Pain Unheard?

Postoperative Pain Assessment in Neonates and Infants.

Onuitgesproken Pijn

Postoperatieve Pijnmeting bij Pasgeborenen en Jonge Kinderen.

CIP- gegevens Koninklijke Bibliotheek. Den Haag

Pain Unheard?
Postoperative Pain Assessment in Neonates and Infants.
ISBN 90-73235-41-3

Gedrukt door: Optima Grafische Communicatie, te Rotterdam

Omslag
ontwerp: Hans Boer
tekening: Sam van der Heijde
gedicht: Nicky van der Heijde

Dit onderzoek is gesubsidieerd door NWO (nr 940-31-031)

Pain Unheard?

Postoperative Pain Assessment in Neonates and Infants.

Onuitgesproken Pijn

Postoperatieve Pijnmeting bij Pasgeborenen en Jonge Kinderen.

Proefschrift

ter verkrijging van de graad van doctor

aan de Erasmus Universiteit Rotterdam,

op gezag van de Rector Magnificus

Prof. dr. ir. J.H.van Bemmel

en volgens besluit van het College voor Promoties

De openbare verdediging zal plaatsvinden op

woensdag 24 januari, 2001 om 15.45 uur

door

Monique van Dijk

geboren te Rotterdam

PROMOTIECOMMISSIE:

Promotoren:	Prof. dr. J.Passchier Prof. dr. D.Tibboel
Co-promotor:	Dr. J.B.de Boer
Overige leden:	Prof. dr. J.N.van den Anker Prof. dr. H.Huijer-Abu-Saad Prof. dr. H.M.Koot

Paranimfen: Tineke Oskam
Maartje Zeldenrust

Aan Bram van Dijk

Table of contents

Chapter 1

Introduction

1.1 Introduction

Anyone who is familiar with infants or toddlers knows that they regularly experience pain, for example when hurting themselves during exploration of their natural environment (Fearon et al., 1996). Subsequently, they generally seek comfort from an available caregiver. Babies who are perceived as having pain from cramps, routine immunisation, or ear infection for instance, are often cradled, massaged, or walked around to ease the discomfort.
How different is the situation in hospital. While everyday pain is acceptable within limits for healthy exploring toddlers, sick hospitalised infants experience various painful procedures sometimes without analgesics not always followed by parental consolation.

Thanks to improved medical and technical possibilities, infants as young as 24-25 weeks gestational age are now able to survive in the Neonatal Intensive Care Unit (NICU) environment. In addition, major surgical procedures are nowadays feasible even in very small newborns while improved perioperative and anaesthetic management have increased the survival rates in infants with major congenital anomalies. As a consequence, infants may be hospitalised for a long period at a very young age and may undergo multiple painful procedures without adequate pain management (Stevens et al., 2000). Furthermore, they are deprived from their home and see their parents only during visits.

The sensory capability of neonates to experience pain was questioned until the late eighties. Not helpful in that respect was the definition of pain from the International Association for the Study of Pain (IASP) (Merskey and Bogduk, 1994):'An unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage.' A note explains that *'pain is always subjective. Each individual learns the application of the word through experiences related to injury in early life*'. This definition does not seem to apply to human beings incapable of self-report, such as neonates, mentally handicapped and demented individuals (Anand and Craig, 1996).

However, pain in young children has received increased attention since the landmark studies of Anand and colleagues (Anand and Hickey, 1987; Anand et al., 1987). One showed that neonates as young as 30 weeks gestational age have the anatomical and functional ability to perceive pain (Anand and Hickey, 1987). The other showed improved postoperative outcome and lower stress responses in premature neonates who received the

analgesic drug fentanyl next to general anaesthesia during surgery compared to premature neonates who only received general anaesthesia (Anand et al., 1987).

Although thanks to Anand's studies the existence of pain in neonates was more broadly acknowledged, a change in pain treatment was not the immediate result. A survey among anaesthetists in the UK revealed that they were reluctant to prescribe analgesia to infants under 1 month of age because of fear of ventilatory depression in this age group (Purcell-Jones et al., 1988). Furthermore, they considered the available clinical signs of pain as potentially misleading. To improve pain treatment, it was essential to develop pain instruments for preverbal infants. Moreover, randomised controlled trials should be performed to determine the efficacy and safety of different analgesic regimens in neonates and young infants.
Since the Dutch situation was not much different, a study was set up at our hospital with Dr. Sunny Anand as our consultant. This was entitled:
The assessment of pain in infants and children less than 3 years: the development of an instrument in relation to hormonal stress responses and morphine plasma levels and was supported by a research grant from NWO (Dutch Organisation for Scientific Research, grant nr. 940-31-031).

1.2 Study

Study aim
The study aimed at answering two questions:

- How reliable, valid, and feasible is the multidimensional COMFORT scale to assess postoperative pain in infants and toddlers 0-3 years of age?
- What is the difference between intermittent morphine administration and continuous intravenous morphine in terms of quality and effectiveness of analgesia for postoperative pain in infants and toddlers 0-3 years of age?

The studies described in this thesis deal primarily with the first question. The differences between the two morphine conditions in relation to hormonal and metabolic plasma levels and morphine plasma levels will be reported elsewhere.
During data collection a third research question came up, inspired by the eventful hospital history of some children combined with the individual differences in pain response and morphine requirement after surgery in our sample. This question was also justified by

publications on subsequent and long-term consequences of neonatal pain. The third, additional question is:

- Are the present postoperative pain and stress response related to past experiences with pain?

Methods

Sample

Between March 1995 and September 1998, a total of 204 children aged 0 to 3 years, who were admitted for major abdominal or thoracic surgery, entered the study after informed consent of the parents had been obtained.

Neonates were included when they were ≥35 weeks gestation and body weight ≥1500 grams.

Exclusion criteria were: use of co-medication (e.g. acetaminophen or midazolam) influencing the measured amount or potency of morphine, use of neuromuscular blockers, hepatic or renal dysfunction, seriously compromised neurological status or altered muscle tone.

Measures

COMFORT scale

To measure postoperative pain we chose the COMFORT scale (Ambuel et al., 1992). This is a multidimensional instrument comprising both behavioural and physiological indicators of pain, which had been developed for the intensive care environment to assess distress / comfort in ventilated children. With the addition of a new item 'Crying' the scale could also be used in non-ventilated infants, which was necessary in our study sample. This item then replaces the item 'respiratory response'. The COMFORT scale comprises eight items, each with five response categories consisting of distinct behavioural descriptions (see appendix). Six of the items are behavioural ones (Alertness, Calmness, Muscle tone, Movement, Facial tension, and Respiratory response/Crying), and two are physiological items: Heart rate (HR) and Mean arterial pressure (MAP).

Visual Analogue Scale (VAS)

The Visual Analogue Scale (Huskisson, 1974) was applied for two reasons. Firstly, to estimate the concurrent validity of the COMFORT scale and secondly, to obtain a criterion for extra pain medication. The VAS is a horizontal continuous ten-centimetre line with the anchors 'no pain' at the left side and 'pain as bad as it could be' or 'worst pain possible' at the right side. Nurses estimate the level of the infant's pain by making a mark on the line.

Surgical Stress Score (SSS)

The Surgical Stress Score (SSS) (Anand and Aynsley-Green, 1988) was originally developed to assess the severity of surgical stress in neonates and includes the following items: Amount of blood loss; Site of surgery; Amount of superficial trauma; Extent of visceral trauma; Duration of surgery; Associated stress factors: a) Hypothermia, b) Infection. The attending anaesthesiologist and surgeon applied the SSS directly after surgery to determine the stressfulness of the surgical procedure.

Blood sampling

Blood samples were drawn from the arterial line before surgery, directly after surgery, and 6,12, and 24h after surgery. Blood analysis included adrenaline, noradrenaline, lactate, insulin, and glucose plasma levels. Furthermore, morphine-, morphine-M3-glucuronide and morphine-M6-glucuronide plasma levels were assessed 5 minutes after loading dosage, and 6,12 and 24h after surgery.

Design

A double-blind, randomised clinical trial was carried out to compare the efficacy of intravenous continuous (CM) and intravenous intermittent morphine (IM) after major abdominal or thoracic surgery in 0 to 3-year-old infants. Prestratification by age was performed because behavioural and physiological differences between age groups were expected to be of importance. Age groups comprised neonates ($\geq$35 weeks gestation and weight $\geq$1500 grams), younger infants (1 to 6 months), older infants (7 to 12 months), and toddlers (1 to 3 years). Infants within age groups were assigned to CM or IM analgesia by random number generation. The hospital pharmacist prepared the study drugs and retained the randomisation schedule until the end of the trial. Pain assessment was performed prior to surgery, after return to the Pediatric Surgical Intensive Care unit (PSICU), and every three hours during the first 36 hours postoperative.

Figure 1 shows a flowchart with for each age group the numbers of infants in each morphine condition.

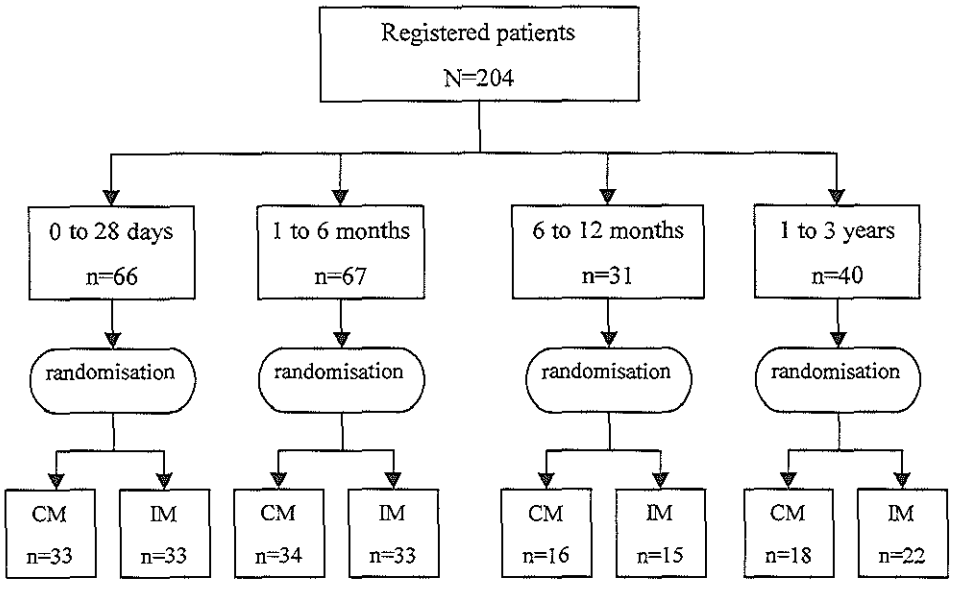


Figure 1.Results of block randomisation for each age group; CM = continuous morphine, IM = intermittent morphine

Procedure

Anaesthetic management was standardised. At the end of surgery, all patients were given an intravenous loading dose of morphine ≥100 μg/kg until they were in minimal pain as indicated by a VAS score <4. Morphine was next administered by protocol. The CM group were given a morphine infusion of 10 μg/kg/h, combined with a three-hourly intravenous placebo bolus (saline). The IM group received a continuous placebo infusion (saline), combined with a three-hourly intravenous morphine bolus of 30 μg/kg. When children were considered to be in pain (VAS ≥4), the protocol provided for additional morphine. Mechanical ventilation was continued after surgery in neonates <37 weeks and after repair of oesophageal atresia or congenital diaphragmatic hernia. In older age groups postoperative ventilation was required depending on the surgical procedure. Table 2 gives an overview of the study design.

Table 2 Schedule of design with check marks representing assessments

	Time										
	Baseline	End of surgery	Return at PSICU	3h	6h	9h	12h	15h	18h	21h	24h
VAS pain[a)]	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓→
COMFORT scale[a)]	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓→
MAP [b)]	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓
HR [b)]	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓
SSS		✓									
IM morphine[c)]		Loading		✓	✓	✓	✓	✓	✓	✓	✓→
30 µg/kg/3 h		dose									
CM morphine[c)]		Loading	✓	→	→	→	→	→	→	→	→
10 µg/kg/hr		dose									
Blood samples[d)]	✓	✓			✓		✓				✓

a) Pain assessment was continued every three hours until 36 hours after surgery.
b) Mean Arterial Pressure and Heart Rate were read from the monitor six times during the two-minute observation during each pain assessment.
c) Children were randomised to receive intermittent (IM) or continuous morphine (CM) for the first 36 hours after surgery.
d) Blood samples to assess adrenaline, noradrenaline, lactate, insulin and glucose levels directly after surgery and 6,12, and 24h after surgery. Furthermore, morphine-, morphine-M3-glucuronide and morphine-M6-glucuronide plasma levels were assessed 5 minutes after loading dosage, and 6,12 and 24h after surgery.
SSS = Surgical Stress Score.

1.3 Scope of this thesis

Chapter 2 describes the psychometric properties of the COMFORT scale as a postoperative pain instrument for neonates and infants. The interrater reliability of the nurses from the pediatric surgical intensive care is determined. In addition, linear structural equation modelling was applied to estimate the internal structure of the COMFORT scale, and the stability over time using the repeated measures after surgery. The VAS was included in the model, to estimate the concurrent validity of the COMFORT scale.

Chapter 3 reviews the VAS as a tool for observational pediatric pain assessment. The psychometric results from different pediatric studies that used the VAS as an observational tool are described and suggestions for the use of the VAS are given.

Chapter 4 analyses the association between behavioural items of the COMFORT 'behaviour' and the actual Heart Rate and Mean Arterial Pressure scores for the repeated measurements. This was initiated by the low associations between the COMFORT behavioural and physiological items. Furthermore, the influence of background characteristics, physical condition and pain-related characteristics on the behavior-physiology correlations is described.

Chapter 5 compares the efficacy of postoperative intermittent and continuous morphine administration in the study sample. The repeated COMFORT 'behaviour' and VAS pain scores were compared between the morphine conditions. In addition the impact of age, severity of stress and mechanical ventilation on the individual pain response was estimated.

Chapter 6 combines the study results of the clinical trial with information from the medical records of the sample. This chapter explores the relationship between past experiences with pain and the postoperative pain and stress response of the 132 infants and toddlers older than 1 months in the current study.

Chapter 7 describes the developments with regard to pain instruments of the period 1995 to October 2000.

Chapter 8 consists of two parts. The first is a general discussion addressing the results from the previous chapters and presenting directives for future research. In the second part we describe some of our experiences during the clinical trial.

1.4 References

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

Anand KJS and Aynsley-Green A. Measuring the severity of surgical stress in newborn infants. J Pediatr Surg 1988;23:297-305.

Anand KJS and Craig KD. New perspectives on the definition of pain. Pain 1996;67:3-6.

Anand KJS and Hickey PR. Pain and its effects in the human neonate and fetus. N Engl J Med 1987;317:1321-1329.

Anand KJS, Sippell WG and Aynsley-Green A. Randomised trial of fentanyl anaesthesia in preterm babies undergoing surgery: effects on the stress response. Lancet 1987;1:243-248.

Fearon I, McGrath PJ and Achat H. 'Booboos': the study of everyday pain among young children. Pain 1996;68:55-62.

Huskisson EC. Measurement of pain. Lancet 1974;2:1127-1131.

Merskey H and Bogduk N. Classification of Chronic Pain: Description of Chronic Pain Syndromes and Definitions of Pain Terms., IASP Press, Seattle, 1994. pp. 209-214.

Purcell-Jones G, Dormon F and Sumner E. Paediatric anaesthetists' perceptions of neonatal and infant pain. Pain 1988;33:181-187.

Stevens BJ, Johnston CC and Gibbins S. Pain assessment in neonates. In: KJS Anand, BJ Stevens and PJ McGrath (Eds.). Pain in neonates, 2nd revised and enlarged edition, Vol. 10, Elsevier, 2000. pp. 101-134.

Chapter 2

The reliability and validity of the COMFORT scale as a postoperative pain instrument in 0 to 3-year-old infants

Based on the article:

The reliability and validity of the COMFORT scale as a postoperative pain instrument in 0 to 3-year-old infants

Monique van Dijk, Josien B. de Boer, Hans M. Koot
Dick Tibboel, Jan Passchier and Hugo J. Duivenvoorden

Pain 84(2000) 367-377

2.1 Abstract

The aim of this study was to test the reliability and validity of the COMFORT scale as a postoperative pain instrument for children aged 0-3 years. Subjects were 158 neonates and toddlers after major abdominal or thoracic surgery. Trained nurses rated the children's pain at 3, 6 and 9 h postoperative on the Pediatric Surgical Intensive Care Unit using the COMFORT and a VAS for pain. Interrater reliability of the COMFORT items proved to be good (Kappa 0.63 to 0.93) for all items with the exception of the item 'Respiratory response', which was moderate (Kappa 0.54). LISREL analyses showed that the structure of the COMFORT data was best represented by three latent variables: COMFORT 'behaviour' with loadings from the behavioural items (Alertness, Calmness, Respiratory response/Crying, Physical movement, Muscle tone and Facial tension) and separate latent variables for 'Heart rate baseline' (HR) and 'Mean arterial blood pressure baseline' (MAP). Factor loadings of the items were invariant across time, indicating stability of the structure.
The latent variables COMFORT 'behaviour' and VAS pain were highly interrelated indicating congruent validity.
Stability of COMFORT 'behaviour' and VAS pain was moderate which might be due to varying painful episodes in this sample. HR and MAP, although stable across time, were weakly related to VAS pain and COMFORT 'behaviour'. These findings support the use of the COMFORT 'behaviour' scale to assess postoperative pain in neonates and infants.

2.2 Introduction

In contrast to a decade ago, there is an increasing awareness among physicians and nurses that pain in neonates and children should be prevented and treated. As a result, there is a growing need for reliable and valid pain instruments that can easily be incorporated into daily care. For the assessment of postoperative pain in preverbal children, a number of observational pain instruments have been developed. Especially for neonates the CRIES (Krechel and Bildner, 1995) and LIDS (Horgan and Choonara, 1996) were developed. For toddlers (from one year), the CHEOPS (McGrath et al., 1985) and the TPPPS (Tarbell et al., 1992) were constructed. Other postoperative pain instruments for infants and toddlers are the POPS (Barrier et al., 1989), the OPS (Hannallah et al., 1987), the FLACC (Merkel et al., 1997) and the MIPS (Buchholz et al., 1998).

Table 1 Overview of psychometric evaluation of postoperative pain instruments for neonates and infants

Pain scale	Content	Reliability		Validity		Age groups
		Interrater	Internal	Criterion	Construct	
CHEOPS	B	Yes	Yes	Compared to VAS after circumcision	CHEOPS 3 min before and after fentanyl bolus, raters not blinded	1 to 7 years
OPS	B and P	No	No	No	Comparison of pain score between controls, caudal and nerve blocks after orchiopexy	18 months to 12 years
POPS	B	No	No	No	Comparison of pain scores between placebo and fentanyl group (after minor surgical procedures)	1 to 7 months
TPPPS	B	Yes	Yes	Comparison between TPPP and VAS nurses and parents after inguinal hernia or hydrocele repair	Comparison between pre- and post-analgesia	1 to 5 years
CRIES	B and P	Yes	No	CRIES compared with OPS	CRIES assessed when analgesia required and 1 h after different surgical procedures	Neonates > 32 weeks
RIPS, POPS, NAPI	B	Yes	Yes	Sensitivity and specificity was examined	Comparison of 'pain' and 'no pain' (based retrospectively on analgesia administration) after different surgical procedures	Neonates to 3 years
LIDS	B	Yes	No	No	Comparison after minor, moderate and major surgery	Neonates
FLACC	B	Yes	No	FLACC compared with OPS and global ratings of nurses	FLACC before and after analgesia, tested after minor surgical procedures	2 months to 7 years
MIPS	B and P	Yes	No	MIPS compared with VAS nurses	Sensitivity and specificity examined	4 to 30 weeks

Abbreviations: CHEOPS, the Children's Hospital of Ontario Pain Scale; OPS, the Objective Pain Scale; POPS, Postoperative Pain Score; TPPPS, Toddler Preschooler Postoperative Pain Scale; CRIES, an acronym for Crying, Requires increased oxygen administration, Expression, Sleeplessness; RIPS, Riley Infant Pain Scale; NAPI, Nursing Assessment of Pain Intensity; LIDS, Liverpool Infant Distress; FLACC, an acronym for Face, Legs, Activity, Cry, and Consolability; MIPS, modified Infant Pain Scale.
B: Behavioural items; P: Physiological items

Schade et al. (1996) employed the RIPS, NAPI and a reduced POPS. Table 1 gives an overview of these postoperative pain instruments and the extent to which the psychometric properties have been tested.

Most instruments include only behavioural items, e.g. facial tension, crying, and body movements, whereas the OPS, CRIES and MIPS also comprise physiological items. Differences between the instruments are mainly confined to the phrasing of the items and the number of response categories. The internal consistency of the instruments was addressed and proved satisfactory for the TPPPS and the RIPS, POPS and NAPI. None of the instruments was tested for stability.

Sensitivity to change, by comparing pain scores before and after analgesics were given, was examined for the CHEOPS, FLACC, CRIES and TPPPS. These analyses were based on small samples (range 20 to 29 cases), because analgesics after minor surgery are only given on demand. Although the COMFORT scale (Ambuel et al,. 1992) was originally designed to assess distress/comfort in ventilated children in an intensive care environment, the items may also be considered indicators of pain, as they are included in other instruments (see Table 1) that were specifically designed to assess pain. We chose the COMFORT scale to assess postoperative pain because the COMFORT comprises both behavioural and physiological items, has five response categories for all items allowing assessment of subtle changes, and is easy to learn for pediatric nurses. Furthermore, the COMFORT was developed to be used in an intensive care environment for children 0-18 years of age. Ambuel et al. (1992) performed preliminary validity and reliability testing of the COMFORT on a limited sample of 37 ventilated infants. Interrater reliability was acceptable and the COMFORT scores correlated 0.75 with an observational VAS for distress. Ambuel et al, (1992) found a two-dimensional structure for the COMFORT; Alertness, Calmness, Movement, Facial tension, and Respiratory response substantially loading on the first factor, and Heart rate, Mean arterial pressure and Muscle tone substantially loading on the second factor.

The COMFORT was predominantly employed to assess level of sedation or distress (El-Khatib et al., 1994; Marx et al., 1994; Reed et al., 1996). Recently, the COMFORT scale was employed to assess procedural pain and proved sensitive to change (Blauer and Gerstmann,1998). As, to our knowledge, the COMFORT scale has never been used to assess postoperative pain, the primary aim of the present study was to assess the psychometric merits of the COMFORT scale as a pain instrument, using repeated measurements.

The main research questions were:

1. What is the reliability of the COMFORT? Specified into: (a) What is the interrater reliability of the PICU nurses on the COMFORT scale? (b) What is the reliability (internal consistency) and stability of the internal structure of the COMFORT scale? (c) What is the stability (test-retest) of the COMFORT scale in the postoperative period?
2. What is the validity of the COMFORT? Confined to: What is the congruent validity of the COMFORT in relation to a VAS for pain?

These research questions were investigated in a large clinical trial that also addressed the efficacy and safety of either intermittent or continuous morphine analgesia.

2.3 Material and methods

Patients

The study sample included neonates of at least 35 weeks gestational age and body weight ≥1500 grams, and infants up to 3 years of age who were admitted for abdominal or thoracic surgery to the Sophia Children's Hospital, Rotterdam. Excluded were children using medication which could influence behavioural assessment e.g. children using muscle relaxants or children with severe neurologic problems.

Table 2 gives the background characteristics of the 158 infants in this study. The sample included a small majority of boys (59%). Most infants (81%) had a major congenital anomaly, not involving the central nervous system, which required surgery. The majority (76%) underwent an abdominal operation. Postoperative mechanical ventilation after surgery was needed in 39% of all cases; 27% required prolonged mechanical ventilation (≥36 hours).

Measures

COMFORT scale

The COMFORT comprises eight items with five response categories each consisting of distinct behavioural descriptions. Six behavioural items (Alertness, Calmness, Muscle tone, Movement, Facial tension, and Respiratory response), and two physiological items: Heart rate (HR) and mean arterial pressure (MAP), are used. For the non-ventilated infants in our study, an item on crying was developed with the response categories 1 'quiet breathing, no

crying', 2 'sobbing or gasping', 3 'moaning', 4 'crying', and 5 'screaming'. Care was taken that the phrasing of this item was similar to the other items of the COMFORT. (See Appendix A for extended COMFORT scale.) In the study, raters observed each child for two minutes at bedside. During this period, they recorded the HR and MAP values from a Hewlett Packard M2350a monitor every 20 seconds, six times in total. Recording of the MAP required an arterial line and was computed as diastolic pressure + {(systolic - diastolic pressure)/3}. Shortly before the end of the observation, the Muscle tone of the child was assessed by lifting an arm or leg of the child. After two minutes, each item was scored. A total score was calculated by coding the scores on the individual items, theoretically ranging from 8 to 40. Translation of the COMFORT items was performed in collaboration with the authors of the original COMFORT scale.

Table 2 Background characteristics of the patients (N=158)

Variable	Number of patients	%[a]
Gender		
Male	94	59
Female	64	41
Age group		
0 to 4 weeks	56	35
1 to 6 months	47	30
7 to 12 months	23	15
1 to 3 years	32	20
Diagnosis[b]		
congenital anomalies	128	81
acquired diseases	30	19
Surgery		
superficial[c]	8	5
abdominal	120	76
thoracic	30	19
Postoperative mechanical ventilation		
None	96	61
Short term (6 to <36 hours)	19	12
Prolonged (≥36 hours)	43	27

[a] Rounded percentages.
[b] Congenital anomalies of which 42% digestive tract obstruction, 10% diaphragmatic hernia, 7.5% Hirschsprung disease, 7.5% malignancies; acquired diseases: 56% necrotizing enterocolitis, 27% intussusception, others 17%.
[c] Superficial: e.g. retroperitoneal surgery.

