tubes	2	1	Acetominophen 60 mg./kg. per day
Recanulation tracheostomy	1	2	Diclophenac 1-2 mg./kg. 3 times a day
Placement endonasal drains	1	2	Acetominophen 90-100 mg./kg. per day
Dentistry:	***2 (3.8%)***		
Dental treatment/teeth extractions	2	2	Acetominophen 90-100 mg./kg. per day
Plastic surgery:	***1 (1.9%)***		
Amputation extra finger	1	2	Acetominophen 90-100 mg./kg. per day
Urology:	***1 (1.9%)***		
Cystoscopy, bilateral ureteral reimplantation	1	2	Acetominophen 90-100 mg./kg. per day
Orthopedics:	***11(21.2%)***		
Adductor tenotomy	7	2	Diclophenac 1-2 mg./kg. 3 times a day & Diazepam 0.15 mg./kg. 3 dd
Correction pes equinovarus	1	3	Morfine 0.1 mg./kg. loading dose, then
Scoliosis correction	3	3	0.01- 0.04 mg./kg./hr. & Diclophenac 1-2 mg./kg. 3 times a day (up till 1 day after stopping morfine)
Ophthalmology:	***1 (1.9%)***		
Enucleation of the eye	1	1	Retrobulbair block (appr. 6hrs pain relieving) Acetominophen 60 mg./kg. per day

Analyses

Data was analyzed using SPSS 10.1. Because of the highly skewed distribution of the data non-parametric tests were used; alpha was set at .05 for all tests.

First, Kruskal-Wallis tests were used to test any difference in occurrence of the 138 indicators across 7 measurements, using graphs to visualize at which measurements the indicators differed.

Second, Wilcoxon signed rank tests (using Fisher's exact test for significance) were used to analyse differences in aggregated scores on the remaining indicators before and after surgery in order to test the sensitivity of these indicators to differences likely (post-surgical) and unlikely (pre-surgical) to induce pain. Spearman's correlations have been calculated between all time points to measure how the indicators are related over time.

Third, item scores were analyzed with the Friedman test for differences between measurements within both episodes separately (before and after surgery) in order to find changes in post-surgical behavior over time.

Fourth, using the Friedman test we tested the significance of differences between scores during the 4 general post-surgical measurements and the one, surgery specific measurement.

Fifth, the sensitivity of indicators to pain intensity related to different types of surgical procedures was analysed using univariate analyses of variance with Bonferroni's Post-hoc test. Finally, the latter analyses were used to test differences in scores on indicators during post-surgical measurements at which VAS ratings were ≥ 4 versus 0-3.

Results

Pain sensitivity

Using any significant difference in occurrence across the 7 measurements as a criterion to keep the indicator included, the item pool was reduced from 138 to 30 indicators (see Table 3). Some indicators that showed significant differences were obviously not attributable to pain (e.g., motor behaviors that were absent directly after surgery (T3) due to anesthetics, but nearly equally shown before surgery and at all time points after surgery other than T3) and were excluded.

Table 5 shows the percentages of children who scored on each indicator of the item pool at least 1 ('sometimes shown') on each of the seven pre- and post-surgical measurements. Seven indicators which turned out to be not significantly different before and after surgery were eliminated from the item pool. Of the 30 indicators 23 showed sensitivity to the pre- and post-surgical situations, most probably related to the absence versus presence of pain. Four behaviors were not shown before surgery but appeared only after surgery ('trembling chin', 'protects sore body part', 'cries hard/loudly', 'breath holding') while 19 behaviors were present before surgery but appeared with increased frequency or intensity after surgery.

Changes in pain over time

Of the indicators that showed significant differences in occurrence between pre- and post-surgical measurements, eight differed significantly among the first 4 post-surgical measurements and thus are most likely to be sensitive to changes in post-surgical pain over time (Table 5). There were no significant differences in appearance of indicators between the two pre-surgical measurements, indicating that potential significant differences in pain indicators within the post-surgical episodes are unlikely to be attributable to coincidental fluctuations.

Post-surgical pain versus directly induced pain

Thirteen indicators were scored significantly different on the post-surgical measurements versus the surgery specific measurement (Table 5). These indicators were all shown less frequently during the post-surgical episodes compared to the direct pain inducing episode, indicating that these are most likely sensitive to differences in post-surgical versus acute pain.

Table 5. Percentage of children (n=52) for whom was scored at least 1 ('sometimes shown') during pre- and post-surgical episodes at each of the 30 indicators

	Pre-surgical			Post-surgical						Analyses		
Indicator	T1	T2	Mean Pre-surgical	T3	T4	T5	T6	T7	Mean Post-surgical	Test 1. Pre- versus post-surgical[c]	Test 2. T3,4,5,6[d]	Test 3. T3,4,5,6 versus T7[c]
Facial expression												
Tense face	8	18.4	**13.1**	57.1[a]	44	38	30.4	94.9	**51.3**	.000*	.004*	.000*
Facial restlessness, tics	12	18.4	**15.2**	26.5[a b]	22[a b]	26[a b]	13	64.1	**29.1**	.001*	.016*	.000*
Grimace	2	6.1	**4**	36.7[a]	26	18	13	61.5	**29.9**	.000*	.005*	.000*
Trembling lips	0	0	**0**	6.1	2	0	0	7.7	**3**	.063	-	-
Pout	2	4.1	**3**	4.1	8	8	8.7	10.3	**7.7**	.078	-	-
Corners of mouth downward	20	16.3	**18.2**	55.1[a b]	54[a b]	38[a b]	37	59	**48.3**	.000*	.026*	.203
Eyes almost closed	24	28.6	**26.3**	49	52	48	50	51	**50**	.000*	.048*	.695
Eyes squeezed	26	12.1	**19.1**	59.2[a]	40.0	44.0	43.4	64.0	**50.1**	.000*	.002*	.249
Looking sad, almost in tears	2	8.2	**5.1**	26.5	18	20	13	30.8	**21.4**	.000*	.456	.027*
Frightened, fearful look	6	6.3	**6.1**	30.6[a]	14	18	8.7	59	**24.8**	.000*	.006*	.000*
Deeper naso-labial furrow	30	24.5	**27.3**	44.9	32	38	30.4	59	**40.2**	.000*	.305	.001*
Trembling chin	0	0	**0**	12.2	6	2	2.2	7.7	**6**	.001*	.137	.789
Motor behavior												
Tensed up	38	30.6	**34.3**	36.7	28	50	43.5	76.9	**45.7**	.000*	.167	.000*
Stretched neck, head backwards	30	28.6	**29.3**	16.3	22	24	30.4	30.8	**24.4**	.706	-	-
Knees pulled up	28	34.7	**31.3**	8.2	28	36	34.8	46.2	**29.9**	.657	-	-
Turning head away	40	46.9	**43.4**	30.6	28	40	41.3	53.8	**38**	.151	-	-
Social behavior/Mood												
Restlessness, being unquiet	34	49	**41.4**	28.6	36	36	37	43.6	**35.9**	.674	-	-
Panics, panic attack	0	2.1	**1**	24.5	16	8	8.7	28.2	**16.7**	.000*	.338	.103
Accepting comfort	0	6.1	**3**	14.3	10	6	4.3	17.9	**10.3**	.010*	.215	.529
Resistant, rebellious	4	2	**3**	6.1	4	8	4.3	28.2	**9.4**	.001*	.703	.002*

Table 5 continues

Table 5. continued

	Pre-surgical			Post-surgical						Analyses		
Indicator	T1	T2	Mean Pre-surgical	T3	T4	T5	T6	T7	Mean Post-surgical	Test 1. Pre- versus post-surgical[c]	Test 2. T3,4,5,6[d]	Test 3. T3,4,5,6 versus T7[c]
Protects sore body part	0	0	**0**	2	0	2	0	15.4	**3.4**	.031*	1.00	.031*
Turns away from sore body part	0	0	**0**	2	0	4	0	10.3	**3**	.125	-	-
Vocalisation												
Cries hard, loudly	0	0	**0**	14.3	8	6	6.5	21.1	**10.7**	.000*	.178	.152
Cries softly	2	4.1	**3**	22.4	12	14	6.5	42.1	**18.5**	.000*	.083	.002*
Moaning, groaning	6	16.7	**11.2**	36.7	28	38	24.4	69.2	**38.2**	.000*	.444	.000*
Penetrating sounds of restlessness	0	6.1	**3**	14.3	14	16	4.4	43.6	**17.6**	.000*	.275	.002*
Physiological signs												
Breath holding, faltering respiration	0	0	**0**	12.2	4	4	2.2	12.8	**6.9**	.000*	.150	.324
Shocking respiration	2	0	**1**	10.2	8	6	4.4	25.6	**10.3**	.000*	.400	.015*
Looks red, turns red	2	10.2	**6.1**	28.6[a]	10	6	2.2	23.1	**13.7**	.000*	.000*	.812
Tears	2	0	**1**	10.2	4	4	2.2	17.9	**7.3**	.000*	.072	.089

[a b] = significant differences between post-surgical episodes
[c] = Wilcoxon test, [d] = Friedman test

Painfulness surgical procedures

Means and SD of scores on each of the 23 indicators on the 4 post-surgical measurements categorized by painfulness of the procedure (low, moderate or highly painful) are given in Table 6. Although most indicators showed increasing mean scores with increasing painfulness of the procedure, only 2 indicators, 'eyes squeezed' and 'trembling chin' were significantly different ($p<.000$) between moderately and highly painful and between low and highly painful. Effect size for 'eyes squeezed' was large (>13.8) and for 'trembling chin' medium (> 5.8) (Cohen 1988).

Table 6. Mean (SD) of scores over the four regular post-surgical measurements (T3, T4, T5, T6) by painfulness of the procedure

	Painfulness		
Indicator	Low (n=16)	Moderate (n=24)	High (n=13)
Facial expression			
Tense face	2.25 (2.67)	3.92 (3.32)	4.50 (4.42)
Facial restless-ness, tics	.94 (1.77)	2.00 (3.08)	2.33 (2.61)
Grimace	1.19 (1.17)	1.71 (2.29)	2.17 (2.17)
Corners of mouth downward	2.75 (3.80)	5.25 (4.92)	4.75 (3.65)
Eyes almost closed	3.38 (2.25)	3.46 (3.11)	3.42 (2.61)
Eyes squeezed	4.31 (3.40)	2.25 (2.36)	5.75 (3.62)
Looking sad, almost in tears	1.44 (2.10)	1.38 (1.84)	1.33 (1.61)
Frightened, fearful look	.50 (1.15)	1.00 (1.18)	1.42 (1.73)
Deeper naso-labial furrow	1.75 (1.44)	2.13 (2.31)	2.83 (2.66)
Trembling chin	.06 (.25)	.29 (.75)	.92 (1.51)
Motor behavior			
Tensed up	2.19 (1.42)	2.42 (2.65)	2.00 (1.95)
Social behavior/Mood			
Panics, panic attack	.38 (.89)	.96 (1.46)	1.08 (1.56)
Accepting comfort	1.37 (2.73)	.67 (1.49)	.58 (1.24)
Resistant, rebellious	.38 (.72)	.13 (.45)	.58 (1.73)
Attitude towards sore body part			
Protects sore body part	.13 (.50)	.00 (.00)	.17 (.57)
Vocalisation			
Cries hard, loudly	.19 (.54)	.54 (1.22)	.58 (1.00)
Cries softly	.44 (.81)	1.29 (2.56)	1.50 (2.07)
Moaning, groaning	1.69 (1.66)	2.17 (2.18)	2.50 (1.78)
Penetrating sounds of restlessness	.44 (.81)	1.04 (1.43)	1.08 (1.93)
Physiological signs			
Breath holding, faltering respiration	.00 (.00)	.37 (.92)	.67 (1.07)
Shocking respiration	.25 (.77)	.50 (1.32)	.50 (1.00)
Looks red, turns red	.94 (1.84)	1.17 (1.88)	2.42 (3.18)
Tears	.25 (.68)	.33 (.82)	.50 (1.00)

Relation VAS ratings and indicator scores

The VAS ratings, used to give an overall impression of the child's pain during the observed episodes did not significantly differ between researchers and parents ($t = -1.45$; $p > .05$), although on average parents' ratings were slightly (i.e., 0.1 to 0.5 cm.) higher than ratings by researchers (Table 7). Parents rated twenty-seven (12.7%) out of 211 episodes as painful episodes (VAS ratings ≥ 4) and researchers 30 out of 333 episodes (9%). Based on the researchers' ratings, the analyses showed that all indicators - with the exception of 'eyes almost closed' - were scored significantly higher on episodes with VAS ratings ≥ 4 compared to the episodes with VAS ratings below 4. When comparing scores on the 23 indicators with VAS ratings below 4 and ≥ 4 for events (surgery specific episode) 10 indicators turned out to be significantly different: 'tense face', 'grimace', eyes squeezed', 'deeper naso-labial furrow', 'panics, panick attack', 'cries softly', 'moaning, groaning', 'penetrating sounds of restlessness' and 'tears' (Tabel 7).

Table 7. Mean scores (SD) on 23 indicators for all episodes and for surgery specific episode by VAS rating 0-3.9 (0) versus ≥ 4 (1)

Indicator	VAS	Mean (SD) all 7 episodes	Mean (SD) T7 = surgery specific episode
Facial expression			
Tense face	0	.65 (1.05)**	1.76 (1.12)*
	1	2.87 (.90)	2.70 (.82)
Facial restlessness, Tics	0	.41 (.87)**	1.24 (1.18)
	1	1.30 (1.46)	1.10 (1.45)
Grimace	0	.25 (.63)**	.86 (.99)*
	1	2.03 (1.24)	2.10 (.99)
Corners of mouth downward	0	.91 (1.36)**	1.52 (1.57)
	1	1.53 (1.30)	1.60 (1.35)
Eyes almost closed	0	.74 (1.01)	.79 (.86)
	1	.93 (1.04)	.90 (1.10)
Eyes squeezed	0	.70 (1.13)**	1.00 (1.16)*
	1	1.93 (1.33)	2.00 (1.15)
Looking sad, almost in tears	0	.19 (.59)**	.34 (86)*
	1	1.70 (1.51)	1.80 (1.48)
Frightened, fearful look	0	.23 (.58)**	1.07 (1.03)
	1	.83 (.98)	.80 (1.03)
Deeper naso-labial furrow	0	.45 (.73)**	.90 (1.05)*
	1	1.97 (1.09)	2.10 (.99)
Trembling chin	0	.03 (.23)**	.14 (.52)
	1	.53 (.93)	.20 (.62)
Motor behavior			
Tensed up	0	.54 (.75)**	1.24 (.91)
	1	1.50 (1.19)	1.70 (1.25)
Social behavior/Mood			
Panics, panic attack	0	.10 (.42)**	.28 (.65)*
	1	1.10 (1.21)	1.30 (1.57)
Accepting comfort	0	.16 (.67)*	.34 (.81)
	1	.50 (1.07)	.40 (.97)
Resistant, rebellious	0	.08 (.36)**	.34 (.72)
	1	.43 (.89)	.90 (1.29)

Table 7. Continued

Indicator	VAS	Mean (SD) all 7 episodes	Mean (SD) T7 = surgery specific episode
Attitude towards sore body part			
Protects sore body part	0	.03 (.29)**	.28 (.84)
	1	.17 (.53)	.30 (.67)
Vocalisation			
Cries hard, loudly	0	.05 (.29)*	.25 (.70)
	1	.70 (1.02)	.70 (1.06)
Cries softly	0	.15 (.54)**	.61 (1.07)*
	1	1.57 (1.30)	1.50 (1.08)
Moaning, groaning	0	.40 (.79)**	1.14 (1.19)*
	1	2.07 (1.20)	2.30 (.95)
Penetrating sounds of restlessness	0	.13 (.46)*	.55 (.91)*
	1	1.33 (1.39)	1.50 (1.27)
Physiological signs			
Breath holding, faltering respiration	0	.03 (.20)**	.14 (.52)
	1	.60 (1.07)	.60 (1.07)
Shocking respiration	0	.06 (.31)**	.34 (.77)
	1	.73 (1.08)	.70 (1.06)
Looks red, turns red	0	.20 (.78)*	.34 (.86)
	1	1.30 (1.55)	.90 (1.20)
Tears	0	.03 (.25)*	.10 (.56)*
	1	.77 (1.04)	1.10 (1.10)

*= $p < .05$ **= $p<.000$

Discussion

Pain sensitivity

This study showed that 23 reliable observable behaviors are sensitive to post-surgical pain in children with profound cognitive impairment. Not surprisingly, most sensitive indicators were from the category 'facial expression' (47.8%). Because most children (84.6%) suffered from motor impairments they were hardly able to express themselves by other meaningful motor behaviors than facial expressions. This finding supports the view of other researchers who reported the importance of facial response as an indicator of pain in these children (Craig 1992; Stallard et al. 2002b). Other investigators used standardized, objective facial coding systems (Grunau et al. 1987; Peters et al. 2002) that provide objective behavioral descriptions of the items. However, the applicability of these systems in children with PCI has not yet been addressed. In addition, this study shows that several other behaviors not included in these systems may be used as reliable indicators of pain in children with PCI.

Within the category 'Guarding the sore body part' only 1 indicator ('protects the body part') appeared to be sensitive to post-surgical pain. Most children were not able to show behaviors within this category because of a lack of motor capacities. Further, four physiological indicators differed significantly before and after surgery, two of which indicated changes in respiration which were corrected for changes as a direct consequence of anesthesia. The indicator 'panics, panic attack' was the most significant indicator from the category 'Social behavior/Mood', probably because pain often goes together with anxiety. It is often difficult to prepare these children for what exactly is going to happen to them when they get admitted to the hospital. Table 5 shows that 2,1% of the children did show frantic behavior when they were admitted to the hospital (before surgery), which probably indicates distress or fear as a result of changes in environment or perhaps remembrance of preceding admissions to the hospital. Panic was observed most often directly after surgery when they awoke from anesthesia (T3) and in the directly pain inducing episode (T7). The same goes for resistant, rebellious behavior which often was shown together with frantic behavior, that is if the child was motorically capable of showing resistance. Indicators within the category 'Social behavior/Mood' might be interpreted as identifiers of anxious, angry or sad children rather than children in pain. Although they appeared at higher rates during the post- versus pre-surgical period, most of these (except 'panicking') appeared to be not very sensitive to potential changes in pain during the post-surgical period or to pain indicated by high VAS scores. The same applies to most physiological signs. Finally, 4 indicators were categorized within 'vocalisation'.

Most indicators were present before as well as after surgery. For example, 'moaning, groaning' was shown before surgery (11.2%) and after surgery (38.2%), indicating that for some children this might be customary behavior because they showed it at comparable rates before and after surgery while for others it seemed an expression of pain because they displayed this behavior more frequent after surgery than before. This indicates the importance of knowledge about customary behavior of the child when he or she is not in pain.

Changes in pain over time

The 8 indicators that turned out to be significantly different within the 4 post-surgical episodes (tense face, facial restlessness, grimace, corner of mouth down, eyes almost closed, eyes squeezed, frightened fearful look, looks or turns red) are most probably sensitive to changes in pain experience over time once pain is present. Ranks of the Friedman test showed that all 8 indicators displayed declining scores over the four general post-surgical measurements. Apparently, these behaviors are most frequently shown when the child suffers from pain right after surgery and indicates that the pain decreases when the child gradually recovers from surgery. Therefore, these indicators may be considered the ones that are not only sensitive to pre- and post-surgical conditions, but also to more subtle changes within the post-surgical episode.

Post-surgical pain versus directly induced pain

Although post-surgical pain is generally a type of acute pain (AAOP and APS 2001; Groenman et al. 1986), a difference can be made between post-surgical pain observed over time (T3-T6) and surgery specific short, intense pain, like treatment of surgical injuries or removal of plaster cast (T7). Thirteen indicators turned out be significantly different between the 4 post-surgical episodes over time and T7. The higher scores at T7 indicate that during this episode a relatively high pain intensity is induced. These items may be regarded to be sensitive to differences between post-surgical pain and more direct acute pain.

Painfulness surgical procedures

Only two indicators turned out to be sensitive to differences in procedures, characterized as implying low, moderate and/or high pain intensity. Therefore, it can be concluded that at least 21 out of 23 indicators are in general suitable for assessing post-surgical pain, regardless of the type of pain inducing surgical procedure.

Relation VAS ratings and indicator scores

In pediatric literature pain a VAS score of 4 or higher is considered to distinguish pain from no-pain conditions (Buchholz et al. 1998) and extra pain medication was proposed after surgery when the score was 4 or higher (Van Dijk et al. 2002). In this study we interpreted the VAS ratings from this perspective. The finding that all indicators, except for one, were scored significantly higher in episodes with VAS ratings above or equal to 4, indicates the sensitivity of these indicators concerning absence (VAS rating < 4) versus presence (VAS ratings ≥ 4) of clinically meaningful levels of pain. When comparing mean item scores only from surgery specific episodes (T7), ten indicators proved to be significantly different at moments with VAS ratings below 4 versus higher or equal to 4. These 10 indicators are probably pain behaviors, mostly facial and vocal expression, that discriminate best between pain and no pain. These indicators were rated significantly different despite a strong reduction in statistical power (only T7 observations were included in this comparison). However, the probability that directly induced surgery-specific pain (T7) might be expressed differently than acute pain lasting several days (post-surgical pain), must be taken into account.

Comparison with other studies

In several recent studies, authors based their research on the concept that children with severe or profound cognitive impairment process information and communicate distress in a different manner than normally developed children and therefore need their own measure to detect pain (Gilbert-MacLeod et al. 2000; Breau et al. 2001, 2002a; Collignon & Giusiano 2001; Hunt 2001). Although in this article we did not directly compare children with PCI with normally developed children, we studied this group of children from this perspective. Our

results are partially consistent with the findings of other studies on this topic. The six pain indicators described as 'core pain cues' by Stallard (crying with or without tears; screaming, yelling, groaning or moaning; screwed up or distress looking face; body appears stiff or tense; difficult to comfort or console; flinches or moves away if touched) (Stallard et al. 2002b) were also found to be sensitive to pain in our study. We also compared our results with those of the study from Breau et al. who provided an observational scale, the Non-Communicating Children's Pain Checklist-Postoperative version (NCCPC-PV), that showed good psychometric properties for this particular group of children (Breau et al. 2002a, 2002b). Twelve out of 27 items included in the NCCPC-PV were also found to be significantly different between our pre- and post-surgical episodes. The fact that 15 significant indicators in their study turned out to be non-significant in this study can partially be explained by a difference in study-designs: the number of observational episodes in the study from Breau et al. counts two (once 30 minutes before surgery and once 30 to 60 minutes after the child had left the recovery room), while in this study children were observed twice before surgery and five times after surgery. Eight out of these 15 indicators turned out to be non-significant in appearance over time in our study because they were shown (nearly) equally over all episodes before and after surgery. The remaining seven of these 15 indicators were also significantly different in our study but not attributable to pain. For example, indicators like 'difficult to distract' and 'less interaction' were scored 'always shown' directly after surgery (T3) which provided the significant difference, due to anaesthetics.