Table 3 Means and standard deviations for COMFORT items and VAS pain at baseline, after installation at the PSICU, and 3, 6 and 9 hours postoperative

COMFORT[a]		Baseline	Postoperative			
			After installation	3h postop.	6h postop.	9h postop.
Alertness	Mean	3.1	2.1	2.3	2.3	2.2
	SD	1.3	1.1	1.1	1.1	1.1
	n	145	157	158	158	158
Calmness	Mean	1.6	1.8	2.0	2.1	2.0
	SD	1.0	1.0	1.1	1.0	1.0
	n	145	157	158	158	158
Respiratory response	Mean	2.0	1.5	1.7	1.8	1.7
	SD	1.0	0.7	0.8	0.8	0.7
	n	19	60	60	61	58
Crying	Mean	1.4	1.7	2.1	2.2	2.0
	SD	0.9	1.0	1.1	1.2	1.2
	n	128	96	98	96	100
Movement	Mean	3.0	2.2	2.5	2.5	2.5
	SD	1.2	1.0	1.0	1.0	1.0
	n	145	157	158	156	158
Muscle tone	Mean	3.0	2.9	3.2	3.2	3.2
	SD	0.6	1.0	0.8	0.7	0.7
	n	144	157	158	157	158
Facial tension	Mean	2.5	2.4	2.7	2.7	2.6
	SD	0.8	0.9	1.0	0.9	0.9
	n	145	157	157	156	157
Blood pressure baseline[b]	Mean	-	2.8	2.6	2.7	2.5
	SD	-	1.5	1.4	1.4	1.4
	n		152	153	153	155
Heart rate baseline[b]	Mean	-	2.7	2.8	2.9	2.8
	SD	-	1.4	1.3	1.4	1.3
	n		158	158	158	158
VAS pain	Mean	0.4	2.1	2.7	2.5	2.2
	SD	1.3	2.3	2.3	2.0	2.1
	n	143	157	157	158	157

[a] All COMFORT item scores range from 1 to 5
[b] At baseline blood pressure and heart rate were assessed to calculate the baseline needed for the COMFORT items (see Appendix A)

Visual Analogue Scale

The VAS is a horizontal continuous 10-cm line with the anchors 'no pain' at the left side and 'extreme pain' at the right side. The VAS rating was done after completing the COMFORT scale. In a subsample (n=26), the VAS was also rated prior to the COMFORT

scale to evaluate the effect of the 2-min observation and COMFORT scoring on the VAS pain.
In pain research, the VAS is frequently used as an observational instrument. Good interrater reliability was found (Varni et al., 1987: Lawrence et al., 1993) and the VAS proved to be highly associated with other postoperative pain instruments (Tarbell et al, 1992; McGrath et al., 1985).

Procedure
The study was approved by the Medical Ethical Committee of the Sophia Children's Hospital. Written informed consent was obtained from the parents by the pediatric intensivist or anaesthetist. Before the operation, the nurse or anaesthetist performed baseline assessments of heart rate (HR) and mean arterial blood pressure (MAP) at the PSICU. Anaesthesia was given in line with standard procedures. After induction, an arterial line was placed from which blood samples were drawn, and, subsequently, HR and MAP were assessed.

Design
The COMFORT scale and the VAS were assessed prior to surgery, after installation of the child at the PSICU, and every three h thereafter up to 36 h after surgery, for a total of 13 assessments. Pain assessment, blood sampling, handling of the child and administration of the morphine or placebo bolus was done in this order every three h during the first 36 hours after surgery.

Training of observers
For adequate use of the COMFORT scale the nurses and anaesthesiologist attended a 2-h training session during which the COMFORT scale was explained by means of videotaped behaviour and in vivo observations of children at the PSICU. Because two to five nurses were trained at the same time, discussion was helpful to solve possible misinterpretations. After the course, each newly trained nurse completed ten COMFORT assessments (scored on the PSICU on an infant after surgery under 3 years of age) with one of the trainers or an experienced colleague. When interrater reliability was acceptable, according to a linearly weighted Cohen's Kappa between 0.40 and 0.60 (Fleiss 1981), the nurse was allowed to score children for the study.

Analytic strategy

To analyse the psychometric qualities of the COMFORT scale several measurement models were built.

The respective measurement models were tested with the SIMPLIS version of LISREL 8.2 for Windows (Jöreskog and Sörbom, 1993).

Parameters were estimated by using the maximum likelihood procedure, based on the covariance matrix of the observed variables. Due to the relatively small sample size (n=158), the number of repeated measurements had to be restricted to 3, 6, and 9 h postoperative assessments. These equally spaced time intervals were chosen because they were considered to be representative of the postoperative period. The assessment immediately after installation at the PSICU was not used because the children might still be under influence of the anaesthesia. (Table 7 in Appendix B gives the correlation matrix of COMFORT items and VAS pain off-diagonal, and mean and standard deviations on the diagonal of the matrix).

Because the VAS consisted of only one item, measurement error for this variable was fixed at 20% of the total variance of the observed VAS score. The loading of the COMFORT item Calmness was fixed at 1.0 in order to measure the latent variable in the same units as the observed variables.

The following performance measures of overall fit were used:

1. χ^2 for model fit: a non-significant value indicates that the model at issue can not be rejected. To account for the effect of sample size on χ^2, the χ^2 /*df* was also employed;
2. Standardized root mean squares of residuals (SRMR): the lower the SRMR the better the model fits;
3. Goodness-of-fit adjusted for *df* (AGFI) which measures how much better the model fits as compared to no model at all, with a theoretical range from 0.0 (no fit at all) to 1.0 (perfect fit);
4. Root mean squares error of approximation (RMSEA): a value of 0.05 indicates a close fit and values up to 0.08 represent reasonable errors of approximation in the population.

To establish the reliability and validity of the COMFORT scale, we tested the following assumptions with LISREL:

1. Invariant error variances across time for corresponding items. This involves comparing models with the equality constraints of equal error variances compared to freely varying error variances for corresponding items;
2. The stability of the factor structure across time (i.e. factorial invariance). This assumption tests whether models with invariant factor loadings across time for corresponding items are favourable compared to less restrictive models in which factor loadings across time for corresponding items may vary;
3. Stability between the latent variables across time. Models with equal stability coefficients are compared with models in which the stability coefficients may differ. In our study, two stability coefficients (i.e. between 3 and 6, and between 6 and 9 hours postoperatively) were compared;
4. Lag one error covariances for corresponding items of the COMFORT scale. The addition of lag one error covariances for corresponding items in a longitudinal design indicates that part of the error is measurement specific. Models without error covariances are more restrictive.

Testing these four assumptions implies comparison of sixteen models (2^4 models). Nested models are compared by means of χ^2 differences. An example of a nested model is when the free parameters of one model are a subset of the free parameters in a second (Bollen, 1989; Jöreskog and Sörbom, 1993). In addition, the two-factor model found by Ambuel et al. (1992) was fitted to the data.
The reliability of the COMFORT scale (composite of items loading on one latent variable) based on the congeneric model was estimated (Reuterberg and Gustafsson, 1992), which is comparable with Cronbach's alpha.
With respect to power, a sample size of >100 for LISREL analysis is considered sufficient (Boomsma, 1985).

Missing data
A total of 175 children entered the study; of these, 17 cases with more than 20% missing data were excluded from analysis. Missing data were mainly due to a failing arterial line. For the remaining 158 cases, missing data (0.7% of all data) were estimated by means of

TWOSTEP regression analysis of BMDPAM Dynamic 7.0 which uses the two best predictor variables to estimate the missing value.

2.4 Results

Interrater reliability

Thirty-nine PSICU nurses and two anaesthesiologists were trained to use the COMFORT scale. Table 4 depicts the linearly weighted Kappa's for the individual items. Interrater reliability, represented by linearly weighted Kappa's, was substantial to almost perfect for almost all items ranging from 0.63 for Facial tension to 0.93 for HR and MAP. Interrater reliability for Respiratory response was moderate (0.54). The median Kappa was 0.70.

Table 4 Linearly weighted Kappa's for COMFORT items[a)]

Comfort Item	Linearly weighted Kappa	No. of paired observations
Alertness	0.74	302
Calmness	0.69	302
Respiratory response	0.54	131
Crying	0.70	170
Physical movement	0.70	302
Muscle tone	0.66	302
Facial tension	0.63	296
Blood pressure baseline	0.93	232
Heart rate baseline	0.93	290

[a)] Ten paired assessments of 39 nurses with trained nurses or trainers were used to calculate the Kappa statistics. The assessments were performed prior to inclusion in the trial and were obtained from 94 infants (0-3 years) staying at the PSICU.

Table 5 Fit indices of models tested[a)]

Model	Model description				Overall fit indices						
	Lag one error covariances	Equal error variances	Equal factor loadings	Equal stability coefficients	χ^2	df	χ^2 /df	p	RMSEA	SRMR	AGFI
1	-	+	+	+	535.76	293	1.8	0.00	0.06	0.05	0.75
2	-	+	+	-	534.94	289	1.8	0.00	0.06	0.05	0.74
3	-	+	-	+	527.27	283	1.9	0.00	0.06	0.05	0.74
4	-	+	-	-	526.59	279	1.9	0.00	0.06	0.05	0.74
5	-	-	+	+	432.96	281	1.5	0.00	0.06	0.05	0.77
6	-	-	+	-	432.18	277	1.6	0.00	0.06	0.05	0.77
7	-	-	-	+	422.85	271	1.6	0.00	0.06	0.04	0.77
8	-	-	-	-	422.23	267	1.6	0.00	0.06	0.04	0.76
9	+	+	+	+	442.40	281	1.6	0.00	0.04	0.05	0.79
10	+	+	+	-	441.64	277	1.6	0.00	0.04	0.05	0.78
11	+	+	-	+	434.51	271	1.6	0.00	0.04	0.05	0.78
12	+	+	-	-	433.85	267	1.6	0.00	0.04	0.04	0.78
13	+	-	+	+	342.74	269	1.3	0.00	0.04	0.05	0.81
14	+	-	+	-	342.07	265	1.3	0.00	0.04	0.05	0.81
15	+	-	-	+	333.75	259	1.3	0.00	0.04	0.04	0.81
16	+	-	-	-	333.14	255	1.3	0.00	0.04	0.04	0.81

[a)] + : included in the model ; - :not included

VAS scoring before and after COMFORT

For 26 children, nurses scored a VAS pain before and after the COMFORT scale during the project, to estimate the effect of the two-minute observation and the scoring of the COMFORT on the VAS pain.

Correlations between VAS 'before' and COMFORT 'behavioural' ranged from 0.64 to 0.73. Correlations between VAS 'after' and COMFORT 'behavioural' ranged from 0.79 to 0.83.

Reliability and validity

The two-factor model of Ambuel et al.(1992) failed to converge. Further analyses were based on the models with three latent variables, evaluating the four assumptions mentioned above.

The longitudinal data analyses revealed the best fit when three latent variables were used for the COMFORT scale; one latent variable ('Comfort behaviour'), for the behavioural items, one for MAP ('MAP') and one for HR ('HR'). The error variance of MAP and HR was a priori fixed at 20%, similar to the VAS pain.

Table 5 shows the results of the 16 tested models. Models 9 to 16 differed from models 1 to 8 in their freeing of lag one error covariances.

The difference χ^2 's for all nested models (model 1 vs model 9, model 2 vs model 10 and so on) were all approximately 90 with *df* 12, ($P<0.001$), favouring the models allowing for lag one error covariance, which were models 9 to 16. Examining the models 9 to 16, the ratio of χ^2 /*df* and the other fit indices indicated models 13 to 16 as the best fitting. The differences between model 13 (more restrictive) and the models 14 and 15 (less restrictive) were not significant. As a result, model 13 was regarded most plausible. The fit indices were satisfactory: χ^2 /*df* ratio of 1.3, RMSEA of 0.04, SRMR of 0.05, and AGFI of 0.81. This model consisted of unequal error variances across time (rejecting assumption 1), equal loadings across time for corresponding items (in accordance with assumption 2), equal stability coefficients between assessments across time (in accordance with assumption 3), and allowing lag one error covariances (as suggested in assumption 4). Figure 1 shows the path diagram for this structural model.

The stability coefficients (fixed invariant across time) were high for the physiological latent variables MAP and HR (0.89 and 0.82, respectively) and moderate for COMFORT 'behaviour' and VAS (0.58 and 0.59, respectively).

Congruent validity of the COMFORT 'behaviour' was implied by high correlations between this latent variable and VAS on all three assessments (0.96, 0.89, and 0.90, respectively) moderately with MAP (0.35, 0.32 and 0.22, respectively) and low with HR (0.08, 0.10 and 0.13, respectively). The correlations between MAP and HR were low (0.05, 0.25 and 0.19, respectively).

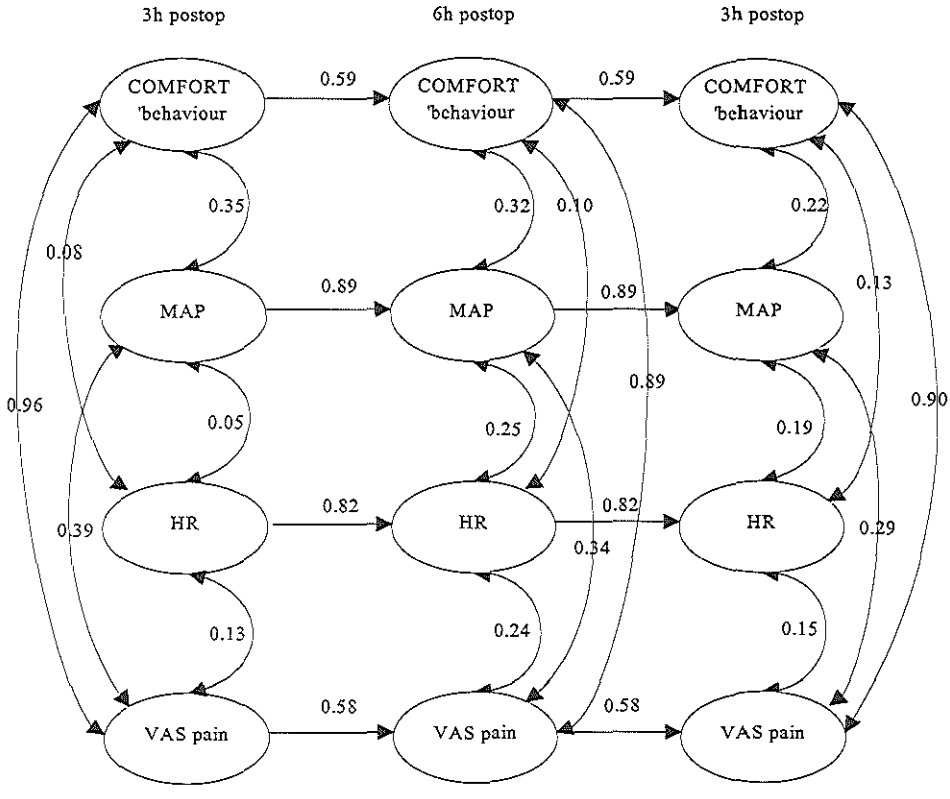


Fig 1. Structural part of model 13 from Table 3; circles represent latent variables, straight unidirectional arrows indicate stability coefficients, curved bidirectional arrows symbolise correlations between latent variables. MAP=Mean Arterial Pressure, HR=Heart Rate.

The internal consistency reliability of the composite score 'COMFORT behaviour' for the three assessments was 0.90, 0.92 and 0.92, respectively.
Table 6 depicts the loadings of the behavioural COMFORT items on the latent variable 'COMFORT behaviour'. The loadings of the behavioural items were consistent across time and were significant. With Calmness fixed at 1.00, the other items had high loadings (0.76 to 0.85) with the exception of Muscle tone, with a loading of 0.51. The items were well represented by the model.

Associations between the latent variables HR and VAS pain at 3 and 9 h postoperative assessments were non-significant (0.13 and 0.15, respectively) and moderate but significant at 6 h (0.24). The associations between VAS pain and MAP were moderate though significant (0.39, 0.34 and 0.29, respectively). These findings suggest that the items Heart rate and Blood pressure have limited validity as measures of postoperative pain.

Table 6 Loadings for COMFORT items on COMFORT 'behaviour

		Loading	
Item	Postoperative hours	Unstandardised	Standardised
Calmness	3	1.00	1.00
	6	1.00	0.99
	9	1.00	0.97
Alertness	3	0.84	0.84
	6	0.84	0.83
	9	0.84	0.81
Crying/Respiratory response	3	0.85	0.86
	6	0.85	0.84
	9	0.85	0.82
Movement	3	0.77	0.78
	6	0.77	0.76
	9	0.77	0.74
Muscle tone	3	0.51	0.52
	6	0.51	0.51
	9	0.51	0.50
Facial tension	3	0.76	0.77
	6	0.76	0.76
	9	0.76	0.73

2.5 Discussion

In comparison with earlier, comparable studies the objective and additional value of this study was the simultaneous estimation of reliability and validity through structural modelling in a large study population.
Using longitudinal data from a sample of 158 neonates and infants aged 0-3 years after major abdominal or thoracic surgery, an empirical statistical model was fitted which justifies the use of the behavioural part of the COMFORT scale as a postoperative pain instrument.
One could argue that the COMFORT scale was developed to asses level of distress on the one end of the continuum and sedation on the other end. However, distress encompasses pain. This is reflected in the similarity of the content of pain instruments and the COMFORT scale. Furthermore, the COMFORT scale was used in a pain related context, excluding children receiving sedative medication or muscle relaxants. In addition, high correlations between COMFORT and VAS pain scores were found. Therefore, it seems unlikely that the COMFORT scores only reflect level of sedation.

Interrater reliability
The interrater reliability for all COMFORT items was good, except for Respiratory response for which it was moderate. Based on nurses' experiences, this might be due to differences in their implicit interpretation of infants' responses towards mechanical ventilation. During future training, it should be emphasised that interpretations should be limited to notations on the records.
The newly incorporated item Crying exhibited good interrater reliability (Kappa 0.70), and validity (high and significant correlations ranging from 0.65 to 0.72 between VAS pain and the item Crying for non-ventilated cases) which makes it possible to extend the COMFORT to postoperative patients who are not ventilated, thus enhancing the usefulness of the COMFORT for daily clinical practice.
Interrater reliability of HR and MAP was excellent (Kappa's were 0.93 for both items), in contrast to Ambuel et al.(1992). This has to be attributed to the fact that the nurses noted the HR and MAP from the monitor six times every 20 s (with the aid of a stopwatch) during the 2-min scoring period.
The fact that the weighted Kappa's in this study were based on 39 nurses and generally remained good, showed that it is possible to train nurses to observe pain-related behaviour

in a reliable way. It was considered relevant to train the PICU nurses, instead of researchers, to stimulate implementation of the pain instrument at a later stage in the hospital.

Reliability of the COMFORT scale

The stability of the latent variables COMFORT 'behaviour' and VAS was distinct though not high. This might be due to the fact that infants in our sample had painful episodes at different intervals. Twenty infants (13%) had their maximum VAS pain score 3 h postoperatively, 27 (17%) at 6 h, and 21 (13%) infants at 9 h after surgery. Additionally, a considerable number of infants (32% never had a VAS >4) were comfortable with the administered morphine dosage during the first 36 h.

The latent variables MAP and HR demonstrated to have considerable stability across time. Correlations between the latent variables MAP and HR were low (ranging from 0.05 to 0.25), similar to the raw correlations between the two items (ranging from 0.04 to 0.20), explaining the lack of fit in the two-factor model. This limited association between vital signs has, to our knowledge, not been reported before and requires further research.

All factor loadings in the selected model were in the 0.50-0.85 range and were invariant across time for corresponding items, which is desirable.

The Muscle tone item had moderate loadings compared to the other items of the COMFORT scale. However, we suggest to maintain this item because very ill infants may have limited energy to manifest their pain whereas muscle tone may be increased. Muscle tone was assessed by lifting an arm or leg. Mere observation of muscle tone would be difficult, because the intermediate categories of muscle tone (e.g. normal muscle tone), require physical examination.

Validity of the COMFORT scale

We were able to show congruent validity between the COMFORT 'behaviour' and the VAS pain. The high correlations may be inflated by the fact that the same nurse assessed the VAS and the COMFORT. This methodological drawback was unavoidable for practical reasons.

The variations in correlation between 'before VAS' and COMFORT compared to 'VAS after' with the COMFORT, might be ascribed to the 2-min observation period of the child. Nurses indicated that the fact that they were stimulated to observe a child for 2 min was valuable in itself. The sensitivity to change of the COMFORT scale could not be estimated

in this study. As the current gold standard for postoperative pain management after major surgery is to prevent pain as opposed to analgesia on demand (Broadman, 1999), all children receive morphine. Therefore, it was not feasible to obtain 'pure' pre-post analgesia data for this sample. Blauer and Gerstmann (1998) found significant changes on the COMFORT scale before and during procedures, such as endotracheal tube suctioning, intubation, intravenous catheter insertion and diaper change. Future study should evaluate the sensitivity to change of the COMFORT after minor surgery when analgesia is given on demand.

MAP and HR showed limited validity with the VAS pain as reference. This might be provoked on the one hand by the construction of the MAP and HR response categories or on the other hand by the complexity of the association between longer lasting pain and physiological outcome. The MAP and HR items contain five response categories that compare six MAP or HR values with a (preoperative) baseline value. Because some infants were stressed prior to surgery due, to e.g. preoperative invasive procedures, this could induce high baseline values, resulting in relatively low postoperative values. The restricted value of physiological measures as postoperative pain indicators is mentioned in literature (Beyer and Wells, 1989; Tyler et al., 1993). It has been argued that autonomic responses adapted to longer lasting pain (Beyer and Wells, 1989), and other factors such as the patient's disease and use of opioids, may modify physiological responses (Tyler et al., 1993). The limited specificity of physiological parameters in pain assessment is also mentioned (Tyler et al. 1993). Research on physiological parameters and pain are restricted to procedural pain, mostly in (premature) neonates (Craig et al., 1993; McIntosh et al., 1993; Johnston et al., 1995). In a review article on pain measurement, Franck and Miaskowski (1997) concluded, that vital signs may not be specific enough to distinguish between painful and non-painful procedures.

Three postoperative instruments contain vital signs, the CRIES, OPS and MIPS (see Table 1). Only Buchholz et al. (1998) examined the contribution of the vital signs within the MIPS and concluded that vital signs did not add to the information given by the behavioural part of the MIPS, which is consistent with our findings.

However, in clinical practice, nurses and physicians use vital signs (heart rate, blood pressure and oxygenation) in their judgement of pain in infants (Burokas, 1985; Purcell-Jones et al. 1988). Because of this contradiction, further research with clinical data on the contribution of vital signs in postoperative pain assessment is required.

Future research

This study addressed pain assessment in neonates and infants after major surgery. Because the sample on which these study results were based included considerably more infants than 1 to 3-year-olds (80 vs. 20%) the conclusions are possibly limited due to the skewness of age in our sample. Since the indices of behavioural distress and their underlying structure may be developmentally sensitive, our findings require replication in a sample including more or exclusively older children.

To extend the applicability of the COMFORT scale it could be tested on other surgical patients (e.g. after minor surgery) in different hospital settings (regional vs. university) and with different judges (more and less experienced nurses and physicians), to distinguish systematic sources of variation.

2.6 Acknowledgements

The authors would like to thank the staff and nurses of the Pediatric Surgical Intensive Care of the Sophia Children's Hospital for their contribution to this study. We also thank the staff and members of the surgical ward and the department of Anesthesiology, especially Dr. J.Bouwmeester, for their contribution to this study. We would also like to thank the parents and children who participated in this study.

This study was supported by a research grant from NWO (Dutch Organisation for Scientific Research, grant nr. 940-31-031).

2.7 References

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

Barrier G, Attia J, Mayer MN, Amiel-Tison C and Shnider SM. Measurement of post-operative pain and narcotic administration in infants using a new clinical scoring system. Intensive Care Med 1989;15:s37-s39.

Beyer JE and Wells N. The assessment of pain in children. Pediatr Clin North Am 1989;36:837-854.

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

Bollen KA. Structural equations with latent variables, Wiley, New York, 1989.

Boomsma A. Nonconvergence, improper solutions, and starting values in LISREL maximum likelihood estimation. Psychometrika 1985;50:229-242.

Broadman LM. Blocks and other techniques pediatric surgeons can employ to reduce postoperative pain in pediatric patients. Semin Pediatr Surg 1999;8:30-33.

Buchholz M, Karl HW, Pomietto M and Lynn AM. Pain scores in infants: a modified infant pain scale versus visual analogue. J Pain Symptom Manage 1998;15:117-124.

Burokas L. Factors affecting nurses' decisions to medicate pediatric patients after surgery. Heart and Lung 1985;14:373-379.

Craig KD, Whitfield MF, Grunau RVE, Linton J and Hadjistavropoulos HD. Pain in the preterm neonate: behavioural and physiological indices. Pain 1993;52:287-299.

El-Khatib MF, Chatburn RL, Potts DL, Blumer JL and Smith PG. Mechanical ventilators optimized for pediatric use decrease work of breathing and oxygen consumption during pressure-support ventilation. Crit Care Med 1994;22:1942-1948.

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

Franck LS and Miaskowski C. Measurement of neonatal responses to painful stimuli: a research review. J Pain Symptom Manage 1997;14:343-378.

Hannallah RS, Broadman LM, Belman AB, Abramowitz MD and Epstein BS. Comparison of caudal and ilioinguinal/iliohypogastric nerve blocks for control of post-orchiopexy pain in pediatric ambulatory surgery. Anesthesiology 1987;66:832-834.

Horgan M and Choonara IA. Measuring pain in neonates: an objective score. Paediatr Nurs 1996;8:24-27.

Johnston CC, Stevens BJ, Yang F and Horton L. Differential response to pain by very premature neonates. Pain 1995;61:471-479.

Jöreskog KG and Sörbom D. LISREL 8: Structural equation modeling with the SIMPLIS command language, Scientific Software International, Chicago, 1993.

Krechel SW and Bildner J. CRIES: a new neonatal postoperative pain measurement score. Initial testing of validity and reliability. Paediatr Anaesth 1995;5:53-61.

Lawrence J, Alcock D, McGrath PJ, Kay J, Brock MacMurray S and Dulberg C. The development of a tool to assess neonatal pain. Neonatal Netw 1993;12:59-66.

Marx CM, Smith PG, Lowrie LH, Hamlett KW, Ambuel B, Yamashita TS and Blumer JL. Optimal sedation of mechanically ventilated pediatric critical care patients. Crit Care Med 1994;22:163-170.

McGrath PJ, Johnson G, Goodman JT, Schillinger J, Dunn J and Chapman JA. CHEOPS: A behavioral scale for rating postoperative pain in children. In: HL Fields, R Dubner and F Cervero (Eds.). Advances in Pain Research and Therapy, Vol. 9, Raven Press, New York, 1985. pp. 395-402.

McIntosh N, Van Veen L and Brameyer H. The pain of heel prick and its measurement in preterm infants. Pain 1993;52:71-74.

Merkel SI, Voepel-Lewis T, Shayevitz JR and Malviya S. The FLACC: A behavioral scale for scoring postoperative pain in young children. Pediatr Nurs 1997;23:293-297.

Purcell-Jones G, Dormon F and Sumner E. Paediatric anaesthetists' perceptions of neonatal and infant pain. Pain 1988;33:181-187.

Reed MD, Yasmashita TS, Marx CM, Myers CM and Blumer JL. A pharmacokinetically based propofol dosing strategy for sedation of the critically ill, mechanically ventilated pediatric patient. Crit Care Med 1996;24:1473-1481.

Reuterberg S and Gustafsson J. Confirmatory factor analysis and reliability: Testing measurement model assumptions. Educ Psychol Meas 1992;52:795-811.

Schade JG, Joyce BA, Gerkensmeyer J, Keck JF. Comparison of three preverbal scales for postoperative pain assessment in a diverse pediatric sample. J Pain Symptom Manage 1996;12:348-359.

Tarbell SE, Cohen TI and Marsh JL. The Toddler-preschooler postoperative pain scale: an observational scale for measuring postoperative pain in children aged 1-5. Preliminary report. Pain 1992;50:273-280.

Tyler DC, Tu A, Douthit J and Chapman CR. Toward validation of pain measurement tools for children: a pilot study. Pain 1993;52:301-309.

Varni JW, Thompson KL and Hanson V. The Varni/Thompson Pediatric Pain Questionnaire. I. Chronic musculoskeletal pain in juvenile rheumatoid arthritis. Pain 1987;28:27-38.

2.8 Appendix A COMFORT Scale

ALERTNESS		
Deeply asleep	1	
Lightly asleep	2	
Drowsy	3	
Fully awake and alert	4	
Hyper-alert	5	
CALMNESS		
Calm	1	
Slightly anxious	2	
Anxious	3	
Very anxious	4	
Panicky	5	
RESPIRATORY RESPONSE		
No coughing and no spontaneous respiration	1	
Spontaneous respiration with little or no response to ventilation	2	
Occasional cough or resistance to ventilator	3	
Actively breathes against ventilator or coughs regularly	4	
Fights ventilator; coughing or choking	5	
CRYING[1)]		
Quiet breathing, no crying	1	
Sobbing or gasping	2	
Moaning	3	
Crying	4	
Screaming	5	
PHYSICAL MOVEMENT		
No movement	1	
Occasional, slight movement	2	
Frequent, slight movements	3	
Vigorous movement limited to extremities	4	
Vigorous movements including torso and head	5	
MUSCLE TONE		
Muscles totally relaxed; no muscle tone	1	
Reduced muscle tone	2	
Normal muscle tone	3	
Increased muscle tone and flexion of fingers and toes	4	
Extreme muscle rigidity and flexion of fingers and toes	5	
FACIAL TENSION		
Facial muscles totally relaxed	1	
Facial muscle tone normal; no facial muscle tension evident	2	
Tension evident in some facial muscles	3	
Tension evident throughout facial muscles	4	
Facial muscles contorted and grimacing	5	
BLOOD PRESSURE (MAP) BASELINE		
Blood pressure below baseline	1	
Blood pressure consistently at baseline	2	
Infrequent elevations of 15% or more above baseline (1-3 during 2 minutes observation)	3	
Frequent elevations of 15% or more above baseline (> 3 during 2 minutes observation)	4	
Sustained elevations of 15% or more	5	
HEART RATE BASELINE		
Heart rate below baseline	1	
Heart rate consistently at baseline	2	
Infrequent elevations of 15% or more above baseline (1-3 during 2 minutes observation)	3	
Frequent elevations of 15% or more above baseline (> 3 during 2 minutes observation)	4	
Sustained elevations of 15% or more	5	

[1)] Newly incorporated item for non-ventilated infants

2.9 Appendix B

Tabel 7 Minimum and maximum correlations between COMFORT-items and VAS pain from the 3, 6 and 9 hours postoperative assessments.[a)]

	Alertness	Calmness	Cry/Respiratory response	Movement	Muscle tone	Facial tension	MAP	HR	VAS
Alertness	**2.2 to 2.3 (1.1)**								
Calmness	0.72 to 0.77	**2.0 to 2.1 (1.0 to 1.1)**							
Cry/Respiratory response	0.60 to 0.64	0.72 to 0.76	**1.9 to 2.0 (1.0 to 1.1)**						
Movement	0.54 to 0.69	0.71 to 0.73	0.53 to 0.66	**2.5 (1.0)**					
Muscle tone	0.40 to 0.52	0.57 to 0.67	0.44 to 0.59	0.50 to 0.54	**3.2 (0.7 to 0.8)**				
Facial tension	0.56 to 0.61	0.74 to 0.76	0.59 to 0.71	0.64 to 0.69	0.55 to 0.63	**2.6 to 2.7 (0.9 to 1.0)**			
MAP	0.11 to 0.32	0.24 to 0.27	0.12 to 0.26	0.07 to 0.32	0.08 to 0.28	0.14 to 0.24	**2.5 to 2.7 (1.4)**		
HR	0.05 to 0.15	0.06 to 0.13	0.01 to 0.11	0.02 to 0.04	0.00 to 0.09	0.11 to 0.17	0.04 to 0.20	**2.8 to 2.9 (1.3 to 1.4)**	
VAS	0.59 to 0.68	0.73 to 0.82	0.66 to 0.66	0.57 to 0.66	0.53 to 0.68	0.70 to 0.75	0.23 to 0.33	0.10 to 0.20	**2.3 to 2.7 (2.0 to 2.3)**

[a)] Means and standard deviations are depicted in bold on the diagonal

Chapter 3

The observational Visual Analogue Scale in pediatric pain assessment: Useful tool or good riddance?

Based on the article:

The observational Visual Analogue Scale in pediatric pain assessment: Useful tool or good riddance?

Monique van Dijk, Hans M. Koot, Huda Huijer Abu Saad, Dick Tibboel, Jan Passchier (submitted)

3.1 Abstract

We reviewed the available English pediatric pain literature and selected those studies that reported quantitative information on reliability and/or validity and optimal cutoff points for the Visual Analogue Scale (VAS), when used as an observational pediatric pain tool. Available psychometric findings concerning the observational VAS (VAS_{obs}) are promising. Further work needs to be done on intraobserver reliability, sensitivity to change, and optimal cutoff points. In conclusion, we argue that the VAS_{obs} is a helpful tool next to a validated pain instrument. While most pain instruments are based on detailed behavioural observations, the global rating on the VAS_{obs} may account for additional knowledge on individual variations in pain sensitivity, idiosyncratic behaviours, and situational influences.

3.2 Introduction

A frequently used tool to quantify pain intensity is the Visual Analogue Scale (VAS) (Ho et al., 1996; Huskisson, 1974; Scott and Huskisson, 1976). Its application extends from self-report in adults and children to observational tool in children below 4 years of age. With the introduction of numerous validated pediatric pain tools during the last decades, further use of the VAS may be unnecessary or even unwanted. In this article we intend to demonstrate that the VAS is still useful.

Originally, the VAS is a tool to measure subjective phenomena like pain, anxiety, and fatigue (Aitken, 1969; Huskisson, 1974; Scott and Huskisson, 1976). It usually consists of a 10 cm line, either vertical or horizontal, that separates extreme boundaries of the phenomenon being measured. At the extremes a verbal description is given of the phenomenon. For pain this implies at the left side 'no pain' and at the right side of the line 'pain as bad as it could be' or 'worst pain possible'. Patients (or observers) estimate the level of pain by making a mark on the line. Figure 1 gives the layout of a VAS for pain.

Figure 1 Example of a Visual Analogue Scale

The VAS for self-report (VAS_{sr}) has been validated both for adults (Huskisson, 1974) and for children over 5 years (Abu-Saad, 1984; McGrath and Unruh, 1994). Different adapted versions of the VAS_{sr} have been developed to assess pain in young children (from 4 years on), for example the Visual Analogue Toy using a koala that can be moved higher on a wooden pole with increasing pain (Arts et al., 1994), a 'do-it-yourself' VAS for children from 7 years of age, using a wooden tongue depressor (Benini et al., 1996), a red and white colour VAS with an increase in red indicating more pain (Maunuksela et al., 1987), and the Coloured Analogue Scale (CAS) with an increase from light pink to deep red at the top for pain intensity (McGrath et al., 1996). The strengths and limitations of the VAS_{sr} in general have been reviewed in several articles (Gift, 1989; McCormack et al., 1988; Miller and Ferris, 1993; Wewers and Lowe, 1990). Its strengths are considered to be its ease of use, good reliability and validity, and its metric that enables parametric testing. Limitations are the difficulty for some subjects to mentally transform a subjective sensation into a mark on a straight line and the unreliability of the use of only a single item representing pain intensity or level of suffering. Furthermore, it may be incorrect to compare VAS scores between subjects for research reasons considering the large idiosyncrasies in VAS ratings that were recently described (Williams et al., 2000). Because of its strength as a self-report measure, use of the VAS was extended to observational pain assessment. In this application, an observer, e.g. a nurse, uses the VAS to rate the intensity of the pain experienced by others (further referred to as VAS_{obs}). Impressions of pain intensity, however, may vary across different observers (Huijer Abu-Saad et al., 1998). This may be due to differences in experience with painful situations/patients/persons (e.g. novice opposed to experienced nurse), differences in ideas and knowledge about pain and pain expression (e.g. cultural differences, family influences), and differences in the relationship with the observed child (e.g. parent versus nurse). Furthermore, it is unclear which observations or clues observers use when applying the VAS_{obs}. It is thus imperative to review the psychometric properties of the VAS_{obs}. Although several publications reported on the reliability and validity of the VAS_{obs}, this information has never been drawn together for a good overview of its strengths and drawbacks. Therefore, we reviewed the available English pediatric pain literature and selected those studies that reported quantitative information on reliability and/or validity of the VAS_{obs}. In addition, we looked for evidence on optimal cutoff points (also known as cut scores, cut-off scores, cutpoints, or standards) on the VAS to discriminate between different pain states.

Table 1 Studies reporting on interobserver reliability of VAS_{obs}

First author, year	Pain situation	Sample size and age range	Interobserver-reliability	Mean (SD) of VAS ratings
	Surgery			
McGrath, 1985	First hour after circumcision	N=30, 1 to 7 years, 127 ratings	VAS_{nurse} with $VAS_{research\ assistant}$ r=0.91	-
O'Hara, 1987	After major orthopaedic surgery	N=25, 7 to 17 years	VAS_{nurse} with VAS_{parent} r=0.50	-
Hendrickson, 1990	After major surgery	N=46, 1 to 16 years	VAS_{nurse} with VAS_{parent} r=0.75	-
LaMontagne, 1991	After surgery	N=13, 8 to 18 years	VAS_{nurse} with $VAS_{physician}$ r= 0.90	Nurse 2.8 (2.0) and Physician 2.1 (2.0)
Romsing,1996	After tonsillectomy	N=100, 3 to 15 years	$VAS_{2\ nurses}$ r=0.52 and 0.60	Before analgesic Nurse I : 3.9 (2.5) Nurse II : 3.3 (2.3) After analgesic Nurse I : 1.6 (1.5) Nurse II : 1.5 (1.4)
Miller, 1996	After surgery	N=20, 7 to 11 years	VAS_{nurse} with VAS_{mother} r=0.36, 0.47 and 0.55	-
	Procedural pain			
Lawrence, 1993	Before, during and after needle invasive procedures	N=38, neonates	$VAS_{2\ nurses}$ r= 0.42 to 0.91	Six paired t-tests comparing VAS of nurses, one significantly different, two reached significance.
	Chronic pain			
Varni, 1987	Chronic pain (JRA)	N=25, 4 to 16 years	$VAS_{physician}$ with VAS_{parent} r=0.85	Parent: 2.9 (2.8) Physician: 2.4 (2.6)
Huijer Abu-Saad, 1995	Chronic pain (JRA)	N=33, 7 to 16 years	$VAS_{physician}$ with VAS_{parent} r*=0.10	Parent: 1.5 (range 0 to 10) Physician: 2.8 (range 0 to 9.5)

r= Pearson product moment correlation coefficient; JRA=juvenile rheumatoid arthritis r*= Spearman rank order correlation coefficient

3.3 Results

Reliability

Results on interobserver reliability of the VAS_{obs} are shown in Table 1.

Correlation coefficients are moderate to high (median 0.55), except that for the parent-physician correlation in the chronic pain situation (Huijer Abu-Saad and Uiterwijk, 1995). The latter might be explained by the fact that the physician scored present pain levels after a potentially painful physical examination of the patients.

The interobserver correlation coefficients between professionals (nurses, physicians or researchers) ranged from 0.42 to 0.91, with a median of 0.75; those between professionals and parent from 0.36 to 0.85, with a median of 0.52. Only one study included neonates in an acute pain situation (Lawrence et al., 1993).

Validity

Studies estimating criterion validity compared VAS_{obs} with the VAS_{sr} of children, because self-report is generally considered the 'gold standard' of pain. Table 2 shows comparisons of VAS_{sr} scores reported by children and VAS_{obs} ratings by professionals or parents for postoperative or chronic pain. The correlation coefficients of VAS_{sr} with the VAS_{obs} of professionals range from 0.23 to 0.85 (median 0.53); those of VAS_{sr} with the VAS_{obs} parents from 0.46 to 0.83 (median 0.70). As far as mean VAS levels were reported, the self-report levels are higher for the postoperative pain and lower for chronic pain compared to the levels reported by parents or caregivers.

The VAS_{obs} has been used to estimate the concurrent validity of newly developed pain instruments. Table 3 gives an overview of the relevant studies which indirectly, also give an indication of the validity of the VAS itself. In all studies the VAS_{obs} was applied by another professional than the one who used the examined pain instrument, except for the study on the COMFORT (Dijk van et al., 2000).

The correlation coefficients between VAS_{obs} and the other pain instruments ranged from 0.42 to 0.86 (median 0.68).

Table 2 Studies reporting on criterion-related validity of VAS_{obs} compared to $VAS_{self\text{-}report}$

First author, year	Pain situation	Sample size and age range	Correlational results $VAS_{self\text{-}report}$ with:	Means and SD
	Surgery			
O'Hara, 1987	After major orthopaedic surgery	N=21, 7-17 years	VAS_{nurse} r=0.52; VAS_{parent} r=0.70	-
Hendrickson, 1990	After major surgery	N=46, 1 to 16 years, n= 30 gave self-report	VAS_{nurse} r=0.85; VAS_{parent} r=0.61	-
LaMontagne, 1991	After surgery	N=13, children 8 to 18 years	VAS_{nurse} r=0.61; $VAS_{physician}$ r=0.59	Child 4.1(2.3); Nurses 2.8 (2.0); Physicians 2.1 (2.0)
Miller 1996	After surgery	N=20, 7 to 11 years	VAS_{nurse} r=0.23, 0.50 and 0.54; VAS_{mother} r=0.46, 0.71 and 0.83	-
	Chronic pain			-
Varni, 1987	Chronic pain (JRA)	N=25, 5 to 16 years	$VAS_{physician}$ r=0.65; VAS_{parent} r=0.72	Present pain: Child 1.6 (2.1); Parent 2.9 (2.8)
Huijer Abu Saad, 1995	Chronic pain (JRA)	N=33, 7 to 16 years	$VAS_{physician}$ r*=0.32; VAS_{parent} r*=0.53 and 0.77	Present pain: Parent 1.5 (range 0 to 10) Physician 2.8 (range 0 to 9.5); Child 1.1 (range 0 to 5.5)

r= Pearson product moment correlation coefficient; r*= Spearman rank order correlation coefficient;JRA=juvenile rheumatoid arthritis

Table 3 Studies reporting on concurrent validity of VAS_{obs} and other pain instruments

First author, year	Pain situation	Sample size and age range	Correlation with pain instrument
McGrath, 1985	First hour after circumcision	N=30, 1 to 7 years	VAS_{nurses} with $CHEOPS_{research\ ass}$ r = 0.86
Tarbell, 1992	After minor surgery	N=74, 12 to 64 months	$TPPPS_{observer}$ with VAS_{nurses} r = 0.42 and 0.55
Lawrence, 1993	Needle invasive procedures	N=38, neonates	VAS_{nurses} with $NIPS_{research\ ass}$ r = 0.53 to 0.84
Taddio, 1995	Directly after immunisation	N=96, 4 to 6 months infants	$VAS_{trained\ observer}$ with videotaped MBPS score r = 0.68 $VAS_{pediatrician}$ with MBPS r = 0.74
Van Dijk, 2000	After major surgery	N=158, 0 to 3 year	Latent variables VAS with COMFORT 'behaviour' correlated 0.89 to 0.96 in LISREL analysis

r= Pearson product moment correlation coefficient

Abbreviations: CHEOPS, the Children's Hospital of Ontario Pain Scale; TPPPS, Toddler Preschooler Postoperative Pain Scale, NIPS, Neonatal Infant Pain Scale

MBPS, Modified Behavioural Pain Scale

Cutoff points

Empirical studies on optimal cutoff points have been performed in adults only. Collins and colleagues (1997) asked 1080 adult patients to rate their initial postoperative pain using the VAS_{sr} and a 4-point verbal rating scale, and concluded that a VAS_{sr} score >3.0 cm reflects at least moderate pain and a VAS_{sr} score >5.4 cm severe pain. Serlin and colleagues (1995) classified VAS_{sr} chronic pain levels based on interference with daily functions as described by the patients as follows: 1-4 reflects mild pain, 5-6 reflects moderate pain and 7-10 reflects severe pain.

In the pediatric literature, various cutoff points are presented that were determined on an a priori basis without further reference to empirical data. Berde et al. (1991) categorised a VAS_{sr} <3 as mild, 3 to <6 as moderate, and 6 or higher as severe. Bray and others (1996) distinguished between a score of 0 to 4.9 cm as 'acceptable' and a score of 5 to 10 cm as 'unacceptable' in a pediatric sample after major surgery. Others distinguished pain from no pain at a VAS score of 4 (Buchholz et al., 1998), or proposed extra pain medication after surgery when the score was 4 or higher (Dijk van et al., 2000).

3.4 Discussion

The findings with regard to interobserver reliability were in general good. However, they were obtained by computing Pearson product moment correlation coefficients between different observers' scores, which method is of limited value to establish interobserver agreement (Bland and Altman, 1986). Correlation coefficients only reflect relative positions of scores, but do not reflect differences in absolute levels of scores. While the correlation may be high, scores of different observers may show large differences at the same time. Thus, although high interobserver correlations were found in most studies, we do not know to what extent pain levels scored by e.g. nurses and parents do differ systematically. However, the available data are promising. Four studies presented mean scores by observers (Table 1), but did not give evidence on systematic differences among professionals and between professionals and parents. The differences between the mean scores are relatively small in comparison with the standard deviation.

Test-retest results reflecting intraobserver reliability are not available for the VAS_{obs}. Therefore we have no estimate of the measurement error due to intraobserver variability.

Observer consistency might be established, for instance, by showing videorecordings of pain behaviour of children to observers and have them score the children at different time intervals and at different times of the day. The variability of their scores could give an indication of intraobserver reliability.

The results on the criterion-related validity reported in this review were mixed. Generally, correlation coefficients between child self-report and proxy reports were high. While some authors found that parents' ratings were more strongly related to children's self-reports than those of nurses' (Miller, 1996; O'Hara et al., 1987), others found the opposite (Hendrickson et al., 1990). Furthermore, we noticed a considerable difference between results obtained in the postoperative pain situation and those in the chronic pain situation. The few available studies suggest that in the postoperative situation, nurses and parents underestimate the pain as compared to the children's reports. On the other hand, parents and physicians tend to 'overreport' in the chronic pain situation. The reported comparisons of VAS_{obs} with children's self-report are restricted to older children. Therefore, they provide no clear evidence on who would be the best observer in case children's self-reports cannot be obtained.

Concurrent validity of the VAS_{obs} with different pain instruments proved to be good, especially considering potential measurement error due to different instruments and different raters.

None of the reviewed studies reported on sensitivity to change for the VAS_{obs}. Sensitivity to change might be assessed by scoring a VAS before and after analgesic treatment by a observer blinded to the treatment, or by scoring before and after a short painful procedure. For the latter situation, videotaped material could make blinding possible. However, sensitivity to change in postoperative or chronic pain will become more difficult to assess in the future because pain treatment is now directed more and more towards pain prevention.

Cutoff points for the VAS_{obs} have not been established scientifically, they are rather based on experience or intuition. Research-based cutoff points should be established for chronic pain and postoperative pain separately. Because we lack a 'gold standard' for pain in preverbal children, we will have to rely on inferential methods to assess the validity of a chosen cutpoint. For instance, using judgements from expert panels to determine cutoff

points scores and further test them empirically by comparing these expert judgements with cutoff scores of professionals and parents.
In summary, available psychometric findings concerning the VAS_{obs} are promising. Further work needs to be done on intraobserver reliability, sensitivity to change, and optimal cutoff points. However, several considerations can be put forward to keep using the VAS_{obs} despite lack of conclusive evidence. First, special attention should be focused on hospitalised (premature) neonates and mentally handicapped children. Given their expressive limitations, the VAS_{obs} may be especially useful for them. Secondly, sound use of the VAS may be enhanced by combining it with age appropriate validated pain instruments, such as the PIPP for premature infants (Stevens et al., 1996), the NIPS for neonates (Lawrence et al., 1993), the CHEOPS (McGrath et al., 1985), the FLACC(Merkel et al., 1997), TPPPS (Tarbell et al., 1992), and COMFORT (Dijk van et al., 2000) for toddlers or young children. In that way, the VAS_{obs} ratings give additional information next to the behavioural (and physiological) information from the standardized pain instruments. For instance, while most pain instruments are based on detailed behavioural observations, the global rating on the VAS_{obs} may account for additional knowledge on individual variations in pain sensitivity, idiosyncratic behaviours, and situational influences. For example, high behavioural pain ratings may in some children express fear, anxiety, or other forms of distress, which can be distinguished with the VAS_{obs}.

Thirdly, using it after an observation period in which separate behavioural and physiological items are scored on a validated instrument will enhance the observational basis of the VAS_{obs} rating. A final consideration is to only use it after substantial experience with pain and pain assessment in children of different ages has been gained, and adequate interobserver reliability has been proven.

In conclusion, we argue that the VAS_{obs} is a helpful tool in observational pain assessment provided the aforementioned considerations are taken into account. Data collected during its continuous use, evidenced by the numerous posters presented at the 5th International Symposium on Paediatric Pain (London, 2000), may provide a sound basis for the much needed psychometric research that was suggested in this review.

3.5 Acknowledgements

This study was supported by a research grant from NWO (Dutch Organisation for Scientific Research, grant nr. 940-31-031). We thank Ko Hagoort for his help in preparing the manuscript.

3.6 References

Abu-Saad H. Assessing children's responses to pain. Pain 1984;19:163-171.

Aitken RC. Measurement of feelings using visual analogue scales. Proc R Soc Med 1969;62:989-993.

Arts SE, Abu-Saad HH, Champion GD, Crawford MR, Fisher RJ, Juniper KH and Ziegler JB. Age-related response to lidocaine-prilocaine (EMLA) emulsion and effect of music distraction on the pain of intravenous cannulation. Pediatrics 1994;93:797-801.

Benini F, Griffith P, Lago P and Gobber D. Evaluating pain in children: experience with a "do-it-yourself" visual analogue scale. Acta Paediatr 1996;85:762.

Berde CB, Lehn BM, Yee JD, Sethna NF and Russo D. Patient-controlled analgesia in children and adolescents: a randomized, prospective comparison with intramuscular administration of morphine for postoperative analgesia. J Pediatr 1991;118:460-466.

Bland JM and Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;:307-310.

Bray RJ, Woodhams AM, Vallis CJ, Kelly PJ and Ward-Platt MP. A double-blind comparison of morphine infusion and patient controlled analgesia in children. Paediatr Anaesth 1996;6:121-127.

Buchholz M, Karl HW, Pomietto M and Lynn AM. Pain scores in infants: a modified infant pain scale versus visual analogue. J Pain Symptom Manage 1998;15:117-124.

Collins SL, Moore RA and McQuay HJ. The visual analogue pain intensity scale: what is moderate pain in millimetres? Pain 1997;72:95-97.

Dijk van M, Boer de JB, Koot HM, Passchier J, Tibboel D and Duivenvoorden HJ. The reliability, stability and validity of the COMFORT scale as a postoperative pain instrument in 0 to 3-year-old infants. Pain 2000;84:367-377.

Gift AG. Visual analogue scales: Measurement of subjective phenomena. Nurs Res 1989;38:286-288.

Hendrickson M, Myre L, Johnson DG, Matlak ME, Black RE and Sullivan JJ. Postoperative analgesia in children: A prospective study of intermittent intramuscular injection versus continuous intravenous infusion of morphine. J Pediatr Surg 1990;25:185-191.

Ho K, Spence J and Murphy MF. Review of pain-measurement tools. Ann Emerg Med 1996;27:427-432.

Huijer Abu-Saad H, Bours GJJW, Bonnie S and Hamers JPH. Assessment of pain in the neonate. Semin Perinatol 1998;22:402-416.

Huijer Abu-Saad HH and Uiterwijk M. Pain in children with juvenile rheumatoid arthritis: a descriptive study. Pediatr Res 1995;38:194-197.

Huskisson EC. Measurement of pain. Lancet 1974;2:1127-1131.

LaMontagne LL, Johnson BD and Hepworth JT. Children's ratings of postoperative pain compared to ratings by nurses and physicians. Issues Compr Pediatr Nurs 1991;14:241-247.

Lawrence J, Alcock D, McGrath PJ, Kay J, Brock MacMurray S and Dulberg C. The development of a tool to assess neonatal pain. Neonatal Netw 1993;12:59-66.

Maunuksela EL, Olkkola KT and Korpela R. Measurement of pain in children with self-reporting and behavioral assessment. Clin Pharmacol Ther 1987;42:137-141.

McCormack HM, Horne DJ and Sheather S. Clinical applications of visual analogue scales: a critical review. Psychol Med 1988;18:1007-1019.

McGrath PA, Seifert CE, Speechley KN, Booth JC and et al. A new analogue scale for assessing children's pain: An initial validation study. Pain 1996;64:435-443.

McGrath PJ, Johnson G, Goodman JT, Schillinger J, Dunn J and Chapman JA. CHEOPS: A behavioral scale for rating postoperative pain in children. In: HL Fields, R Dubner and F Cervero (Eds.). Advances in Pain Research and Therapy, Vol. 9, Raven Press, New York, 1985. pp. 395-402.

McGrath PJ and Unruh AM. Measurement and assessment of paediatric pain. In: PD Wall and R Melzack (Eds.). Textbook of pain., Churchill Livingstone, Edinburgh, 1994. pp. 303-313.