Heterogeneous group

Individuals with profound cognitive impairment form an extremely heterogeneous group in terms of behavioral repertoire and functional abilities. Our results emphasize the heterogeneity of this group of children, which is consistent with the findings of several investigators (Oberlander et al. 1999a; Collignon & Giusiano 2001; Breau et al. 2002a; Stallard et al. 2002a). Although all children in this study were non-verbal, they differed in origin of cognitive impairment and a wide range of capacities and impairments within several developmental areas was documented. It should be emphasized that this heterogeneous group of children described in this study is an average representation of children with PCI in general, which makes it possible to work on the development of a generic observational pain scale, not one based on specific diseases or syndromes within this population.

Generic pain indicators

These findings lead us to suggest that, although the specific expression of pain may be very individual, there appears to be a set of shared generic indicators. The identification of this item pool of observable pain behaviors enables the development of an observational pain assessment scale that can easily be used in daily practice to assess and monitor pain within this particular group of children.

Future research

Further research is required to address several issues. First, it is quite possible that post-surgical medication has affected pain expressions. This might explain the relatively low number of episodes with VAS-ratings higher or equal to 4. Therefore, in future studies the impact of anaesthetic and analgesic medication on the occurrence of indicators should be studied by observing children before and after the administration of medication in a standardized way. As in other patient groups, our attitude for pain management is pre-emptive analgesia.

Second, it is important to validate the VAS for indicating pain intensity in children with PCI. Due to the lack of a valid tool, we used the VAS assuming a score above 4 reflects pain, which in fact is set as the cut-off point for normally developed children (Van Dijk et al. 2002).

Third, it needs to be studied whether these indicators allow additive scaling to develop a standardized instrument that is sensitive to individual and time-related differences in pain expression. Finally, new observations directly at the bedside are needed to examine the clinical usefulness of the resulting instrument.

From this study it can be concluded that at least 21 out of the 23 indicators as described in Table 6 are in general suitable for assessing post-surgical pain in children with profound cognitive impairment despite the heterogeneity of this group of children.

Chapter 4

Psychometric properties of the Checklist Pain Behavior: evaluating post-surgical pain in children with profound cognitive impairment

Chantal Terstegen
Hans M. Koot
Josien B. de Boer
Dick Tibboel

(Submitted for publication)

Abstract

In this study a pool of pain indicators described in our preceding study was further developed into an observational scale - the Checklist Pain Behavior (CPG) - that can provide assistance to caregivers in assessing and monitoring pain in children with profound cognitive impairment (PCI) in a post-surgical setting. Fifty-two children with PCI (aged 3-19 yr), who were admitted to a level III children's hospital for an elective surgical procedure, were included. The children were video-recorded twice pre-surgical and five times post-surgical. Directly after each observational episode, researchers and parents (if present) scored pain intensity using the Visual Analogue Scale (VAS). All video-recordings (n=333) were scored using the CPG and the VAS. Findings from the PRINCALS analyses showed a highly satisfactory fit to the data of a four-dimension solution, containing 'Facial expression' (9 items), 'Social behavior/Mood' (4 items), 'Vocalisation' (4 items) and 'Physiological signs' (4 items). This CPG showed excellent internal consistency (Cronbach alpha .89) and interrater reliability (ICC .97). Interscale correlations were moderate to high for the resulting subscales and the total scale. Mean scores on the CPG subscales and the total scale differed significantly before and after surgery. Regression analyses showed that all subscales, except 'Social behavior/Mood', contributed independently to the prediction of VAS scores. These results provide evidence that post-surgical pain in children with PCI can be reliably assessed and monitored using the Checklist Pain Behavior (CPG). However, it is a generic pain assessment tool and in view of the fact that these children display differences in type of pain behavior, one should be aware of the child's customary behavior before observing their post-surgical behavior. We recommend, therefore, to first apply the CPG at a moment before surgery when the child is not in pain, and to use the obtained score as a reference score.

Introduction

In the last decade research on pediatric pain has increased substantially. Only until recently, studies became also focussed on a special group of children, i.e. those with severe or profound cognitive impairment (Breau et al. 2002a, 2002b; Stallard et al. 2002a, 2002b; Terstegen et al. 2003). Children with profound cognitive impairment (PCI) probably experience many episodes of pain due to a variety of painful somatic conditions such as gastro-oesophageal reflux, contractures and epilepsy. These children have a higher risk of experiencing pain while their impaired communication skills increase the likelihood that pain will remain unrecognised due to wrongly interpreted observations of behavior and therefore remains untreated (Anand et al. 2000; Stallard et al. 2001; Terstegen et al. 2003). The recognition and assessment of these probable pain experiences is hindered by both the limited expressive repertoire of the child due to a number of cognitive, verbal and motor limitations and uncertainty about the meaning of available behavior as a sign or expression of pain.

Validated instruments to assess pain in these children were not available for a long period of time. Fortunately, two pain rating scales for children with cognitive impairment were developed recently. Breau et al. (2002a) developed a pain rating scale, the Non-Communicating Children's Pain Checklist-Revised (NCCPC-R), which displayed excellent psychometric properties in several types of pain settings (e.g., injury; chronic condition; illness). In addition, for post-surgical pain an adapted version, the NCCPC-PV, was shown to be sensitive to surgery-induced pain (Breau et al. 2002b). Terstegen et al. (2003) developed the Checklist Pain Behavior (CPG) specifically for children with profound cognitive impairment. We had several reasons to develop a separate instrument for children of this level of intellectual functioning. First, due to severe motor limitations, especially in those with profound multiple impairments, children with PCI are not able to display certain pain behaviors that are shown by children with mild, moderate or even severe CI. Second, they may show certain behaviors which are not shown by children without CI that are indicative of the presence of pain. Third, although the NCCPC-R is sensitive to painful versus non-painful conditions, it lacks the potential to monitor changes in post-surgical pain. To be certain not to miss any potential indicator of pain, we derived 206 behaviors from several sources (Terstegen et al. 2003). This way, we made a complete inventory of all possible candidate pain behaviors displayed in this specific group of children. During the study, it became evident that children with PCI who also suffered from profound motor impairments (thus, children with profound multiple impairments) displayed different behavior when in pain compared to children with PCI and mild or severe motor impairments. For example, seven behaviors described in the NCCPC were not included in the CPG because participants did not have the motor capacities to display these behaviors.

For a pain instrument to be useful in a surgical setting, it has to be reliable and sensitive to differences in pre- and post-surgical signs of pain, and sensitive to changes in pain expression during the post-surgical period. In addition, given the limited expressive potential of children with PCI and profound motor impairments, it would be helpful to know which type of expression (e.g., facial, vocal) is most informative. Finally, since no gold-standard for pain assessment in these children exists, it is important to know whether ratings of children's pain are related to independent indicators of the painfulness of the surgical procedure.

We observed children with PCI several times after surgery in order to assess the sensitivity of pain indicators to changes in pain across the post-surgical period, when adequate pain management is critical. In this study the pool of pain indicators described in a preceding study (Terstegen et al. 2003) was further developed into an observational scale - the Checklist Pain Behavior (CPG) - that can provide assistance to caregivers (e.g., nurses, parents, physicians) in assessing pain and help to improve the administration of analgesia in children with PCI in a post-surgical setting.

The first goal of this study was to investigate whether all 23 indicators which were listed in the pool of pain indicators in the preceding study (Terstegen et al. 2003), could be summarized into meaningful and reliable clusters as a basis for specific and overall observational subscales to be included in the CPG.

Secondly, we assessed the reliability of the CPG. To this end, internal consistency, inter-rater agreement and interscale correlations for the resulting subscales and overall scale were determined.

Thirdly, the construct and convergent validity of the CPG were assessed in four steps. The first step studied its power to discriminate between pain versus no pain conditions by comparing CPG scores in pre- and post-surgical episodes and by comparing CPG scores in episodes with VAS ratings lower than 4 (absence of pain) versus episodes with VAS ratings higher than or equal to 4 (presence of pain). The second step studied the potential of the CPG to discriminate between continuous versus acute pain by comparing CPG scores obtained in the post-surgical episodes with those obtained during an acutely painful surgery specific procedure. The third step studied the sensitivity of the CPG to changes in pain behavior over time, by comparing the CPG subscale and total scores across subsequent post-surgical episodes. Finally, the fourth step concerned the sensitivity of the CPG subscale and total scale scores to pain intensity by comparing scores obtained from children admitted to surgical procedures that were assumed to differ in the degree of painfulness (low, moderate or high) as rated by medical professionals.

Instruments and methods

Participants

This cross-sectional study was undertaken to assess post-surgical pain and included 52 children with PCI (27 boys, 25 girls; aged 3-19 yr, mean age 8.0 (SD 4.7), who were admitted to a level III children's hospital for an elective surgical procedure. To be included in the study, children must have lived in their present situation for at least one month, to ensure they did not recently experience major environmental changes. Children with abnormal renal/adrenal function or abnormal liver function were excluded, because of an possibly altered drug metabolism.

Since no diagnostic tools are available which can precisely determine a child's IQ level below 40, children were included if their cognitive development was estimated to be below a calendar age of 2 years old (equivalent IQ 0-20). In addition we assessed language development using 'minimal speech standards' (Goorhuis & Schaerlaekens 2000). Standards for PCI include the ability of making sounds but not expressing oneself through words. None of the children's abilities exceeded these standards, indicating that the participants were unlikely to express pain verbally.

The origins of the included children's cognitive impairments were diverse. Physiological impairments and chronic physical diseases were frequently documented. All children had

previously undergone one or more surgical procedures (mean=4; range=1-7) under general anaesthesia during their lifetime up till the present study (for detailed sample information, see Terstegen et al. 2003).

Instruments

Items in the Checklist Pain Behavior (CPG)

The CPG contains 23 observable behaviors which were selected for further analysis because they were found to be sensitive to post-surgical pain in children with PCI in our preceding study (Terstegen et al. 2003). Using behavioral descriptions these behaviors were rated on a 5-point scale, with 0=never shown, 1=sometimes shown, 2=regularly shown, 3=often shown, 4=always shown. In this study, an initial item pool consisting of 138 potential pain indicators was reduced to 30 indicators, using any significant difference in occurrence across seven pre- and post-surgical measurements as a criterion to keep the indicator included. Twenty-three of these 30 indicators showed significant increases from the pre- to the post-surgical episodes. Eight of these remaining 23 indicators differed significantly among the four post-surgical measurements and thus are most likely to be sensitive to more subtle changes in pain over time. Thirteen indicators were scored significantly less frequently on the four post-surgical measurements compared to the surgery specific measurement, indicating that these behaviors are most likely to be sensitive to differences between post-surgical pain and more direct acute (surgery specific) pain. For further analysis each of the 23 items was assigned to one of four a priori distinctive subscales: Facial expression, Social behavior/Mood, Vocalisation, Physiological signs.

Visual Analogue Scale (VAS)

A 100 mm VAS was used to obtain a general impression concerning the severity of the children's pain during the observed episodes. Both researchers and parent(s) rated the child's pain from 0 (no pain) to 10 (worst pain) (Huskisson 1974). As a valid tool for children with PCI is still lacking, we used the VAS as a criterion measure, assuming a score equal to or above 4 to reflect pain, like in normally developed preverbal children (Buchholz et al. 1998; Van Dijk et al. 2002). Interrater agreement (intraclass correlation, two-way mixed effect model) for VAS ratings, completed for 52 children with PCI, from 2 researchers over all 333 episodes was excellent (.79).

Procedures

The study was approved by the medical ethical committee of the Erasmus MC - Sophia. Parents were asked for written informed consent when their child was scheduled for elective surgery in the hospital. Information on the children's medical diagnoses associated with their cognitive impairment level was obtained from their medical records. Information about speech behavior was obtained from (1) a parental questionnaire, (2) descriptions of the

child's verbal capacities in medical records, and (3) interpretation of the video recordings by the investigators.

We video-recorded the children at seven time points: twice pre-surgical and five times post-surgical. Pre-surgical recordings consisted of one episode at home within one month before surgery, when the child was not in pain or distress (T1), and one with the child already hospitalized (T2). The post-surgical recordings were made at 30 minutes (T3), at 3 (T4), 24 (T5), and 48 hours (T6) after extubation, respectively, with the fifth during an intervention-specific pain-inducing moment that differed in type and timing depending on the type of surgery (T7). Recordings lasted six minutes, and focused successively on the whole body (2 min.), the face and trunk (2 min.) and the face (2 min.). In total we recorded 333 of the scheduled 364 episodes (52 children x 7 episodes).

After each episode, researchers and parents (if present) scored pain intensity using the VAS. Video recordings of pre-surgical episodes were only included in the analysis if the child was not acutely ill or in distress and VAS scores, as rated by both researchers and parents, were 0.

All observations were done blind to medication and the VAS and CPG were completed in a random order. It was not feasible, however, to make observations blind with respect to pre- versus post-surgical episodes. We considered presenting the video episodes in a random order to the observers, however, it was still obvious which episodes were videotaped before or after surgery (e.g., due to the state of the child, being on a monitor in ICU, being in a cast or bandage).

Painfulness of surgical procedures

In this study participants underwent 18 different types of surgical procedures, including ENT surgery, dentistry, plastic surgery, gastro-intestinal tract, urology, orthopedics and ophthalmology. These were rated on painfulness (low (1), moderate (2) or high (3)) by 11 experienced medical specialists all working fulltime in the children's hospital with an average number of 18.0 years of experience with the included procedures. The surgical procedure was assigned to one of the three categories of painfulness based on the average ratings from the medical specialists, with priority given to the rating of the executive specialist (e.g., the rating of the orthopedic surgeon was counted double regarding 'adductor tenotomy' in comparison with ratings from the other specialists whose mean ratings were counted only once).

All intra-operative anaesthetic and analgesic medication was standardized for each type of surgery and was documented. Post-surgical medication was registered as well (see Terstegen et al. 2003).

Statistical analyses

Data were analyzed using SPSS 9.0. Because of the highly skewed distribution of the data non-parametric tests (setting alpha at 0.05) were used.

First, a non-parametric Principal Component Analysis (PRINCALS) was conducted for detecting relationships among the 23 indicators and to test the goodness of fit of several models representing the 23 indicators which were listed in the item pool in the preceding study (Terstegen et al. 2003).

Second, internal consistency reliability of the scales was assessed using Cronbach's alpha. Spearman correlation coefficients were calculated among the 4 CPG subscales and between the subscales and the total score.

Third, to assess construct validity the Wilcoxon signed rank test (using Fisher's exact test for significance) was used (1) to compare mean scores of episodes before (T1-T2) and after (T3-T7) surgery, and (2) to compare mean scores of the post-surgical non-surgery specific episodes (T3-T6) and the post-surgical surgery specific episode (T7).

To test their sensitivity to post-surgical pain fluctuation, CPG subscale and total scale means at each of the post-surgical episodes were compared using Friedman's test.

To assess convergent validity, scores on the subscales and total scale score obtained from the post-surgical measurements in which VAS ratings were < 4 were compared to those in which VAS ratings were ≥ 4, using the Kruskal-Wallis test with Fisher's exact test for significance. Because the VAS has not been validated for post-surgical pain assessment in children with PCI in preceding studies, interrater reliability of the VAS was calculated. Furthermore, to investigate to which extent CPG scores are predictive of VAS ratings, univariate and multivariate regression analyses were conducted, in which the CPG subscale scores and total scale score were entered as predictors.

Finally, to assess their sensitivity to pain severity, mean scores for the subscales and total scale for children undergoing low, moderate, and highly painful procedures were compared using the Friedman test for significant differences.

Results

Internal structure

The list of 23 observational items used in the analysis and their component loadings are presented in Table 1. We tested a one-dimensional solution and an a priori four dimension solution. The items were treated as ordinal variables. Component loadings, indicating the measure of fit within the dimensions, are presented for items on which at least 1 ('sometimes shown') was scored, during the pre- and post-surgical episodes. The number of scoring categories was set at 3 (1='sometimes shown'; 2='often shown'; 3='always shown'). The analysis performed on these items resulted in a solution with 4 optimal dimensions that together explain 59% of the variance. The associated eigenvalues, the goodness-of-fit of the entire solution, were 0.43 for dimension 1, 0.38 for dimension 2, 0.35 for dimension 3 and 0.33 for dimension 4. Based on the content of the items, four scales were constructed. Although the item 'eyes almost closed' had the highest component loading in dimension 1, it

was not included in the scale because of the large discrepancy with the other loadings. A similar decision was made for the item 'accepting comfort' in dimension 2.

Table 1. CPG items, component loadings and marginal frequencies obtained from PRINCALS

Component loadings	Dimension 1	Dimension 2	Dimension 3	Dimension 4
Tense face	**.59***	.25	-.01	-.25
Facial restlessness, tics	**.71***	.40	-.31	-.10
Grimace	**.81***	.30	-.18	.17
Corners of mouth downwards	**.43***	.18	-.11	.14
Eyes almost closed	**.29***	.03	-.02	.17
Eyes squeezed	**.61***	.21	-.16	.44
Looking sad, almost in tears	**.68***	.40	.33	.04
Frightened, fearful look	**.60***	-.34	-.01	-.34
Deeper naso-labial furrow	**.64***	.28	-.27	.21
Trembling chin	**.41***	-.32	.35	.15
Tensed up (whole body)	.25	**.41****	-.22	-.33
Panics, panicky reaction	.18	**.57****	.32	-.53
Accepting comfort	-.29	**-.25****	-.22	-.28
Resistant, rebellious	.06	**.71****	.23	.05
Protects sore body part	.06	**.35****	-.08	.27
Cries hard, loudly	.06	-.22	**.72*****	-.08
Cries softly	-.45	-.07	**.86*****	-.12
Moaning, groaning	-.07	.24	**.83*****	-.12
Penetrating sounds of restlessness	.03	-.79	**.80*****	-.16
Breath holding	.02	-.28	.41	**.52******
Faltering respiration	.10	-.34	-.15	**.75******
Looks red, turns red	.25	-.41	**.46*****	.32
Tears	.04	.17	**.70*****	.28

Iteration number 18; total fit: .59; total loss 3.4; multiple loss 3.3; Single loss .05

* Indicator belongs to (*) dimension 1, (**) dimension 2, (***) dimension 3 and (****) dimension 4

The item scores appeared to be highly skewed. For this reason, and also because observation in clinical practice benefits from fewer response categories, we tested the goodness of fit of the models using four response categories by combining the 'regularly shown' and 'often shown' responses into one 'often shown' category (initially there were 5 response categories: 0=never shown, 1=sometimes shown, 2=regularly shown, 3=often shown, 4=always shown). However, since the models based on four and five response categories had an equal fit to the data, we maintained this four-point scale in the further analyses.

Internal consistency and interrater agreement

Internal consistency of the CPG was determined by computing Cronbach's alpha for each subscale (Table 2). The presence of two indicators, 'eyes almost closed' in the subscale 'Facial expression' and 'accepting comfort' in the subscale 'Social behavior/Mood' resulted in lower alphas for these scales in almost all episodes. As they also showed low component

loadings (.29 and -.25 respectively), they were excluded from further analyses. For 'Facial expression' and 'Vocalisation' internal consistency was adequate (.82 and .83, respectively), and for 'Physiological signs' and 'Social behavior/Mood' moderate (alphas were .62 and .56, respectively). The total scale had excellent internal consistency (.89).

Interrater intraclass correlations were .96 for 'Facial expression', .92 for 'Social behavior/Mood', .95 for 'Vocalisation', .93 for 'Physiological signs' and .97 for the total scale, indicating excellent interrater agreement for all CPG subscales and the total scale.

Spearman correlation coefficients showed moderate correlations between all subscales, except between 'Facial expression' and 'Vocalisation', where the correlation was high (Table 2). These correlations indicate that the subscales are not independent, but that the overlap is limited. The correlations between the subscales and the total scale show that 'Facial expressions' contributes by far the most (r=0.94) to the total score variance.

Table 2. CPG internal consistency, interrater reliability and interscale correlations of the CPG scales

CPG subscales	Cronbach's Alpha	Interrater ICC	Interscale Correlations			
			Social behavior/Mood	Vocalisation	Physiological signs	Total scale
Facial expression	0.82	0.96	0.47	0.61	0.41	0.94
Social behavior/Mood	0.56	0.95		0.44	0.40	0.57
Vocalisation	0.83	0.93			0.43	0.70
Physiological signs	0.62	0.92				0.62
Total scale	0.89	0.97				

* For interscale relations, Spearman correlation coefficients were used, p<.001

Construct and convergent validity

Discrimination between absence versus presence of pain

Mean scores on the CPG subscales and the total scale differed significantly before and after surgery (Table 3). On the total scale, 98% of the children scored higher after surgery. On 'Facial expression', the proportion of children with higher scores after surgery was 94%, on 'Social behavior/Mood' 69% (26% were ties), on 'Vocalisation' 79% (13% were ties), and on 'Physiological signs' 85%.

Further, mean scores on the subscales 'Facial expression' and 'Vocalisation' and mean scores on the total scale different significantly between the regular post-surgical episodes and the surgery specific episodes.

Table 3. CPG mean scores (SD) during pre-surgical and post-surgical, and surgery specific episodes

	Pre- versus Post-surgical episodes (total n = 333)		Post-surgical episodes versus surgery specific episode	
CPG subscales	Pre-surgical N= 99	Post-surgical N=234	Post-surgical N=195	Surgery specific N = 39
Facial expression	0.18 (0.19)	0.68 (0.38)*	0.59 (0.38)	1.10 (0.70)*
Social behavior/Mood	0.01 (0.07)	0.18 (0.18)*	0.14 (0.17)	0.44 (0.73)
Vocalisation	0.06 (0.15)	0.38 (0.37)*	0.30 (0.35)	0.86 (0.80) *
Physiological signs	0.10 (0.11)	0.30 (0.22)*	0.25 (0.20)	0.38 (0.60)
Total scale	0.09 (0.08)	0.39 (0.21)*	0.38 (0.22)	0.70 (0.52)*

* $p<0.0001$

CPG scores obtained from the four post-surgical episodes all declined from 30/60 minutes till 48 hours after surgery. The Friedman test showed significant differences over time for the mean scores on the subscales 'Facial expression', 'Vocalisation', 'Physiological signs' and the total scale (Table 4), but not for subscale 'Social behavior/Mood'.