Merkel SI, Voepel-Lewis T, Shayevitz JR and Malviya S. The FLACC: A behavioral scale for scoring postoperative pain in young children. Pediatr Nurs 1997;23:293-297.

Miller D. Comparisons of pain ratings from postoperative children, their mothers, and their nurses. Pediatr Nurs 1996;22:145-149.

Miller MD and Ferris DG. Measurement of subjective phenomena in primary care research: the Visual Analogue Scale. Fam Pract Res J 1993;13:15-24.

O'Hara M, McGrath PJ, D'Astous J and Vair CA. Oral morphine versus injected meperidine (demerol) for pain relief in children after orthopedic surgery. J Pediatr Orthop 1987;7:78-82.

Romsing J, Moller-Sonnergaard J, Hertel S and Rasmussen M. Postoperative pain in children: Comparison between ratings of children and nurses. J Pain Symptom Manage 1996;11:42-46.

Scott J and Huskisson EC. Graphic representation of pain. Pain 1976;2:175-184.

Serlin RC, Mendoza TR, Nakamura Y, Edwards KR and Cleeland CS. When is cancer pain mild, moderate or severe? Grading pain severity by its interference with function. Pain 1995;61:277-284.

Stevens BJ, Johnston CC, Petryshen P and Taddio A. Premature Infant Pain Profile: development and initial validation. Clin J Pain 1996;12:13-22.

Taddio A, Nulman I, Koren BS, Stevens BJ and Koren G. A revised measure of acute pain in infants. J Pain Symptom Manage 1995;10:456-463.

Tarbell SE, Cohen TI and Marsh JL. The Toddler-preschooler postoperative pain scale: an observational scale for measuring postoperative pain in children aged 1-5. Preliminary report. Pain 1992;50:273-280.

Varni JW, Thompson KL and Hanson V. The Varni/Thompson Pediatric Pain Questionnaire. I. Chronic musculoskeletal pain in juvenile rheumatoid arthritis. Pain 1987;28:27-38.

Wewers ME and Lowe NK. A critical review of visual analogue scales in the measurement of clinical phenomena. Res Nurs Health 1990;13:227-236.

Williams AC, Davies HT and Chadury Y. Simple pain rating scales hide complex idiosyncratic meanings. Pain 2000;85:457-463.

Chapter 4

The association between physiological and behavioural pain measures in 0 to 3-year-old infants after major surgery

Based on the article:

The association between physiological and behavioural pain measures in 0 to 3-year-old infants after major surgery

Authors: Monique van Dijk, Josien B. de Boer, Hans M. Koot, Hugo J. Duivenvoorden, Jan Passchier, Nancy Bouwmeester, Dick Tibboel

Journal of Pain and Symptom Management; in press

4.1 Abstract

Purpose of the study: To estimate the association between behavioural and physiological pain measures and to identify determinants predicting the level of association.
Measures and design: The COMFORT 'behaviour' scale as well as heart rate (HR), mean arterial pressure (MAP), and the variability of HR and MAP (HRV and MAPV) were assessed every three hours after major abdominal or thoracic surgery. Subjects were 204 infants aged 0-3 years.
Results: the within-subject correlations, using the repeated measures, were 0.37, 0.44, 0.48 and 0.49 for COMFORT 'behaviour' with HRV, HR, MAP and MAPV, respectively. Neonates had lower behaviour-physiology correlations than the older infants, due to low pain scores. Pain characteristics significantly predicted the COMFORT 'behaviour'-HR/MAP correlations, suggesting that the behaviour-physiology correlations increase with increasing pain. The behaviour-physiology correlations were not greatly affected by physical condition.
Conclusion: There were large interindividual differences in behaviour-physiology correlations after major surgery in 0 to 3-year-old infants who should be further explored in future research.

4.2 Introduction

In neonates and preverbal children, behaviour is used as a substitute for self-report of pain (Anand and Craig, 1996; McGrath, 1998). Moreover, a multidimensional approach is preferred, combining behavioural and physiological measures of pain (Franck and Miaskowski, 1997; Huijer Abu-Saad et al., 1998; Stevens, 1998). The three most frequently used behavioural measures are body movement, cry and facial expression (Franck and Miaskowski, 1997; McGrath, 1998) which are implemented in the majority of pain instruments (McGrath, 1998). Multidimensional instruments, for example the COMFORT scale (Ambuel et al., 1992; Dijk van et al., 2000), the Modified Infant Pain Scale (Buchholz et al., 1998), CRIES (acronym for Crying, Requires increased oxygen administration, Expression and Sleeplessness;)(Krechel and Bildner, 1995) and the Premature Infant Pain Profile (Stevens et al., 1996), also incorporate physiological measures of pain.

Physiological measures of pain, particularly heart rate (HR), respiratory rate (RR) and oxygen saturation (SaO_2) and to a lesser extent blood pressure (BP) have been extensively examined in neonates during acute pain caused by heel lance (Johnston et al., 1995; McIntosh et al., 1993) and circumcision (Arnett et al., 1990; Benini et al., 1993; Howard et al., 1994; Weatherstone et al., 1993; Williamson and Williamson, 1983). When no analgesia was given, HR, RR and BP increased and SaO_2 decreased during heel lance or circumcision, suggesting that these changes reflect acute pain (Benini et al., 1993; Owens and Todt, 1984; Williamson and Williamson, 1983).

The variability of physiological indicators, especially heart rate variability (HRV), in relation to pain has been examined in two studies. In premature neonates (26 to 34 weeks' gestation) HRV was significantly higher for real heel prick than for sham heel prick (McIntosh et al., 1993). In contrast, Lindh (1999) found decreased HRV during the most stressful part (squeezing) of the heel lance procedure in healthy neonates. Both behavioural and physiological pain measures lack sensitivity and specificity (Franck and Miaskowski, 1997; Stevens, 1998). The behavioural pain response resembles behaviour caused by other sources of distress like anxiety or anger. Moreover, the absence of pain-related behaviour (e.g. not crying or no body movement) does not necessarily imply the absence of pain (Berde, 1989; Beyer et al., 1990; Morton, 1997; Terndrup, 1996). Also, behaviour can be influenced by physical condition. For instance, the severity of illness was found to affect the acoustic cry variables in premature neonates during heel prick (Stevens et al., 1994). Barr (1998) suggested that ill children on a NICU may lack the strength to cry.

Physiological responses to pain may lack sensitivity because they are also related to other factors like infection, anaemia, trauma, use of opioids (Eland, 1990; Tyler et al., 1993) or influenced by medical interventions. Age affects both behavioural and physiological pain responses. Cognitive development matures with increasing age, reflected by more differentiated behaviour with increasing age (Davis, 1990; McGrath and Unruh, 1994). Age-related physiological changes are for instance the decrease in RR and HR and increasing BP with increasing age. Because both behaviour and physiology lack sufficient specificity and sensitivity, they are preferably combined in pain measures to increase validity (Burrows and Berde, 1993; Franck and Miaskowski, 1997; Huijer Abu-Saad et al., 1998; Porter, 1993; Stevens, 1998). However, the question remains as to how strongly related behavioural and physiological pain measures are, and whether they manifest themselves at the same time.

A limited number of studies on the association between physiological and behavioural pain measures during acute pain show that correlations between these measures are often significant though moderate. Johnston and colleagues (1995) found correlations of 0.12 to 0.55 between facial expressions and HR and HRV, respectively, during heel prick in 48 premature neonates. Within-subject correlations, ranging from 0.48 to 0.84 were found between HR and crying during heel lance (Owens and Todt, 1984). The correlation between a composite behavioural response and salivary cortisol level during inoculation at 2, 4 and 6 months of age was 0.29, 0.29 and 0.21, respectively, in a sample of 55 infants (Lewis and Ramsay, 1995).

In contrast to acute sharp pain, the association between physiological and behavioural measures during non-acute pain, e.g. postoperative pain is largely unknown. The importance of acquiring more insight into the behaviour-physiology association during pain was recently emphasised by Barr (1998), who addressed the problem of contradictory information from behavioural and physiological pain responses for clinical decision-making. More knowledge about determinants that influence the behaviour-physiological association could give directives for clinical practice whether both or any of the two measures are valuable for pain assessment.

To address this issue, the aims of the present study were to estimate the correlations between behavioural and physiological pain measures after major surgery, and to identify determinants of this association. The following research questions will be addressed:

1. To which degree are behavioural and physiological pain measures associated after major surgery in 0 to 3-year-old infants?
2. Which determinants are predictive of the level of association between the behavioural and the physiological pain measures?

4.3 Methods

Patients

A total of 204 children aged 0 to 3 years, who were admitted for major abdominal or thoracic surgery entered the study. Excluded were children with renal or hepatic dysfunction, neurological damage, altered muscle tone, preoperative anaemia (haematocrit less than 30%) and children who had received neuromuscular blockers. The setting of the

study was the Pediatric Surgical Intensive Care Unit (PSICU) of the University Sophia Children's Hospital Rotterdam. The Medical Ethical Committee of the Hospital approved the study. Written informed consent was obtained from the parents by the pediatric intensivist or anaesthesiologist.

Design

This study is part of a double-blind randomised clinical trial aimed at examining the best way to assess postoperative pain and to compare the efficacy and safety of administering morphine continuously or intermittently after major abdominal or thoracic surgery in 0 to 3-year-old infants. Stratification by age was performed because behavioural and physiological differences between age groups were considered important. Age groups comprised neonates (>35 weeks gestation and weight >1500 grams), young infants (1 to 6 months), infants (7 to 12 months) and toddlers (1 to 3 years). Pain assessment was performed at baseline, after return to the PSICU, and every three hours during the first 24 hours postoperative.

Instruments

The COMFORT scale (Ambuel et al., 1992) was originally developed to assess distress in ventilated children of all ages in an intensive care environment. It comprises six behavioural items and two physiological items. The behavioural part of the COMFORT scale (the COMFORT 'behaviour') consists of the summation of six behavioural items; Alertness, Calmness, Muscle tone, Movement, Facial tension, and Respiratory response (for ventilated children) or Crying (for non-ventilated children) with response categories ranging from 1 (low distress/no pain) to 5 (high distress/pain). The internal consistency, stability and validity of the COMFORT scale items were analysed by structural equation modelling with the statistical package of LISREL. The COMFORT 'behaviour' score may range from 6 to 30 and proved a reliable and valid instrument to assess postoperative pain in neonates and infants (Dijk van et al., 2000). The remaining two items include Heart Rate and Mean Arterial Pressure and require an indwelling arterial line. During the two-minute interval period needed to assess the COMFORT scale, six HR and six MAP values are registered from the monitor and compared with the baseline range of the child, assessed prior to surgery. Although HR and MAP were originally included in the COMFORT total score (Ambuel et al., 1992) they appear to constitute a component separate from the behavioural part of the COMFORT scale.

In addition to the COMFORT scale, the nurses used a Visual Analogue Scale (VAS) for clinical rating of pain. The VAS is a horizontal continuous ten-centimetre line with the anchors 'no pain' on the left side and 'extreme pain' on the right side. The score ranges from 0 to 10 (McGrath et al., 1985; Varni et al., 1987).

Heart rate and Mean Arterial Pressure

Heart rate and Mean Arterial Pressure were read from the Hewlett Packard M2350a monitor, six times during the two-minute interval of each pain assessment. Means and standard deviations of the six HR and MAP values were estimated. This yielded four physiological pain measures: mean HR and MAP, and HR and MAP variability (HRV and MAPV, respectively).

Surgical Stress Score (SSS)

This score was derived from the SSS developed by Anand and Aynsley-Green (1988). This scoring system was originally developed to assess severity of surgical stress in neonates. In this study, we summed the following items from the scale: percentage of blood loss (score range 0 to 3), amount of blood loss (score range 0 to 3), site of surgery (score range 0 to 2), amount of superficial trauma (score range 1 to 3), extent of visceral trauma (score range 1 to 4), duration of surgery (score range 1 to 5). These items are clearly described in the scoring list; information needed for scoring of the SSS is obtained from the surgical and anaesthesia records, which are noted minute-by-minute during the surgical procedure. The SSS was scored jointly by the attending anaesthesiologist and surgeon immediately after surgery.

Systemic Inflammatory Response Syndrome (SIRS)/sepsis

A dichotomous variable (scored 0 or 1), indicating if SIRS/sepsis was present or not, based on the classification of Hayden (1994).

This score was determined by one of the co-authors (NB), who is an experienced anaesthesiologist, and confirmed by the consultant pediatric intensivist (DT).

Cardiorespiratory insufficiency

A dichotomous variable (0,1) indicating severe cardiorespiratory insufficiency due to e.g. congenital heart anomalies or congenital or acquired lung anomalies.

Training of nurses

For adequate use of the COMFORT scale thirty-nine nurses and two anaesthesiologists attended a two-hour training session during which the COMFORT scale was explained by means of videotaped behaviour and in vivo observations of children at the PSICU. Because two to five nurses were trained at the same time, discussion was helpful to solve possible misinterpretations. After the course, each newly trained nurse completed ten COMFORT assessments (scored at the PSICU on infants after surgery under three years of age) with one of the trainers or an experienced colleague. When interrater reliability was acceptable, according to a linearly weighted Cohen's Kappa between 0.40 and 0.60 (Fleiss, 1981), the nurse was allowed to score children for the study. Linearly weighted Cohen's Kappa of the individual items ranged from 0.54 to 0.93, with a median Kappa of 0.70 (Dijk van et al., 2000). Trained nurses rated the infants included in the study that were on the ward during their 8 hours shift. Thus, infants were rated by at least three different nurses during a 24h-period.

Procedure

After anaesthetic induction, an arterial line was placed to enable non-invasive blood sampling and MAP monitoring. After surgery, the children returned to the PSICU and pain assessment started after installation of the child. Pain assessment comprised a two-minute interval with notation of HR and MAP every 20 seconds (totalling six times) from the monitor.

Pain assessment was performed prior to handling of the child and morphine or placebo bolus administration. When children were considered to be in pain (VAS $\geq$4) at any time after surgery, additional morphine could be given according to the protocol.

Statistical analysis

Only cases with at least six out of nine postoperative measurements were included in the analysis.

To estimate the association between the repeated COMFORT 'behaviour' scale scores and each of the series of physiological scores, a within-subject correlation, r_{within} was estimated using dummy variables to represent subjects in a multiple regression approach (Bland and Altman, 1995).

To identify determinants of these correlations, the following procedure was used: For each subject, r_{within} between the COMFORT 'behaviour' scale scores and each of the

physiological indicators was entered in SPSS 8.0 and z-transformed. Z-transformation was employed to normalise the distribution of correlation coefficients (Cohen and Cohen, 1983). The transformed score was entered as an outcome variable in multiple regression analysis. The following potentially relevant determinants, age, gender, surgical stress, SIRS/sepsis, cardiorespiratory insufficiency, morphine condition, morphine dosage, and the average COMFORT 'behaviour' score across the first 24 hours were simultaneously entered into the analysis. The pre-stratified age groups were entered as dummy variables, because age distribution was not linearly related to the outcome variables. Background characteristics were compared among age groups by one-way ANOVA (with Bonferroni correction), χ^2 test, and Fisher exact test or Kruskal-Wallis test, as appropriate. The significance of the difference between dependent correlation coefficients was calculated by a formula described by Cohen (1983, page 56-57).

Table 1 Background characteristics of the age groups

	Neonates (n=66)	Young infants (n=67)	Infants (n=31)	Toddlers (n=40)	p
Baseline characteristics					
Gender (%male)	59%	64%	58%	47.5%	0.41
Mean Baseline HR (SD)	136 (16)	139 (17)	133 (15)	121 (24)	0.001
Mean Baseline MAP (SD)	52 (11)	64 (12)	75 (13)	77 (13)	0.001
Perioperative characteristics					
Mean Surgical stress score (SD)	9 (2)	9 (3)	9 (3)	10 (3)	0.29
SIRS /sepsis (%)	17%	4%	3%	7.5%	0.07
Postoperative characteristics					
Ventilatory support (%)	77	33	13	25	0.001

SD= standard deviation, HR= Heart Rate, MAP= Mean Arterial Pressure, SIRS= Systemic inflammatory response syndrome

4.4 Results

Patients

Table 1 shows the background characteristics of the 204 children randomised onto this trial. Baseline HR was significantly different across age groups, with toddlers having a lower HR than the other age groups. Baseline MAP increased with age. The surgical stress score was not different between age groups. SIRS/sepsis occurred most frequently in neonates, although not significantly more often than in other age groups. Neonates required ventilatory support after surgery more often than children in the other age groups.

Table 2 Pain assessments of the age groups

Pain assessments[1] first 24 hrs after surgery	**Neonates (n=66)**	**Young infants (n=67)**	**Infants (n=31)**	**Toddlers (n=40)**	**p**
VAS pain	1.3 (0.7)	2.7 (1.1)	2.4 (1.1)	2.1 (1.3)	0.000
COMFORT 'behaviour'	12 (2)	14 (6)	14 (5)	13 (4)	0.13
HR	131 (36)	126 (52)	121 (62)	130 (33)	0.76
MAP	45 (19)	58 (28)	60 (34)	71 (26)	0.000
HRV	4 (2)	7 (3)	7 (3)	7 (3)	0.000
MAPV	2 (1)	4 (1)	3 (1)	3 (1)	0.000
Morphine dosage µg/kg/h					
Median	10.0	11.9	11.9	10.9	0.000
IQR	10.0 to 10.7	10.4 to 16.4	10.3 to 14.9	10 to 14.3	

[1] Values in mean (sd) unless otherwise stated

HR= Heart Rate, MAP= Mean Arterial Pressure, HRV= Heart Rate variability, MAPV= Mean Arterial Pressure variability, IQR= Interquartile Range

Table 2 shows the pain characteristics for the four age groups. Neonates had significantly lower VAS pain scores than young infants (p=0.000), infants (p=0.000) and toddlers (p=0.005) after surgery. Toddlers had significantly lower VAS pain scores than young infants (p=0.013). MAP was significantly lower in neonates than in young infants (p=0.03) and toddlers (p=0.000). HRV and MAPV were significantly lower in neonates compared to all other age groups (all p- values <0.000). The significant differences in morphine dosage

(Kruskal-Wallis χ^2 =42.4, p<0.000) were due to the relatively low morphine dosage given to neonates compared to other age groups and the higher morphine dosages given to young infants, who required more morphine than neonates and toddlers.

Missing data

Thirty cases had >3 missing values on MAP due to: a failing missing line/missing notation on record (n=20), use of muscle relaxants or sedatives (n=7), deceased 3 hrs after surgery (n=1), ventilatory depression (n=1). The excluded patients were not significantly different from the included group with respect to gender, age, surgical stress, SIRS/sepsis, cardiorespiratory insufficiency, morphine condition, and average morphine dosage. Twenty-one of these thirty cases had >3 missing values on HR. As a result, the sample size for HR was 183, and for MAP 174.

Table 3 Within-subject correlations (r_{within}) of COMFORT 'behaviour' scores with physiological indicators of pain

	r_{within}	95% CI	% of variance explained	N
HR	0.44	0.31 to 0.55	19.3	183
MAP	0.48	0.35 to 0.58	22.7	174
HRV	0.37	0.24 to 0.49	13.8	183
MAPV	0.49	0.36 to 0.59	23.5	174

HR= Heart Rate, MAP= Mean Arterial Pressure, HRV= Heart Rate variability,
MAPV= Mean Arterial Pressure variability, CI=Confidence Interval, N= Number of patients

Within-subject correlations

To estimate the average within-subject correlation, the multiple regression analyses with COMFORT 'behaviour' as outcome variable, N-1 subjects as dummy variables and HR, MAP, HRV and MAPV, respectively as predictor variables resulted in the average within-subject correlations, r_{within}, and percentage of explained variance, presented in Table 3. All r_{within} were significant (p<0.001, two-tailed). On average, an increase in COMFORT 'behaviour' scores was associated with an increase in the physiological indicators. The correlation was highest for MAPV with COMFORT 'behaviour' ($r_{MAPV\ and\ COMFORT}$ =

0.49), and lowest for HRV with COMFORT 'behaviour' ($r_{HRV\ and\ COMFORT} = 0.37$). The correlation coefficients from table 3 were not significantly different from each other. Because all r_{within} for each case were entered, the median and range of the between-subject variability in r_{within} could be calculated. The median value and range of the individual r_{within} were for $r_{HR\ and\ COMFORT}$ 0.48 (range: –0.74 to 0.98), for $r_{MAP\ and\ COMFORT}$ 0.48 (range: –0.78 to 0.98), for $r_{HRV\ and\ COMFORT}$ 0.41 (range:-0. 81 to 0.95) and for $r_{MAPV\ and\ COMFORT}$ 0.48 (range: –0.63 to 0.98). With the exception of $r_{HRV\ and\ COMFORT}$, all median values of the r_{within} were equal.

The r_{within} of the different physiological indicators are given in Table 4. Correlations ranged from 0.04 to 0.45 indicating low to moderate correlations between physiological measures.

Table 4 Within subject correlations (r_{within}) of physiological indicators of pain

	HR	MAP	HRV
HR	-		
MAP	0.23 (0.08 to 0.36)	-	
HRV	0.04 (-0.10 to 0.19)	0.25 (0.11 to 0.38)	-
MAPV	0.20 (0.05 to 0.34	0.35 (0.22 to 0.48)	0.45 (0.33 to 0.56)

95% confidence intervals are given in brackets

Bivariate regression coefficients

The standardized bivariate regression coefficients of the z-transformed correlations (within-subject correlations of COMFORT 'behaviour' with each of four physiological indicators) with the potential determinants are depicted in Table 5.

Age group contributed significantly to the $r_{MAP\ and\ COMFORT}$ and $r_{HR\ and\ COMFORT}$, indicating larger correlations with increasing age group.

The $r_{MAP\ and\ COMFORT}$ was negatively and significantly correlated with the presence of SIRS/sepsis.

The average dosage of morphine (per kg/h) during the first 24 hours was positively and significantly associated with the $r_{MAP\ and\ COMFORT}$ and $r_{MAPV\ and\ COMFORT}$.

The average COMFORT 'behaviour' score of the first 24 hours after surgery was positively and significantly associated with the correlations between COMFORT 'behaviour' and HR, MAP and MAPV.

Multiple regression analyses

Table 6 shows the standardised coefficients obtained from multiple regression analyses, in which all potential determinants were entered simultaneously. Several age group differences were found, as well as relations with sepsis, and level of behavioural pain scores.

The $r_{HR\ and\ COMFORT}$ was significantly higher for infants (7 to 12 months) compared to neonates. Further, the $r_{MAP\ and\ COMFORT}$ was higher for all age groups compared to neonates. Finally, the $r_{MAPV\ and\ COMFORT}$ was significantly higher for the young infants compared to the neonates. The presence of SIRS/sepsis decreased the $r_{MAP\ and\ COMFORT}$ significantly. A higher COMFORT 'behaviour' was associated with higher $r_{HR\ and\ COMFORT}$ and higher $r_{MAP\ and\ COMFORT}$.

The $r_{HRV\ and\ COMFORT}$ was not significantly influenced by any of the determinants.

None of the other potential determinants had a significant influence on behaviour-physiology correlations. Thus, behaviour-physiology correlations appeared to be independent of gender, as well as to the physical-condition related variables surgical stress and cardiorespiratory insufficiency and the pain-related variables morphine condition and morphine dosage.

Table 5 Bivariate regression coefficients[1)] between determinants and behaviour-physiology correlations

Determinants	$r_{HR \text{ and } COMFORT}$ [2)] (n=183)	$r_{MAP \text{ and } COMFORT}$ [2)] (n=174)	$r_{HRV \text{ and } COMFORT}$ [2)] (n=183)	$r_{MAPV \text{ and } COMFORT}$ [2)] (n=174)
Background characteristics				
Gender	0.03 (-0.12 to 0.17)	0.03 (-0.12 to 0.18)	-0.03 (-0.17 to 0.12)	-0.03 (-0.18 to 0.12)
Age group	0.19 (0.05 to 0.33)	0.28 (0.14 to 0.41)	0.08 (-0.06 to 0.22)	0.15 (0.001 to 0.29)
Physical condition				
Surgical stress	-0.06 (-0.20 to 0.08)	-0.02 (-0.17 to 0.13)	0.09 (-0.05 to 0.23)	-0.04 (-0.19 to 0.11)
SIRS/sepsis	-0.11 (-0.25 to 0.03)	-0.26 (-0.39 to –0.12)	-0.02 (-0.16 to 0.12)	-0.11 (-0.25 to 0.04)
Cardiorespiratory insufficiency	0.01 (–0.13 to 0.15)	0.04 (-0.11 to 0.19)	0.02 (-0.12 to 0.16)	0.14 (-0.01 to 0.28)
Pain-related characteristics				
Morphine condition (CM vs. IM)	0.07 (-0.07 to 0.21)	0.04 (-0.11 to 0.19)	0.06 (-0.08 to 0.20)	0.03 (-0.12 to 0.18)
Morphine dosage (per kg/h)	0.09 (–0.05 to 0.23)	0.18 (0.03 to 0.32)	0.06 (-0.08 to 0.20)	0.22 (0.07 to 0.36)
COMFORT 'behaviour' (average)	0.26 (0.12 to 0.39)	0.40 (0.27 to 0.52)	0.06 (-0.08 to 0.20)	0.27 (0.13 to 0.40)

[1)] Standardised bivariate regression coefficients were presented for comparability of the different determinants, statistically significant coefficients are underlined

[2)] The within-subject correlations are z-transformed

SIRS= systemic inflammatory response syndrome, CM= continuous morphine administration, IM= intermittent morphine administration, n= number of patients

95% confidence intervals are given in brackets

Table 6 Multiple regression coefficients[1] of determinants with behaviour-physiology correlations

	$r_{HR\ and\ COMFORT}$ [2]		$r_{MAP\ and\ COMFORT}$ [2]		$r_{HRV\ and\ COMFORT}$ [2]		$r_{MAPV\ and\ COMFORT}$ [2]	
Determinants	β	p	β	p	β	p	β	p
Background characteristics								
Gender	0.10	0.18	0.07	0.27	-0.05	0.51	-0.02	0.82
Young infants [3]	0.06	0.53	0.31	0.001	0.04	0.68	0.20	0.05
Infants [3]	0.17	0.05	0.18	0.03	-0.02	0.81	0.07	0.45
Toddlers [3]	0.05	0.52	0.23	0.004	0.10	0.27	0.08	0.34
Physical condition								
Surgical stress	-0.04	0.53	-0.04	0.92	0.08	0.27	-0.02	0.82
SIRS/sepsis	-0.06	0.45	-0.23	0.003	0.00	1.00	-0.11	0.15
Cardiorespiratory insufficiency	-0.02	0.76	0.01	0.88	0.01	0.90	0.10	0.19
Pain-related characteristics								
Condition (CM vs. IM)	0.04	0.55	0.03	0.61	0.05	0.54	0.03	0.67
Morphine dosage (per kg/h)	-0.05	0.56	0.01	0.96	0.05	0.57	0.13	0.12
COMFORT 'behaviour' (average)	0.24	0.01	0.22	0.01	0.03	0.75	0.09	0.32

[1] Standardised multivariate regression coefficient were presented for comparability; statistically significant coefficients are underlined

[2] The within-subject correlations are z-transformed

[3] Age groups entered as dummy variables, β's compared to neonates

SIRS= systemic inflammatory response syndrome, CM= continuous morphine administration, IM= intermittent morphine administration, n= number of patients

4.5 Discussion

This study is the first to address the association between behavioural and physiological pain measures during the postoperative course in a large sample of 0 to 3-year-old infants after major surgery. The correlations between COMFORT 'behaviour' and the physiological indicators HR, MAP, HRV and MAPV (0.44, 0.48, 0.37 and 0.49, respectively) were significant. In particular, the correlations between the behavioural measure, COMFORT 'behaviour' and the MAP and MAPV, respectively, were fairly high although not significantly higher than the other correlations. The within subject correlation between the four physiological indicators ranged from 0.04 for the within subject correlation of HR with HRV to 0.45 for HRV with MAPV. These low to moderate correlations between physiological measures show that each of these are relatively independent and may have a different meaning.