Table 4. CPG mean (SD) scores across post-surgical episodes

	Episode 3	Episode 4	Episode 5	Episode 6	Significant contrasts from post hoc comparisons*
Facial expression	8.10 (6.20)	5.24 (5.10)	4.42 (5.05)	3.33 (3.97)	E3>E6; E4>E6
Social behavior/Mood	0.59 (1.24)	0.22 (0.40)	0.24 (0.72)	0.22 (0.73)	
Vocalisation	1.80(1.13)	1.06 (2.35)	1.18 (2.21)	0.70 (1.50)	
Physiological signs	1.55 (2.46)	0.56 (1.53)	0.32 (1.02)	0.20 (0.72)	E3>E4>E5>E6
Total scale	12.04 (10.51)	7.10 (8.38)	6.16 (7.41)	4.43 (6.14)	E3>E5>E6

$p< 0.05$ (post-hoc comparisons)

Convergent validity was assessed by relating the CPG subscales and total scale with VAS ratings from researcher and parents on all episodes. Mean CPG scores for all scales were significantly different for episodes with VAS ratings < 4 versus VAS ratings ≥ 4 (Table 5). Univariate regressions indicated that VAS ratings increased significantly with increasing CPG subscale scores and total scale scores. The multivariate regression analysis using the four subscales as predictors, showed that subscales 'Facial expressions', 'Vocalisation' and 'Physiological signs' contributed independently to the prediction of VAS scores. In Table 4 standardized Beta coefficients are shown, which represent the strength of the CPG's prediction on VAS ratings. The regression analyses also showed that while 49% of the variance in VAS scores was predicted by the combined CPG subscales, 53% was predicted by the CPG total score alone.

Table 5. CPG mean scores (SD) for episodes with low versus high VAS scores and regression of VAS on CPG scales

	Episodes VAS <4 (267) versus VAS ≥4 (55) scored by *researchers*		Episodes VAS <4 (180) versus VAS ≥4 (47) scored by *parents*		Regression analyses of VAS on CPG scale over all episodes	
	Kruskal-Wallis test		Kruskal-Wallis test		R^2	Beta's
Subscales	VAS <4	VAS ≥4	VAS <4	VAS ≥4		
Facial expression	2.81 (3.17)	8.84 (4.86)*	2.89 (3.37)	7.48 (4.59)*	0.42*	0.22*
Social behavior/Mood	0.26 (0.77)	1.24(1.32) *	0.24 (0.67)	1.52 (1.58)*	0.20*	0.09
Vocalisation	0.54 (1.24)	3.31 (2.75)*	0.61 (1.33)	2.74(2.73)*	0.37*	0.14*
Physiological signs	0.71 (0.98)	2.98 (2.13)*	0.83 (1.11)	2.59 (2.10)*	0.43*	0.37*
Total scale	4.32 (4.96)	16.36 (9.18)*	4.58 (5.29)	14.33(9.27)*	0.53*	-

*= P<0.0001, adjusted R^2 and standardized Betas are reported

CPG scores and differences in severity of pain

Mean scores for the post-surgical episodes (T3-T6) differed significantly from those for the surgery specific episodes (T7) with regard to 'Facial expression', 'Vocalisation' and the total scale score (Table 3).

Means and SD of the subscale scores and total scale score based on the four post-surgical measurements categorized for the low, moderate and highly painful procedures respectively, are given in Table 6. CPG scores on all subscales as well as on the total scale increased by painfulness. The Kruskal-Wallis test showed significant differences for 'Facial expression', 'Vocalisation' and the total scale for low versus high pain intensity.

Table 6. CPG mean (SD) scores for children undergoing surgical procedures of different painfulness

	Low N=16	Moderate N=24	High N=13	Exact significance post hoc comparisons
Facial expression	15.19 (9.70)	19.92 (13.76)	26.00 (13.23)	low < high
Social behavior/Mood	0.88 (1.50)	1.08 (1.56)	1.83 (2.52)	
Vocalisation	2.75 (2.65)	5.04 (6.45)	5.67 (6.20)	low < high
Physiological signs	1.44 (2.85)	2.38 (3.12)	4.08 (4.12)	
Total scale	20.25 (13.02)	28.42 (20.45)	37.58 (19.79)	low < high

p< 0.05 (post-hoc comparisons)

Discussion

Internal structure

Findings from the PRINCALS analyses showed a highly satisfactory fit to the data of both a four-dimension solution and a one-dimension solution. This shows that the four subscales were clearly distinguishable, as was confirmed by the moderate inter-subscale correlations, and that these may be combined into a total score without loss of information. Although two items ('looks red, turns red' and 'tears') showed higher loadings on dimension 3 (subscale Vocalisation) we choose to categorize them within dimensions 4 (subscale Physiological signs) because of a more appropriate fit based on their content. These two items probably showed higher loadings on the vocal subscale because they are often shown together with other items in this subscale (e.g. crying).

Reliability

Both PRINCALS and internal consistency and interrater reliability analyses provide evidence that the Checklist Pain Behavior, consisting of 21 of the initial 23 indicators, has good internal reliability when used for assessing post-surgical pain in children with PCI. The subscales 'Facial expression', 'Vocalisation' and the total scale showed high internal showed somewhat lower internal consistencies. However, because of their adequate percentages of explained variance, the latter two subscales appear to contribute substantially to the total scale score. Interrater reliability was excellent for all subscales and for the total score. Interscale correlations were moderate for most subscales, but high between the facial and vocal subscales.

Construct and convergent validity

Discrimination between absence versus presence of pain

The CPG construct validity relies on differing scores before and after surgery. The Wilcoxon signed rank test showed that almost all children (98%) scored significantly higher on the CPG 'Total score' after surgery compared to the pre-surgical episodes. This finding is consistent with a statement of Breau et al. (2002a), who indicate that the most important quality of pain assessment tools is the ability to discriminate between episodes with likely presence versus absence of pain.

The convergent validity of the CPG turned out to be good; regression analyses showed that all subscales, except 'Social behavior/Mood', contributed independently to the prediction of VAS scores. This shows that each of the CPG components 'Facial expression', 'Vocalisation' and 'Physiological signs' is related to an overall pain rating. The exception for 'Social behavior/Mood' confirms our previous hypothesis (Terstegen et al. 2003) stating that behaviors included in this scale ('panicking', 'resistant/rebellious', 'protecting sore body part') might be more indicative of anxiety or anger rather than pain, as indicated by VAS ratings. However, all mean CPG subscale scores and the total scale score were significantly higher for

the post-surgical episodes when VAS ratings were higher or equal to 4, compared to VAS < 4. Because this study also showed excellent interrater reliability of the VAS ratings, it can be concluded that the Checklist Pain Behavior indeed measures pain in this particular group of children.

CPG scores and differences in severity of pain

We analyzed the CPG scores on the four regular post-surgical measurements categorized by a low, moderate or highly painfulness of surgery. Scores on the subscales 'Facial expression', 'Vocalisation' and the total scale score appeared to differ significantly between children undergoing procedures that differed in painfulness, indicating that these indicators are most probably sensitive to differences in severity of pain.

Scores on the subscales 'Facial expression' and 'Vocalisation' and the total scale were also significantly higher for the surgery specific episode compared to scores for the four other post-surgical episodes. This finding is consistent with results from our preceding study (Terstegen et al. 2003), in which scores for ten indicators proved to be significantly different at surgery specific episodes with VAS ratings < 4 versus higher or equal to 4. Five of these ten indicators referred to 'Facial expression' and three to 'Vocalisation'. In each of the subscales 'Social behavior/Mood' and 'Physiological signs', for one indicator significant differences were found.

Thus we may conclude that 'Facial expression' and 'Vocalisation' are the most sensitive subscales for differences in pain intensity. Nevertheless, the other two subscales contribute substantially to the total scale score, judging from their explained variances (respectively 13.9% and 17.3%).

Changes in pain over time

Statistical analyses showed significantly different, declining mean scores across the post-surgical episodes for all subscales, except 'Social behavior/Mood'. This subscale displayed nearly similar mean scores, and thus seemed to be equally present over time. An explanation could be that some of the behaviors included in this scale ('panicking', 'resistant/rebellious', 'protecting sore body part') might be more sensitive to anxiety or anger rather than pain (Terstegen et al. 2003).

In summary, our results provide evidence that post-surgical pain in children with PCI can be reliably assessed and monitored using the Checklist Pain Behavior (CPG). It is a generic pain assessment tool, however, and in view of the fact that these children display differences in type of pain behavior (Giusiano & Jimeno 1995; Terstegen et al. 2003), one should be aware of the child's customary behavior before observing their post-surgical behavior. We recommend, therefore, to first apply the CPG at a moment before surgery when the child is not in pain, and to use the obtained score as a reference score.

Furture research

Several relevant issues were not yet addressed in this study. First, although the CPG appears to be sensitive to pain versus no-pain situations, and to changes in pain expression, its validity would be further corroborated by evidence of sensitivity to the effects of pain medication. Second, we need to know to what extent the behavioral pain responses of children with PCI to specific painful procedures are typical. This might be elucidated by comparing responses of children with PCI to those of other children when submitted to the same procedure, e.g., adenotonsillectomy. Third, it has been suggested that some children with severe or profound cognitive impairment might be insensitive to pain. Recently, Breau et al. reported that caregivers with greater knowledge about children with CI, had a stronger belief that these children experience pain less than children without CI (2003a). In addition, children with cognitive impairment of different origin might be differentially sensitive. This issue of differential pain sensitivity may be addressed using the CPG in a broad sample of children with different levels and causes of cognitive impairment. Fourth, feasibility in daily practice is of great importance for successful implementation of a pain assessment tool. It would be interesting to learn whether the CPG's 21 pain sensitive indicators, which already were derived from 209 possible pain indicators, could be further reduced and thus improve its use in daily practice. Currently, we use the CPG for observations directly at the bedside to examine the clinical usefulness of the resulting instrument in daily caregiving to this specific group of children.

The children in this study differed in origin of cognitive impairment and showed a wide range of capacities and impairments within several developmental areas. Notwithstanding this heterogeneity, we detected a set of shared generic indicators which are used by these children to express pain. However, this study also reports that each individual child shows an individual type of pain behavior. Hence, reducing the number of indicators may well reduce the number of children for which the CPG would be appropriate (Breau et al. 2002a). Furthermore, it would be of interest to study whether application of the CPG by children with specific diseases with a well determined genetic cause, would enable to relate changes in pain scores with the underlying neuro-anatomical abnormalities.

Chapter 5

Comparing post-surgical pain responses in children with profound cognitive impairment and normally developed preverbal children

Chantal Terstegen
Josien B. de Boer
Hans M. Koot
Dick Tibboel

(Submitted for publication)

Abstract

Children with cognitive impairment (CI) express their pain in a different manner than normally developed children do (Gilbert-MacLeod et al. 2000; Breau et al. 2001). However, in which way they differ is a question that still remains to be answered. This study investigated differences in post-surgical pain expressions between children with profound cognitive impairment (PCI) and children with a normal development. Sixteen children with PCI and 12 normally developed preverbal children were included. All observations were performed with the Checklist Pain Behavior (CPG), which includes 20 behavioral indicators. The CPG has exhibited adequate psychometric performance and proved to be sensitive to pain in children with PCI (Terstegen et al. 2003). Children were observed after admission for elective ENT-surgery at predetermined time points. Visual Analogue Scale (VAS) ratings were used to obtain a general impression of the severity of their pain. Within the group of normally developed children 12 behaviors - and for children with PCI 4 behaviors - were not shown before surgery (VAS=0, absence of pain) but appeared only after surgery (presence of pain; post-surgical observations with VAS ≤ 4 were excluded from this calculations). Four behaviors - and for children with PCI 16 - were present before surgery and appeared with increased frequency or intensity after surgery. Four behavioral indicators were only scored in children with PCI. Several pain expressions displayed by children with PCI were also observed in children without CI, but often less attenuated or, on contrary, more frequent. To our knowledge, this is the first study directly comparing pain behavior in children with and without PCI undergoing comparable surgical procedures with potentially high pain intensity, while they received the same standardized analgesia. For the assessment of pain in children with PCI we advocate the use of measures, such as the CPG, that are specifically designed and validated in this group of children.

Introduction

While recent advances in pain assessment and management have made it a realistic goal to minimize post-surgical discomfort, children with CI remain vulnerable to under-treatment of post-surgical pain (Malviya et al. 2001). In comparison with normally developed children or even children with a mild or severe CI, children with PCI have a limited behavioral repertoire to express their pain. Several studies reported that these children process information and communicate distress and pain in a different manner than normally developed children do (Gilbert-MacLeod et al. 2000; Breau et al. 2001; Stallard et al. 2001; Terstegen et al. 2003). However, in which way they differ is a question that still remains to be answered.

For assessing post-surgical pain in preverbal children, a pool of measures with adequate psychometric properties is available (Van Dijk 2001). However, because children with PCI exhibit different pain behavior, scales validated for normally developed children may not be appropriate for children with PCI (Terstegen et al. 2003). A study of Malviya and colleagues

(Malviya & Voepel-Lewis 2001) revealed that it may be questionnable to apply pain measures developed for normally developed children in children with CI. They indicated that children with CI received significantly smaller doses of opioids after surgery than children without CI and suggested that this might be due to the fact that pain was assessed with measures developed for children without CI. It is not until recently that several measures have been developed to assess post-surgical pain behavior in children with PCI (Stallard et al. 2002a, 2002b; Breau et al. 2002a, 2002b;Terstegen et al. 2003).

One study has been conducted comparing pain responses of children with mild cognitive impairment versus normally developed children (Gilbert-MacLeod et al. 2000). To our knowledge, no studies have been conducted in which pain behavior of children with and without PCI was directly compared. However, this comparison is of crucial importance to demonstrate the need for separate pain instruments and the development of tailored analgesic regimen for children with PCI.

The purpose of the current study was to investigate differences in pain behavior between children with and without PCI in a standardized post-surgical setting using the Checklist Pain Behavior (CPG). Some of the psychometric properties of the CPG were evaluated in a previous study (Terstegen et al, submitted). Because assessing validity is an ongoing process, in the present study we further examined its construct validity by comparing CPG scores from children with and without PCI.

Toward this end, we compared the children's scores on the CPG indicators and scales in painful and non-painful situations. In order to distinguish painful from non-painful situations, we tested differences in CPG scores within and between both groups at times when VAS ratings were 0-3.9 (no pain) versus ≥ 4 (pain). Second, we studied the differential sensitivity of the CPG in both groups to changes in pain over time across two post-surgical episodes. Third, in order to evaluate post-surgical pain as well as sharp short-lasting pain, we compared the CPG scores obtained during two post-surgical pain measurements and directly induced, surgical specific pain.

Methods

Participants

Sixteen children with PCI (11 boys, 5 girls; mean age 6.0 years) and 12 normally developed preverbal children (5 boys, 7 girls; mean age 18 months) were included in the current study. They were observed on predetermined time points after admission for elective Ear Nose Throat-surgery (ENT procedures) such as middle ear tubes (PCI n=4, without PCI n=2), adenotomy (PCI n=5, without PCI n=4) and adenotonsillectomy (ATE) (PCI n=7, without PCI n=6). The origins of the included children's cognitive impairments were diverse (Table 1).

Table 1. Origin of cognitive impairment

Congenital/ metabolic anomalies:	
Syndrome:	
- chromosomal disorder	4
- identified metabolic disorder	2
Major structural cerebral abnormalities	2
Intra-uterine asphyxia	1
Cognitive impairment during birth:	
Perinatal asphyxia	2
Cognitive impairment after birth: (between 3 days and 4 years old)	
Status epilepticus 3 days post partum	1
Pneumococcal sepsis	2
Origin remained unknown	2

The children with CI who participated in this study were enrolled in a pain study aiming at the development and validation of an observational scale, the Checklist Pain Behavior (CPG), for assessing and monitoring post-surgical pain in these children (Terstegen et al. 2003). Although there is no clear definition available of children with PCI, they can be described as children with a cognitive development estimated to be below a calendar age of 2 years (equivalent IQ 0-20). Although recently an instrument (GTI) was developed to scale the capacities of persons with severe multiple impairments (Vlaskamp et al. 2002), no diagnostic tools are available which can precisely determine a child's IQ level below 40. In addition, we assessed language development using 'minimal speech standards' (Goorhuis & Schaerlaekens 2000). Standards for PCI include the ability of making sounds but not expressing oneself through words. None of the children's abilities exceeded these standards, indicating that the participants were unlikely to express pain verbally.
For optimal comparison concerning verbal capacities to express emotions such as pain, a group of normally developed children with a mean age of 18 months were included.

Measures

Visual Analog Scale (VAS)

Researchers and parents gave an overall impression of the severity of the children's pain directly after the observations and at the bedside using the VAS (Huskisson 1974), a 10 cm line with the anchors 0 (no pain) and 10 (worst pain). VAS ratings ≥ 4 were considered clear indications of pain (Van Dijk 2001). Kappa between parents' and researchers' ratings was 0.69 indicating good inter-observer reliability (Landis & Koch 1977).

Checklist Pain Behavior (CPG)

All observations were scored with the CPG, which consists of 20 observable behaviors. The CPG indicators are aggregated into 4 behavioral categories ' Facial expression' (9 items), 'Social behavior/Mood' (3 items), 'Vocalisation' (4 items) and 'Physiological signs' (4 items). Four scale scores are computed and a total score can be derived by summing the indicator scores from each of the 4 categories. Using behavioral descriptions, indicators were rated on a 4-point scale, with 0=never shown, 1=sometimes shown, 2=often shown, 3=always shown.

In a prior study we evaluated the psychometric properties of the CPG. Analyses showed that the CPG proved to be valid and reliable to assess post-surgical pain in children with PCI. Inter-rater reliability (intraclass correlations) was excellent: 0.92-0.96 for the subscales and 0.97 for the total scale. Internal consistency (Cronbach's alpha) was adequate: 0.56-0.83 for the subscales and 0.89 for the total scale. All mean CPG scores were significantly different (1) before versus after surgery and (2) with VAS ratings of 0 to 3.9 (absence of pain) versus VAS ratings 4 to 10 (Terstegen et al. 2003).

Procedure

The study was approved by the Medical Ethical Committee of the Erasmus-MC. Parents were asked for informed consent when their child was scheduled for elective surgery in the hospital. Information on medical diagnoses associated with their cognitive impairment level was obtained from medical records. Information about speech behavior of the children was gathered by (1) sending a questionnaire to the parents of the children with PCI, (2) checking written descriptions of the child's verbal capacities in medical records for both groups of children, and (3) interpretation of the video recordings by the investigators.

Six-minute video recordings were made in order to observe the child's behavior during 4 predetermined episodes, one pre-surgical episode when the child was not suffering from pain or distress (T1), and three post-surgical episodes at 30 minutes (T2) and 3 hours (T3) after extubation, and for patients undergoing adenotonsillectomy (ATE), during their first drink after surgery (T4).

After each episode pain intensity was scored on the VAS by researchers and parents (when available). Video recordings of the pre-surgical episode were only included if the child was not acutely ill or in distress at that time and, as a consequence, if VAS scores were 0 as rated by both researchers and parents.

All intra-operative anesthetic and analgesic medication was standardized for each type of surgery and was documented in anesthetic and surgical records. Anesthetic medication penthotal (5 mg. per kg. bodyweight) was administered iv. Before intubating, sufenta 0,2 mcg./kg. iv was administered, if necessary repeated with 1/3 of the first dose. During surgery, anesthesia was maintained with O2, NO2 and isoflurane 0,5 MAC. Mivacron (0,2 mg/kg) was used as muscle paralysis.

Two hours before surgery paracetamol 40-50 mg/kg sup was adminstered. Post-surgical medication and doses were also registered; for ATE surgery acetominophen 90-100 mg./kg./24 hrs was administered and for middle ear tubes acetominophen 60 mg./kg./24 hrs.

Statistical Analyses

Data were analyzed using SPSS 10.1. Because of the highly skewed distribution of the data non-parametric tests were used; alpha was set at .05 for all tests.

Before we investigated differences in pain behavior, we determined the internal consistency of the CPG for children with and without PCI by computing Cronbach's alpha for each subscale and the total scale.

First, we studied differences in behaviors showed in painful situations within and between both groups. The sensitivity of the CPG for absence versus presence of pain was tested. Wilcoxon signed ranks tests (using Fisher's exact test for significance) were used to analyze differences in CPG scores on each of the indicators and subscales before and after surgery *within* both groups of children. The Mann-Whitney U test (using Fisher's exact test for significance) was conducted to analyze differences *between* both groups.

Second, we evaluated the ability of the CPG to detect changes in pain over time across the two post-surgical episodes (T2 and T3) by comparing CPG scores across these episodes using Wilcoxon signed ranks test for two related samples (with the exact test for significance).

Third, using the Wilcoxon signed ranks test we tested the significance of differences between scores during the two general post-surgical measurements and the one, surgery specific measurement (T4).

Results

The *internal consistency reliability* of the CPG for both groups of children is shown in Table 2. The total scale for the group of children with PCI showed a good internal consistency (.89). The Cronbach's alpha of the total scale for the group of children without PCI is based on 16 indicators, as 4 out of the original 20 behaviors were not shown at all within this group ('facial restlessness, tics', 'frightened fearful look' and 'trembling chin' from the subscale Facial expression and 'penetrating sounds of restlessness' from the subscale Vocalisation) and had also good internal consistency (.88). For the children without CI 'Facial expression' alphas indicated good internal consistency, for 'Social behavior/Mood' and 'Vocalisation' internal consistency was moderate and for the subscale 'Physiological signs' Cronbach's alpha was low (0.82, 0.63, 0.61 and 0.44, respectively).