These findings fill part of the gap in our knowledge on relations between behaviour and physiology in the postoperative pain situation. Although the behaviour-physiology correlations in our study were significant, the percentage of explained variance (the squared correlations), ranging from 14 to 23%, revealed that the association was far from perfect. This can be explained by the fact that with regard to stress, physiological systems tend to be loosely, rather than closely related to behavioural responsive systems (Barr, 1998; Lacey, 1967). Moreover, in the ICU setting physiological measures are influenced by medical interventions, such as raising inspiratory O_2 fraction, fluid resuscitation and the use of inotropic drugs. In addition, these associations decrease due to the large individual differences in pain response as described in literature (Barr et al., 1994; Gunnar et al., 1995; Lewis, 1992) and confirmed by our data. In our study, the within-subject behaviour-physiology correlations were between 0.37 and 0.49 but the range of correlations was enormous (-0.81 to 0.98) reflecting large inter-individual differences. These inter-individual differences may be related to determinants addressed in the second part of this study.
The relevance of demographic characteristics, physical and pain-related characteristics on the behaviour-physiology correlations, was examined and proved of varying importance in the explanation of inter-individual differences found. Only behavioural pain intensity and morphine dosage were consistently related to these associations. First, the average COMFORT 'behaviour' score, representing pain intensity during the first 24 hours,

significantly increased the association between COMFORT 'behaviour' and HR and MAP, respectively, both in bivariate and multiple analyses, though MAPV only bivariately. Second, morphine dosage, including protocol dosage and extra morphine given 'on demand', was positively related to the association between MAP and MAPV, respectively, and COMFORT 'behaviour' in the bivariate analyses.
The contribution of both characteristics suggests that the behaviour-physiology correlations increase with increasing pain. Therefore, it is likely that behavioural and physiological indicators tend to be more interdependent in highly painful situations and may corroborate the diagnosis of pain. However, all children in our study received at least 10 μg/kg/h morphine during the first 36 hours after surgery, which explains the relatively low pain scores. This was particularly true for the neonates who had significantly lower pain scores than the other age groups. More neonates were satisfied with the morphine dosage of the trial (10 μg/kg/h) and had the lowest pain scores compared to the older age groups, which could be explained by the fact that neonates metabolise morphine slower than older infants (Lynn et al., 1998). The highest pain scores were seen in the young infants. The reasons for these relatively high pain scores in this age group are unknown. We speculate that developmental aspects may play a role. Young infants are less able to show differentiated expressions of pain versus e.g. anxiety, which may influence even standardised pain ratings. However, they are able to sense differences in environmental input (noise, NICU environment) which makes them more vulnerable to non-pain-related distress, whereas they are not able to cope with pain or distress through distraction or comfort-seeking, like toddlers do.

In general when pain scores were lower, behaviour-physiology correlations decreased. Physical condition, represented by surgical stress, SIRS/sepsis and cardiorespiratory insufficiency had little effect on the behaviour-physiology correlations. Only SIRS/sepsis had a decreasing effect on the correlation between MAP and COMFORT 'behaviour', partly explained by the changes in metabolism and hormone profiles during this metabolic stress response. We expected physical condition to blur the behaviour-physiology correlations. However, the results from this study suggest that the physical condition variables have a limited influence on the behaviour-physiology correlations.

4.6 Clinical implications

Our data show that heart rate and blood pressure can be useful for postoperative pain assessment in an ICU environment when monitoring of physiological measures is standard procedure and when pain is relatively high. Under certain conditions we have to rely solely on physiological measures, for example during surgery or when neuromuscular blockers are used during mechanical ventilation. After minor surgical procedures, without an arterial line, behavioural measures are preferable, because manually assessed heart rate and blood pressure measurement can be threatening for a frightened child and, consequently, may be less reliable as a measure of pain. When possible, the combination of physiological and behavioural measures is preferable, especially when behavioural pain seems high.
The large inter-individual variability in the relation between physiology and behaviour, and the limited association with demographic and illness characteristics, suggest that each individual has his/her own unique way of manifesting pain (Anand and Craig, 1996).

Future research
Because most studies have investigated acute painful procedures, future research should focus on the postoperative period. Research on the behavioural and physiological association with regard to pain could be extended to the association between behavioural and biochemical indicators (e.g. cortisol, adrenaline and noradrenaline).
The predictability of individual differences in pain expression could be an interesting area of research, as the influence of cognitive development, temperament or medical pain history on (postoperative) pain expression is still an unexplored area.

4.7 Acknowledgements

The authors would like to thank the staff and nurses of the Pediatric Surgical Intensive Care of the Sophia Children's Hospital for their efforts in performing the pain assessments for this study. We also thank the staff and members of the Operation Room and the department of Anaesthesiology, for their contribution to this study. In particular, we thank the parents and children who participated in this study.
This study was supported by a research grant from NWO (Dutch Organisation for Scientific Research, grant nr. 940-31-031)

4.8 References

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

Anand KJS and Aynsley-Green A. Measuring the severity of surgical stress in newborn infants. J Pediatr Surg 1988;23:297-305.

Anand KJS and Craig KD. New perspectives on the definition of pain. Pain 1996;67:3-6.

Arnett RM, Jones JS and Horger EO. Effectiveness of 1% lidocaine dorsal penile nerve block in infant circumcision. Am J Obstet Gynecol 1990;163:1074-1080.

Barr RG. Reflections on measuring pain in infants: dissocation in responsive systems and "honest signalling". Arch Dis Child Fetal Neonatal Ed 1998;79:F152-F156.

Barr RG, Boyce WT and Zeltzer LK. The stress-illness association in children: A perspective from the biobehavioural interface. In: RJ Haggerty (Ed.). Stress, risk and resilience in children and adolescents: process, mechanisms, and interventions, Cambridge University Press, Cambridge, 1994. pp. 183-224.

Benini F, Johnston CC, Faucher D and Aranda JV. Topical anesthesia during circumcision in newborn infants. JAMA 1993;270:850-853.

Berde CB. Pediatric postoperative pain management. Pediatr Clin North Am 1989;36:921-941.

Beyer JE, McGrath PJ and Berde CB. Discordance between self-report and behavioural pain measures in children aged 3-7 years after surgery. J Pain Symptom Manage 1990;5:350-356.

Bland JM and Altman DG. Calculating correlation coefficients with repeated observations: Part 1- - Correlation within subjects. BMJ 1995;310:446.

Buchholz M, Karl HW, Pomietto M and Lynn AM. Pain scores in infants: a modified infant pain scale versus visual analogue. J Pain Symptom Manage 1998;15:117-124.

Burrows FA and Berde CB. Optimal pain relief in infants and children. BMJ 1993;307:815-816.

Cohen J and Cohen P. Applied multiple regression/correlation analysis for the behavioural sciences, Lawrence Erlbaum, New York, 1983. pp. 490.

Davis KL. Postoperative pain in Toddlers: nurses' assessment and intervention. In: DC Tyler and EJ Krane (Eds.). Advances in pain research and therapy, Vol. 15, Raven Press, New York, 1990. pp. 53-61.

Dijk van M, Boer de JB, Koot HM, Passchier J, Tibboel D and Duivenvoorden HJ. The reliability, stability and validity of the COMFORT scale as a postoperative pain instrument in 0 to 3-year-old infants. Pain 2000;84:367-377.

Eland JM. Pain in children. Nurs Clin North Am 1990;25:871-883.

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

Franck LS and Miaskowski C. Measurement of neonatal responses to painful stimuli: a research review. J Pain Symptom Manage 1997;14:343-378.

Gunnar MR, Porter FL, Wolf CM, Rigatuso J and Larson MC. Neonatal stress reactivity: predictions to later emotional temperament. Child Dev 1995;66:1-13.

Hayden WR. Sepsis terminology in pediatrics. J Pediatr 1994;124:657-658.

Howard CR, Howard FM and Weitzman ML. Acetaminophen analgesia in neonatal circumcision: the effect on pain. Pediatrics 1994;94:641-646.

Huijer Abu-Saad H, Bours GJJW, Bonnie S and Hamers JPH. Assessment of pain in the neonate. Semin Perinatol 1998;22:402-416.

Johnston CC, Stevens BJ, Yang F and Horton L. Differential response to pain by very premature neonates. Pain 1995;61:471-479.

Krechel SW and Bildner J. CRIES: a new neonatal postoperative pain measurement score. Initial testing of validity and reliability. Paediatr Anaesth 1995;5:53-61.

Lacey JI. Somatic response patterning and stress: some revisions of activation theory. In: MH Appley and R Trumbull (Eds.). Psychological stress: issues in research, Appleton-Century-Crofts, New York, 1967. pp. 14-42.

Lewis M. Individual differences in response to stress. Pediatrics 1992;90:487-490.

Lewis M and Ramsay DS. Developmental change in infants' responses to stress. Child Dev 1995;66:657-670.

Lindh V, Wiklund U and Håkannson S. Heel lancing in term new-born infants: an evaluation of pain by frequency domain analysis of heart rate variability. Pain 1999;80:143-148.

Lynn AM, Nespeca M, Bratton SL, Strauss SG and Shen DD. Clearance of morphine in postoperative infants during intravenous infusion: the influence of age and surgery. Anesth Analg 1998;86:958-963.

McGrath PJ. Behavioural measures of pain. In: GA Finley and PJ McGrath (Eds.). Progress in pain research and management, Vol. 10, IASP Press, Seattle, 1998. pp. 83-102.

McGrath PJ, Johnson G, Goodman JT, Schillinger J, Dunn J and Chapman JA. CHEOPS: A behavioural scale for rating postoperative pain in children. In: HL Fields, R Dubner and F Cervero (Eds.). Advances in Pain Research and Therapy, Vol. 9, Raven Press, New York, 1985. pp. 395-402.

McGrath PJ and Unruh AM. Measurement and assessment of paediatric pain. In: PD Wall and R Melzack (Eds.). Textbook of pain., Churchill Livingstone, Edinburgh, 1994. pp. 303-313.

McIntosh N, Van Veen L and Brameyer H. The pain of heel prick and its measurement in preterm infants. Pain 1993;52:71-74.

Morton NS. Pain assessment in children. Paediatr Anaesth 1997;7:267-272.

Owens ME and Todt EH. Pain in infancy: neonatal reaction to a heel lance. Pain 1984;20:77-86.

Porter FL. Pain assessment in children: infants. In: NL Schechter, CB Berde and M Yaster (Eds.). Pain in infants, children, and adolescents, Williams & Wilkins, Baltimore, 1993. pp. 87-96.

Stevens B. Composite measures of pain. In: GA Finley and PJ McGrath (Eds.). Progress in pain research and management, Vol. 10, IASP Press, Seattle, 1998. pp. 161-178.

Stevens BJ, Johnston CC and Horton L. Factors that influence the behavioural pain responses of premature infants. Pain 1994;59:101-109.

Stevens BJ, Johnston CC, Petryshen P and Taddio A. Premature Infant Pain Profile: development and initial validation. Clin J Pain 1996;12:13-22.

Terndrup TE. Pediatric pain control. Ann Emerg Med 1996;27:466-470.

Tyler DC, Tu A, Douthit J and Chapman CR. Toward validation of pain measurement tools for children: a pilot study. Pain 1993;52:301-309.

Varni JW, Thompson KL and Hanson V. The Varni/Thompson Pediatric Pain Questionnaire. I. Chronic musculoskeletal pain in juvenile rheumatoid arthritis. Pain 1987;28:27-38.

Weatherstone KB, Rasmussen LB, Erenberg A, Jackson EM, Claflin KS and Leff RD. Safety and efficacy of a topical anesthetic for neonatal circumcision. Pediatrics 1993;92:710-714.

Williamson PS and Williamson ML. Physiologic stress reduction by a local anaesthetic during newborn circumcision. Pediatrics 1983;71:36-40.

Chapter 5

Efficacy of continuous versus intermittent morphine administration after major surgery in 0 to 3-year-old infants: a double-blind randomised controlled trial

Based on the article:

Efficacy of continuous versus intermittent morphine administration after major surgery in 0 to 3-year-old infants: a double-blind randomised controlled trial
Authors: Monique van Dijk, Nancy Bouwmeester, Hugo J.Duivenvoorden, Hans M.Koot, Dick Tibboel, Jan Passchier, Josien B.de Boer
(submitted)

5.1 Abstract

A randomised double-blind clinical trial compared the efficacy of 10 μg/kg/h morphine continuous IV infusion (CM) with 30 μg/kg morphine (IM) every three hours after major abdominal or thoracic surgery, in 181 infants aged 0 to 3 years. Efficacy was assessed with the COMFORT 'behaviour' and VAS every three hours in the first 24 hours after surgery. Random regression modelling was used to simultaneously estimate the effect of morphine condition, actual morphine dosage (protocol dosage plus extra morphine when required), age group, SSS and the time-varying covariate mechanical ventilation on COMFORT 'behaviour' and VAS pain. Overall, CM and IM morphine administration were equally effective in reducing postoperative pain. A significant interaction effect of condition with age group showed that the CM condition was favourable for the oldest age group (1 to 3 year old infants). Actual morphine dosage and age group significantly predicted the repeated pain assessments. The greatest differences in pain response and actual morphine dosage were between neonates and infants aged 1 to 6 months, with lower pain response in neonates who were on average satisfied with the protocol dosage of 10 μg/kg/h. Surgical stress score and mechanical ventilation were not related to postoperative pain or morphine dosages, leaving the inter-individual differences in pain response and morphine requirement largely unexplained.

5.2 Introduction

Reports on pediatric postoperative pain advocate good pain management comprising both pain assessment by validated instruments and adequate analgesic treatment (Berde, 1989; Beyer and Bournaki, 1989; Cohen, 1993; Glass, 1998; Goddard and Pickup, 1996; Morton, 1997). To accomplish this, more evidence-based knowledge about postoperative pain and related issues in pediatric samples is needed (McGrath, 1998; McIntosh, 1997).
One area of research examines the psychometric properties of (newly developed) postoperative pain instruments (Boelen et al., 1999; Buchholz et al., 1998; Dijk van et al., 2000; Gilbert et al., 1999; McGrath et al., 1985; Merkel et al., 1997; Tarbell et al., 1992). These latter studies suggest that behavioural pain measures are the preferred substitute measures when self-report is not possible, as is the case in preverbal infants. Another area of research describes the efficacy and safety of analgesic treatments after surgery in infants. After major surgery, opioids (especially morphine) are the most frequently

employed analgesics. The efficacy and safety of intravenous (IV) continuous morphine after major surgery in children was examined in three studies. Beasley and Tibbals (1987) found that 20 to 25 μg/kg/h morphine was effective after major surgery in 121 non-ventilated children 0 to >14 years of age. Millar and colleagues (1987) concluded that 14 to 21 μg/kg/h morphine was effective in 85% of 20 children, 3 months to 12 years of age. Finally, Esmail et al. (1999) evaluated the efficacy of morphine after major surgery in 110 non-ventilated children aged 3 months to 16 year old children, with infusion rates ranging from 10 to 40 μg/kg/h; the 65.5% inadequate analgesia in the latter study was related to the lower infusion rates. In all three studies, pain was the major outcome variable, assessed by either the Visual Analogue Scale (VAS) or Graphic Rating Scale (GRS) (Scott and Huskisson, 1976) in the studies of Beasley and Tibbals (1987) and Millar et al. (1987), and age-appropriate validated pain instruments in the study of Esmail and colleagues (1999). Clinical signs of ventilatory depression, an important safety outcome, were not observed in these latter studies. All three studies have some methodological drawbacks. Firstly, they are based on a broad age range, making it difficult to establish age-related differences in pain and morphine requirement. Secondly, morphine dosages were not administered according to standardised protocols, which makes the rationale for varying dosages unclear. Thirdly, because ventilated infants were excluded, neonates were underrepresented in these studies, which is unfortunate because, nowadays, infants with congenital anomalies are often operated at an early age (Jona, 1998).

Studies comparing different routes of administration in pediatric samples, reported that IV morphine administration gives better pain relief than intramuscular morphine injections (Bray, 1983; Hendrickson et al., 1990). In addition, for intramuscular injections, needle pain and fear makes this an undesirable route of administration (Hendrickson et al., 1990). To our knowledge, there are no double-blind randomised clinical trials which have compared the efficacy of IV morphine administration and intermittent IV morphine administration in both neonates and infants. In evaluating the efficacy of these two modes of administration it is important to account for effects of individual patient or procedural characteristics. First, there may be large age-related differences in response to standardised levels of morphine. Second, the efficacy of morphine after major surgery is often determined without consideration of differences in surgical procedures. One study described slower morphine clearance after cardiac than after non-cardiac surgery in neonates and infants (Lynn et al., 1998), which may implicate that smaller morphine dosages may be required in cardiac surgery. Although there are no objective measures to

determine the painfulness of a surgical procedure, there is a measure developed to assess the severity of surgical stress (SSS) in neonates (Anand and Aynsley-Green, 1988), but the relationship between postoperative pain and surgical stress has not yet been explored. Third, the effects of mechanical ventilation on postoperative pain assessment are unknown. This study addressed the postoperative analgesic efficacy of two modes of morphine administration, i.e. equal dosages of morphine either through continuous IV infusion or through IV bolus injections every 3 hours following major non-cardiac surgery.
In addition we examined the impact of age, severity of stress and mechanical ventilation on the individual pain response.

5.3 Methods

Patients
Between March 1995 and September 1998 a total of 204 children aged 0 to 3 years, who were admitted for major abdominal or thoracic surgery, entered the study after informed consent of the parents was obtained.
Included were: neonates (≥35 weeks gestation and bodyweight ≥1500 grams) and infants aged up to 3 years undergoing major thoracic or abdominal surgery.
Exclusion criteria were: use of co-medication (e.g. acetaminophen or midazolam) influencing the measured amount or potency of morphine, use of neuromuscular blockers, hepatic or renal dysfunction, seriously compromised neurological status or altered muscle tone.
The Medical Ethical Committee of the Hospital approved the study.

Instruments

COMFORT scale
The behavioural part of the COMFORT scale (further referred to as COMFORT 'behaviour') proved a reliable and valid instrument to assess postoperative pain in 158 neonates and infants, using trained nurses as observers (Dijk van et al., 2000). The COMFORT 'behaviour' score consists of the summation of six behavioural items: Alertness, Calmness, Muscle tone, Movement, Facial tension, and Respiratory response (for ventilated children) or Crying (for non-ventilated children) with response categories

ranging from 1 (low distress/no pain) to 5 (high distress/pain). The COMFORT 'behaviour' score ranges from 6 to 30.

Visual Analogue Scale (VAS)
The nurses completed a VAS for a clinical rating of pain in each child. The VAS is a horizontal continuous ten-centimetre line with the anchors 'no pain' on the left side and 'extreme pain' on the right side. Observers estimate the level of pain by making a mark on the line. The score ranges from 0 to 10 (McGrath et al., 1985; Varni et al., 1987).

Surgical Stress Score (SSS)
The SSS (Anand and Aynsley-Green, 1988) was originally developed to assess the severity of surgical stress in neonates and includes the following items: Amount of blood loss (score range 0 to 3); Site of surgery (score range 0 to 2); Amount of superficial trauma (score range 1 to 3); Extent of visceral trauma (score range1 to 4); Duration of surgery (score range 1 to 5); Associated stress factors: a) Hypothermia (score range 0 to 3), b) Infection (score range 0 to 3). The items Prematurity and Cardiac surgery of the SSS were not applicable in our study and therefore the score could range from 3-22.

Design
A double-blind randomised clinical trial was carried out to compare the efficacy of intravenous continuous morphine (CM) and intravenous intermittent morphine (IM) after major abdominal or thoracic surgery in infants aged 0 to 3 years. Prestratification by age was performed because behavioural and physiological differences between age groups were expected to be of importance. Age groups comprised neonates (≥35 weeks gestation and weight ≥1500 grams), younger infants (1 to 6 months), older infants (7 to 12 months) and toddlers (1 to 3 years). Infants within age groups were assigned to CM or IM analgesia by random number generation. The hospital pharmacist prepared study drugs and the randomisation schedule was only known to the pharmacist and retained until the end of the trial.

Procedure
Anaesthetic management was standardised. At the end of surgery, all patients were given an intravenous loading dose of morphine ≥100 μg/kg until they were in minimal pain as indicated by a VAS score <4.

Protocol morphine dosage, following the loading dose, was administered in the following way. The CM group started with a morphine infusion of 10 μg/kg/h, combined with a three-hourly intravenous placebo bolus (saline). The IM group received a continuous placebo infusion (saline), combined with a three-hourly intravenous morphine bolus of 30 μg/kg. When children were considered to be in pain (VAS ≥4), additional morphine could be given according to the following decision rules:
The first hour after surgery and VAS ≥4: 30 μg/kg morphine (= approx. 1/3 of the loading dose) as required every 15 minutes. More than 1 h after surgery and VAS ≥4: 5 μg/kg morphine as required every 10 minutes. If this did not result in adequate pain control, the anaesthesiologist was consulted for additional analgesic treatment.
Pain assessment was performed prior to surgery, after return to the Pediatric Surgical Intensive Care unit (PSICU), and every three hours during the first 36 hours postoperative. Mechanical ventilation was continued after surgery in neonates <37 weeks and after repair of oesophageal atresia or congenital diaphragmatic hernia. In older age groups, postoperative ventilation was required depending on the surgical procedure.

Statistical analysis
Random regression modelling was used to simultaneously estimate the effect of morphine condition, actual morphine dosage (protocol dosage plus extra morphine when required), age group, SSS and the time-varying covariate mechanical ventilation on COMFORT 'behaviour' and VAS pain. Actual morphine dosage was highly skewed and therefore discretised into three categories (1: 10 μg/kg/h, 2: >10 to ≥15 μg/kg/h and 3: >15 μg/kg/h). The outcome variables COMFORT 'behaviour' and VAS pain were skewed to the right and logtransformed (base 10) to achieve normality.
Random regression modelling has many advantages: it allows for missing data or an unequal number of data for each subject, and for the inclusion of fixed and time-varying covariates. Furthermore, a realistic covariance structure (as opposed to compound symmetry or independence between repeated measures) can be implemented (Gibbons et al., 1993).
The intention-to-treat rationale was considered inappropriate because additional non-opiate analgesia or sedatives may considerably influence the pain assessments and dosage requirements for morphine. This resulted in the exclusion of cases from analysis that received non-opiate analgesia or sedatives.

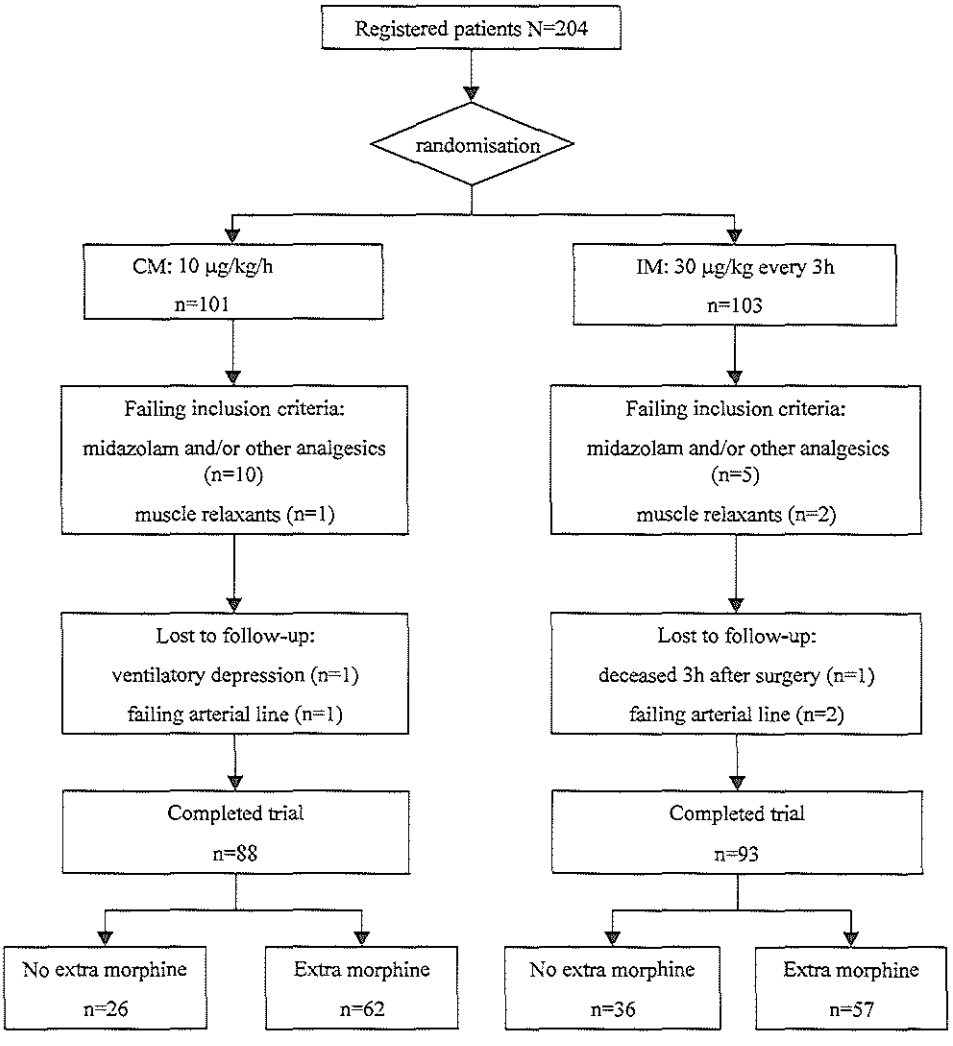


Figure 1.Flow chart of study sample; CM=continuous morphine, IM= intermittent morphine

5.4 Results

Figure 1 depicts the progress of patients through the trial. Of the 204 patients, 23 (13 CM;10 IM) were excluded, primarily due to the administration of other analgesics and/or midazolam (n=15; 10 CM, 5 IM).