Table 2. Internal consistency (Cronbach's alpha) of CPG scales

	Children with PCI (from original study)		Children without PCI	
	Alpha	No. items	Alpha	No. items
Facial expression	0.82	9	0.83	6*
Social behavior/Mood	0.56	3	0.63	3
Vocalisation	0.83	4	0.61	3**
Physiological signs	0.62	4	0.44	4
Total scale	0.89	20	0.88	16

* Three indicators ('facial restlessness, tics', 'frightened fearful look' and 'trembling chin') were deleted because of zero variance.
** One indicator ('penetrating sounds of restlessness') was deleted because of zero variance.

First, we studied *differences in pain behavior* between both groups. Painful and non-painful episodes were indicated by a VAS score 0 to 3.9 (absence of pain) versus ≥ 4 (presence of pain). VAS ratings did not significantly differ between researchers and parents for both groups ($t = -3.74$; $p > .05$). On average, parents' ratings were slightly higher (i.e., 0.1 to 0.7 cm.) than ratings from researchers for children with PCI and slightly lower for those without PCI (i.e., 0.1 to 0.8 cm.). For children with PCI, researchers rated 14 out of 56 episodes (25%) as painful episodes (VAS ratings ≥ 4) and for those without PCI 10 out of 39 episodes (25.6%) were rated as painful. For both groups, based on the researchers' ratings, the analyses showed that on episodes with VAS ratings ≥ 4, scores on all 4 subscales and the total scale were significantly higher compared to the episodes with VAS ratings below 4 ('Facial expression' with and without PCI p=.000; 'Social behavior/Mood' PCI p=.020, without PCI p=.015; 'Vocalisation' with and without PCI p=.000; 'Physiological signs' PCI p=.024, without PCI .008; total scale with and without PCI p=.000).

Table 3 presents percentages of children who scored at least 1 ('sometimes shown') on each CPG indicator for each of the 4 pre- and post-surgical measurements and mean scores for the pre-surgical measurement and for the combined post-surgical measurements. For both groups of children, all subscales and the total scale were significantly different when pre- and post-surgical CPG scores were compared, which indicated sensitivity to absence versus presence of pain. Four behaviors were shown only by children with PCI:'facial restlesness, tics', 'frightened fearful look', 'trembling chin' and 'penetrating sounds of restlessness'. Further, within the group of normally developed children 12 behaviors - and for children with PCI four behaviors - were not shown before surgery (VAS=0, absence of pain) but appeared only after surgery (presence of pain; post-surgical observations with VAS ≤ 4 were excluded from these calculations). Four behaviors – and 16 for children with PCI - were present before surgery and appeared with increased frequency or intensity after surgery (Table 3, test 1).

Further, several types of behavior within each CPG subscale were scored significantly different in children with PCI compared with normally developed children (Table 3, test 2).

First, children with PCI had higher scores on 'tense face', 'corners of mouth downward', 'frightened, fearful look' and 'trembling chin' from the subscale 'Facial expression' and lower scores on 'looking sad, almost in tears'. Within the subscale 'Social behavior/ Mood' children without CI had higher scores on 'protects sore body part'. For the subscale 'Vocalisation' children with PCI had higher scores on 'penetrating sounds of restlessness' and 'crying softly' and lower scores on 'crying hard, loudly'. Finally, for the subscale 'Physiological signs', the indicator 'breath holding' was scored higher for children with PCI.

Second, on examination of the *changes in pain over time,* for children with PCI the mean scores on the subscale 'Facial expression' and the total scale were significantly higher on T2 (30 minutes after surgery) compared to T3 (3 hours after surgery). This is in contrast with the group of children without CI, for whom no significant differences in CPG scores between the two post-surgical measurements were found on any of the subscales.

Third, we studied *post-surgical pain versus directly induced pain* (drinking for the first time after ATE). For the group of children with PCI, only the indicator 'grimace' from the subscale 'Facial expression' and the total CPG score were scored significantly different on the post-surgical measurements versus the surgery specific measurement. These were shown less frequently during the post-surgical episodes compared to the direct pain inducing episode, indicating that these are most likely to be sensitive to differences in post-surgical versus acute pain. Again, in the group of children without cognitive impairment, no significant differences were found.

Table 3. Percentage of children with PCI (group 1) and without PCI (group 2) who scored at least 1 ('sometimes shown') during pre- and post-surgical episodes

Indicators	Pre-surgical		Post-surgical								Tests of significance (p-values)			
	Mean pre-surgical **(T1)**		**T2**		**T3**		**T4**		Mean post-surgical **(T2,3,4)**		Test 1. Pre- vs post-surgical within groups [a]		Test 2. Pre- vs post-surgical between groups[b]	
	1	2	1	2	1	2	1	2	1	2	1	2	Pre	Post
Facial expression														
Tense face	18.8	20	81.4	54.6	56.3	63.7	75	50	75	51.7	.000*	.020*	.770	.158
Facial restlessness, tics	12.5	0	31.3	0	31.3	0	50	0	35	0	.031*	1.000	.508	.003*
Grimace	6.3	0	37.6	54.6	37.5	54.6	37.5	14.3	50	51.7	.000*	.002*	1.000	.089
Corners of mouth downward	18.8	10	62.6	36.4	43.8	27.3	87.5	14.3	60	27.6	.000*	.250	.639	.000*
Eyes squeezed	18.8	0	50.1	54.6	18.9	18.2	62.5	14.3	40	40	.047*	.016*	.262	.332
Looking sad, almost in tears	18.9	0	50.1	63.7	6.3	54.6	62.5	28.6	35	51.7	.052	.002*	.262	.017*
Frightened, fearful look	0	0	37.5	0	6.3	0	50	0	27.5	0	.016*	1.000	.385	.017*
Deeper naso-labial furrow	18.8	10	75	63.7	43.8	81.9	62.5	57.2	60	69	.000*	.000	.639	.953
Trembling chin	6.3	0	31.3	0	6.3	0	12.5	0	17.5	0	.116	1.000	1.000	.035*
Subscale 1 score	**68.8**	**40**	**94.2**	**81.9**	**62.9**	**91**	**87.5**	**85.8**	**82.5**	**86.2**	**.000***	**.002***	**.096**	**.107**
Social behavior/Mood														
Panics, panic attack	12.6	0	18.8	27.3	20	36.4	25	0	20.5	24.1	.406	.063	.508	.946
Resistant, Rebellious	0	0	6.3	18.2	12.5	27.3	75	14.3	17.5	20.6	.031*	.063	1.000	.988
Protects sore body part	0	0	6.3	27.3	0	27.3	25	14.3	7.5	24.1	.040*	.031*	1.000	.023*
Subscale score	**12.6**	**0**	**25.1**	**45.5**	**31.3**	**50.5**	**62.5**	**14.3**	**35**	**41.2**	**.037***	**.008***	**.508**	**.344**
Vocalisation														
Cries hard, loudly	0	0	18.8	45.5	18.8	45.5	25	28.6	20	41.3	.016*	.004*	1.000	.006*
Cries softly	0	0	43.8	0	31.4	9.1	75	0	45	3.4	.000*	1.000	1.000	.000*
Moaning, groaning	0	0	37.5	45.5	12.5	45.5	50	42.9	30	44.8	.008*	.008*	1.000	.618

Table 3. continued

Penetrating sounds of restlessness	6.3	0	18.8	0	18.8	0	62.5	0	27.5	0	.031*	1.000	1.000	.019*
Subscale 3 score	**6.3**	**0**	**62.7**	**54.6**	**37.7**	**54.6**	**87.5**	**57.2**	**60**	**55.1**	**.000***	**.004***	**1.000**	**.427**
Physiological signs														
Breath holding	6.3	0	37.6	9.1	12.6	0	62.5	14.3	32.5	6.8	.008*	.500	1.000	.050*
Faltering respiration	0	0	6.3	9.1	0	18.2	50	14.3	12.5	13.8	.045*	.056	1.000	.746
Looks red, turns red	18.8	10	37.6	45.5	25.1	36.4	37.5	28.6	32.5	37.9	.023*	.034*	.940	.729
Tears	0	0	0	36.4	12.6	9.1	37.5	14.3	12.5	20.6	.041*	.031*	1.000	.352
Subscale 4 score	**18.8**	**10**	**62.7**	**63.7**	**37.6**	**36.4**	**87.5**	**42.9**	**60**	**48.3**	**.004***	**.008***	**.403**	**.482**
Total score	**81.4**	**40**	**94.2**	**72.8**	**75**	**91**	**62.5**	**85.8**	**87.5**	**93.1**	**.000***	**.002***	**.031***	**.185**

Note: [a] Test 1: Wilcoxon signed ranks test
[b] Test 2: Mann-Whitney U test

Discussion

To our knowledge, this is the first study directly comparing pain behavior in children with and without PCI undergoing the same amount of surgical trauma, while they received equivalent and standardized analgesia. It is clear that for children with PCI, different pain behaviors are indicative of pain. Furthermore, they show certain behaviors in response to painful stimulation similar to normally developed preverbal children, however some of these behaviors were shown more subtle and less obvious (e.g., they cry softly instead of loudly) or, conversely, more frequent or intense (e.g., 'tense face', 'faltering respiration') compared to children without PCI.

Four behavioral indicators were only shown by children with PCI: 'facial restlessness, tics', 'frightened, fearful look', 'trembling chin' and 'penetrating sounds of restlessness'. Furthermore, 4 behaviors in children without PCI – and 16 for children with PCI – were present before surgery (VAS = 0) and appeared with increased frequency or intensity after surgery (VAS ≥ 4). These findings suggest that certain behaviors in children with PCI are sensitive to pain but are not specific pain indicators. Children with PCI are in general very limited in possibilities to express themselves, which leaves them very little alternatives in "choosing" behaviors to express their pain. Some of the few behaviors they are capable of showing, will be used both when they are not in pain and in painful situations, but then more frequent or intense. This indicates that it is recommendable to observe the child with the CPG before surgery when he or she is not in pain (baseline measure) so these scores can be used as an individual point of reference.

These findings also suggest that children with PCI will substantially benefit from an observational scale, like the CPG, which is validated for this specific group of children. The importance of a valid, reliable and clinically useful pain assessment tool for use in this population is also emphasized by Malviya & Voepel-Lewis (2001) and Breau et al. (2001). In Malviya's study, children with PCI had poor pain assessment and received fewer opioids after surgery compared to their sample of children without CI. In this study, they relate the differences in pain management partially to the children's inability to effectively communicate pain and therefore concluded that specific pain assessment tools are necessary. For this reason, observations should be done very accurate using an appropriate instrument.

The psychometric performance of the CPG, is clearly better in children with PCI than in children without PCI. Four indicators were not observed in children without CI and the internal consistency proved to be better in the group of children with PCI, except for the subscale 'Social behavior/ Mood' which had higher Cronbach's alpha in children without CI (although both alpha's were below the standard of 0.70). This could be due to the small sample size causing less variability in scores. It also can be explained by the fact that children with PCI, compared to normally developed children, have less motor capacities to show behaviors in the Social behavior/Mood scale, such as 'protects sore body part'. For this reason, we should re-examine the importance of these behaviors when assessing pain in children with PCI.

Consequently, we advocate specific measures and expect that the use of the CPG together with a tailored analgesic regimen will increase the possibility to assess and monitor pain more adequately and minimize the possibility of over- or undertreatment of pain in children with PCI.

There are some points of debate concerning our study. The small sample size may limit the ability to generalize the findings from this study. Despite this limitation, we feel that our data remains important as they identify new differences in pain behavior between children with and without PCI.

Next, the CPG is validated on post-surgical pain observations and needs to be tested in daily practice in a variety of conditions apart from post-surgical pain. Currently, we are using the CPG for observations directly at the bedside to examine the clinical usefulness of this instrument in daily care giving to this specific group of children. Chronic pain of several different origins will be observed (e.g., pain due to obstipation, recurrent urinary infections, gastro-oesophageal reflux) using the CPG and the VAS.

The standards of analgesia used in this study were based on pain studies in normally developed children (Bouwmeester et al. 2001, 2003, 2004; Van der Marel et al. 2003) and need to be translated to children with PCI. The development of an algorithm of analgesia using the predetermined cutoff points of the CPG and VAS (for presence versus absence of pain) and the same principles of pre-emptive analgesia in this specific group of children, is an essential step to minimize pain in children with PCI.

Chapter 6

The Checklist Pain Behavior in daily practice: assessing everyday pain responses in children with profound cognitive impairment

Chantal Terstegen
Hans M. Koot
Patricia L. van Deventer
Esmée Lanjouw
Josien B. de Boer
Dick Tibboel

(submitted for publication)

Abstract

The Checklist Pain Behavior (CPG) was developed and validated for assessing and monitoring post-surgical pain in children with profound cognitive impairment (PCI). Since the instrument was tested in the context of a research project to assess post-operative pain, its applicability in clinical and daily practice for the assessment of everyday pain by trained nurses is largely unknown. The purpose of this study was to use de CPG in daily practice and to test its feasibility, reliability and validity when used for observations of everyday, often more chronic pain, related to physiological impairments and diseases. The CPG was used by nurses working on (1) a neurology ward of an academic children's hospital and (2) a residential healthcare facility for children with PCI. Forty-one nurses were trained to use the CPG autonomously and conducted a total of 348 bed-side observations on 48 children with PCI (23 boys and 25 girls). The moments to observe were chosen at random and after 2 minutes of observation the CPG and the VAS were completed. Convergent validity was tested and sensitivity and specificity of the CPG were examined. For all nurses interrater reliability was satisfactory (kappa ≥ .60). Internal consistency and interrater agreement (ICC) of the CPG rated at bed-side were adequate (.87 and .80, respectively). All CPG scores were significantly higher for observations with VAS ratings ≥4 compared to observations with VAS ratings <4. Univariate regressions indicated that all subscales, except for 'Social behavior/Mood', contributed independently to the prediction of the VAS scores. ROC curves showed a sensitivity of .71 and a matching specificity of .68 (area under the curve .94) by a CPG score of 18 or higher, thus, this CPG score turned out to be a fair indication of pain in children with PCI. This study provides evidence that the Checklist Pain Behavior can be used to assess everyday, more chronic pain related to physiological impairments and diseases in children with profound cognitive impairment, in both a hospital setting and in the healthcare facility where these children live. Nurses who were unfamiliar with these children were also able to detect signs of pain using the CPG. This study reports the possibility to succesfully train groups of nurses to observe pain behaviors in children with PCI in everyday pain events in a reliable way, using the CPG as a validated pain assessment scale.

Introduction

The assessment of pain in individuals with limited expressive ability presents a persistent clinical problem. Standardized measures for children who lack the ability to express their pain verbally or through alternative means (e.g., pictorials) were not available for a long period of time. However, in the last decade research on pediatric pain has increased substantially, and recently, studies also focused on a special group of children, i.e. those with severe or profound cognitive impairment (McGrath et al. 1998; Breau et al. 2001; Hunt 2001; Stallard et al. 2002a, 2002b).

Recently, the Checklist Pain Behavior (CPG) was developed and validated for assessing and monitoring post-surgical pain in children with profound cognitive impairment (PCI) (Terstegen et al. 2003). However, some issues in the assessment of pain with the CPG which have not been addressed yet, will be investigated in this study.

The primary purpose of the current study was to investigate whether the CPG could detect everyday, more chronic pain related to physiological impairments and diseases, in children with PCI in both a hospital setting (non-surgical) and in a residential healthcare facility where these children live. During its development, the CPG was only tested in an academic children's hospital setting including children with PCI following elective surgery. Besides experiencing pain as a consequence of medical interventions, chronic conditions associated with their syndromes or disorders often cause pain in these children (Breau et al. 2003a). As post-surgical pain may differ from everyday, more chronic pain (Breau et al. 2002b) it is not clear to which extent the behavioral and physiological items of the CPG are sensitive to everyday pain originating from chronic physical conditions in these children.

Besides, the original study of the reliability of the CPG was conducted with researchers who scored video fragments displaying children's post-surgical behavior. It is unknown whether nurses involved in the caretaking of these children in different settings can be trained to score the child's behavior reliably at the bed-side using the CPG. Therefore, the psychometric qualities of the CPG were evaluated by testing its reliability, validity and feasibility in daily practice in a variety of conditions apart from post-surgical pain.

First, to enhance the reliable use of the CPG in daily practice, interrater agreement (Kappa) between researcher and nurses was calculated during the nurses' training to gain full competence in autonomously completing the CPG.

Second, to address the reliability of its actual use in daily practice, the internal consistency and interrater agreement of the CPG were calculated for observations completed by the trained nurses. In addition, we do not know if scoring with the CPG is influenced by knowledge of the child's common behavior. Therefore we investigated if there are any differences regarding reliability of scoring the CPG between nurses who are familiar with the observed child (and with children with PCI in general) and nurses who are not (or less) familiar with the child.

Third, we examined the convergent validity of the CPG when used in daily care giving. To this end, children with everyday, chronic pain of different origins were observed (e.g., constipation, recurrent urinary tract infections, gastro-oesophageal reflux, epilepsy) using both the CPG and the Visual Analogue Scale (VAS) for assessing pain. These data were used to test convergence between CPG scores and VAS ratings. For construct validity, sensitivity and specificity of the CPG were examined.

Methods

Participants

For the current study, 48 children with PCI (23 boys; ages 1.1 to 12.4 years old, mean age 8.7, and 25 girls; ages 2.1 to 14.1 years old, mean age 6.9) were included. Twenty-one of them were observed in a residential healthcare facility for children with PCI, 27 children were observed while admitted to an academic children's hospital (Erasmus MC - Sophia). To be included in the study, children needed to be living in their present situation for at least one month, to make sure they did not recently go through any major environmental changes that might influence their common behavior. Children with abnormal renal/adrenal function or abnormal liver function were excluded because of a potentially altered drug metabolism.

Since no diagnostic tools are available to precisely determine a child's cognitive level below an IQ of 40, children were included if their cognitive development was estimated to be below a calendar age of 2 years old (equivalent IQ 0-20). In addition we assessed language development using 'minimal speech standards' (Goorhuis and Schaerlaekens 2000). Standards for PCI included the ability of making sounds but not expressing oneself through words. None of the children's abilities exceeded these standards, indicating that the participants were unable to express pain verbally.

The origins of the included children's cognitive impairments were diverse (Table 1).

Table 1. Origin of the children's cognitive impairment (N=48)

Origin of cognitive impairment		N	% of total N
Congenital/ metabolic anomalies:			
Syndrome:			
- *chromosomal disorder*		6	12.5
- *identified metabolic disorder*		8	16.6
- *without identified cause*		3	6.2
Major structural cerebral abnormalities		10	20.8
Infections		1	2.0
Metachromatic leukodystrofy		1	2.0
Intra-uterine asphyxia		1	2.0
	Total	**30**	**62.5**
Cognitive impairment during birth:			
Perinatal asphyxia		4	8.3
Posthypoxic encephalopathy		2	4.1
	Total	**6**	**12.5**
Cognitive impairment after birth: (between 2 months and 4 years old)			
Encephalopathy		3	6.2
Meningo- or pneumococcal sepsis		3	6.2
Near-drowning-syndrome		1	2.0
	Total	**7**	**14.5**
Origin remained unknown			
	Total	**5**	**0.4**

Physiological impairments and chronic physical diseases were frequently documented (Table 2). All children underwent one or more surgical procedures (mean=4.2; range=2-7) under general anesthesia during their lifetime up till the time of the present study.

Table 2. Chronic physical disorders in children with cognitive impairment (N=48)

Type of disorder	N	% of total N
Epilepsy	41	85.4
Spastic tetraplegia	34	70.8
Poor vision or blindness	31	64.5
Fed by nasogastric / gastrostomy tube	29	60.4
Constipation	26	54.1
Gastroesophageal reflux	24	50.0
Bronchial and pulmonary infections	21	43.7
Ear Nose Throat infections	21	43.7
Hearing loss or deafness	21	43.7
Malnutrition	19	39.5
Scoliosis	19	39.5
Urinary tract infection	16	33.3
Asthma, COPD*	9	18.7
Hyper- and hypothermia	6	12.5
Cardiac disease	5	10.4
Anaemia	3	6.2
Metabolic disorder	3	6.2
Eczema	2	4.1
Hypersensitivity perioral	2	4.1
Hemangiomas	1	2.0
Overweight	1	2.0

*COPD=chronic obstructive pulmonary disease
Note: more than one disorder per child is possible (mean=7.1 disorders per child, range=1-12)

Measures

Checklist Pain Behavior (CPG)

The CPG contains 21 observable behaviors which were found to be sensitive to post-surgical pain in children with PCI in a preceding study (Terstegen et al. 2003). Using behavioral descriptions these behaviors were rated on a 4-point scale, with 0=never shown, 1=sometimes shown, 2=often shown, 3=always shown. Based on its content, each of the 21 behaviors was assigned to one of four distinctive subscales derived from principal component analysis: Facial expression, Social behavior/Mood, Vocalisation and Physiological signs. Scale scores are computed by adding the scores in these subscales, as well as a total scale score based on all items in the Checklist.

Statistical analyses provided evidence that the CPG has good psychometric properties when used for assessing post-surgical pain in children with PCI (Terstegen et al., submitted) with excellent interrater reliability (ICCs ranging from .92 to .97) and adequate internal consistency (alphas ranging from .56 to .89). Both the subscale scores and the total scale score were shown to be sensitive to the presence and severity of pain, changes in post-

surgical pain over time, and the difference between post-surgical pain and acute directly induced pain in these children.

Visual Analogue Scale (VAS)

A 10 cm VAS was used to obtain a general rating of the severity of the children's pain during the observed episodes (Huskisson 1974). Both researchers and nurses rated the child's pain from 0 (no pain) to 10 (worst pain). A VAS ≥ 4 is considered to indicate the presence of pain (Buchholz et al. 1998; Van Dijk et al. 2002). In a previous study we reported a good validity of the VAS for indicating pain intensity in children with PCI. Intraclass correlation (two-way mixed effect model) for VAS ratings from 2 researchers over 333 episodes was good (.79).

Procedure

The study was approved by the Medical Ethical Committee of the Erasmus Medical Center. Parents were informed about the use of the CPG in clinical practice. Information on medical diagnoses associated with the children's cognitive impairment level was obtained from their medical record. Information on speech behavior was gathered by written descriptions of the child's verbal capacities in medical records and through interpretation of observations of the child at the bedside.