In the CM condition, 26 (29%) of the infants were without pain with the 10 μg/kg/h morphine as opposed to 36 (39%) of the infants in the IM condition (Figure 1). This was not significantly different for the two conditions ($\chi^2_{\text{Yates' corrected}} = 1.31$, p=0.26).
Table 1 depicts sample characteristics for both the original sample (n=204) and the selected sample (n=181).

Table 1 Background characteristics of the original sample and the group completing the trial

	Original (n=204)		Completed trial (n=181)	
	n	%	n	%
Age groups				
Neonates	66	32	65	36
1 to 6 months	67	33	54	30
6 to 12 months	31	15	27	15
1 to 3 years	40	20	35	19
Gender				
Boys	119	58	104	57
Girls	85	42	77	43
Location of surgery				
Superficial	13	6	12	7
Low abdominal	91	45	86	48
High abdominal	62	30	53	29
Thoracic	31	15	24	13
Thoracic + abdominal	7	3	6	3
Mechanical ventilation postoperative	85	42	75	42
Yes	119	58	106	58
No				
Severity of surgical stress				
Mean (SD)	9.7 (3.0)		9.6 (3.0)	

n; number of patients

Table 2 shows the summary statistics for pain assessments, morphine dosage and postoperative mechanical ventilation for the two conditions. There were no significant differences between the CM and IM condition.
The effects of the predictor variables on COMFORT 'behaviour' and VAS pain as outcome variables are given in Table 3. Both models incorporated random intercepts and random slopes.
The two morphine conditions did not have differential effects on the repeated COMFORT 'behaviour' or VAS pain scores. However, actual morphine dosage was significantly related to COMFORT 'behaviour' and VAS. The interaction between condition and morphine

dosage was not significant, indicating that the relation between the actual morphine dosage and the level of VAS pain and COMFORT 'behaviour' score did not depend on the morphine condition.

Table 2 Data on pain assessment, morphine dosage and mechanical ventilation after surgery, in the two conditions: CM (n=88) and IM (n=93)

Pain assessment [1)]	**CM**	**IM**
VAS Median (IQR)	1.9 (1.1 to 2.8)	1.8 (1.2 to 2.6)
#times VAS ≥ 4 Median (IQR)	1 (0 to 2)	1 (0 to 2)
COMFORT 'behaviour' Median (IQR)	14.2 (2.9)	14.3 (2.5)
Morphine dosage [2)]		
Actual morphine dosage (μg/kg/h)		
Median (IQR)	10.8 (10 to 12.2)	10.4 (10 to 12.5)
Actual morphine dosage	**n (%)**	**n (%)**
10 μg/kg/h	26 (30%)	36 (39%)
>10 to ≤ 15 μg/kg/h	54 (61%)	43 (46%)
> 15 μg/kg/h	8 (9%)	14 (15%)
Mechanical ventilation		
Yes	38 (43%)	37 (40%)
No	50 (57%)	56 (60%)

[1)] assessed nine times in the first 24 hours
[2)] maximum dosage 36.9 in the CM condition, 26.7 in the IM condition
CM=continuous morphine condition, IM=intermittent morphine condition;
IQR=Interquartile range, n =number of patients

Age group significantly predicted both outcome variables, with the largest difference between neonates and the younger infants. The interaction between age groups and morphine condition was significant for VAS pain (p=0.0006) and almost reached significance (p=0.055) for the COMFORT 'behaviour'. This was primarily explained by the difference between neonates and toddlers with the CM condition being more favourable than the IM condition for the toddlers, but not for the neonates. The VAS pain score significantly decreased over time, whereas the COMFORT 'behaviour' score showed no significant time trend.

Table 3a Predictability of COMFORT 'behaviour' (n=181)[1]

Outcome variable : COMFORT 'behaviour'		
	F	P
Morphine condition	1.41	0.24
Morphine dosage[2)]	**4.06**	**0.02**
Age group [3)]	**7.68**	**0.0001**
Morphine condition* dosage	0.03	0.97
Morphine condition * age group	2.54	0.055
Time trend	0.01	0.93
Surgical Stress Score	2.23	0.14
Mechanical ventilation	0.06	0.81

Table 3b Predictability of VAS pain (n=181)[1]

Outcome variable : VAS pain		
	F	p
Morphine condition	0.74	0.39
Morphine dosage [2)]	**15.49**	**0.0001**
Age group [3)]	**8.65**	**0.0001**
Morphine condition* dosage	0.97	0.38
Morphine condition * age group	**5.78**	**0.0006**
Time trend	**38.81**	**0.0001**
Surgical Stress Score	0.06	0.81
Mechanical ventilation	0.98	0.32

[1)] Random regression modelling for repeated measurements.
[2)] Age groups were discretised as: 1= neonates, 2= 1 to 6 months, 3= 6 to 12 months 4= 1 to 3 years
[3)] Morphine dosage 1: 10 μg/kg/h, 2: >10 to ≤15 μg/kg/h, 3: >15 μg/kg/h
Significant predictor variables ($p < 0.05$) printed boldfaced.

The SSS and mechanical ventilation did not significantly predict the level of COMFORT 'behaviour' and VAS pain. The most significant predictor variables are shown in Figures 2 and 3 which give the means and standard errors of COMFORT 'behaviour' and VAS pain scores over time for the four age groups, split by the overall median dosage (10.7 μg/kg/h). As was confirmed by the random regression analyses, average levels of COMFORT 'behaviour' and VAS were significantly different for the groups split by overall median morphine dosage level. Age effects were clearly seen, especially between neonates and young infants. The VAS showed a moderate decline in scoring over time. For the COMFORT 'behaviour' the toddler group showed a slight increase in levels of COMFORT 'behaviour', with limited differences between those below and above the overall median morphine dosage level.

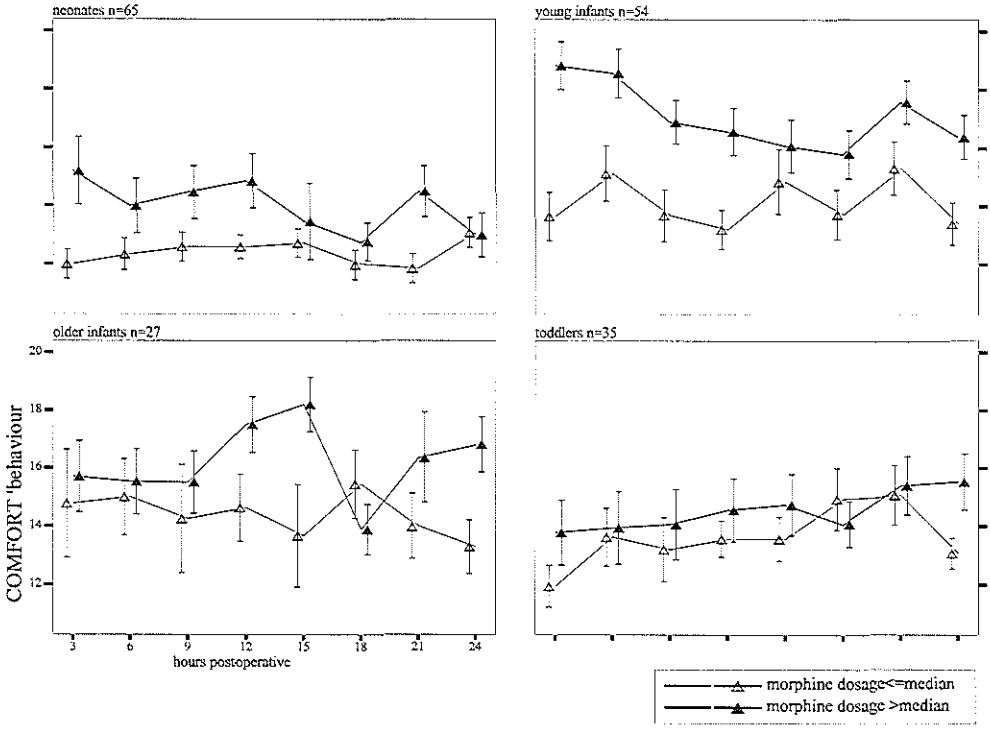


Figure 2 Mean and SE of COMFORT 'behaviour' scores for the four age groups split by median morphine dosage: 77% of the neonates (50 of 65) had morphine dosage<=median, 30% of young infants (16 of 54), 33% of the older infants (9 of 27) and 46% of the toddlers (16 of 35)

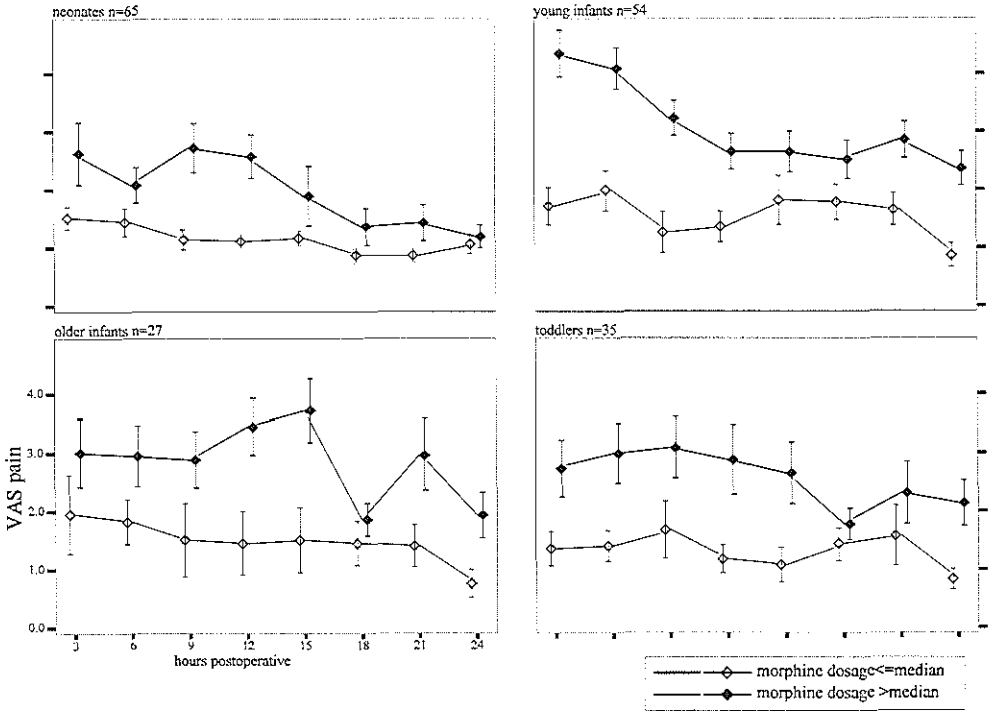


Figure 3 Mean and SE of VAS pain scores for the four age groups split by median morphine dosage: of the neonates (50 of 65) had morphine dosage<=median, 30% of young infants (16 of 54), 33% of lder infants (9 of 27) and 46% of the toddlers (16 of 35)

5.5 Discussion

Morphine condition

This randomised double-blind controlled trial showed that continuous and intermittent IV administration of morphine are equally effective in reducing postoperative pain in infants aged 0 to 3 years. Only for the 1 to 3-year-old infants, was the CM condition somewhat

favourable. This could be explained by the fact that the half-life of morphine reaches adult values with half-life of 2h ± 1.8 for infants older than 2 months (Kart et al., 1997a). Therefore, the 3-hour period between IV morphine bolus injections in the intermittent condition could result in morphine plasma levels below the therapeutic range in the 1 to -3 year old age group. The unexpected equal results for the two routes of morphine administration in relation to postoperative pain are promising, especially for clinical settings with a limited availability of infusion pumps (e.g. developing countries).

Inter-individual variability

This trial also showed inter-individual variability in analgesic need after major surgery. Although the protocol dosage of morphine was relatively low (10 μg/kg/h), 62 infants (34%) seemed satisfied with this dosage according to the pain assessment by the nurses, 97 infants required between 10.1 and 15 μg/kg/h, whereas 22 infants (12%) received more than 15 μg/kg/h with a maximum of 36.9 μg/kg/h. Another 15 infants were excluded from analyses, because other analgesics and/or sedatives were considered necessary based on repeated pain responses. Table 4 shows the characteristics of the 15 (excluded) infants who received additional analgesics or sedatives.

In two of the patients (no. 4 and 5) midazolam was administered because they were 'fighting the ventilator' and were not considered to be in pain. The only excluded neonate (no.1) was given midazolam once as sedative. One child (no.15) was given midazolam with alimemazine once, because he was very restless. The other 11 excluded infants received midazolam and/or acetaminophen (APAP) or diclofenac, because they remained in pain or were described as 'tense' despite extra morphine (dosage ranging from 12.8 to 40.2 μg/kg/h). (The decisions to use additional analgesics or sedatives were made at varying time points explaining the range in morphine dosage).

Table 4 Overview of patients excluded due to use of analgesic or sedative medication other than morphine

No.	Age (in days)	Morphine condition[a]	COMFORT[b] 'behaviour'	VAS[b]	Morphine dosage (µg/kg/hr)	Midazolam[c]	APAP or diclofenac	Mechanical Ventilation	SSS	Location of surgery	Reasons for other medication
1	4	IM	12	1.5	11.2	+		-	15	Superficial	To settle for the night
2	61	CM	17	4	21.3		+	-	10	High abdominal	Painful despite extra morphine
3	61	CM	17	2.2	22.2		+	+	11	Low abdominal	Painful despite extra morphine
4	90	IM	12	1.7	10.4	+		+	7	High abdominal	Fighting the ventilator
5	91	CM	13	2.8	24.2	+	+	+	13	Thoracic	Fighting the ventilator
6	95	IM	17	3.6	34.2	+		+	10	Low abdominal	Painful and tense (peritonitis)
7	157	CM	20	5.7	17.2	+C		-	6	Low abdominal	Painful despite extra morphine
8	163	CM	16	4.1	21.7	+C		-	10	Thoracic	Painful despite extra morphine
9	170	CM	16	3.0	22.4		+	-	12	Low abdominal	Painful despite extra morphine
10	185	CM	17	4.1	30.4	+		-	13	Low abdominal	Very tense, despite extra morphine
11	203	IM	13	1.5	12.8		++	-	13	Thoracic	Painful
12	301	CM	17	2.5	40.2	+C	+	-	7	High abdominal	Painful despite extra morphine
13	498	CM	17	3.8	16.2	+	+	-	14	Thoracic+abdominal	Jittery
14	528	IM	20	5.4	32.7	+C	+	-	6	High abdominal	Painful and anxious
15	757	CM	12	1.0	10.00	+[d]		-	12	Thoracic	Very restless

[a] CM=continuous morphine infusion, IM=intermittent morphine boluses
[b] Average COMFORT 'behaviour' and VAS for each person across first 24 h after surgery
[c] += One or more boluses midazolam, +C =followed by continuous infusion of midazolam
[d] Received midazolam and alimemazine once

Age

Age was related to pain and actual morphine dosage, although not in a linear way. Neonates had lower COMFORT 'behaviour' scores and VAS pain scores compared to infants aged 1 to 6 months. Neonates metabolise morphine slower than older infants (Kart et al., 1997a; Lynn et al., 1998) which could explain the lower pain scores of the neonates. Why the young infants had the most pain (followed by the older infants) could not be explained by surgery-related characteristics and we have no explanation for this phenomenon. However, on average postoperative pain was moderate to low, as was confirmed by the median VAS below 2. A matter of consideration with infants and toddlers is the problem of differentiating between pain, anger and anxiety. With increasing age infants are likely to become more aware of their environment and to respond in their unique way. Furthermore, infants become more lively/active when they feel better, which might unintentionally increase the behavioural pain scores. This could explain the slight increase in COMFORT 'behaviour' scores across time for the toddlers (Figure 2), and the non-significant time trend for COMFORT 'behaviour'. During training and implementation of the COMFORT scale, it should be emphasised to score distress behaviour only. The observational VAS is less sensitive for such misinterpretation as it asks directly to assess the pain intensity. On the other hand, the quality of the observational VAS depends strongly on the observer's experience and knowledge of infants in pain. By contrast, the COMFORT 'behaviour' is based solely on a 2-minute behavioural observation and can improve by training.

Of the 62 infants who received 10 μg/kg/h morphine after surgery, 42 were neonates (CM: n=20; IM: n=22). Only one neonate was included in the group who received >15 to 36.9 μg/kg/h morphine. For term neonates 10 μg/kg/h, and 15 μg/kg/h for infants, seems an adequate dose to begin with after major surgery. In the literature, recommended dosages for morphine IV infusion range from 10 to 40 μg/kg/h, mostly depending on age. In a review by Kart et al. (1997a) the recommended dosage for term neonates was calculated to be 7 μg/kg/h and for infants 20 μg/kg/h. Considering the variability within the age groups in the present study, our results confirm the current opinion that dosages should be determined or adjusted for each individual separately (Kart et al., 1997b; Pokela et al., 1993).

Surgical stress score (SSS)

In our study the severity of surgical stress was not significantly related to postoperative pain. In other studies, postoperative metabolic and hormonal stress responses were positively correlated with the SSS in neonates (Anand and Ward Platt, 1988), and metabolic and hormonal stress responses decreased thanks to analgesic treatment after surgery in premature neonates (Anand et al., 1987). Since ours is the first study to relate SSS to behavioural pain measures, we may conclude that more research is required to establish knowledge about the stress-pain relationship (Aynsley-Green, 1996). Another explanation for the apparent lack of association between SSS and pain response might be that the SSS in our study was relatively low with limited variability across patients, partly due to the exclusion of premature neonates and cardiac surgery. In our opinion, a score designed to rank the painfulness of surgical procedures, in addition to the stressfulness, would be useful.

Mechanical ventilation

In this study, the time-varying covariate mechanical ventilation did not predict the repeated pain assessments. In our sample 42% (75 of 182 cases) required postoperative ventilation, equally divided across the two conditions. The majority of the ventilated cases were neonates (51 of 75 cases). Ventilated cases are usually excluded either because safety is an outcome variable or because pain assessment of ventilated cases is too complex. Although the safety of morphine for the ventilated cases could not be determined, pain assessment for the ventilated cases was not a problem in our study, because we used the COMFORT 'behaviour' which has one item specifically developed for ventilated cases, which replaces the item 'crying' used for non-ventilated cases. In the present study, one patient with clinical signs of ventilatory depression was excluded from the trial. The inclusion of ventilated cases in our study had the advantage that our findings may be generalised to the population of the PSICU. Our results on the pharmacodynamic and pharmacokinetic properties of morphine, comprising blood gas analyses, will be the topic of another paper. In summary, the results of this study show that continuous and intermittent IV morphine administration after major surgery, were equally effective in infants up to 1 year of age. Differences in pain response and morphine dosage were most prominent between neonates and infants 1 to 6 months old, with lower pain response in neonates (who were on average satisfied with the protocol dosage of 10 μg/kg/h) and higher pain response in infants aged 1 to 6 months, who required higher dosages of morphine. Individual differences in pain

response and morphine requirement remained largely unexplained. Surgical stress score was not related to postoperative pain or morphine dosages. These findings expand our knowledge on postoperative pain and may contribute to better postoperative pain management for infants aged 0 to 3 years after major surgery.

5.6 Acknowledgements

The authors would like to thank the staff and nurses of the PSICU , the members of the surgical ward and Anaesthesiology of the Sophia Children's Hospital for their contribution to this study. We would also like to thank the parents and children who participated in this study. And finally, we thank Laraine Visser for her language corrections. This study was supported by a research grant from NWO (Dutch Organisation for Scientific Research, grant nr. 940-31-031).

5.7 References

Anand KJS and Aynsley-Green A. Measuring the severity of surgical stress in newborn infants. J Pediatr Surg 1988;23:297-305.

Anand KJS, Sippell WG and Aynsley-Green A. Randomised trial of fentanyl anaesthesia in preterm babies undergoing surgery: effects on the stress response. Lancet 1987;1:243-248.

Anand KJS and Ward Platt MP. Neonatal and pediatric stress responses to anesthesia and operation. Int Anesthesiol Clin 1988;26:218-225.

Aynsley-Green A. Pain and stress in infancy and childhood--where to now? Paediatr Anaesth 1996;6:167-172.

Beasley SW and Tibbals J. Efficacy and safety of continuous morphine infusion for postoperative analgesia in the paediatric surgical ward. Aust N Z J Surg 1987;57:233-237.

Berde CB. Pediatric postoperative pain management. Pediatr Clin North Am 1989;36:921-941.

Beyer JE and Bournaki M. Assessment and management of postoperative pain in children. Pediatrician 1989;16:30-38.

Boelen WJC, Scheffer E, Haan de RJ and Groot de CJ. Clinimetric evaluation of the pain observation scale for young children in children aged between 1 and 4 years after ear, nose, and throat surgery. J Dev Behav Pediatr 1999;20:14-19.

Bray RJ. Postoperative analgesia provided by morphine infusion in children. Anaesthesia 1983;38:1075-1078.

Buchholz M, Karl HW, Pomietto M and Lynn AM. Pain scores in infants: a modified infant pain scale versus visual analogue. J Pain Symptom Manage 1998;15:117-124.

Cohen DE. Management of postoperative pain in children. In: NL Schechter, CB Berde and M Yaster (Eds.). Pain in infants, children, and adolescents, Williams & Wilkins, Baltimore, 1993. pp. 87-96.

Dijk van M, Boer de JB, Koot HM, Passchier J, Tibboel D and Duivenvoorden HJ. The reliability, stability and validity of the COMFORT scale as a postoperative pain instrument in 0 to 3-year-old infants. Pain 2000;84:367-377.

Esmail Z, Montgomery C, Court C, Hamilton D and Kestle J. Efficacy and complications of morphine infusions in postoperative paediatric patients. Paediatr Anaesth 1999;9:321-327.

Gibbons RD, Hedeker D, Elkin I, Waternaux C, Kraemer HC, Greenhouse JB, Shea MT, Imber SD, Sotsky SM and Watkins JT. Some conceptual and statistical issues in analysis of longitudinal psychiatric data. Application to the NIMH treatment of Depression Collaborative Research Program dataset. Arch Gen Psychiatry 1993;50:739-750.

Gilbert CA, Lilley CM, Craig KD, McGrath PJ, Court CA, Bennett SM and Montgomery CJ. Postoperative pain expression in preschool children: validation of the child facial coding system. Clin J Pain 1999;15:192-200.

Glass NL. Pediatric postoperative pain management. Anesth Analg 1998;Suppl:28-31.

Goddard JM and Pickup SE. Postoperative pain in children. Anaesthesia 1996;51:588-590.

Hendrickson M, Myre L, Johnson DG, Matlak ME, Black RE and Sullivan JJ. Postoperative analgesia in children: A prospective study of intermittent intramuscular injection versus continuous intravenous infusion of morphine. J Pediatr Surg 1990;25:185-191.

Jona JZ. Advances in neonatal surgery. Pediatr Clin North Am 1998;45:605-617.

Kart T, Christrup LL and Rasmussen M. Recommended use of morphine in neonates, infants and children based on a literature review: Part 1--Pharmacokinetics. Paediatr Anaesth 1997a;7:5-11.

Kart T, Christrup LL and Rasmussen M. Recommended use of morphine in neonates, infants and children based on a literature review: Part 2--Clinical use. Paediatr Anaesth 1997b;7:93-101.

Lynn AM, Nespeca M, Bratton SL, Strauss SG and Shen DD. Clearance of morphine in postoperative infants during intravenous infusion: the influence of age and surgery. Anesth Analg 1998;86:958-963.

McGrath PJ. Behavioral measures of pain. In: GA Finley and PJ McGrath (Eds.). Progress in pain research and management, Vol. 10, IASP Press, Seattle, 1998. pp. 83-102.

McGrath PJ, Johnson G, Goodman JT, Schillinger J, Dunn J and Chapman JA. CHEOPS: A behavioral scale for rating postoperative pain in children. In: HL Fields, R Dubner and F Cervero (Eds.). Advances in Pain Research and Therapy, Vol. 9, Raven Press, New York, 1985. pp. 395-402.

McIntosh N. Pain in the newborn, a possible new starting point. Eur J Pediatr 1997;156:173-177.

Merkel SI, Voepel-Lewis T, Shayevitz JR and Malviya S. The FLACC: A behavioral scale for scoring postoperative pain in young children. Pediatr Nurs 1997;23:293-297.

Millar AJW, Rode H and Cywes S. Continuous morphine infusion for postoperative pain in children. S Afr Med J 1987;72:396-398.

Morton NS. Pain assessment in children. Paediatr Anaesth 1997;7:267-272.

Pokela M, Olkkola KT, Seppälä T and Koivisto M. Age-related morphine kinetics in infants. Dev Pharmacol Ther 1993;20:26-34.

Scott J and Huskisson EC. Graphic representation of pain. Pain 1976;2:175-184.

Tarbell SE, Cohen TI and Marsh JL. The Toddler-preschooler postoperative pain scale: an observational scale for measuring postoperative pain in children aged 1-5. Preliminary report. Pain 1992;50:273-280.

Varni JW, Thompson KL and Hanson V. The Varni/Thompson Pediatric Pain Questionnaire. I. Chronic musculoskeletal pain in juvenile rheumatoid arthritis. Pain 1987;28:27-38.

Chapter 6

Postoperative pain and stress response in 1 to 36 month old infants in relation to hospital history.

6.1 Introduction

The growing interest in long-term consequences of pediatric pain and stress response is instigated by findings suggesting that prior (neonatal) pain experiences affect behavioural and physiological pain response. One of the few studies in formerly full term healthy neonates was performed by Taddio et al. (1997). Infants who were circumcised at neonatal age showed stronger pain responses to subsequent routine vaccination at 4 or 6 months than uncircumcised infants did. Among the circumcised group, preoperative treatment with lidocaine-prilocaine cream (EMLA) attenuated the pain response to vaccination.