For reliable use of the CPG, the nurses attended a two-hour training session during which the CPG was explained by means of videotaped behavior of children who were admitted to the children's hospital. Because 3 to 8 nurses were trained simultaneously, discussions were used to solve possible misinterpretations of the description of some behaviors included in the CPG. After this theoretical training, each nurse completed 5 CPG assessments (including a VAS rating) directly at the bedside of a child with the trainer (research nurse) or an already qualified colleague. The scores were compared and discussed. Nurses were trained in their own environment (hospital or healthcare facility) for both the theoretical and practical part. When interrater reliability was adequate, the nurse gained her autonomous competence to score with the CPG. Altogether, 41 nurses used the CPG autonomously and together conducted a total of 348 observations (186 observations in the healthcare facility, 162 observations in the children's hospital). After the training session nurses were asked for their opinion about the content of the CPG and its feasibility in clinical practice.

In this study nurses were trained how to observe and interpret the child's behavior within the framework of the use of the CPG in clinical practice for which 2-minute observations were performed in both painful and non-painful situations, chosen randomly. Children (n=27) were admitted to the neurology ward of the children's hospital for several reasons (status epilepticus, choreadystonic movements, observation developmental retardation/crying without obvious reason, near drowning, MRI under anaesthesia, EEG, subdural haematoma, PH-measurement/gastro-oesophageal reflux). All 21 children living in a residential healthcare facility were observed during potential non-painful moments (e.g., when lying on a aquarius

mattress, listening to music or watching bright colored lights in their wheelchair or resting in bed) and potential everyday painful moments (e.g., pain during and after eating/swallowing due to oesophageal reflux, abdominal pain due to constipation, during and after an epileptic seizure, when bathing and clothing the child due to contractures, hip problems). After each observation the CPG and the VAS were completed.

Statistical analyses

Data was analyzed using SPSS 10.1. Because of the highly skewed distribution of the data, non-parametric tests were used; alpha was set at .05 for all tests.

First, to test the reliability of the use of the CPG after training, interrater agreement between the nurse and the trainer (or an already qualified colleague) was calculated using linearly weighted Cohen's Kappa's, which was regarded acceptable from .60 (Fleiss 1981).

Second, internal consistency of the CPG scales was computed using Cronbach's alpha based on all observations. Interrater agreement for everyday use of the CPG and the VAS was calculated on 25 paired observations (a researcher and a nurse), using Intraclass Correlation Coefficients (ICC) based on the two-way mixed effect model.

Further, we examined if the CPG can be used in a reliable way by different types of observers. Possible differences in observing children with PCI using the CPG between nurses who are and who are not (or less) familiar with the observed child, were studied by comparing the ICCs from all nurses who scored together with a researcher.

Third, for assessing convergent validity, the sensitivity of the CPG to everyday, more chronic pain was tested by comparing CPG ratings with VAS ratings of 0-3.9 to those with VAS ratings ≥4, using the Kruskal-Wallis test with Fisher's exact test for significance. Furthermore, to investigate to which extent CPG scores are predictive of VAS ratings in everyday pain situations, univariate regression analyses and linear regression analyses were conducted, in which the CPG subscale scores were entered as predictors.

In order to examine construct validity, we assessed the sensitivity and the specificity of the CPG for detecting everyday pain, using Receiver Operating Characterics (ROC) curves, computed non-parametrically (Zweig & Campbell 1993).

Results

First, for all 41 nurses who were trained to use the CPG autonomously, Kappa's indicating reliability to criterion were satisfactory (Table 3). Mean Kappa for all nurses was .83, for nurses in the children's hospital .82 and .86 for nurses working in the residential healthcare facility.

Table 3. Distribution of Kappa values (%) for nurses acquainted (healthcare facility, n=18) versus non-acquainted (children's hospital, n=23) with the observed children

Kappa	Healthcare facility	Children's hospital
60-70	0%	18.2%
71-80	40.7%	22.7%
81-90	45.0%	36.4%
91-100	14.3%	22.7%

Second, internal consistency of the CPG total score was excellent for observations in the children's hospital (Cronbach's alpha .91) and good for observations completed in the healthcare facility (.81). As indicators in the subscale 'Social behavior/Mood' were never scored by nurses from the healthcare facility, alpha's in this group were calculated for the three remaining subscales. Cronbach's alpha for the two participating groups together was .87. Alphas for each subscale for the different groups are shown in Table 4.

Table 4. Internal consistency (α) for the CPG subscales and total scale

CPG scales	All observations (n=348)	Residential healthcare facility (n=186)	Children's hospital (n=162)
Facial expression	.81	.72	.86
Social behavior/ Mood	.43	**	.43
Vocalisation	.54	.49	.60
Physiological signs	.64	.47	.72
Total scale	.87	.81	.91

** indicators in this subscale were never scored by nurses from the healthcare facility

When comparing interrater agreement (ICC) from nurses who are acquainted (healthcare facility, n=18) and who are not acquainted (children's hospital, n=23) with the observed children, we found that for nurses working in the healthcare facility the ICC was .74 and for nurses who observed children in the children's hospital the ICC was .82 (Table 5). For both participating groups together, the ICC of the CPG total scale was good (.80).

Table 5. Interrater agreement (ICC) for the CPG subscales and total scale

CPG scales	All observations (n=25)	Residential healthcare facility (n=14)	Children's hospital (n=11)
Facial expression	.63	.43	.70
Social behavior/Mood	.75	**	.75
Vocalisation	.90	.85	.89
Physiological signs	.95	.90	.97
Total scale	.80	.74	.82

** Zero variance, indicators in this subscale were never scored by nurses from the healthcare facility

Third, convergent validity was assessed by relating the CPG scores to VAS ratings on all observations. For both groups of nurses, CPG subscale and total scale scores were significantly higher for observations with VAS ratings ≥ 4 (CPG total score range from 2 to 52; mean 17.9) compared to observations with VAS ratings 0 – 3.9 (CPG total score range from 0 to 27; mean 3.8). Univariate regressions indicated that VAS ratings increased significantly with increasing CPG subscale scores and total scale scores. Linear regression analyses (stepwise) showed that all subscales, except for 'Social behavior/Mood', contributed independently to the prediction of VAS scores. In Table 6 standardized Beta coefficients are shown, which represent the strength of the CPG's prediction on VAS ratings. These analyses showed that the largest proportion of variance in VAS ratings, independent of the influence of the other variables, was explained by subscale 'Facial expression' (58%).

Table 6. Regression analyses of VAS on CPG scale over all episodes

CPG subscales	R^2 obtained from univariate regression	Beta's obtained from multivariate regression
Facial expression	0.58	0.43*
Social behavior/Mood	0.26	0.05
Vocalisation	0.37	0.30*
Physiological signs	0.34	0.19*
Total scale	0.79	-

* P<0.0001
Adjusted R^2 and standardized Betas are reported

Further, for the CPG's construct validity, ROC analyses showed a sensitivity of 0.71 and a matching specificity of 0.68 by a CPG score of 18 or higher. The level of the area under the curve (AUC) was .94. Thus, a CPG score ≥ 18 turned out to be a fair indication of pain in children with PCI. For nurses in the children's hospital who were not (or less) familiar with the child, sensitivity and specificity were slightly lower (.67 and .63 respectively, AUC .93) compared to nurses from the healthcare facility who were acquainted with the observed children (.72 and .69 respectively, AUC .94).

Clinical implications

During the use of the CPG in daily clinical practice, we received feedback from several nurses participating in this study and new questions emerged concerning the content and the feasibility of the CPG. In this section we evaluate the CPG with respect to these topics based on the nurses' opinions. Some of the CPG indicators were questioned to be relevant when assessing pain in these children. First, the relevance of 'protect one's sore body part' from the subscale 'Social behavior/Mood' was questioned, because it was only scored 6 times within 348 observations, but all in situations with presence of pain. Second, several behaviors included in the CPG were considered not to be pain specific (e.g., 'facial restlessness, tics', 'corners of the mouth downwards', 'nasolabial furrow' and 'tears') and were all described as individual characteristics of the child. However, in a previous study we reported that 16 out of the 21 behaviors in the CPG were found sensitive to pain but were not specific pain indicators

for children with PCI. The four behaviors mentioned in the current study were among these 16 behaviors. Third, the subscale 'Physiological signs' was suggested to need further attention because approximately a third of our sample suffered from airway obstructions. This causes rhonchi which can be mistaken by the indicator 'breath holding' in this subscale.

An important complicating factor when using the CPG was found in the complexity of clinical practice, which made changing the daily routines of the nurses in a healthcare facility, though less than in a children's hospital according to some participants, rather difficult. Some nurses who expressed scepticism about working with the CPG were convinced that knowing a child for many years does lead to well interpreted pain behaviors without using a pain assessment scale. However, most participating nurses indicated that structured and unambiguous observations are essential in order to assess and monitor pain adequately and, consequently, improve prescription and administration for analgesia in children with PCI.

Discussion

For all nurses who were trained to use the CPG autonomously, Kappa's were generally shown to be very satisfactory. This finding indicates that it was possible to train the nurses to observe pain related behaviors in everyday pain events using the CPG, in a reliable way and thus feasibility of the CPG turned out to be satisfactory.

The CPG turned out to be internally consistent when used in both participating institutions, although subscale 'Social behavior/Mood' showed a low Cronbach alpha. Interrater reliability was adequate for both participating groups of nurses, although they were somewhat higher for observations completed in the children's hospital (.82) compared to those completed in the healthcare facility (.74). Further, the highest percentage of kappa values (indicating interrater reliability to score autonomously with the CPG) was for both groups between .81 and .90. Kappa values for nurses from the children's hospital started from .60 while those from the healthcare facility started from .71.

The convergence between CPG scores and VAS ratings was satisfactory: VAS ratings increased significantly with increasing CPG subscale and total scores. VAS ratings were significantly predicted by the CPG subscale scores, except for subscale 'Social behavior/Mood'. In a previous study (Terstegen et al., submitted) this same subscale also turned out to be a rather poor predictor of VAS ratings. Further, analyses showed that the largest proportion of variance, independent of the influence of the other variables, was explained by subscale 'Facial expression' (54%).

Additionally, the CPG was further tested on its sensitivity and specificity. A value of .94 for the area under the curve (AUC) means that for 94% of the time, a random selection from the positive group (those who are correctly diagnosed as painful) will have a greater score on the CPG than a random selection from the negative group (those who are incorrectly classified as painful). Considering the fact that a test (here the CPG) with an AUC value of .50 can be interpreted as performing no better than change and value 1.00 as perfectly accurate (Zweig & Campbell 1993), it can be concluded that the CPG is very acurate in detecting pain.

The finding that sensitivity was better than specificity for both groups of nurses is appropriate for an observational pain assessment scale designed to supplement clinical judgement and to alert healthcare professionals to the possibility that a child is in pain (Breau et al. 2002b).

Further, nurses from the children's hospital who were not (or less) familiar with the child, had slightly lower sensitivity and specificity values compared to the nurses from the healthcare facility who were acquainted with the children, indicating that knowing the child and its customary behavior may improve the sensitivity and specificity of the CPG. Although the CPG, as indicated by ROC curves, correctly classified 71% of the children in pain, one should stay alert towards the child's customary behavior before observing their pain behavior. We recommend therefore to first apply the CPG at a moment (or, if possible several moments) when the child is not in any pain or distress, and to use this obtained score as a reference score.

In summary, the results from this study provide evidence that the CPG can reliably detect everyday pain, related to physiological impairments and diseases and thus in most cases more chronic pain, in children with PCI in both a hospital setting and in a healthcare facility. Nurses who were unfamiliar with these children were also able to detect signs of pain using the CPG.

Future research

It seems that knowing the child, and thereby having the capacity to recognize changes in their behavior, is an important issue when assessing pain in these children. To explore this topic more thoroughly, paired observations from observers who are and are not familiar with the child should be gathered and further evaluated.

When an appropriate and valid pain assessment scale is available, the development of an algorithm of analgesia, using the appropriate cutoff points of the CPG and the VAS and principles of pre-emptive analgesia in this specific group of children, is an essential step to minimize pain in children with PCI. Currently we are exploring the construction of an initial algorithm based on combined information from the CPG and the VAS. In the near future, CPG scores will be linked to analgesic drug regimes for an appropriate analgesic treatment in these children. Finally, the effect of combining a validated pain assessment scale (CPG) with an algorithm for pain management needs to be constructed and evaluated in daily practice.

Some aspects of the feedback concerning the content of the CPG which was given during the use of the CPG in daily clinical practice, should be addressed in future research. Especially the indicators from the subscale 'Social behavior/Mood' should be re-examined because they seem to be identifiers of anxiety, anger or sadness rather than pain. Further, the relevance of the indicator 'protect one's body part' from this subscale was questionable, because it was only scored 6 times in 348 observations. The subscale 'Physiological signs' needs further attention because approximately a third of our sample suffered from airway obstructions which may blur an adequate interpretation of observing respiration of these children.

Chapter 7

General discussion and conclusions

General discussion and conclusions

Although children with profound cognitive impairment (PCI) generally go through many episodes of pain, caused by somatic conditions such as gastro-oesophageal reflux, contractures, hip problems and epilepsy, their pain often goes unrecognized. There are two possible reasons: the children's cognitive, verbal and motor limitations narrow their expressive repertoire and, on the other hand, caregivers may fail to interpret the displayed behavior as a sign or expression of pain.

The main aims of this study were to investigate indicators of pain and to develop an observational scale – including those indicators that were found to be valid – for the assessment and monitoring of post-surgical pain in children with PCI. The purpose of this final chapter is to integrate the results and conclusions from the previous chapters and to give suggestions for future research.

Cortisol as physiological parameter to assess pain

As an objective criterion for the assessment of pain in children with severe or profound cognitive impairment is still lacking, we examined in this study whether salivary cortisol levels might serve as an objective parameter to be used in addition to behavioral observations. Cortisol has been shown to be a marker of stress in normally developed children (Kiess et al. 1995). The question to be answered was whether diurnal cortisol levels in children with CI are comparable to those in normally developed children.

We, therefore, measured saliva cortisol levels three times a day (morning, afternoon and evening) in 49 children with severe and profound cognitive impairment and 44 children without cognitive impairment (CI), and determined mean levels and the circadian rhythm of cortisol secretion.

Although numerous sources of stress are commonly present in cognitively impaired children, we found no significant differences in mean cortisol levels between the children with CI and the normally developed children. Across all measurements, the mean cortisol levels for the cognitively impaired children were on average 0.5 nmol/L higher, and showed more variation than in the group of normally developed children.

On average, the cortisol levels of the children with CI showed a circadian rhythm similar to that of normally developed children. However, when looking at intra-individual variation of cortisol levels, 20 of the 49 cognitively impaired children (41%) did not show the expected circadian rhythm, i.e. either their cortisol production did not decrease from early morning to evening, or cortisol levels were low in the morning and peaked in the afternoon or evening. Fourteen of the 44 normally developed children (32%) also showed no circadian rhythm, in all cases because cortisol levels peaked at 12:30 hrs. However,

the absence of circadian rhythm did not significantly differ between the two groups (X^2= 0.68, α > 0.05).

Our findings suggest that mean salivary cortisol levels for normally developed children are also representative for children with severe or profound cognitive impairment. However, as a sizeable portion of the children – two fifths of the children with CI and one third of the normally developed children – do not show the expected circadian curve in cortisol levels, individual baseline data seem essential for the assessment of cortisol reactivity to acute stimuli. Finally, although it may be useful to assess salivary cortisol for research purposes, it is not suitable for everyday pain assessment.

Developing an observational scale: the Checklist Pain Behavior

The studies described in chapters 3 and 4 aimed at identifying behavioral post-surgical pain indicators shown by children with profound cognitive impairment (PCI) and at constructing an observational scale that could assist parents and professional caregivers in assessing and monitoring pain behavior in this specific group of children. First, we conducted a qualitative study to gain insight into non-verbal pain expressions in children with PCI, by means of focussed semi-structured interviews (n=29) with parents and caregivers, and by observations (n= 32) of behavior during several possibly painful events. This qualitative study yielded 169 possible pain indicators. Next we added other pain expressions mentioned in the literature, which resulted in an item pool of 209 indicators [van Dongen et al, submitted]. These indicators were used in a pilot study in 15 children with PCI who had to undergo any of a variety of surgical interventions. This yielded a number of 71 expressions that were discarded because (1) they turned out to be not suitable for clinical observation of pain or (2) overlapped others. Finally, the remaining 138 indicators were organised into six behavioral categories: Facial expression, Motor behavior, Social behavior/Mood, Attitude towards sore body part, Vocalisation and Physiological signs.

Next, a cross-sectional study was undertaken to assess post-surgical pain in 52 children who were admitted to the Erasmus-MC Sophia Children's Hospital for an elective surgical procedure. In this study we further tested the 138 behavioral indicators from the item pool composed in the preceding qualitative study. Twenty-three indicators turned out to be sensitive to post-surgical pain in children with PCI. The most sensitive indicators belonged to the category 'Facial expression' (47.8%), probably because most children (84.6%) suffered from severe motor impairments and therefore were hardly able to express themselves by other behaviors than facial expressions. The same applies to most indicators within the category 'Attitude towards sore body part'. Again, lack of motor capacities prevented most children from showing behaviors within this category, except for the item 'protects sore body part', which appeared to be sensitive to post-

surgical pain. Furthermore, four physiological indicators significantly differed before and after surgery, but two of which indicated changes in respiration which were corrected as changes directly resulting from anesthesia. The indicator 'panics, panic attack' was the most significant indicator from the category 'Social behavior/Mood', probably because pain often goes together with anxiety. Finally, four indicators were categorised within 'Vocalisation'.

It is remarkable that so few from an initially large pool of potential indicators of pain proved to be sensitive to pain. This suggests that, in the absence of a clear standard for pain, caregivers as well as (para)medical personnel tend to attribute pain to a large number of behaviors that apparently are not at all related to pain.

The remaining 23 items were further tested for reliability and validity, and finally served to construct an observational scale - the Checklist Pain Behavior (CPG). PRINCALS analysis resulted in a solution with four optimal dimensions that together explained 59% of the variance in children's scores. Based on the 23 items within these dimensions, four subscales were constructed: Facial expression, Social behavior/Mood, Vocalisation and Physiological signs.

The total scale had an excellent interrater reliability (0.97) and internal consistency (0.89). While the total scale as well as the subscales 'Facial expression', 'Vocalisation' showed adequate internal consistencies over all episodes, the subscales 'Social behavior/Mood' and 'Physiological signs' showed somewhat lower consistencies (alphas were 0.56 and 0.62, respectively). However, as their explained variance is adequate (13.9% and 17.3%, respectively), these subscales appear to contribute substantially to the total scale score. Further, the presence of two indicators, 'eyes almost closed' in the subscale Facial expression and 'accepting comfort' in the subscale Social behavior/Mood resulted in lower Cronbach alphas for these subscales in almost all episodes. As they also showed low component loadings in PRINCALS analyses (0.29 and -0.25 respectively), they were excluded from further analyses.

Correlations between the four subscales were moderate to high, indicating that the subscales are not independent, but that overlap is limited. The correlations between the subscales and the total scale showed that 'Facial expression' contributes by far the most ($r=0.94$) to the total score variance.

This finding is consistent with previous studies reporting the importance of facial response as an indicator of pain in these children. Stallard et al (2002b) found 'screwed up or distressed looking face' to be the strongest individual predictor. Breau et al (2002a) showed that only the Facial subscale of researchers' NCCPC-PV significantly predicted nurses' ratings on the Visual Analoge Scale (VAS).

Examining the extent to which the subscales and the total scale will *discriminate between absence versus presence of pain*, we found that mean scores on all CPG subscales and the total scale differed significantly before and after surgery. On the total scale, 98% of the children scored higher after surgery. On 'Facial expression', the proportion of children with higher scores after surgery was 94%, on 'Social behavior/Mood' 69% (26% were ties), on 'Vocalisation' 79% (13% were ties), and on 'Physiological signs' 85%.

Most indicators were present before as well as after surgery. For example, 'moaning, groaning' was shown before surgery by 11.2% of the children and after surgery by 38.2%. However, while for some children this might be customary behavior because they showed this behavior at comparable rates before and after surgery, for others it seemed an expression of pain because they displayed this behavior more frequently after surgery than they did before. This ambiguity indicates the importance of having knowledge of the individual child's customary behavior when not in pain.

Comparing regular post-surgical episodes (30/60 minutes, 3 hrs, 24 hrs and 48 hrs after extubation) *with the more acutely painful, surgery-specific episode* (directly induced pain, depending on type of surgery), we found mean scores on the subscales 'Facial expression' and 'Vocalisation' and mean scores on the total scale to differ significantly between the regular post-surgical episodes and the surgery-specific episode.

Studying the sensitivity of the CPG for *changes in pain over time*, we found that scores obtained from the four post-surgical episodes all declined from 30/60 minutes till 48 hours after surgery. Significant differences over time were found for the mean scores on the subscales 'Facial expression', 'Vocalisation', 'Physiological signs' and the total scale, but not for the subscale 'Social behavior/Mood'. An explanation could be that some of the behaviors included in this scale ('panicking', 'resistant/rebellious', 'protecting sore body part') might be more sensitive to anxiety or anger rather than pain.

Finally, when studying the CPG's sensitivity to *differences in severity of pain*, we found that CPG scores on all subscales as well as on the total scale increased with increasing painfulness of the surgical procedure. Significant differences were found for 'Facial expression', 'Vocalisation' and the total scale for low versus high pain intensity.

The convergent validity proved to be good. As a valid tool for children with PCI was still lacking at the start of this study, we used the VAS as a criterion measure, assuming a score equal to or above 4 to reflect pain, like in normally developed preverbal children (Buchholz, Karl et al. 1998; Van Dijk et al. 2002). Mean CPG scores for all subscales were significantly different for episodes with VAS ratings < 4 versus VAS ratings ≥ 4. Univariate regressions indicated that VAS ratings increased significantly with increasing CPG subscale scores and total scale scores. The multivariate regression analysis using the four subscales as predictors, showed that subscales 'Facial expression', 'Vocalisation'

and 'Physiological signs' contributed independently to the prediction of VAS scores. This makes clear that each of these CPG components is related to an overall pain rating. The exception for the subscale 'Social behavior/Mood' confirms our previous hypothesis stating that behaviors included in this scale might be more indicative of anxiety or anger rather than of pain, as indicated by VAS ratings. The regression analyses also showed that while 49% of the variance in VAS scores was predicted by the combined CPG subscales, 53% was predicted by the CPG total score alone.

In sum, in this study we identified a core set of behaviors displayed by children with PCI to express post-surgical pain. We provided evidence that the CPG reliably assesses post-surgical pain in children with PCI, and that it is sensitive to painful versus non-painful conditions, to post-surgical changes in pain over time, as well as to induced intense pain of short duration.

Differences in pain behavior between children with and without profound cognitive impairment

The manner in which children with PCI express their pain differs from that of normally developed children (Gilbert-MacLeod et al. 2000; Breau et al. 2001). However, the question in which way both groups differ, i.e. in type and intensity of pain behavior, still remained to be answered. This is an important issue, since the answer might give us additional clues to justify the use of a separate pain instrument for the children with profound cognitive impairment.

In order to investigate the extent to which CPG indicators in response to painful stimuli are more frequently, or even exclusively, seen in children with PCI, we applied the CPG in a standardised post-surgical setting – i.e. after ENT-surgery – and tested its reliability in both groups as well.

The internal consistency of the CPG proved to be better in the group of children with PCI, except for the subscale 'Social behavior/ Mood'. The exception can be explained by the possibility that children with PCI, compared to normally developed children, have less motor capacities to show behaviors in this subscale.

Four indicators, i.e. 'facial restlessness, tics', 'frightened fearful look', 'trembling chin' and 'penetrating sounds of restlessness', were not observed in children without PCI, and thus seem to be behaviors specific for children with PCI. Sixteen behaviors for children with PCI - and four for children without PCI - were present before surgery (VAS = 0) and appeared with increased frequency or intensity after surgery (VAS $\geq$ 4). This suggests that certain behaviors in children with PCI are sensitive to pain, but nevertheless are not specific pain indicators. An explanation might be that children with PCI generally have limited possibilities to express themselves, which leaves them few alternatives in "choosing" behaviors to express their pain. Some of the few behaviors they are capable

of showing, will be used both in painless and in painful situations, but in the latter case more frequently or intensely. This, again, underlines the importance of applying the CPG before surgery when the child is not in pain, and use the outcome as a baseline measure.

In sum, it is clear that for children with PCI different pain behaviors are indicative of pain compared to children without PCI. Furthermore, certain behaviors in response to painful stimulation were similar to those of normally developed preverbal children, although some were displayed more subtly (e.g., they cry softly rather than loudly), and others, conversely, more frequently or intensely (e.g., tense face, faltering respiration).

Applicability of the CPG in everyday, more chronic pain situations

In chapter 6 we reported a study testing whether the Checklist Pain Behavior can be used for children with PCI to assess everyday, more chronic pain related to their physiological impairments and diseases, in both a hospital setting (in non-surgical situations) and in a residential healthcare setting, i.e. the accommodation where they live.

Since up to this point the CPG had been tested only in a research context and only for its sensitivity to post-surgical pain, the results from this study were considered paramount to justify its use in everyday practice. This part of the investigation addressed the reliability and validity of the CPG when used by trained nurses, as well as its sensitivity to everyday pain.

We studied if nurses caring for children with PCI in different settings can be trained to reliably score the children's behavior at the bed-side using the CPG. Also, as it was still unknown whether CPG scoring is influenced by knowledge of a child's customary behavior, we investigated possible differences between nurses who are familiar with the observed child (and with children with PCI in general) and nurses who are not (or less) familiar with the child.

The nurses who were trained to use the CPG autonomously, generally reached satisfactory Kappa values. This finding indicates that nurses can be trained to reliably observe pain related behaviors in everyday pain events. While Kappa values for nurses in the children's hospital started from 0.60, those for nurses in the healthcare facility started from 0.71. In both groups, however, most values were between 0.81 and 0.90.

In both settings the CPG's interrater reliability of bed-side observations was adequate, and internal consistency was good. However, Cronbach's alpha for the subscale 'Social behavior/Mood' was rather low for the group of nurses in the children's hospital, and indicators in this subscale were never scored by nurses in the healthcare facility. It seems that indicators within this subscale might be interpreted as identifiers of anxious, angry or sad children rather than of children in pain. Anxiety often goes together with

pain. It is difficult to prepare children with PCI for the events involved in admission to a hospital. Panic was observed most often directly after surgery, i.e. on awaking from anesthesia, and in the directly pain-inducing episode. The same holds true for resistant, rebellious behavior, which often was shown together with panicking behavior, if the child was physically capable of showing resistance. It can be concluded that the internal consistency of the Social subscale was rather low, as some of its indicators seem to represent other emotions, such as anxiousness independently or in combination with pain.

The convergence between CPG scores and VAS ratings was good. VAS ratings were significantly predicted by the CPG subscale scores, except for the subscale 'Social behavior/Mood'. In Chapter 3 this subscale was also described as a rather poor predictor of VAS ratings. This, again, can be explained by the fact that this subscale probably represents other emotions besides pain.

For testing construct validity, we used ROC analyses to examine the sensitivity and the specificity of the CPG regarding everyday pain. We found a sensitivity value of 0.74 and a specificity value of 0.68 for CPG scores of 18 or higher. The value for the area under the curve (AUC) was 0.94, which means that for 94% of the time, a random selection from the positive group (those who are correctly diagnosed as painful) will score higher on the CPG than does a random selection from the negative group (those who are incorrectly classified as painful). Considering that a test with an AUC value of 0.50 can be interpreted as performing no better than chance, and one with an AUC value of 1.00 as perfectly accurate (Zweig & Campbell 1993), the CPG turned out to be very accurate in detecting pain.

In summary, these results provide evidence that the Checklist Pain Behavior can be used to assess everyday, more chronic pain in children with PCI, both in a hospital setting and in a residential healthcare setting. Nurses who were unfamiliar with these children were also able to detect signs of pain using the CPG. This study demonstrates that it is possible to successfully train groups of nurses in both settings to apply the CPG for observing pain related behaviors in everyday pain events in a reliable way.

Comparison with other studies

Validated instruments to assess pain in children with PCI have not become available until recently. Since the start of the current investigation in 1998, several other studies were performed in Canada, New-Zealand and England addressing pain behavior in children with severe or profound cognitive impairment. The results of this thesis are partially consistent with the findings of these other studies on this topic, and some of these results will be compared in this section. Stallard et al. (2002a, 2002b) developed the Pain Indicator for Communicatively Impaired Children (PICIC). First, they asked

parents to identify six core cues as pain indicators in their child (i.e. crying with or without tears'; 'screaming, yelling, groaning or moaning'; 'screwed up or distressed looking face'; 'body appears stiff or tense'; 'flinches or moves away if touched'; 'difficult to comfort or console'). Next, they obtained data from 67 caregivers relating to 49 severely cognitively impaired children. The caregivers were asked to score the six core cues on a four-point rating scale during a set observation period. Five of these cues (all except 'crying with or without tears') turned out to be significantly associated with the presence of pain as indicated by the caregivers. These five were initially also found to be sensitive to pain in our study, but we later discarded one of them ('difficult to comfort or console') for its very low loading on the subscale (0.25), as indicated by PRINCALS analyses.

Further, Breau (2002a) developed a pain rating scale, the Non-Communicating Children's Pain Checklist-Revised for several pain settings (e.g., injury; chronic condition; illness). In addition, an adapted version for post-surgical pain, the NCCPC-PV, was shown to be sensitive to surgery-induced pain (Breau et al. 2002b). There are several differences between Breau's studies and ours. First, while Breau and colleagues started the development of the NCCPC with 31 possible pain behaviors extracted from semi-structured interviews with primary caregivers, we derived 206 behaviors from several sources (Chapter 3). This way, we made a complete inventory of all possible candidate pain behaviors displayed in this specific group of children, excluding the possibility of missing important pain behaviors as much as possible (e.g., eight behaviors included in the CPG after intensive testing were never included in the NCCPC and thus not tested on their possible sensitivity to pain). Second, not unimportantly, the groups of participants in both studies differ. In our study, it was evident that children with PCI who also suffered from profound motor impairments displayed different behavior when in pain compared to children with PCI with mild or severe motor impairments. For this reason, seven behaviors described in the NCCPC were excluded from the CPG because our participants did not have the motor capacities to display these behaviors.

Although the NCCPC-PV is sensitive to painful versus non-painful conditions, it had not yet been shown to be able to monitor changes in post-surgical pain. Ideally, a pain measure should not only simply detect pain at given moments, but also monitor the presence of pain over time. Having data from six post-surgical observational episodes at our disposal, we were able to assess the construct validity more thoroughly by comparing post-surgical mean scores over time, and found the CPG to be sensitive to changes in pain behavior over time. With pain slowly decreasing after surgery, CPG subscale scores declined, except for the subscale 'Social behavior/Mood', which displayed nearly similar mean scores over time. Concerning the convergent validity, multiple regression analyses showed that three of the four CPG subscales contributed independently to the prediction

of VAS scores (by researcher and parents), whereas within the NCCPC-PV only the Facial subscale significantly predicted VAS scores (by nurses).

In sum, the CPG is the only observational scale which has been validated for both assessing and monitoring pain in children with PCI.

Further, our findings pertain to a group of children with profound cognitive impairment, whereas in the other studies groups consisted of children with severe cognitive impairment, or of a combination of children with severe and with profound cognitive impairment. Children with PCI are not able to display certain pain behaviors that are shown by children with mild, moderate or even severe CI. This is often due to the lack of motor capacities which tends to increase with a lower level of cognitive functioning. This lack of motor capacities limits their ability to express themselves physically. Besides, children with PCI show lower speech development, i.e. they are not able to express their pain verbally through one or more words, in contrast to mildly or severely cognitively impaired children. Although the decision to focus only on children with PCI reduced the number of children for which the observational instrument would be appropriate, the behaviors included in the CPG are optimally tuned to this specific group of children.

Strengths and limitations

This thesis encompasses a set of generic pain indicators which were thoroughly investigated on their sensitivity to pain in children with PCI using a structured and standardised method. A total of 72 children with PCI were included in this study: 52 for the study described in Chapters 3 and 4, and 20 other children for the study described in chapter 6. We decided to only analyse data from the children with profound cognitive impairment, and therefore excluded children with severe cognitive impairment who were originally included. Our decision was based on the fact that the severely and profoundly cognitively impaired children seem to differ in their capacities to express themselves, mainly with regard to motor and speech functions. Thus we were able to develop an observational scale which is reliable, valid and clinically useful in children with PCI in particular, both in a post-surgical setting and in everyday practice.

Futhermore, we made a complete inventory of all possible candidate pain behaviors displayed in this specific group of children and excluded the possibility of missing important pain behaviors as much as possible

We tested the CPG on its sensitivity for absence versus presence of pain, for changes in pain over time, for differences in pain severity, and for discrimination between post-surgical pain and acute, directly induced pain. The children in this study underwent any

of 18 types of surgery, and it can be concluded that all indicators are generally suitable for assessing post-surgical pain, regardless of the type of surgical procedure.

Finally, the results described in this thesis provide evidence that the CPG can be used to assess post-surgical pain as well as everyday, more chronic pain in children with PCI, both in a hospital setting and in a residential healthcare setting. Nurses who were unfamiliar with these children were also able to detect signs of pain using the CPG. This study demonstrates that it is possible to train nurses and caregivers to reliably use the CPG for observing pain related behaviors in children with PCI in everyday pain events.

Besides the above strengths, our study has a number of limitations.
Even though the participating children were observed very thoroughly and intensively, the small sample sizes could have blurred the outcomes of this study. E.g., the fact that we found four behaviors to be exclusive pain expressions for children with PCI – they were not observed in the group of children without PCI – might be a consequence of the small groups.

Next, although this study involved many observations, the number of actually observed painful episodes was rather small. This can be explained by the fact that the children were given standardised analgetics based on the concept of pain prevention as in other situations of tissue damage (Van Dijk 2000; Peters 2001; Bouwmeester et al. 2001). On the other hand, our findings on procedure-specific pain were quite convincing.

Furthermore, the children in this study differed in origin of cognitive impairment and showed a wide range of capacities and impairments within several developmental areas. Notwithstanding this heterogeneity, we detected a set of shared generic indicators which these children use to express pain. However, we report that each individual child may show an individual type of pain behavior and suggest, therefore, that assessment of pain behavior in these children always must be 'made-to-measure'; child-specific pain behaviors should be supplemented to the generic behaviors included in the CPG.

Finally, the lack of a gold standard or criterion to assess pain in this particular group of children might evoke discussion on the interpretation of pain behavior and the use of the CPG in clinical practice. However, we performed several analyses to test the convergent validity of the CPG, and are confident that the results indeed indicate satisfactory validity.

In sum, validation of any pain instrument requires repeated tests of validity and reliability across samples, settings and observers (Breau et al. 2002a). This thesis reports on the reliability and validity of the CPG for different aspects of pain – presence, severity, type, changes over time – in different samples and different settings – surgical, neurological, healthcare facility – and with different observers – nurses familiar and nurses unfamiliar with the observed children.

Clinical implications

The findings from this study indicate that, although the specific expression of pain may be very individual, there appears to be a set of shared generic indicators. Identifying this item pool of observable pain behaviors enabled us to develop an observational pain assessment scale that can easily and reliably be used in daily practice to assess and monitor pain in children with PCI. However, in view of the fact that these children do display differences in type of pain behavior (Giusiano & Jimeno 1995; Terstegen et al. 2003), we recommend, therefore, to first apply the CPG at a moment before surgery when the child is not in pain, and to use the obtained score, which reflects the child's customary behavior, as a reference score.

After being trained in applying the CPG, nurses and caregivers will be able to reliably assess post-surgical pain situations as well as everyday, more chronic painful situations in different settings – children's hospital or residential healthcare facility.

Implications for future research

Some aspects of the feedback concerning the content of the CPG, generated when it was applied in daily clinical practice (Chapter 6), should be addressed in future research. Especially the indicators from the subscale 'Social behavior/Mood' should be re-examined because they seem to be identifiers of anxiety, anger or sadness, rather than of pain. Children in pain often experience the former three types, and behaviors used to express these feelings are usually hard to distinguish from pain behavior. In order to be able to differentiate between behaviors resulting from these emotions, and to be certain that the CPG should detect pain behavior only, future research should compare the CPG with scales developed for assessing emotions such as anger or anxiety in children with PCI. To our knowledge, this issue has not been adressed so far.

The relevance of the indicator 'protects one's body part' from the subscale Social Behavior/Mood was questionable, because it was only scored six times in 348 observations. The subscale 'Physiological signs' needs further attention because approximately one third of our sample suffered from airway obstructions which may blur adequate interpretation of respiration of these children.

We detected a set of shared generic indicators used by these children to express their pain, but also reported that each individual child shows an individual type of pain behavior. Therefore, we propose to explore the development of individual pain profiles (Hunt 2001) containing child-specific pain behaviors supplementary to the generic behaviors included in the CPG.

Our results emphasise the heterogeneity of this group of children, which is consistent with the findings of other investigators (Vlaskamp 1997; Oberlander et al. 1999a; Collignon & Giusiano 2001; Breau et al. 2002a; Stallard et al. 2002). The heterogeneous group of children described in this thesis is an average representation of children with PCI in general, and this makes it possible to work on the development of an observational pain scale that is generic, rather than based on specific diseases or syndromes within this population. However, it would be of interest to study whether application of the CPG in children with specific diseases with a well-determined genetic cause would enable us to relate changes in pain scores to the underlying neuro-anatomical substrates. This might give us clues on the neuro-anatomical and neurophysiological basis of pain behavior in children with PCI.

It has been suggested that children with severe or profound cognitive impairment might be insensitive to pain. Recently, Breau et al. reported that caregivers with more learning regarding children with CI had a stronger belief that these children experience pain less than children without CI (Breau et al. 2003a). In addition, children with cognitive impairment of different origin might be differentially sensitive to pain. This issue of differential pain sensitivity may be addressed using the CPG in a broad sample of children with different levels and causes of cognitive impairment.

The standards of analgesia used in this study were based on pain studies in normally developed children (Bouwmeester et al. 2001, 2003, 2004; Van der Marel et al. 2003) and need to be translated to children with PCI. Now that a valid pain assessment scale is available, the next essential step to minimize pain in children with PCI is to develop an analgesic treatment algorithm based on the appropriate CPG and VAS cut-off points in the context of pre-emptive analgesia. Currently, we are exploring the construction of an initial algorithm based on combined information from the CPG and the VAS. In the near future, CPG scores will be linked to analgesic drug regimens in order to enhance analgesic treatment in these children. The effect of combining a validated pain assessment scale such as the CPG with an algorithm for pain management needs to be evaluated in daily practice. Perhaps this combination could become the integral standard of care in children with PCI.

References

References

A.A.P. and A.P.S. (2001). The assessment and management of acute pain in infants, children and adolescents. *Pediatrics, 108,* 793-797.

Anand KJS, Stevens B & McGrath PJ (2000). *Pain in neonates.* 2nd revised and enlarged edition. Amsterdam: Elsevier.

Berger J & Cunningham CC (1981). The development of eye contact between mothers and normal versus Down's syndrome infants. *Developmental Psychology, 17,* 678-689.

Biersdorff KK (1991). Pain insensitivity and indifference: Alternative explanations for some medical catastrophes. *Mental Retardation, 29,* 359-362.

Biersdorff KK (1994). Incidence of significantly altered pain experience among individuals with developmental disabilities. *American Journal of Mental Retardation, 98,*619-31.

Bollen K (1989). *Structural equations with latent variables.* New York: Wiley.

Born J, Hansen K, Marshall L, Moller M & Fehm HL (1999). Timing the end of nocturnal sleep. *Nature, 397,* 29-30.

Bouwmeester NJ, Anand KJS, Van Dijk M, Hop WCJ, Boomsma F & Tibboel D (2001). 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. *British Journal of Anaesthesia, 87,* 390-399.

Bouwmeester NJ, Van den Anker JN, Hop WCJ, Anand KJS & Tibboel D (2003). Age- and therapy-related effects on morphine requirements and plasma concentrations of morphine and its metabolites in postoperative infants. *British Journal of Anaesthesia, 90,* 642-652.

Bouwmeester NJ, Anderson BJ, Tibboel D & Holford NHG (2004). Developmental pharmacokinetics of morphine and metabolites in neonates, infants and young children. *British Journal of Anaesthesia, 92,* 208-217.

Breau LM, McGrath PJ, Camfield C, Rosmus C & Finley GA (2000). Preliminary validation of an observational pain checklist for persons with cognitive impairments and inability to communicate verbally. *Developmental Medicine and Child Neurololgy, 42,* 609-616.

Breau LM, Camfield CS, McGrath PJ, Rosmus C & Finley GA (2001). Measuring pain accurately in children with cognitive impairments: Refinement of a caregiver scale. *The Journal of Pediatrics, 138,* 721-727.

Breau LM, McGrath PJ, Camfield CS & Finley GA (2002a). Psychometric properties of the non-communicating children's pain checklist-revised. *Pain, 99,* 349-357.

Breau LM, Finley GA, McGrath PJ & Camfield CS (2002b). Validation of the Non-Communicating Children's Pain Checklist - Postoperative Version. *Anesthesiology, 96,* 528-535.

Breau LM, MacLaren J, McGrath PJ, Camfield CS & Finley GA (2003a). Caregivers' beliefs regarding pain in children with cognitive impairment: Relation between pain sensation and reaction increases with severity of impairment. *The Clinical Journal of Pain, 19,* 335-344.

Buchholz M, Karl HW, Pomietto M & Lynn AM (1998). Pain scores in infants: A modified infant pain scale versus visual analogue. *Journal of Pain and Symptom Management, 15,* 117-124.

Bufkin LJ & Altman R (1995). A developmental study of nonverbal pragmatic communication in students with and without mild mental retardation. *Education and Training in Mental Retardation and Developmental Disabilities, 30,* 199-207.

Cohen, J (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale: Lawrence Erlbaum.

Collignon P & Giusiano B (2001). Validation of a pain evaluation scale for patients with cerebral palsy. *European Journal of Pain, 5,* 433-442.

Coolens JL, Van Baelen H & Heyns W (1987). Clinical use of unbound cortisol as calculated form total cortisol and corticosteroid-binding globulin. *Journal of Steroid Biochemistry and Molecular Biology, 26,* 197-202.

Craig KD, Hadjistavropoulos HD, Grunau RVE & Whitfield MF (1994). A comparison of two measures of facial activity during pain in the newborn child. *Journal of Pediatric Psychology, 19,* 305-318.

Del Giudice E, Staiano A, Capano G, Romano A, Florimonte L, Miele E, Ciarla C, Campanozzi A & Crisanti AF (1999). Gastrointestinal manifestations in children with cerebral palsy. *Brain Development, 21,* 307-311.

Fanurik D, Koh JL, Schmitz ML, Harrison RD, Conrad TM (1999). Children with cognitive impairment: Parent report of pain and coping. *Journal of Developmental and Behavioral Pediatrics, 20,* 228-234.

Finley GA & McGrath PJ (1998), Introduction: the roles of measurement in pain management and research. In: Finley GA, McGrath PJ, editors. *Measurement of pain in infants and children*. Seattle,WA: IASP Press, p.104.

Fleiss JL (1981). Statistical methods for rates and proportions. Wiley, New York.

Gauvain-Piquard A, Rodary C, Rezvani A & Lemerle J (1987). Pain in children aged 2-6 years: a new observational rating scale elaborated in a pediatric oncology unit – preliminary report. *Pain, 31,* 177-188.

Gilbert-MacLeod CA, Craig KD, Rocha EM & Mathias MD (2000). Everyday pain responses in children with and without developmental delays. *Journal of Pediatric Psychology, 25,* 301-308.

Giusiano B & Jimeno MT (1995). Utilization of neural network in the elaboration of an evaluation scale for pain in cerebral palsy. *Methods of Information in Medicine, 34,* 498-502.

Goorhuis SM & Schaerlaekens AM (2000). *Handboek taalontwikkeling, taalpathologie en taaltherapie bij nederlands sprekende kinderen.* Leusden: De Tijdstroom.

Gorodischer R & Koren G (1992). Salivary excretion of drugs in children. *Developmental Pharmacology and therapeutics, 19,* 161-177.

Groenman N, Schuerman J, Vlaeyen J & Van Eek H (1986)Chronic pain. *Behavioral Medicine, 17,* 65-82.

Grunau R & Craig KD (1987). Pain expression in neonates: Facial action and cry. *Pain, 28,* 395-410.

Gunnar MR (1992). Reactivity of the Hypothalamic-Pituitary-Adrenocortical System to stressors in normal infants and children. *Pediatrics, 90,*491-497.

Gunnar MR, Brodersen L, Krueger K & Rigatuso J (1996). Dampening of adrenocortical responses during infancy: Normative changes and individual differences. *Child Development, 67,* 877-889.