Johnston and Stevens (1996) compared infants of 32 weeks postconceptional age (PCA) born within the past 4 days, with infants of the same PCA who had been born 4 weeks earlier and had spent that time in a neonatal intensive care. During heel stick the latter showed less behavioural pain responses and higher maximum heart rate. The blunted behavioural response was primarily explained by the higher number of invasive procedures, the increased physiological responses by perinatal factors. At 4 months corrected conceptional age (CCA), 21 former extreme low birth weight (ELBW) and 24 full term infants were compared during finger lance (Oberlander et al., 2000). Overall, behavioural and cardiac autonomic responses to the lance were similar between groups. These overall findings were the same in a study in which 19 former ELBW infants at 8 months CCA were compared with 20 term born controls during finger lance (Grunau et al., 2000a). However, in this study former ELBW infants had significantly higher baseline mean heart rate. In addition, the number of invasive procedures from birth to 8 months CCA was related to higher baseline heart rate and less facial pain score at finger lance in the ELBW (Grunau et al., 2000).
Parent ratings of pain sensitivity of ELBW at 18 months CCA were lower compared to controls (Grunau et al., 1994a). Furthermore, the relationship between pain sensitivity and temperament varied systematically across the groups. Mothers of ELBW infants of 4 ½ years CCA, gave higher somatization scores to their children than mothers of full term controls (Grunau et al., 1994b). In the ELBW infants, 9 of 36 (25%) had a somatization score above the clinical cut-off, compared to none in the control group. Maternal factors and temperament at age 3 were related to this somatization score (Grunau et al., 1994b). Finally, at 8 to10 years of age, former ELBW and full birthweight (FBW) children themselves rated several pain situations with the Pediatric Pain Inventory (Grunau et al., 1998). There were no overall differences between the groups. However, the ELBW

children rated medical pain intensity higher than psychosocial pain, unlike the FBW group. Furthermore, duration of NICU stay was related to increased pain affect ratings in recreational and daily living settings (Grunau et al., 1998).

Studies on the long-term consequences of pain on subsequent stress response are scarce. The effect of birth condition (optimal vs non-optimal) on salivary cortisol response to routine inoculation at 2, 4 and 6 months was examined by Ramsay and Lewis (1995). Non-optimal birth condition was associated with lower cortisol response to inoculation at 2 months and higher cortisol response at 4 and 6 months. Salivary cortisol response to inoculation at 8 weeks was highest in babies born by assisted delivery, and lowest in those born by elective caesarean section (Taylor et al., 2000). Adrenocorticol responses to heelstick tended to increase or sensitise in term neonates receiving a second heelstick after 24 hours (Gunnar et al., 1991).

Recent reviews further confirm the importance of long-term consequences of pain (Anand, 2000a; Anand et al., 1997; Grunau, 2000b; Porter et al., 1999). In this context, Anand (2000) proposed the following hypothesis: 'the plasticity of the developing pain system provides a critical window for producing long-term changes in subsequent behaviour, responses to stress, and susceptibility to psychosomatic complaints and psychiatric disorders in later life.
There are no follow-up studies relating both behavioural and stress responses to prior painful experiences. Moreover, there are no studies evaluating the pain and stress responses after major surgery in both prematurely and full term infants under 3 years of age with a hospital past on account of premature birth or major congenital anomalies. During a randomised controlled trial in which two analgesic regimens were compared, we noticed considerable inter-individual differences in pain response, as measured by pain instruments (Dijk van et al., 2000) and stress hormones (Bouwmeester et al., submitted) not related to the analgesic treatment. The sample included many children with major multiple congenital anomalies and also prematurely born infants, who had been long hospitalised with adverse events. The inter-individual differences and hospital history on the one hand and literature findings on long-term effects on the other hand, led us to draw up the following hypotheses:

1. After a surgical procedure infants with more severe past (painful) experiences show on average more intense pain responses

2. Stress responses expressed as levels of adrenaline and noradrenaline will reflect past experiences with pain.

   To test the hypotheses, we combined the results from the aforementioned clinical trial with information from the medical records of the sample.

6.2 Methods

Patients

Between March 1995 and September 1998 a total of 204 children aged 0 to 3 years, who were admitted for major abdominal or thoracic surgery, entered a trial performed at the Pediatric Surgical Intensive Care (PSICU) of the Sophia Hospital Rotterdam. It encompassed a double-blind randomised clinical trial, comparing the efficacy of intravenous continuous morphine (CM) and intravenous intermittent morphine (IM) after major abdominal or thoracic surgery. Stratification in 4 groups: (I) neonates (≥35 weeks gestation and weight ≥1500 grams), (II) infants 1 to 6 months, (III) infants 7 to 12 months, and (IV) toddlers 1 to 3 years. More detailed information of the original design is give elsewhere (Dijk van et al., 2000).

Data collection

From the medical records and the Hospital Computer System we collected the following data at the end of the trial:

- Total number of days hospitalised before present surgery;
- Number of procedures under anaesthesia (surgery, diagnostic procedures, insertion of central lines etc) before present surgery
- Total number of days at mechanical ventilation prior to present surgery;
- Use of previous morphine (number of days of continuous infusion of morphine, including days of weaning);
- Relevant co-variables: gestational age, postnatal age, sex, and postoperative morphine dosage

Pain response

The children's pain responses were measured every three hours after surgery for the first 24 hours postoperative. The COMFORT scale (Ambuel et al., 1992) was originally developed to assess distress in ventilated children of all ages in an intensive care

environment. The COMFORT 'behaviour' score may range from 6 to 30 and proved a reliable and valid instrument to assess postoperative pain in neonates and infants (Dijk van et al., 2000).
In addition the nurses used a Visual Analogue Scale (VAS) for clinical rating of pain. The used VAS is a horizontal continuous ten-centimetre line with the anchors 'no pain' on the left side and 'extreme pain' on the right side. The score ranges from 0 to 10 (McGrath et al., 1985; Varni et al., 1987).

Stress hormones
Blood samples were obtained prior to surgery, directly after, 6, 12 and 24 hours after surgery. Plasma concentrations of adrenaline and noradrenaline were measured by HPLC using fluorimetric detection (van der Hoorn et al., 1989).

Surgical Stress Score (SSS)
The SSS (Anand and Aynsley-Green, 1988) was originally developed to assess the severity of surgical stress in neonates and comprises a score between 3 and 22. This score includes the following surgical characteristics: amount of blood loss, site of surgery, amount of superficial trauma, extent of visceral trauma, duration of surgery and infection. The score is calculated directly after surgery by the attending anaesthesiologist and surgeon.

Statistical analysis
Continuous data with non-normal distribution were analysed with Kruskal-Wallis H-test. The Spearman rank correlation coefficient was used to determine the association between variables.
Multiple regression analysis was applied to test for statistical significance of explanatory variables with regards to the different outcome variables. To gain insight into the relative importance of the explanatory variables, the standardised regression coefficient (β) was estimated. This enabled us to compare the importance of the co-variables with each other. Stress hormone data had to be transformed logarithmically in order to obtain approximate normal distributions. The postoperative mean value of all available postoperative measurements was used for the outcome variables. Extremely skewed explanatory variables were dichotomised at their median value and transformed to a dummy variable (0 or 1 coding). As a rule of thumb, the ratio of the number of explanatory variables and sample size should be 1:10 at least (Harrell et al., 1984).

Table 1 Patient characteristics of study sample (N=132)

Variables	Number	%
Age in days		
median (IQR)	195 (92 to 439)	dna [a]
range	29 to 1070	dna
Sex (male/female)	76/56	57.6/42.4
Gestational age (in weeks) at birth		
24 to 31.6	16	12.1
32 to 35.6	21	15.9
term	95	72.0
Diagnosis		
Congenital anomalies [b]		
Small bowel obstruction	16	12.1
Hirschsprung disease	13	9.8
Lung anomalies	8	6.1
Anorectal atresia	7	5.3
Nephro-urogenital anomalies	7	5.3
Esophageal atresia	6	4.5
Biliary atresia	4	3.0
Diaphragmatic hernia	4	3.0
Esophageal + anorectal atresia	2	1.5
Heart anomalies	2	1.5
Others	8	6.1
Acquired Diseases		
Necrotizing enterocolitis (old)	17	12.9
Intussusception	10	7.6
Other acquired diseases	7	5.3
Gastro-esophageal reflux	3	2.3
Malignancies	18	13.6

[a] dna= Does not apply
[b] Twenty infants in the sample of 132 had multiple congenital anomalies
IQR=Interquartile Range

6.3 Results

From our original sample of 204 infants, all 66 neonates (postnatal age 0 to 28 days) were excluded because 82% of them had been operated on within 8 days after birth and consequently lack hospital history. Six others were excluded because of missing data; one

child died 3 hours after surgery due to therapy resistant pulmonary hypertension, three infants had a failing arterial line and two infants required muscle blockers during postoperative mechanical ventilation. The remaining sample included 62 infants of 1 to 6 months old, 31 infants of 6 to 12 months, and 39 infants of 1 to 3 years.
Patient characteristics of the included 132 infants are given in Table 1.

Table 2 Surgery-related characteristics of the patients (n=132)

Variables	Descriptives
Postoperative Pain	
VAS pain	
Mean (SD)	2.5 (1.2)
Range	0.4 to 5.7
COMFORT 'behaviour'	
mean (SD)	15.3 (2.5)
range	10.7 to 21.9
Morphine dosage (µg/kg/h)	
median (IQR)	11.5 (10.2 to 15.1)
range	10 to 40.2
Postoperative stress response [1)]	
Adrenaline (nmol litre^{-1})	
Median (IQR)	1.0 (0.6 to 1.9)
Range	0.1 to 5.5
Noradrenaline (nmol litre^{-1})	
Median (IQR)	2.53 (1.76 to 3.27)
Range	0.55 to 13.58
Surgery	Number
Superficial	9
Low abdominal	68
Upper abdominal	35
Thoracic	15
Thoracic + abdominal	5
Surgical stress score	
Mean (SD)	9.7 (3.1)
Range	3 to 17

[1)] mean values were calculated from the levels directly after surgery, 6, 12 and 24 h after surgery

Seventy-seven children had a major congenital anomaly, twenty of whom had more than one anomaly. Thirty-seven infants had an acquired disease for which they required surgery, and eighteen a malignancy. Table 2 lists the surgery-related characteristics.

Mean postoperative VAS pain was low and well below 4, the value which was considered to represent pain. Most patients (78%) underwent abdominal surgery.

Table 3 Overview of previous hospital experiences

	1 to 6 month old infants n=62	6 to 12 month old infants n=31	1 to 3 year old toddlers n=39
Prior hospital stay in (days)			
Median (IQR)	40 (15 to 75)	35 (16 to 74)	29 (11 to 55)
Range	0 to 135	0 to 234	1 to 400
Period at home since last hospital stay[1)]			
Median (IQR)	7 (0 to 51)	52 (27 to 153)	82 (24 to 164)
Range	0 to 172	0 to 326	6 to 520
Number of prior procedures under anesthesia [2)]			
Median (IQR)	1 (0 to 2)	1 (0 to 2)	1 (1 to 3)
Range	0 to 5	0 to 9	0 to 9
Prior morphine intake (in days)			
Median (IQR)	0 (0 to 4)	3 (0 to 7)	0 (0 to 2)
Range	0 to 37	0 to 25	0 to 33
Period of prior mechanical ventilation (in days)			
Median (IQR)	0 (0 to 8)	0 (0 to 6)	0 (0 to 0)
Range	0 to 40	0 to 56	0 to 44

1) n=105 because 27 infants were never hospitalised before (n=12 in 1-6 months, n=6 in 6-12 months, n=9 in 1-3 year group). 20 infants from 1 to 6 months group were hospitalised their entire life, opposed to 2 in 6 to 12 months and none in 1 to 3 year old group.

2) Totally 207 procedures under anesthesia were documented of which 47% surgery, 18% insertion of central venous line, 13% endoscopies, 16% biopsies and 6% others (cardiac catheterisation, ventriculography or manometry)

Table 3 gives an overview of previous hospital experiences. Findings are stratified by age groups.

The hospital stay in days is not significantly different between age groups (Kruskal-Wallis p=0.43). Twenty-seven infants had never been hospitalised prior to the present hospital stay, whereas twenty-two infants had been hospitalised their entire life. For the remaining 105 children the period since last hospital stay is significantly different for the three age groups (Kruskal-Wallis test, $\chi^2 = 26{,}20$, p=0.000), primarily due to the difference between the 1 to 6 months old infants compared to the other two age groups. In our sample, 93 (71.5%) infants had undergone 1 to 9 procedures (median of 2 procedures), the remaining 39 infants had no prior procedures under anaesthesia in the past. Fifty-five (42%) infants received morphine infusion in the past for a period of 1 to maximally 37 days (median of 5 days). Fifty-four infants (41%) had been mechanically ventilated in the past between 1 to maximally 56 days (median of 8 days).

Table 4 Bivariate Spearman rank correlations (with 95% CI) of postoperative pain response with explanatory variables

	Postoperative pain response			
	COMFORT 'behaviour' [1]		VAS [1]	
	r	95% CI	r	95% CI
Background characteristics				
Postnatal age	-0.31	-0.45 to –0.14	-0.21	-0.37 to –0.03
Sex [2]	-0.14	-0.31 to 0.03	-0.09	-0.26 to 0.08
Gestational age	-0.18	-0.34 to –0.01	-0.07	-0.24 to 0.10
Surgery characteristics				
Postoperative mean morphine dosage	0.49	0.35 to 0.61	0.68	0.57 to 0.76
Surgical Stress	-0.14	-0.30 to 0.04	0.01	-0.16 to 0.19
Hospital history				
Prior hospital stay (in days)	0.15	-0.03 to 0.31	0.10	-0.07 to 0.27
# of prior procedures under anesthesia	0.10	-0.07 to 0.27	0.08	-0.09 to 0.25
Prior mechanical ventilation (in days)	0.20	0.02 to 0.36	0.13	-0.05 to 0.30
Prior morphine intake (in days)	0.19	0.02 to 0.36	0.10	-0.07 to 0.27
Period since last hospital stay (in days)[3]	-0.10	-0.29 to 0.09	-0.14	-0.33 to 0.05

[1] mean value of pain assessments across first 24 hours postoperatively
[2] dummy coding: 0=boys, 1=girls
[3] n=105, because not applicable for infants who were hospitalised entire life
CI= Confidence Interval

The Spearman rank correlation coefficients between prior hospital stay, prior procedures under anesthesia, prior morphine use, and mechanical ventilation were significant and varied from 0.46 to 0.76.

Table 4 gives the (bivariate) Spearman rank correlation coefficients (with 95% CI) of average postoperative VAS, COMFORT 'behaviour', with the explanatory variables.

Table 5 Multiple regression analyses of mean VAS and COMFORT 'behaviour' (N=127)

	Dependent variables			
	COMFORT 'behaviour' after surgery[1)]		VAS after surgery[1)]	
Background characteristics	β	95% CI	β	95% CI
Age group	-0.24	-0.40 to –0.08	-0.17	-0.23 to –0.01
Sex	-0.12	-0.27 to 0.03	-0.06	-0.21 to 0.09
Gestational age	-0.04	-0.24 to 0.16	0.05	-0.15 to 0.20
Surgery characteristics				
Postoperative mean morphine dosage	0.43	0.28 to 0.58	0.51	0.36 to 0.66
Surgical Stress	-0.15	-0.30 to 0.00	-0.03	-0.18 to 0.12
Hospital history variables				
Total hospital stay	0.04	-0.17 to 0.25	0.01	-0.20 to 0.22
Prior procedures under anesthesia	0.12	-0.09 to 0.33	0.15	-0.06 to 0.36
Prior mechanical ventilation	-0.12	-0.35 to 0.11	-0.05	-0.28 to 0.18
Prior morphine intake	0.03	-0.21 to 0.27	-0.01	-0.25 to 0.23
	$R=0.56, R^2_{adj}=0.26$		$R=0.56, R^2_{adj}=0.26$	

Dummy coding: age group (0=1 to 6 months, 1=6months to 3 years), sex (0=boy, 1=girl), Postoperative mean morphine dosage, hospital stay, prior mechanical ventilation , prior procedures, and prior days of morphine infusion dichotomised at median value

1) mean value of pain assessments across first 24 hours postoperatively

Pain response was higher in the younger infants reflected in r_s of –0.31 for the COMFORT 'behaviour' , and r_s of -0.21 for VAS pain with postnatal age. Higher COMFORT 'behaviour' scores are significantly correlated with postoperative mean morphine dosage (r_s=0.49). The mean scores on VAS pain were significantly related to mean morphine

dosage (r_s=0.68). Hospital history variables were not significantly correlated with the pain responses, except moderately for mechanical ventilation (r_s=0.20) , and prior morphine intake (r_s=0.19) with COMFORT 'behaviour'.
Multiple regression analyses performed on the pain response outcome variables are shown in Table 5.
The multiple regression analyses show that the postoperative pain response after the present surgery was poorly explained by the hospital history, controlling for surgical and patient characteristics simultaneously.

Table 6 Bivariate correlations (with 95% CI) of postoperative stress response with explanatory variables

	Postoperative stress response			
	Mean adrenaline [a]		Mean noradrenaline[a]	
	r	95% CI	r	95% CI
Background characteristics				
Postnatal age	0.45	0.30 to 0.58	-0.26	-0.41 to –0.09
Sex [b]	-0.03	-0.20 to 0.14	-0.18	-0.34 to –0.01
Gestational age	0.01	-0.16 to 0.18	-0.15	-0.31 to 0.02
Surgery characteristics				
Postoperative mean morphine dosage	-0.01	-0.18 to 0.16	-0.06	-0.23 to 0.11
Surgical Stress Score	0.29	0.12 to 0.44	0.29	0.12 to 0.44
Hospital history				
Prior hospital stay	0.14	-0.03 to 0.30	0.11	-0.06 to 0.28
Prior procedures under anaesthesia	0.30	0.14 to 0.45	0.11	-0.06 to 0.28
Prior mechanical ventilation	0.005	-0.17 to 0.17	0.09	-0.08 to 0.26
Prior morphine intake	0.19	0.02 to 0.35	0.13	-0.04 to 0.29

[a] mean value of all four postoperative plasma levels
[b] 0=boys, 1=girls.
CI= Confidence Interval

Table 6 gives the (bivariate) correlation coefficients (and 95% CI) between mean adrenaline and noradrenaline plasma levels and the explanatory variables.
The following correlations were significant. Firstly, mean postoperative adrenaline was positively correlated with postnatal age (r_s= 0.45), surgical stress of present surgery

(r_s=0.29), number of previous procedures under anesthesia (r_s=0.30) and prior morphine intake (r_s=0.19). Noradrenaline was negatively correlated with postnatal age (r_s=-0.18), and positively correlated with surgical stress (r_s=0.29). Overall, the contribution of hospital history variables separately on present postoperative pain and stress response was limited. Table 7 gives the results of multiple regression analyses with adrenaline and noradrenaline respectively as outcome variables.

Table 7 Multiple regression analyses of mean adrenaline and mean noradrenaline (N=125)

	Outcome variables[a]			
Explanatory variables	Mean adrenaline after surgery[b]		Mean noradrenaline after surgery[b]	
	β	95% CI	β	95% CI
Baseline adrenaline	0.30	0.15 to 0.44		
Baseline noradrenaline			0.58	0.44 to 0.72
Background characteristics				
Age group	0.37	0.22 to 0.52	-0.21	-0.35 to –0.07
Sex	-0.03	-0.18 to 0.12	-0.13	-0.27 to 0.01
Gestational age	0.09	-0.10 to 0.28	-0.04	-0.22 to 0.14
Surgery characteristics				
Postoperative mean morphine dosage	-0.02	-0.17 to 0.13	-0.001	-0.14 to 0.14
Surgical Stress Score	0.28	0.13 to 0.43	0.13	-0.01 to 0.27
Hospital history variables				
Prior hospital stay	0.15	-0.05 to 0.35	0.11	-0.08 to 0.30
Prior procedures under anaesthesia	0.03	-0.16 to 0.22	0.07	-0.11 to 0.26
Prior mechanical ventilation	-0.03	-0.25 to 0.19	-0.21	-0.42 to 0.003
Prior morphine intake	0.19	-0.03 to 0.41	0.18	-0.03 to 0.39
	$R=0.64$, $R^2_{adj}=0.35$		$R=0.69$, $R^2_{adj}=0.43$	

[a] Outcome variables and baseline stress hormones were log10 transformed
[b] mean value of all four postoperative plasma levels
Dummy coding: age group (0=1 to 6 months, 1=6months to 3 years), sex (0=boy, 1=girl), Postoperative mean morphine dosage, hospital stay, prior mechanical ventilation , prior procedures, and prior days of morphine infusion dichotomised at median value

The adjusted R^2 , 0.35 and 0.43 for adrenaline and noradrenaline, respectively, had to be primarily attributed to the baseline plasma levels. Next to these, age was related to the levels of both stress hormones. The 1 to 6 month old infants as a group showed relatively higher adrenaline and lower noradrenaline plasma levels after surgery than the older infants. Surgical stress was positively related to higher adrenaline values.
The multiple regression analyses show that the postoperative stress response after the present surgery was poorly explained by the hospital history, controlling for surgical and patient characteristics simultaneously.

6.4 Discussion

We have examined the influence of previous hospital experiences on postoperative pain and stress responses, controlling for patient characteristics and surgical characteristics. We demonstrated in our sample that previous hospital experiences were not predictive of postoperative pain response or stress response in a clinically substantial manner. This was against our expectations based on our personal clinical experience and literature findings.

Several explanations for our findings are possible. Firstly, after major surgery, pre-emptive analgesics are given as a standard in our hospital. This shift from pain treatment toward pain prevention (Broadman, 1999), resulting in low pain scores, might explain the limited impact of previous hospital experiences (Oberlander et al., 2000; Whitfield and Grunau, 2000). On the other hand, many short painful procedures like insertion of peripheral lines, venapunctures, heel prick, and endotracheal suctioning are not performed under standard analgesia. We were not able to collect this information which could have been relevant for the prematurely born infants and the infants who were operated on as a neonate. Secondly, the heterogeneity of our sample with regard to, prior hospital experiences, and diagnosis may have influenced the results. Hospital history was not an issue in 27 infants who experienced their first hospital admittance of which eleven were acutely operated on after admittance (intussusception, perforated appendix, malrotation), with preoperative pain and tissue damage that may affect postoperative pain and analgesic requirements.

A special subset are the eighteen infants who were operated on for malignancies. The prior chemotherapy and anxiety of the parents and the children may play a role in postoperative pain (distress) behaviour in this group.

With respect to the available studies, described in the introduction, the following should be considered. Firstly, premature and term neonates show important developmental differences. Prior pain experience appear to increase subsequent behavioural response to pain in healthy term babies, whereas in premature babies it appears to diminish these responses (Anand, 2000b; Grunau, 2000b; Whitfield and Grunau, 2000). Secondly, the studies differ considerably in time span between pain procedure(s) and actual assessment. For instance, Johnston and Stevens (1996) assessed the premature neonates directly after either a period of NICU stay or shortly after birth, whereas Oberlander et al. (2000) assessed ELBW and full term controls during finger poke at 4 months (CCA). Thirdly, the past pain experiences vary in number and intensity. A history of NICU stay with many daily painful procedures differs from that of 'just' a circumcision or a stressful delivery. And finally, studies which evaluate long-term consequences of pain on behavioural pain response differ in design from those on stress response, e.g. cortisol levels.
Apart from these methodological considerations, critical evaluation of the findings from the studies , leads us to conclude that long-term consequences of early pain experiences seem not as detrimental as expected. The follow up studies of ELBW infants reveal moderate differences in pain response compared to FBW controls, but other factors, e.g. poorer motor co-ordination could also explain these differences (Grunau, 2000b). Perhaps human brain and neurobiological systems are capable of 'resetting' themselves, provided enough 'resting ' time is given.

However, we should be cautious, because a recent animal study showed that localised inflammation during the neonatal period in rat pups permanently alters the neuronal circuits that process pain in the spinal cord (Ruda et al., 2000). This knowledge calls for continuous patient related research. Future research should incorporate follow-up studies of large cohorts of neonates, both term and premature. A prospective design will enable us to carefully document all invasive painful procedures which take place at NICU's, including attempted procedures that are usually not documented. In addition, temperament and mother-child interaction (e.g. period parent visits NICU, period parent has bodily contact) could be assessed across time as well, in order to obtain a greater insight in the interplay of those variables with the painful experiences.
Only time can tell if the expression 'Time cures all things' is appropriate in this context at anatomical, physiological, and behavioural levels.

6.5 References

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

Anand KJ. Effects of perinatal pain and stress. Prog Brain Res 2000a;122:117-129.

Anand KJ. Pain, plasticity, and premature birth: a prescription for permanent suffering? Nat Med 2000b;6:971-973.

Anand KJS and Aynsley-Green A. Measuring the severity of surgical stress in newborn infants. J Pediatr Surg 1988;23:297-305.

Anand KJS, Grunau RVE and Oberlander TF. Developmental character and long-term consequences of pain in infants and children. Child Adolesc Psychiatr Clin N Am 1997;6:703-724.

Bouwmeester NJ, Anand KJS, Dijk van M, Hop WCJ, Boomsma F and Tibboel D. hormonal and metabolic stress response after major surgery in children aged 0-3 years: a double-blind, randomized trial comparing the effects of continuous versus intermittent morphine, submitted.

Broadman LM. Blocks and other techniques pediatric surgeons can employ to reduce postoperative pain in pediatric patients. Semin Pediatr Surg 1999;8:30-33.

Dijk van M, Boer de JB, Koot HM, Passchier J, Tibboel D and Duivenvoorden HJ. The reliability, stability and validity of the COMFORT scale as a postoperative pain instrument in 0 to 3-year-old infants. Pain 2000;84:367-377.

Grunau RVE, Oberlander TF, Whitfield MF, Fitzgerald C and Saul P, Prior pain experience and pain reactivity in former extremely low birthweight infants and term born controls at corrected age 8 months. Pedatriac Academic Societies and the American Association of Pediatrics year 2000 Joint Meeting, Boston, Massachusetts, 2000a.

Grunau RVE. Long-term consequences of pain in human neonates. In: KJS Anand, BJ Stevens and PJ McGrath (Eds.). Pain in neonates, 2nd revised and enlarged edition, Vol. 10, Elsevier, 2000b. pp. 55-76.