Gunnar MR & Vasquez DM (2001). Low cortisol and a flattening of expected daytime rhythm: Potential indices of risk in human development. *Development and Psychopathology, 13,* 515-538.

Heim C, Ehlert U & Hellhammer DH (2000). The potential role of hypocortisolism in the pathophysiology of stress-related bodily disorders. *Psychoneuroendocrinology, 25,* 1-35.

Hadden KL & Von Baeyer CL (2002). Pain in children with cerebral palsy: Common triggers and expressive behaviors. *Pain, 99,* 281-288.

Horgan M & Choonara IA (1996). Measuring pain in neonates: An objective score. *Pediatric Nursing, 8,* 24-27.

Hunt A (2001). Towards an understanding of pain in the child with severe neurological impairment. PhD thesis. Oxford, UK.

Hunt A & Goldman A (2002). *Validation of the Pediatric Pain Profile: A behaviour rating scale to assess pain in children with severe neurological impairment.* 10th World Congress on Pain, San Diego, USA, A1703-P1251.

Hunt A, Mastroyannopolou K, Goldman A & Seers K (2003). Not knowing - the problem of pain in children with severe neurological impairment. *International Journal of Nursing Studies, 40,* 171-183.

Huskisson EC (1974). Measurement of pain. *Lancet, 2,* 1127-1131.

Kiess W, Meidert A, Dressendorfer RA, Schriever K, Kessler U & Konig A (1995). Salivary cortisol levels troughout childhood and adolescence: Relation with age, pubertal stage and weight. *Pediatric Research, 37, 502-506.*

Kirschbaum C & Hellhammer DH (1989). Salivary cortisol in psychobiological research: An overview. *Neuropsychobiology, 22,* 150-169.

Kirschbaum C (1989). Cortisol and behaviour: Adaptation of a radioimmunoassay kit for reliable and inexpensive salivary cortisol determination. *Pharmacology, Biochemistry and Behaviour, 34,* 747-751.

Kirschbaum C & Hellhammer DH (1994). Salivary cortisol in psychoneuroendocrine research: Recent developments and applicants. *Psychoneuroendocrinology, 19,* 313-333.

LaChapelle DL & Hadjistavropoulos T (1999). Pain measurement in persons with intellectual disabilities. *Clinical Journal of Pain, 15,* 13-23.

Landis JR & Koch GG (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159-174.

Larson MC, White BP, Cochran A, Donzella B & Gunnar MR (1998). Dampening of the cortisol response to handling at 3 months in human infants and its relation to sleep, circadian cortisol activity and behavioral distress. *Developmental Psychobiology, 33,* 327-337.

Lawrence J & Alcock D (1993). The development of a tool to assess neonatal pain. *Neonatal Network, 12,* 59-66.

Luckasson R (1992). *Mental Retardation, Definition, Classification and Systems of Supports.* Washington D.C.: American Association on Mental Retardation.

Magnano CL, Diamond EJ & Gardner JM (1989). Use of salivary cortisol measurements in young infants: A note of caution. *Child Development, 60,* 1099-1101.

Malviya S & Voepel-Lewis T (2001). Pain management in children with and without cognitive impairment following spine fusion surgery. *Pediatric Anaesthesia, 11,* 453-458.

Maurer H & Newbrough J (1987). Facial expressions of mentally retarded and nonretarded children 2: Recognition by nonretarded adults with varying experience with mental retardation. *American Journal of Mental Deficiency, 91,* 511-515.

McGrath PJ & Johnson CP (1985). CHEOPS: A behavioral scale for rating postoperative pain in children. *In: Fields HL, Dubner R & Cervero F (Eds), Advances in Pain Research and Therapy* (pp. 395-402). New York: Raven Press.

McGrath PJ, Rosmus C, Camfield C, Campbell MA & Hennigar A (1998). Behaviours caregivers use to determine pain in non-verbal, cognitively impaired individuals. *Developmental Medicine and Child Neurology, 40,* 340-343.

Nolan J, Chalkiadis GA, Low J, Olesch CA & Brown TC (2000). Anaesthesia and pain management in cerebral palsy. *Anaesthesia, 55,* 32-41.

Oberlander TF, O'Donnell ME & Montgeomey CJ (1999a). Pain in children with significant neurological impairment. *Journal of Developmental and Behavioral Pediatrics, 20*, 235-43.

Oberlander TF, Gilbert CA, Chambers CT, O'Donnell ME & Craig KD (1999b). Biobehavioral responses to acute pain in adolescents with a significant neurologic impairment. *Clinical Journal of Pain, 15*, 201-209.

Peters JWB (2001). *Facing pain in infancy and childhood.* Dept. of anaesthesiology, PhD thesis, Rotterdam, The Netherlands, Erasmus MC: 153.

Peters JWB, Duivenvoorden HJ, Grunau R, de Boer JB, Tibboel D & Koot HM, Validity of the NFCS for assessment of postoperative pain., Pediatric Academic Societies' Annual Meeting, Baltimore, 2002.

Pignatelli D, Magalhaes MM & Magalhaes MC (1998). Direct effects of stress on adrenocortical function. *Hormone and Metabolic Research, 30,* 464-474.

Putignano P, Kaltsas GA, Satta MA & Grossman AB (1998). The effects of anti-convulsant drugs on adrenal function. *Hormone and Metabolic Research, 30,* 389-97.

Stallard P, Williams L, Lenton S & Velleman R (2001). Pain in cognitively impaired, non-communicating children. *Archives of Disease in Childhood, 85*, 460-462.

Stallard P, Williams L, Velleman R, Lenton S & McGrath PJ (2002a). Brief report: Behaviors identified by caregivers to detect pain in noncommunicating children. *Journal of Pediatric Psychology, 27,* 209-214.

Stallard P, Williams L, Velleman R, Lenton S, McGrath PJ & Taylor G (2002b). The development and evaluation of the pain indicator for communcatively impaired children (PICIC). *Pain, 98,* 145-149.

Terstegen CM, Koot HM, De Boer JB & Tibboel D (2003). Measuring pain in children with cognitive impairment: Pain response to surgical procedures. *Pain, 103,* 187-198.

Van der Marel CD, Anderson BJ, Van Lingen RA, Holford NH, Pluim MA, Van den Anker JN & Tibboel, D (2003). Paracetamol and metabolite pharmacokinetics in infants. *European Journal of Clinical Pharmacology, 59*, 243-251.

Van Dijk M, De Boer JB & Koot HM (2000). The reliability and validity of the COMFORT scale as a postoperative pain instrument in 0 to 3-year-old infants. *Pain, 84,* 367-377.

Van Dijk M. (2001). Pain unheard? Postoperative pain assessment in neonates and infants. Sophia Children's Hospital, Dept. Child- and Adolescentpsychiatry and Pediatric Surgery. *PhD thesis*, Rotterdam, The Netherlands, Erasmus University Rotterdam: 147.

Van Dijk M, Koot HM, Huijer Abu-Saad H, Tibboel D & Passchier J (2002). The observational Visual Analoge Scale in pediatric pain assessment: Useful tool or good riddance? *Clinical Journal of Pain, 18,* 310-316.

Van Dongen KAJ, Huijer Abu-Saad H & Hamers J (1999). *On the development of an observational scale to measure pain in non-verbal children with severe or profound cognitive impairment.* Proceedings of the 9th world congress on pain in Vienna, Seattle, IASP Press.

Van Herk R, Baar F, et al. (2004). Is het mogelijk om pijn te meten bij ouderen met een uitingsbeperking? *Pijninfo, in press.*

Vlaskamp C (1997). The implementation of a care program for individuals with profound multiple disabilities. *European Journal on Mental Disability, 4,* 3-12.

Vlaskamp C, Van der Meulen BF, Zijlstra HP (2002). De instrumentele realisering van het GedragsTaxatieInstrument. *Tijdschrift voor Orthopedagogiek, 41,* 22-31.

Wary B & Doloplus C (1999). Doloplus-2, a scale for pain measurement. *Soins Gerontologie, 19,* 25-27.

Zweig M & Campbell G (1993). Receiver operating characteristics (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clinical Chemistry, 39,* 561-577.

Summary / Samenvatting

Summary

Pain is a subjective phenomenon and therefore complex in its assessment. Self-report is considered to be the gold standard for assessing pain, but there are several groups of persons who lack this ability due to their limited expressive repertoire. They were excluded from pain studies because they are a rather difficult population to investigate. Only until recently, studies became more focused on the group of persons who have difficulties in communicating their experiences to their caregivers, such as children with profound cognitive impairment (PCI). These children process information and communicate distress and pain in a different way than normally developed children. Therefore, it can be concluded that observational scales which are available to assist the observer in post-surgical pain assessment in normally developed children, probably may not be appropriate for children with PCI. Thus, the assessment of pain in this population might profit substantially from empirical evidence showing the validity of observable behaviors as indicators of pain. The objective of the present study was to assess pain indicators in children with PCI. The aims of the study were:

1. To determine the value of salivary cortisol levels used as objective physiological parameter in evaluating pain in children with PCI next to behavioral observations.
2. To test the sensitivity of a large pool of non-verbal candidate indicators for the assessment of surgical pain in children with PCI.
3. To construct and validate an observational scale containing indicators which are sensitive to post-surgical pain in children with PCI.
4. To compare pain behavior displayed by children with PCI with that of normally developed children.
5. To determine the observational scale's applicability and reliability in daily practice for the assessment of everyday, more chronic pain responses in children with PCI.

In *Chapter 2*, information is provided about baseline values of circadian salivary cortisol levels as a physiological pain parameter. Rises in cortisol levels are indicators of distress such as pain. However, reference values of circadian salivary cortisol levels have only been published for normally developed children. The aim of this study was to provide these reference values for children with severe and profound cognitive impairment and assess differences with normally developed children. Since children with cognitive impairment tend to experience a number of physical problems and associated stress or pain, their average cortisol levels may be elevated even if they are in a stable condition. Further, the circadian rhythm of cortisol levels may be altered in children with CI due to their unusual sleep-wake pattern, which may be most evident in suppressed early morning cortisol levels. Using a cross-sectional design, saliva samples were collected at 7.30 AM, 12.30 PM and 17.30 PM in 49 children with severe and profound cognitive impairment and in 44 normally developed children matched by age (1-18 years).

Circadian cortisol levels in children with cognitive impairment were not significantly different from those of normally developed children, although in more than one third of the impaired children a circadian rhythm was lacking. However, in the group of normally developed, healthy children a lack of circadian rhythm was also demonstrable. Our results indicate that reference values of salivary cortisol for normally developed children can also be used for children with severe and profound cognitive impairment. In the absence of a gold standard for the assessment of pain in children with PCI, it could be informative to use a more objective physiological parameter such as cortisol levels, which can be addressed when evaluating pain next to behavioral observations.

The study reported in *Chapter 3* investigated post-surgical pain in children with PCI, searching for a core set of cues these children use to express their pain. The development of an observational scale is described, from an item pool consisting of 138 items, reduced to 23 items. Fifty-two children were observed twice before surgery and five times after surgery within the period of their admission to the Erasmus MC-Sophia Children's Hospital. All 138 possible pain indicators were scored at each point of observation, using a five point scale ranging from 0 (never shown) to 4 (always shown). We used the Visual Analogue Scale (VAS) to give a general impression of the severity of the children's pain during the episodes they were observed. Several analyses provided evidence that 23 observable behaviors are sensitive to post-surgical pain in children with PCI, regardless of the pain intensity of the surgical procedures they underwent. The finding that all indicators, except for one, were scored significantly higher on episodes with VAS ratings ≥ 4 (which were considered clear indications of pain), confirms the sensitivity of these indicators concerning absence versus presence of pain. This study reveals that, although the specific expression of pain may be very individual, there appears to be a set of shared generic indicators used by children with PCI. The identification of these 23 observable pain behaviors enables the development of an observational pain assessment scale for this particular group of children.

Chapter 4 presents how this item pool was further developed into an observational scale - the Checklist Pain Behavior – and the scale's psychometric properties were examined. We investigated whether all 23 indicators could be summarized into meaningful and reliable subscales to be included in the CPG. Findings from PRINCALS analyses showed a highly satisfactory fit to the data of both a four-dimension solution and a one-dimension solution. This shows that the four subscales were clearly distinguishable, as was confirmed by moderate inter-subscale correlations, and that these may be combined into a total score without loss of information. However, two items ('eyes almost closed' and 'accepting comfort') were removed from the scale because of the large discrepancy with loadings from the other items in the same subscale. Further, the item scores appeared to be highly skewed. For this reason, and also because observation in clinical practice benefits from fewer response categories, we combined the 'regularly shown' and 'often shown' response categories into one 'often shown' category.

Because the models based on four and five response categories had an equal fit to the data, we maintained this four-point scale.

The Checklist Pain Behavior showed good reliability; interrater reliability was excellent for all subscales and for the total scale. Interscale correlations were moderate for most subscales, but high between the facial and vocal subscales. The convergent validity of the CPG turned out to be good; regression analyses showed that all subscales, except 'Social behavior/Mood', contributed independently to the prediction of VAS scores. This shows that each of the CPG subscales 'Facial expression', 'Vocalisation' and 'Physiological signs' is related to an overall pain rating. The exception for 'Social behavior/Mood' might be explained by the finding that behaviors included in this scale ('panicking', 'resistant/rebellious', 'protecting sore body part') might be more indicative of anxiety or anger rather than pain, as indicated by VAS ratings.

Further, all mean CPG subscale scores and the total scale score were significantly higher for the post-surgical episodes when VAS ratings were higher or equal to 4, compared to VAS < 4. Because this study also showed excellent interrater reliability for the VAS ratings used in this group of children, it can be concluded that the Checklist Pain Behavior indeed adequately assesses pain in children with PCI.

Next, we analyzed the CPG scores on the four regular post-surgical measurements (30/60 minutes and 3 hrs, 24 hrs and 48 hrs after extubating) categorized by a low, moderate or highly *painfulness of surgery*. Scores on the subscales 'Facial expression', 'Vocalisation' and the total scale appeared to differ significantly between children undergoing procedures that differed in painfulness, indicating that these items are most probably sensitive to differences in severity of pain.

In this study, we also reported the CPG's sensitivity for *changes in pain over time*. A pain measure should not only detect pain but also should be able to monitor the presence of pain over time. Because we had data from 5 post-surgical observational episodes at our disposal, we were able to assess the construct validity more thoroughly by comparing post-surgical mean scores over time. Significantly different, declining mean CPG scores across the post-surgical episodes were found for all subscales, except 'Social behavior/Mood'. This indicated that the CPG is sensitive to changes in pain behavior over time.

Further, mean scores on the subscales 'Facial expression' and 'Vocalisation' and on the total scale were significantly higher on the surgery specific, directly pain induced episode compared to the regular post-surgical episodes.

This study provided evidence that the Checklist Pain Behavior can assist caregivers (e.g., nurses; parents; physicians) in assessing and monitoring pain responses and help to improve the prescription and administration of analgesia tuned to children with PCI in a post-surgical setting.

Chapter 5 reports on the comparison of post-surgical pain responses between children with and without PCI. Children with PCI express their pain in a different manner than normally

developed children. However, in which way they differ is a question that still remained to be answered. Sixteen children with PCI and 12 normally developed preverbal children were observed after admission for elective Ear-Nose-Throat surgery at predetermined time points using the CPG. Again, the VAS was used to obtain a general impression of the severity of the children's pain.

Analyses showed that several pain expressions displayed by children with PCI, were also observed in children without cognitive impairment, but often attenuated or, conversely, more frequent. Four behavioral indicators were only scored in children with PCI. Within the group of normally developed children 12 behaviors – and 4 behaviors for children with PCI - were not shown before surgery (VAS score = 0, absence of pain) but appeared only after surgery (presence of pain; post-surgical observations with VAS < 4 were excluded from these calculations). Four behaviors – and 16 for children with PCI - were present before surgery and appeared with increased frequency or intensity after surgery. These findings suggest that certain behaviors in children with PCI are sensitive to pain but are not specific pain indicators. Children with PCI are in general very limited in possibilities to express themselves, which leaves them very little alternatives in "choosing" behaviors to express their pain. Some of the few behaviors they are capable of showing, apparently are used both in non-painful *and* in painful situations, but more frequent or, conversely, more subtle in the latter.

Chapter 6 examined the use of the CPG in clinical practice. The primary purpose of this study was to investigate whether this observational measure could detect everyday, more chronic, pain in children with PCI. The CPG was completed by 41 nurses working on a neurology ward of an academic children's hospital and by nurses in a healthcare facility for children with PCI. They were trained to use the CPG autonomously and conducted a total of 348 bed-side observations on 48 children with PCI. For all trained nurses, Kappa's were satisfactory (≥ 60). The moments to observe were chosen at random and after 2 minutes of observation the CPG and the VAS were completed.

Internal consistency and interrater agreement (ICC) of the CPG were adequate (.87 and .80, respectively). For both groups of nurses, the CPG scores were significantly higher for observations with VAS ratings ≥4 compared to observations with VAS ratings < 4. Univariate regressions indicated that all subscales, again except for 'Social behavior/Mood', contributed independently to the prediction of VAS ratings.

Further, a CPG score of 18 or higher provided a sensitivity of 0.71 and a matching specificity of 0.68 and therefore turned out to be a fair indication of pain in children with PCI. In 94% of the time (AUC value .94), a random selection from the positive group (those who are correctly diagnosed as painful) will have a greater score on the CPG than a random selection from the negative group (those who are incorrectly classified as painful). Considering the fact that a test with an AUC value of .50 can be interpreted as performing no better than change and value 1.00 as perfectly accurate, it can be concluded that the CPG is very accurate in detecting pain.

These results provide evidence that the Checklist Pain Behavior can be applied to assess everyday, more chronic pain in children with PCI in both a hospital setting (in non-surgical situations) and in the healthcare facility where these children live. Nurses who were unfamiliar with these children were also able to detect signs of pain using the CPG. This study reveals good prospects to train both groups of nurses to observe pain related behaviors in everyday pain events in a reliable way, using the CPG.

Finally, in *Chapter 7*, the main findings and conclusions of the chapters in this thesis were discussed, and strengths and limitations of the study were described. The results of this research project are partially consistent with findings of other studies addressing pain responses in children with PCI. The design and the results of this study will be compared briefly with some of the other studies. Further clinical implications and directives for future research were given.

Samenvatting

Pijn is een subjectief fenomeen and daardoor complex in zijn beoordeling. Zelfrapportage staat bekend als de gouden standaard voor pijnbeoordeling, maar er zijn verschillende groepen mensen die niet in staat zijn tot zelfrapportage vanwege hun beperkte uitingsmogelijkheden. Omdat zij een tamelijk moeilijke onderzoeksgroep vormen, werden zij bij pijnonderzoeken onder de exclusiecriteria geschaard. Sinds een aantal jaar richten verschillende studies zich meer op de groep mensen die moeilijkheden hebben met het communiceren van hun ervaringen naar degenen die hen verzorgen, zoals kinderen met een zeer ernstige verstandelijke handicap. Deze kinderen verwerken informatie en communiceren stress en pijn op een andere manier dan normaal begaafde kinderen dat doen. Het lijkt dan ook logisch dat observationele schalen die beschikbaar zijn om de observator te assisteren bij het beoordelen van postoperatieve pijn in normaal begaafde kinderen, waarschijnlijk niet geschikt zijn voor zeer ernstig verstandelijke gehandicapte kinderen. Geconcludeerd kan worden dat pijnbeoordeling binnen deze populatie aanzienlijk zou kunnen profiteren van empirisch bewijs dat valide observeerbare pijnindicatoren naar voren brengt.

Het hoofddoel van deze studie was het beoordelen van pijngedrag bij kinderen met een zeer ernstig verstandelijke handicap. De doelstellingen kunnen als volgt worden omschreven:

1. Het bepalen van de waarde van cortisol in speeksel gebruikt als objectieve fysiologische parameter bij het evalueren van pijn in kinderen met een zeer ernstige verstandelijke handicap, te hanteren naast gedragsobservaties.
2. Het testen van de sensitiviteit van een uitgebreide itempool met non verbale mogelijke pijnindicatoren voor het beoordelen van chirurgische pijn bij kinderen met een zeer ernstige verstandelijke handicap.
3. Het ontwikkelen en valideren van een observationele schaal met indicatoren sensitief voor postoperatieve pijn bij kinderen met een zeer ernstige verstandelijke handicap.
4. Het vergelijken van pijngedrag getoond door kinderen met een zeer ernstige verstandelijke handicap en normaal begaafde, gezonde kinderen.
5. Het beoordelen van de observationele schaal op toepasbaarheid en betrouwbaarheid in de dagelijkse praktijk voor het beoordelen van alledaagse, meer chronische pijn reacties van kinderen met een zeer ernstige verstandelijke handicap.

In *hoofdstuk 2*, wordt informatie gegeven over nulmeting waarden van circadiane cortisol spiegels in speeksel als een fysiologische pijn parameter. Verhoogde cortisol waarden indiceren stress, zoals pijn. Echter, referentiewaarden van circadiane cortisol spiegels zijn alleen bekend voor normaal begaafde, gezonde kinderen. Het doel van deze studie was het verkrijgen van deze nulmeting waarden voor kinderen met een (zeer) ernstige verstandelijke handicap en het beoordelen van verschillen tussen hen en normaal begaafde kinderen. Omdat kinderen met een (zeer) ernstige verstandelijke handicap vaak met diverse fysiologische

problemen te maken hebben en daaraan gerelateerde stress en pijn, kunnen hun gemiddelde cortisol spiegels verhoogd zijn, ook in een stabiele toestand. Verder is het mogelijk dat het circadiane ritme van de cortisol spiegels anders is als gevolg van een vaak verstoord dag en nacht ritme bij deze kinderen, welke het meest duidelijk naar voren kan komen in onderdrukte ochtend spiegels.

Een cross-sectionele methode werd gebruikt om cortisol monsters te verzamelen om 7.30 uur, 12.30 uur en 17.30 uur bij 49 kinderen met een (zeer) ernstige verstandelijke handicap en 44 normaal begaafde gezonde kinderen, gekoppeld op leeftijd (1-18 jaar).

Circadiane cortisol spiegels in kinderen met een (zeer) ernstige verstandelijke handicap waren niet significant verschillend dan die van normaal begaafde kinderen, alhoewel in meer dan een derde van de gehandicapte kinderen het circadiane ritme ontbrak. Echter, in de groep normaal begaafde kinderen was de afwezigheid van het circadiane ritme ook aantoonbaar. Onze resultaten indiceren dat de referentiewaarden (nulmetingen) van cortisol in speeksel zoals bekend voor normaal begaafde, gezonde kinderen ook geraadpleegd kunnen worden voor kinderen met een (zeer) ernstige verstandelijke handicap. Omdat een gouden standaard voor pijnbeoordeling bij deze groep gehandicapte kinderen ontbreekt, kan het informatief zijn om een meer objectieve maat zoals cortisol spiegels te gebruiken bij het beoordelen van pijn als aanvulling op gedragsobservaties.

De studie beschreven in *hoofdstuk 3* onderzocht postoperatieve pijn bij kinderen met een zeer ernstige verstandelijke handicap, zoekende naar een set met generieke indicatoren die deze kinderen gebruiken om hun pijn te uiten. De ontwikkeling van een observationele schaal is beschreven waarbij een itempool bestaande uit 138 mogelijke pijnindicatoren uiteindelijk gereduceerd is naar 23 indicatoren. Er zijn 52 kinderen twee maal preoperatief en vijf maal postoperatief geobserveerd rondom hun opname in het Erasmus MC-Sophia Kinderziekenhuis. Alle 138 pijnindicatoren zijn gescoord op ieder observatiemoment op een vijfpunt-schaal van 0 (nooit getoond) tot 4 (continu getoond). De Visueel Analoge Schaal (VAS) werd gebruikt om een algemene indruk te geven van de ernst van de pijn van het kind op de observatiemomenten. Aan de hand van verschillende analyses werden 23 indicatoren sensitief bevonden voor postoperatieve pijn bij zeer ernstig verstandelijk gehandicapte kinderen, ongeacht de pijnintensiteit van de operatie die zij ondergingen. De bevinding dat alle indicatoren, op één na, significant hoger werden gescoord op momenten waarbij de VAS ≥ 4 was (wat aanwezigheid van pijn indiceert) bevestigt de sensitiviteit van deze indicatoren aangaande het kunnen differentiëren in aan- en afwezigheid van pijn. Deze studie onthult dat, ondanks de zeer individuele aard van pijnexpressie, er een set generieke indicatoren aan te wijzen is welke gebruikt wordt door kinderen met een zeer ernstige verstandelijke handicap. De identificatie van deze 23 indicatoren maakt het mogelijk om een observationele pijnbeoordelingsschaal te ontwikkeling voor deze specifieke groep kinderen.

Hoofdstuk 4 rapporteert hoe de itempool verder is ontwikkeld tot een observationele schaal - de Checklist PijnGedrag (CPG) - en de psychometrische kwaliteiten van de schaal worden toegelicht. We onderzochten of alle 23 indicatoren konden worden samengevoegd in betekenisvolle and betrouwbare subschalen binnen de CPG. Bevindingen van PRINCALS analyses toonden een heel bevredigende fit op de data van zowel een vier-dimensie oplossing als een één-dimensie oplossing. Dit indiceert dat de vier subschalen duidelijk te onderscheiden waren, als ook was bevestigd door de matige interne subschaal correlaties, maar dat ze ook kunnen worden gecombineerd tot één totale schaal score zonder verlies van informatie. Twee items ('ogen in spleetjes' en 'accepteren van troost') zijn verwijderd uit de schaal in verband met hun grote discrepantie met ladingen van de andere items in dezelfde subschaal.

Omdat de item scores zeer scheef verdeeld bleken, en omdat observatie in de klinische praktijk beter gedijt onder minder antwoordcategorieën, combineerden we de categorieën 'regelmatig aanwezig' en 'vaak aanwezig' in één categorie 'vaak aanwezig'. Aangezien de modellen gebaseerd op vier en vijf antwoordcategorieën een gelijke fit op de data hadden, hebben we de vierpunt-schaal verder gehanteerd.

De CPG toonde een goede betrouwbaarheid; interbeoordelaarsbetouwbaarheid was zeer goed voor alle subschalen en voor de totale schaal. Interne schaal correlaties waren matig, maar hoog tussen de subschalen 'Gezichtsexpressie' en 'Vocalisatie'. De convergente validiteit van de CPG is goed; regressie analyses laten zien dat de subschalen, met uitzondering van 'Sociaal Gedrag/Stemming', onafhankelijk bijdragen aan de voorspelling van VAS scores. Dit betekent dat de subschalen 'Gezichtsexpressie', 'Vocalisatie' en 'Fysiologische signalen' gerelateerd zijn aan een algehele pijn score. Dat 'Sociaal Gedrag/Stemming' geen onafhankelijke bijdrage levert kan verklaard worden door het gegeven dat de gedragingen in deze subschaal ('paniekerig', 'zich verzetten, niet meewerken', beschermen van aangedaan lichaamsdeel') meer emoties zoals angst of boosheid vertegenwoordigen en niet zozeer pijn (zoals geïndiceerd door VAS scores).

Alle gemiddelde CPG subschaal scores en de totale schaal score waren significant hoger voor de postoperatieve observaties wanneer een VAS score ≥ 4 was vergeleken met een VAS score < 4. Omdat deze studie ook een uitstekende interbeoordelingsbetrouwbaarheid rapporteert voor VAS scores wanneer toegepast in deze groep kinderen (.97), kan geconcludeerd worden dat de CPG op adequate wijze pijn beoordeelt bij kinderen met een zeer ernstige verstandelijke handicap.

Vervolgens hebben we CPG scores geanalyseerd afkomstig van de vier reguliere postoperatieve observatie momenten (30/60 minuten en 3 uur, 24 uur, 48 uur na extubatie) gecategoriseerd naar een lage, matige of hoge pijnintensiteit. Scores op de subschalen 'Gezichtsexpressie', 'Vocalisatie' en de totale schaal verschilde significant tussen kinderen die operaties ondergingen met verschillende pijnintensiteit, wat indiceert dat deze items hoogst waarschijnlijk sensitief zijn voor de ernst van pijn.

In deze studie rapporteerde we tevens the CPG's sensitiviteit voor veranderingen in pijn over de tijd. Een pijn schaal dient niet alleen pijn te beoordelen maar dient tevens in staat te zijn om de aanwezigheid van pijn over een periode te monitoren. Daar wij over gegevens beschikten van 5 postoperatieve momenten (vier reguliere momenten, zie hierboven, en één operatie specifiek, direct pijn inducerend moment) gestandaardiseerd uitgezet in de tijd, was het mogelijk om de construct validiteit meer grondig te beoordelen door de CPG scores over deze momenten heen te vergelijken. Significant verschillende, aflopende gemiddelde CPG scores over de postoperatieve observatiemomenten waren gevonden voor alle subschalen, echter met uitzondering van 'Sociaal Gedrag/Stemming'. Dit geeft aan de CPG sensitief is voor veranderingen in pijngedrag over de tijd. Verder waren gemiddelde scores op de subschalen 'Gezichtsexpressie' en 'Vocalisatie' en op de totale schaal significant hoger op het ingreep specifieke observatie moment vergeleken met de reguliere postoperatieve observaties.

Deze studie toont bewijs dat de CPG voorziet in het bieden van assistentie aan hulpverleners (verpleegkundigen, ouders, artsen) in het beoordelen en monitoren van pijn reacties en te helpen in het verbeteren van het voorschrijven en toedienen van pijnmedicatie afgestemd op deze speciale groep kinderen in een postoperatieve setting.

Hoofdstuk 5 beschrijft de vergelijking van postoperatieve pijn reacties tussen kinderen met en zonder verstandelijke handicap. Kinderen met een zeer ernstige verstandelijke handicap geven op een andere wijze uiting aan pijn dan normaal begaafde kinderen dat doen. Echter, op welke manier ze dan met elkaar verschillen is een nog onbeantwoorde vraag. Zestien kinderen met een zeer ernstige verstandelijke handicap en 12 normaal begaafde preverbale kinderen zijn met gebruik van de CPG geobserveerd na opname voor een KNO (keel, neus en oor) ingreep op van te voren vastgestelde momenten. Wederom was de VAS toegepast om een algemene indruk van de ernst van de pijn te kunnen geven.

Analyses toonden aan dat verschillende pijnexpressies getoond door kinderen met een zeer ernstige verstandelijke handicap, ook geobserveerd zijn bij de groep normaal begaafde kinderen, echter veelal subtieler of juist het tegengestelde, veel frequenter. Vier gedragingen werden alleen gescoord bij de gehandicapte kinderen. Binnen de groep normaal begaafde kinderen werden 12 gedragingen – en 4 bij de gehandicapte kinderen – niet voor de operatie getoond (VAS score = 0, absence of pain) maar verschenen alleen postoperatief (aanwezigheid van pijn, VAS scores < 4 werden bij deze berekeningen geëxcludeerd). Vier gedragingen – en 16 bij de gehandicapte kinderen – waren zowel voor als na de operatie aanwezig. Deze bevindingen suggereren dat bepaalde gedragingen bij de zeer ernstig verstandelijk gehandicapte kinderen wel sensitief zijn voor pijn maar geen indicatoren zijn specifiek voor pijn. Deze kinderen zijn over het algemeen zeer beperkt in hun uitingsmogelijkheden, wat hen weinig keuzemogelijkheden geeft in gedragingen die ze willen gebruiken om hun pijn te uiten. Sommige van die weinige gedragingen die zij kunnen tonen, worden klaarblijkelijk gebruikt in zowel niet pijnlijke als pijnlijke situaties, maar in dat laatste geval op een andere manier dan gewoonlijk (frequenter of juist milder, subtieler).

Hoofdstuk 6 onderzocht het gebruik van de CPG in de klinische praktijk. De hoofddoelstelling was te bezien of deze observationele schaal ook alledaagse, meer chronische pijn in zeer ernstig verstandelijk gehandicapte kinderen kan detecteren. De CPG was ingevuld door 41 verpleegkundigen werkzaam op de neurologie afdeling van een academisch kinderziekenhuis en door verpleegkundigen werkzaam in een residentiele zorginstelling speciaal toegerust op deze kinderen. Zij werden getraind in het zelfstandig gebruik van de CPG en voerden samen een totaal van 348 observaties uit direct aan het bed bij 48 zeer ernstig verstandelijk gehandicapte kinderen. Voor alle verpleegkundigen bleken de Kappa's acceptabel (vanaf .60). The momenten waarop geobserveerd werd zijn willekeurig gekozen en na 2 minuten observeren zijn de CPG en de VAS ingevuld. Interne consistentie en interbeoordelaarsbetrouwbaarheid (ICC) van de CPG waren adequaat (.87 en .80, respectievelijk). Voor beide groepen verpleegkundigen waren de CPG scores significant hoger voor observaties met VAS scores ≥ 4 vergeleken met observaties waarbij de VAS scores < 4 waren. Univariate regressies indiceerde dat alle subschalen, weer met uitzondering van de subschaal 'Sociaal Gedrag/Stemming' , onafhankelijk hebben bijgedragen aan de voorspelling van de VAS scores.

Bij een CPG score van 18 of hoger behoort een sensitiviteit van .71 en een bijbehorende specificiteit van .68, en is daardoor een goede indicatie van pijn bij zeer ernstig verstandelijk gehandicapte kinderen. Hiernaast werd gerapporteerd dat in 94% van de gevallen (AUC waarde .94), een willekeurige selectie uit de positieve groep (diegenen die terecht als pijnlijk gediagnostiseerd zijn) een hogere CPG score heeft dan een willekeurige selectie uit de negatieve groep (diegenen die onterecht als pijnlijk geclassificeerd zijn). Gezien het gegeven dat een test (zoals de CPG) met een AUC waarde van .50 geïnterpreteerd kan worden als 'niet meer dan toeval' en een waarde 1.00 als zeer accuraat, kan geconcludeerd worden dat de CPG erg accuraat is in het detecteren van pijn.

De resultaten beschreven in dit hoofdstuk voorzien in bewijs dat de CPG toegepast kan worden voor het beoordelen van alledaagse, meer chronische pijn bij kinderen met een zeer ernstige verstandelijke handicap in zowel een (niet chirurgische) ziekenhuis setting als in een residentiele zorginstelling waar deze kinderen woonachtig zijn. Verpleegkundigen die het te observeren kind niet (of nauwelijks) kende waren ook in staat om met de CPG pijnsignalen te detecteren. Deze studie toont goede vooruitzichten voor het trainen van verpleegkundigen, in verschillende settingen, in het observeren van pijn gerelateerd gedrag in alledaagse pijn situaties, op een betrouwbare manier, aan de hand van de CPG.

Uiteindelijk, in *hoofdstuk 7,* worden de belangrijkste bevindingen en conclusies van de hoofdstukken in dit proefschrift bediscussieerd en sterke punten en beperkingen van de studie zijn beschreven. De resultaten van dit onderzoek zijn gedeeltelijk overeenkomstig met bevindingen uit andere studies naar pijngedrag bij kinderen met een zeer ernstige verstandelijke handicap. De onderzoeksprocedure en de resultaten van dit onderzoek worden beknopt vergeleken met enkele andere studies. Tot slot worden instructies voor de klinische praktijk en voor toekomstig onderzoek gegeven.

Dankwoord/
Curriculum vitae

Dankwoord

Bij het tot stand komen van dit proefschrift zijn veel mensen belangrijk voor me geweest. Een aantal van hen wil ik hier bedanken en daarmee begin ik bij het thuisfront.

Lieve Mirthe en Esmee, ik beloof jullie dat ik vanaf nu niet meer zoveel boven achter de computer zal werken, maar dat we samen veel leuke dingen gaan doen! Lieve Maarten, eindelijk is het boekje klaar! Heel erg bedankt dat je me alle ruimte heb gegeven om dit voor elkaar te krijgen, vanaf nu zal ik minder gestresst zijn en gaan we in de weekenden weer doen waar ze voor bedoeld zijn!
Pap en mam, heel erg bedankt voor jullie steun! Ik vind het een groot voorrecht te weten dat ik altijd op jullie kan rekenen. Oma Lia, lieve mam, zonder jou was dit proefschrift nog lang niet af geweest! Dankzij het ontelbaar aantal uurtjes dat je met de meisjes hebt doorgebracht, heb ik zonder al te veel schuldgevoel kunnen werken aan dit onderzoek.
Martine, zusje en beste vriendin, jij was mijn redder in nood toen mijn computer ermee ophield twee weken voordat het manuscript bij de commissie moest liggen. Bedankt dat je het niet erg vond dat ik vaak tot 's avonds laat achter jouw computer werkte!
Marcel, wat is het toch handig om een broer te hebben die alles van computers weet en videomateriaal kan editen. Heel erg bedankt voor al je hulp tijdens het onderzoek en bij het maken van de omslag van dit boekje! Astrid en jij zijn bij deze uitgenodigd om nu eens bij ons te komen 'rendangen'!

Onmisbaar was de begeleiding van verschillende mensen bij het uitvoeren van dit onderzoek en het schrijven van dit proefschrift.
Prof. Dr. H.M. Koot, beste Hans, jouw kennis op statistisch gebied en het vermogen om helderheid en structuur aan te brengen in de artikelen, heeft mijn proefschrift gemaakt zoals het nu is. Heel veel dank voor de tijd die je altijd voor me vrij hebt gemaakt, ondanks je eigen overvolle agenda.
Prof. Dr. D. Tibboel, beste Dick, enorm bedankt voor je steun en getoonde vertrouwen. Op momenten dat me alles wat boven 't hoofd groeide, was jij degene die dat opviel en mij bemoedigende woorden toesprak. Ik vind het bijzonder deel uit te hebben gemaakt van één van je vele projecten binnen de onderzoekslijn 'Pijn bij kinderen'.
Dr. J.B. de Boer, allerbeste Josien, bedankt dat je mijn co-promotor wilt zijn! Je bent er altijd voor me geweest en ik vond het erg leuk om samen aan dit onderzoek te werken.
Prof. F. Verhulst, beste Frank, heel hartelijk dank dat ik al die jaren een plekje kreeg op jouw afdeling Kinder en Jeugdpsychiatrie. In combinatie met de aanwezigheid van ontzettend veel leuke collega's kijk ik met veel plezier terug op deze tijd! Marielle, Gonneke, Ilja, Suzan, Dennis, Rhodee, Victor, Wendy, Karen, Patricia, Pol, Jolanda, Laura, Inge, Monique, Jeroen, Donna, Gerri-Janne, Laureen, Helène en vele anderen, hartelijk dank voor jullie gezelligheid, collegialiteit en hulp!

Jolanda, zonder jouw redigeerwerk hadden er in dit proefschrift nog aardig wat foutjes gestaan! Bedankt dat je mijn proefschrift de moeite waard vond om er zoveel vrije tijd in te stoppen. Beste Ko Hagoort, u heeft me in de laatste fase fase enorm geholpen met het redigeren van een paar hoofdstukken en de handleiding. Bedankt voor uw snelle en onmisbare hulp!
Patricia van Deventer, als iemand verpleegkundigen weet te motiveren tijdens een implementatieproject, ben jij het wel! Ik heb veel van je geleerd. Bedankt voor de prettige samenwerking en de enorme hoeveelheid werk die je hebt verzet in het kader van dit onderzoek!
Prof. Dr. J. Passchier, al vanaf het begin bent u betrokken geweest bij dit onderzoek, nu is het eindelijk af! Bedankt voor uw tijd om het proefschrift te lezen en de rol van secretaris van de kleine commissie te vervullen.
Prof. Dr. C. Vlaskamp, beste Carla, ik vind het een hele eer iemand met zoveel deskundigheid en enthousiasme voor de doelgroep in de kleine commissie te hebben! Dank voor je tijd en energie om het proefschrift te lezen en te beoordelen.
Prof. Dr. A.J. van der Heijden, ik wil u hartelijk danken voor uw deelname aan de kleine commissie en het lezen en beoordelen van mijn proefschrift.
Dr. J.P.H. Hamers, beste Jan, na een lange tijd zien we elkaar toch weer. Bedankt dat je zitting wilt nemen in de grote commissie.
Prof. Dr. H. Evenhuis, Prof. Dr. F. Verhulst, beste Heleen, beste Frank, dank dat jullie in de grote commissie plaats willen nemen.
Dr. Lynn Breau, thank you for reading my thesis and for flying over to take part in the committee. I hope you will enjoy your stay in The Netherlands!

Hans Calf en Els Weijers, dank voor jullie enthousiasme en inzet tijdens het uitvoeren van een deel van dit onderzoek bij het ASVZ, locatie Lombardijen. Ook Astrid van Slavenburg en alle groepsleiding, dank voor jullie hulp. Ik ben altijd met veel plezier bij jullie geweest!

Ook wil ik graag Katinka van Dongen bedanken voor de ontzettend leuke tijd die we hebben gehad tijdens 3 jaar data verzamelen. Je was een heel fijne collega en ik vind het oprecht jammer dat het niet gelopen is zoals we graag wilden. Ik wens je heel veel goeds toe!
Verder wil ik nog een aantal geneeskunde studenten bedanken (die inmiddels al arts zijn). Suzanne Koudijs, jij was de eerste student die bij mij het afstudeeronderzoek kwam doen. Ik heb bewondering voor jouw enorme kennis over en betrokkenheid bij deze groep kinderen, het werkt aanstekelijk. Ik vind je een fantastisch mens en ben blij dat je straks als paranimf naast me staat.
Anke de Boer, Jeske van Waterschoot en Esmee Lanjouw, dank voor jullie inzet en bijdrage aan het onderzoek. Jeske, ik vond het heel gezellig dat je nog iets langer kon blijven na je afstudeeronderzoek. Dankzij jou is de onderzoeksgroep uiteindelijk nog uitgebreid, bedankt!

En Sinno, tijdens mijn zwangerschapsverlof heb je het onderzoek draaiende gehouden en nu promoveer je twee maanden voor mij. Bedankt en gefeliciteerd!!

Veel vrienden die hebben meegeleefd, met name in de laatste fase waarin het toch altijd even doorwerken is, wil ik ook hartelijk danken. In het bijzonder Sandra, je bemoedigende mailtjes en kaarten kwamen vaak precies op een goed moment. Beer en Esther, bedankt voor jullie steun en de zorg voor de kinderen, jullie zijn vrienden voor het leven! En natuurlijk Mariëlla, Jel, bedankt dat je mijn paranimf wilt zijn! Het feit dat jij straks naast me staat, stelt me gerust.

Ik wil alle ouders en kinderen bedanken die hebben deelgenomen aan dit onderzoek. Dankzij hen heeft dit onderzoek kunnen plaatsvinden.
Ook het Erasmus MC - Sophia personeel, hartelijk dank voor jullie onmisbare medewerking. De dames van het opnameburo, de anesthesisten, Dr. J. Bouquet, Dr. M. Sinaasappel, Dr. A. van Diepstraten, Dr. G. Madern, Dr. Hoeve en velen anderen, dank voor jullie inzet en interesse in dit onderzoek.
Ik ben MW-NWO veel dank verschuldigd voor de subsidiering van dit onderzoek.

Curriculum vitae

Chantal Terstegen werd op 12 september 1971 geboren te Zwolle. In 1987 behaalde ze haar MAVO diploma te Leiden en in 1989 haar HAVO diploma. In 1993 behaalde zij haar HBO-V diploma aan de Hogeschool Leiden. Van 1993 tot en met 1997 studeerde zij aan de Faculteit Sociale Wetenschappen te Leiden en in mei 1997 legde zij het doctoraal-examen Orthopedagogiek af. Van 1997 tot en met 1998 werkte zij als wetenschappelijk onderzoeker aan de Rijksuniversiteit Leiden, afdeling Orthopedagogiek, aan een studie over behandeling en begeleiding van verstandelijk gehandicapte plegers en slachtoffers van seksueel misbruik. Van 1998 tot en met februari 2003 was ze aangesteld bij het Erasmus MC - Sophia Kinderziekenhuis op de afdelingen Kinderheelkunde en Kinder- en Jeugdpsychiatrie voor het onderzoek waarvan de resultaten zijn beschreven in dit proefschrift.
Vanaf de aanvang van haar studie Orthopedagogiek tot heden is zij werkzaam als wijkverpleegkundige in de Intensieve Thuiszorg voor chronisch zieke kinderen alwaar zij ernstig meervoudig gehandicapte kinderen in de thuissituatie verpleegt. Momenteel is zij werkzaam bij het Landelijk Platform voor Ernstig Meervoudig Gehandicapten te Utrecht.
De auteur is getrouwd met Maarten Wirtz en samen hebben zij twee dochters: Mirthe (2000) en Esmee (2003).