Grunau RVE, Whitfield MF and Petrie JH. Pain sensitivity and temperament in extremely low-birth-weight premature toddlers and preterm and full-term controls. Pain 1994a;58:341-346.

Grunau RVE, Whitfield MF and Petrie JH. Children's judgements about pain at age 8-10 years: do extremely low birthweight (< 1000g) children differ from full birthweight peers? J Child Psychol Psychiatr 1998;39:587-594.

Grunau RVE, Whitfield MF, Petrie JH and Fryer LE. Early pain experience, child and family factors, as precursors of somatization: a prospective study of extremely premature and fullterm children. Pain 1994b;56:353-359.

Gunnar MR, Hertsgaard L, Larson M and Rigatuso J. Cortisol and behavioral responses to repeated stressors in the human newborn. Dev Psychobiol 1991;24:487-505.

Harrell FE, Jr., Lee KL, Califf RM, Pryor DB and Rosati RA. Regression modelling strategies for improved prognostic prediction. Stat Med 1984;3:143-152.

Johnston CC and Stevens BJ. Experience in a neonatal intensive care unit affects pain response. Pediatrics 1996;98:925-930.

McGrath PJ, Johnson G, Goodman JT, Schillinger J, Dunn J and Chapman JA. CHEOPS: A behavioral scale for rating postoperative pain in children. In: HL Fields, R Dubner and F Cervero (Eds.). Advances in Pain Research and Therapy, Vol. 9, Raven Press, New York, 1985. pp. 395-402.

Oberlander TF, Eckstein Grunau R, Whitfield MF, Fitzgerald C, Pitfield S and Saul JP. Biobehavioral pain responses in former extremely low birth weight infants at four months' corrected age. Pediatrics 2000;105:e6.

Porter FL, Grunau RE and Anand KJ. Long-term effects of pain in infants. J Dev Behav Pediatr 1999;20:253-261.

Ramsay DS and Lewis M. The effects of birth condition on infants' cortisol response to stress. Pediatrics 1995;95:546-549.

Ruda MA, Ling QD, Hohmann AG, Peng YB and Tachibana T. Altered nociceptive neuronal circuits after neonatal peripheral inflammation. Science 2000;289:628-631.

Taddio A, Katz J, Ilersich AL and Koren G. Effect of neonatal circumcision on pain response during subsequent routine vaccination. Lancet 1997;349:599-603.

Taylor A, Fisk NM and Glover V. Mode of delivery and subsequent stress response. Lancet 2000;355:120.

van der Hoorn FA, Boomsma F, Man in 't Veld AJ and Schalekamp MA. Determination of catecholamines in human plasma by high-performance liquid chromatography: comparison between a new method with fluorescence detection and an established method with electrochemical detection. J Chromatogr 1989;487:17-28.

Varni JW, Thompson KL and Hanson V. The Varni/Thompson Pediatric Pain Questionnaire. I. Chronic musculoskeletal pain in juvenile rheumatoid arthritis. Pain 1987;28:27-38.

Whitfield MF and Grunau RE. Behavior, pain perception, and the extremely low-birth weight survivor. Clin Perinatol 2000;27:363-379.

Chapter 7

Pain instruments in preverbal infants: An overview of the period 1995 to October 2000

7.1 Definition of pain

The debates concerning the IASP pain definition have been initiated by the fear that those without the ability of self-report, would be undertreated as a consequence (Anand and Craig, 1996). Although self-report is generally considered the only 'gold standard' of pain assessment (McGrath and Unruh, 1994), others consider observational pain assessment in preverbal infants just as valid (Anand and Craig, 1996). However, the limitations of observation pain assessment can not be neglected. In this respect, as early as 1974 Huskisson stated truthfully:"It is difficult to accept that an observer, no matter how experienced, could ever measure another person's pain (Huskisson, 1974).
Anyhow, there was a need for a pain instrument in preverbal infants, either as a 'gold standard' or 'silver' one.
The developments with regard to pain assessment instruments of the past five years will be described in this chapter.

7.2 Pain instruments

Approximately thirty pain instruments for either procedural or postoperative pain in neonates and infants have been developed in the past fifteen years.
In general, pain instruments can be either unidimensional or multidimensional. Unidimensional pain instruments focus on one type of indicator (e.g. behaviour) or a unitary dimension of pain assessment. Examples of contrasting unidimensional instruments are the Visual Analogue Scale (VAS) (Huskisson, 1974) and the Neonatal Facial Coding System (NFCS) (Grunau and Craig, 1990). The VAS gives an estimate of the infant's pain intensity by a professional or parent and is widely used because of its great feasibility. The NFCS evaluates pain, on the basis of facial activity shown on videotape. Multidimensional pain instruments combine behavioural, physiological, and sometimes contextual indicators of pain. Examples of multidimensional instruments are the COMFORT scale (Ambuel et al., 1992; Dijk van et al., 2000) and the PIPP (Stevens et al., 1996). Multidimensional instruments are generally preferred because they are considered to best represent the complexity of pain (Franck et al., 2000; Huijer Abu-Saad et al., 1998; Stevens et al., 2000).

Table 1a lists the published unidimensional pain instruments from the period 1995 until October 2000 that comprise behavioural items only, Table 1b the multidimensional

instruments. The grey areas show which indicators are incorporated and the number of check marks indicates the number of items that are related to that indicator.

The response categories of the separate items range from two to six in the different instruments. For example the item 'crying' may have two response categories (0= no cry, 1= moan, scream) or five (0=no cry, 1=whimpering, groaning, fussiness, 2=intermittent crying, 3=sustained crying, 4=screaming). In most instruments the number of response categories and scoring range is equal for all items, suggesting that all items are considered equally important. Other instruments have varying numbers of response categories and in addition a different weighting of items. Extra weight can also be given to an indicator when several items for the same indicator are used. For instance, the POCIS (Boelen et al., 1999) has three items on body movement: Torso, Arms/fingers and Legs/toes, and the PIPP has three on facial activity (Stevens et al., 1996). In general, the authors of the instrument give no explanation for their choices of response categories and weighting of the items. Tables 1a and 1b further show that the various instruments use the same indicators. This is further emphasised by the knowledge that some instruments have been derived from existing instruments which are not that much different either. For instance, the MBPS (Taddio et al., 1995) was derived from the Behavioural Pain Scale (Robieux et al., 1991) with minor adaptations in response categories. The POCIS was adapted from both the CHEOPS (McGrath et al., 1985) and the NIPS (Lawrence et al., 1993) and has items from both instruments but with dichotomised response categories. The MIPS (Buchholz et al., 1998) was derived from the Clinical Scoring System (Attia et al., 1987; Barrier et al., 1989) with only a more detailed item for facial expression. And finally, the SUN (Blauer and Gerstmann, 1998) strongly resembles the COMFORT scale (Ambuel et al., 1992) with the exclusion of one item.

The design of the procedural and postoperative pain instruments show no distinct differences. However, while acute procedural pain, as caused by the heelstick, will result in an immediate pain response, observation of subacute postoperative pain is more complex, because the pain response may be diminished due to duration of the pain or lack of energy (Barr, 1998). Furthermore pain management has fortunately shifted from pain treatment towards pain prevention, resulting in less vehement pain responses after surgery in general.

7.3 Physiological indicators

The additional value of physiological indicators for postoperative pain assessment is under debate for the following reasons. Firstly, they are generally considered not specific for pain. Other reasons for fluctuations in e.g. heart rate (HR), blood pressure (BP), and oxygen saturation (O_2 sat), are e.g. blood loss, fluid intake, body temperature, and medical interventions, all of which are variable after surgery. Secondly, from a practical point of view, these indicators should be read from a monitor, which is less feasible after day surgery. Manually measuring heart rate and blood pressure could induce fear, which might increase the unreliability of such pain indicators.
Despite their drawbacks, HR, MAP, BP and O_2 sat have been implemented in many postoperative pain instruments. Three studies however concluded that physiological pain indicators are less valuable as postoperative pain measures (Buchholz et al., 1998; Buttner and Finke, 2000; Dijk van et al., 2000).

7.4 Psychometric testing

Several reviews evaluate the various instruments (Franck et al., 2000; Franck and Miaskowski, 1997; Hain, 1997; Huijer Abu-Saad et al., 1998; McGrath, 1998; Morton, 1997; Stevens, 1998; Stevens et al., 2000). They documented the reliability and validity testing of the different instruments. The various authors further conclude that most instruments have been used for research purposes only and that the feasibility (i.e. how easy it is to use) and the clinical utility (of which usefulness is the key element) of the instruments are neglected areas.
Two instruments from Table 1b, NPAT (Friedrichs et al., 1995) and DSVNI (Sparshott, 1996), have not been tested on psychometric properties. Two others, the CRIES (Krechel and Bildner, 1995) and LIDS (Horgan and Choonara, 1996) used 24 and 16 neonates respectively, which are considered small samples for psychometric testing.

Interrater reliability testing has been performed for most instruments from Table 1. However, some authors still use the Pearson correlation coefficient to estimate the agreement among raters (Buchholz et al., 1998; Buttner and Finke, 2000; Krechel and Bildner, 1995; Schultz et al., 1999). This may be misleading because correlation coefficients only reflect relative positions of scores, but do not reflect differences in

absolute levels of scores. While the correlation may be high, scores of different observers may show large differences at the same time (Bland and Altman, 1986). A more appropriate statistic tool is the weighted kappa coefficient, which provides for varying gravity of disagreement and is chance corrected (Cohen, 1968).

Validity testing of most instruments was performed by one or more of the following procedures: 1) concurrent validity by correlating the instrument with other tools or instruments, 2) criterion validity by comparing the instrument with the global rating of an expert, 3) construct validity by comparing pain scores between pre and post procedure or between pre and post analgesic medication, 4) factorial validity by determining the internal structure using explorative factor analysis or principal component analysis.
Surprisingly, none of the psychometric articles used structural equation modelling (SEM) that for many reasons has become popular in the psychometric area. SEM enables one to simultaneously estimate the validity (complexity of the structure and concurrent validity) and reliability (internal consistency and stability) of instruments when repeated measurements have been collected (Jöreskog and Sörbom, 1993). A further advantage of SEM is that relationships among factors are free of measurement error because measurement error has been removed. Another advantage (over explorative factor analysis) is the possibility to test the adequacy of a model and to compare the fit indices between models.

7.5 Things to be done

One problem with implementing pain instruments is perhaps the call from the working-floor for cutoff points to distinguish between pain and non-pain states, so as to be able to guide analgesic treatment. Some researchers are reluctant to give general cutoff points, which is understandable considering the individual differences in pain expression and circumstances. Stevens (2000) suggests to use previous observations and pain scores of a particular infant to compare that infant's current pain score with.
Cutoff scores should be further evaluated in clinical practice on their usefulness in pain treatment.
Another unsolved problem is how to differentiate between pain and other forms of distress, such as anger, anxiety, or stress in preverbal infants. The wish to distinguish these is dictated by the need of adequate treatment and requires further research.

7.6 Overall conclusion

The steady flow of new pain instruments reflects the dissatisfaction with the available ones.
However, it seems unnecessary to develop new pain instruments unless new indicators for pain are added. The available pain instruments should be tested in different environments, for instance day-care surgery, general ward, intensive care environment, and recovery, and in other pain situations, for instance necrotizing enterocolitis or stressful situations such as suctioning. This in order to obtain more information about the usefulness of the various instruments in different settings.
Pediatric pain assessment is not routinely performed in most clinical settings (Franck et al., 2000) and will be slowed down if one has to choose between numerous pain instruments.

Table 1a behavioural pain instruments (with contextual indicators in some) developed or validated since 1995 with checkmarks representing included items.

Instrument, (author, year)	Age groups	Pain situation	Behavioural indicators				Contextual indicators
			Facial	Cry	Movements or activity [1]	Muscle tone	
MBPS [2], (Taddio et al., 1995)	4 to 6 months	Procedural	✓	✓	✓		
LIDS, (Horgan and Choonara, 1996)	Neonates	Postop	✓	✓✓	✓✓✓	✓	Sleep pattern /amount
FLACC, (Merkel et al., 1997)	2 months to 7 year	Postop	✓	✓	✓✓		Consolability
MIPS[3], (Buchholz et al., 1998)	Infants 4 to 30 weeks	Postop	✓	✓	✓✓✓	✓	Sleep preceding hour, response to stimulation, sucking, consolability, sociability
CHIPPS, (Buttner and Finke, 2000; Buttner et al., 1998)	Neonates and infants to 5 years	Postop	✓	✓	✓✓✓		

[1] This indicator represents either subtle (flexion of fingers and toes) or more gross body movements
[2] A zero MBPS score is not neutral but positive (e.g. smiling and playing).
[3] High scores indicate no pain or more comfort
Grey areas indicate inclusion of indicator in pain instrument
Abbreviations: MBPS, Modified Behavioral Pain Scale; LIDS, Liverpool Infant Distress Score; FLACC, an acronym for Face, Legs, Activity, Cry, and Consolability; MIPS, Modified Infant Pain Scale; CHIPPS, Children's and Infants' Postoperative Pain Scale.

Table 1b Multidimensional pain instruments (with contextual indicators in some) developed or validated since 1995 with checkmarks representing included items.

Instrument, (author, year)	Age groups	Pain situation	Behavioural indicators				Physiological indicators	Contextual indicators
			Facial	Cry	Movements or activity	Muscle Tone		
CRIES, (Krechel and Bildner, 1995)	Neonates> 32 weeks	Postop	✓	✓			HR/BP Requiring O_2	Sleepless
NPAT, (Friedrichs et al., 1995)	Neonates	Procedural		✓	✓		HR, RR, BP, O_2 sat	Behavioural state
PIPP, (Ballantyne et al., 1999; Stevens et al., 1996)	Premature and term neonates	Procedural	✓✓✓				HR, O_2 sat	Behavioural state, gestational age
DSVNI, (Sparshott, 1996)	Ventilated Neonates	Procedural	✓		✓		HR, BP, O2 sat, colour, body temperature	
SUN, (Blauer and Gerstmann, 1998)	Neonates	Procedural	✓	✓	✓	✓	HR, MAP, breathing	Behavioural state
PEPPS[1], (Schultz et al., 1999)	12 to 24 months	Postop	✓	✓	✓		HR	Consolability, sociability, sucking/feeding
POCIS, (Boelen et al., 1999)	1 to 4 year	Postop	✓	✓	✓✓✓		breathing pattern	Arousal
COMFORT, (Dijk van et al., 2000)	0 to 3 year	Postop	✓	✓	✓	✓	HR, MAP, respiratory response	Alertness, calmness

[1] High scores indicates no pain or more comfort

Abbreviations: CRIES, an acronym for Crying, Requires increased oxygen administration, Expression, Sleeplessness; NPAT, Neonatal Pain Assessment Tool; PIPP, Premature Infant Pain Profile; DSVNI, Distress Scale for Ventilated Newborn Infants; SUN, Scale for Use in Newborns; PEPPS, Pre-Verbal, Early Verbal Pediatric Pain Scale; MIPS, modified Infant Pain Scale; POCIS, Pain Observation Scale for Young Children,

HR= Heart rate, BP=Blood Pressure, RR= Respiratory Rate, MAP= Mean Arterial Pressure, O_2 sat= Oxygen saturation

7.7 References

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

Anand KJS and Craig KD. New perspectives on the definition of pain. Pain 1996;67:3-6.

Attia J, Amiel-Tison C, Mayer MN, Shnider SM and Barrier G. Measurement of postoperative pain and narcotic administration in infants using a new clinical scoring system. Anesthesiology 1987;67:A532.

Ballantyne M, Stevens B, McAllister M, Dionne K and Jack A. Validation of the premature infant pain profile in the clinical setting. Clin J Pain 1999;15:297-303.

Barr RG. Reflections on measuring pain in infants: dissocation in responsive systems and "honest signalling". Arch Dis Child Fetal Neonatal Ed 1998;79:F152-F156.

Barrier G, Attia J, Mayer MN, Amiel-Tison C and Shnider SM. Measurement of post-operative pain and narcotic administration in infants using a new clinical scoring system. Intensive Care Med 1989;15:s37-s39.

Bland JM and Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;:307-310.

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

Boelen WJC, Scheffer E, Haan de RJ and Groot de CJ. Clinimetric evaluation of the pain observation scale for young children in children aged between 1 and 4 years after ear, nose, and throat surgery. J Dev Behav Pediatr 1999;20:14-19.

Buchholz M, Karl HW, Pomietto M and Lynn AM. Pain scores in infants: a modified infant pain scale versus visual analogue. J Pain Symptom Manage 1998;15:117-124.

Buttner W and Finke W. Analysis of behavioural and physiological parameters for the assessment of postoperative analgesic demand in newborns, infants and young children: a comprehensive report on seven consecutive studies. Paediatr Anaesth 2000;10:303-318.

Buttner W, Finke W, Hilleke M, Reckert S, Vsianska L and Brambrink A. [Development of an observational scale for assessment of postoperative pain in infants]. Anasthesiol Intensivmed Notfallmed Schmerzther 1998;33:353-361.

Cohen J. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. Psych Bull 1968;70.

Dijk van M, Boer de JB, Koot HM, Passchier J, Tibboel D and Duivenvoorden HJ. The reliability, stability and validity of the COMFORT scale as a postoperative pain instrument in 0 to 3-year-old infants. Pain 2000;84:367-377.

Franck LS, Greenberg CS and Stevens B. Pain assessment in infants and children. Pediatr Clin North Am 2000;47:487-512.

Franck LS and Miaskowski C. Measurement of neonatal responses to painful stimuli: a research review. J Pain Symptom Manage 1997;14:343-378.

Friedrichs JB, Young S, Gallagher D, Keller C and Kimura RE. Where does it hurt? An interdisciplinary approach to improving the quality of pain assessment and management in the neonatal intensive care unit. Nurs Clin North Am 1995;30:143-159.

Grunau RVE and Craig KD. Facial activity as a measure of neonatal pain expression. Vol. 15, Raven Press, New York, 1990. pp. 147-155.

Hain RD. Pain scales in children: a review. Palliative Medicine 1997;11:341-350.

Horgan M and Choonara IA. Measuring pain in neonates: an objective score. Paediatr Nurs 1996;8:24-27.

Huijer Abu-Saad H, Bours GJJW, Bonnie S and Hamers JPH. Assessment of pain in the neonate. Semin Perinatol 1998;22:402-416.

Huskisson EC. Measurement of pain. Lancet 1974;2:1127-1131.

Jöreskog KG and Sörbom D. LISREL 8: Structural equation modeling with the SIMPLIS command language, Scientific Software International, Chicago, 1993.

Krechel SW and Bildner J. CRIES: a new neonatal postoperative pain measurement score. Initial testing of validity and reliability. Paediatr Anaesth 1995;5:53-61.

Lawrence J, Alcock D, McGrath PJ, Kay J, Brock MacMurray S and Dulberg C. The development of a tool to assess neonatal pain. Neonatal Netw 1993;12:59-66.

McGrath PJ. Behavioral measures of pain. In: GA Finley and PJ McGrath (Eds.). Progress in pain research and management, Vol. 10, IASP Press, Seattle, 1998. pp. 83-102.

McGrath PJ, Johnson G, Goodman JT, Schillinger J, Dunn J and Chapman JA. CHEOPS: A behavioral scale for rating postoperative pain in children. In: HL Fields, R Dubner and F Cervero (Eds.). Advances in Pain Research and Therapy, Vol. 9, Raven Press, New York, 1985. pp. 395-402.

McGrath PJ and Unruh AM. Measurement and assessment of paediatric pain. In: PD Wall and R Melzack (Eds.). Textbook of pain., Churchill Livingstone, Edinburgh, 1994. pp. 303-313.

Merkel SI, Voepel-Lewis T, Shayevitz JR and Malviya S. The FLACC: A behavioral scale for scoring postoperative pain in young children. Pediatr Nurs 1997;23:293-297.

Morton NS. Pain assessment in children. Paediatr Anaesth 1997;7:267-272.

Robieux I, Kumar R, Radhakrishnan S and Koren G. Assessing pain and analgesia with a lidocaine-prilocaine emulsion in infants and toddlers during venipuncture. J Pediatr 1991;118:971-973.

Schultz AA, Murphy E, Morton J, Stempel A, Messenger-Rioux C and Bennett K. Preverbal, Early Verbal Pediatric Pain Scale (PEPPS): development and early psychometric testing. J Pediatr Nurs 1999;14:19-27.

Sparshott MM. The development of a clinical distress scale for ventilated newborn infants. J Neonatal Nursing 1996;April:5-11.

Stevens B. Composite measures of pain. In: GA Finley and PJ McGrath (Eds.). Progress in pain research and management, Vol. 10, IASP Press, Seattle, 1998. pp. 161-178.

Stevens BJ, Johnston CC and Gibbins S. Pain assessment in neonates. In: KJS Anand, BJ Stevens and PJ McGrath (Eds.). Pain in neonates, 2nd revised and enlarged edition, Vol. 10, Elsevier, 2000. pp. 101-134.

Stevens BJ, Johnston CC, Petryshen P and Taddio A. Premature Infant Pain Profile: development and initial validation. Clin J Pain 1996;12:13-22.

Taddio A, Nulman I, Koren BS, Stevens BJ and Koren G. A revised measure of acute pain in infants. J Pain Symptom Manage 1995;10:456-463.

Chapter 8 Part I

Discussion and future directives: Pain unheard?

8.1 Introduction

The aims of this thesis were 1) to analyse the psychometric properties of the COMFORT scale as a postoperative pain instrument in 0 to-3- year old infants and 2) to compare the efficacy of continuous intravenous morphine and intermittent morphine administration postoperative pain in this sample. A third aim was to investigate the relationship between past (painful) experiences and the pain and stress response after surgery.
The purpose of this final chapter is to integrate the results from the previous chapters and give suggestions for future research. In the second part of this chapter we will summarise our experiences during the trial.

8.2 Reliability and validity of the COMFORT scale

In chapter 2 the interrater reliability of 39 nurses in our study was described with linearly weighted kappas, and was generally shown to remain good. This finding showed that it was possible to train the nurses to observer pain-related behaviour in a reliable way.
The COMFORT scale scores were analysed with structural equation modelling (SEM) using three repeated measurements, 3, 6 and 9 hours postoperative. In this way a model was fitted that incorporated significant parameter estimates, stability over time, and congruent (concurrent) validity of the COMFORT factors with VAS pain. The COMFORT scale was best represented by three factors, one for the behavioural items, and two for heart rate (HR) and mean arterial pressure (MAP) separately. HR and MAP showed limited association with the latent variables COMFORT 'behaviour' and VAS pain, but were stable across time. The stability of the latent variables COMFORT 'behaviour' and VAS pain was distinct though not high. All factor loadings in the selected model were significant and in the 0.50-0.80 range. The correlation between the latent variables VAS pain and COMFORT 'behaviour' was high, indicating congruent validity.

We found the COMFORT 'behaviour' to be a reliable and valid instrument to assess postoperative pain in neonates and toddlers. With leaving the physiological items aside, the scale shows similarity to other behavioural instruments. However, the fine-graded response categories and the inclusion of an item for ventilated infants makes it a useful tool for the intensive care environment. Further testing will have to show if the COMFORT scale is valid and feasible in different settings (e.g.day-care surgery) and for different groups (e.g. premature neonates, children after minor surgery).

8.3 The Visual Analogue Scale

The Visual Analogue Scale (VAS) has been often used as a self-report tool to quantify pain intensity (Scott and Huskisson, 1976). Because of its good psychometric properties and ease of use as a self-report measure, the use of the VAS was extended to observational pain assessment. In this application an observer, e.g. a nurse, rates the intensity of the pain experienced by a preverbal infant. In our study the VAS was used to estimate the concurrent validity of the COMFORT scale. It has been used for this application in several other studies as well (Lawrence et al., 1993; McGrath et al., 1985; Taddio et al., 1995; Tarbell et al., 1992). Furthermore, in pediatric samples it has been applied as a criterion for cutoff points (Berde et al., 1991; Buchholz et al., 1998;) to discriminate between pain and no pain states and to guide analgesic treatment. We reviewed the psychometric properties of the observational VAS in the English literature. We concluded that the available psychometric findings were promising but that further work needed to be done on intraobserver reliability, sensitivity to change and optimal cutoff points. Because behavioural pain ratings may express other forms of distress in some infants, the observational VAS comprises a global rating which may account for additional knowledge on individual variations in pain sensitivity, idiosyncratic behaviours, and situational influences. Preferably the observational VAS is only used after a observation period and when substantial experience with pain and pain assessment in preverbal infants has been gained.
The application of solely the observational VAS is not recommended unless the score has been motivated by objective arguments.

8.4 Association of physiological and behavioural pain indicators

Although the structural equation modelling with the LISREL statistical program had shown a limited association between the behavioural and physiological items of the COMFORT, the question was if this was partly caused by the format of the items, which reduces the scores to five response categories. Therefore the six HR and six MAP values at each assessment were used to determine the relationship between COMFORT 'behaviour' and the mean and standard deviation of HR and MAP at repeated measurements.
The correlations (within subjects) between these physiological measures and the COMFORT 'behaviour' were moderate, and ranged from 0.37 to 0.49. Furthermore, there

were large interindividual differences in behaviour-physiology correlations. In addition, the influence of background characteristics, physical condition and pain-related characteristics on the behaviour-physiology correlations was determined. Neonates showed lower behaviour-physiology correlations than the older infants did and pain characteristics significantly predicted some of the behavior-physiology correlations, suggesting that the behaviour-physiology correlations increase with increasing pain. The behaviour-physiology correlations were not greatly affected by physical condition. These findings suggest that physiological pain indicators are not really valid for postoperative pain. However, their validity and usefulness have to be determined for situations in which behavioural pain assessment is not possible (e.g. ventilated infants receiving neuromuscular blockers) or less reliable (e.g. in very ill children).

8.5 Postoperative analgesic treatment

Intravenous morphine infusion is still one of the preferred analgesics for pain treatment after major surgery (Kart et al., 1997; Truog and Anand, 1989). When our study started in 1995, there was little knowledge on the efficacy of different routes of postoperative morphine administration in young infants. Several studies had shown that intravenous continuous infusion was preferred over intramuscular injections (Bray, 1983; Hendrickson et al., 1990). Considering the additional pain of intramuscular injection, this result was not surprising.
A recent study compared the efficacy and safety of continuous morphine infusion to a targeted morphine concentration of 20 ng ml^{-1} with those of intermittent bolus doses as needed (0.05 mg kg^{-1} ordered every 1-2h) in 0 to–1-year old infants (Lynn et al., 2000). Intravenous continuous morphine infusion resulted in better analgesia as scored with the MIPS than in the bolus group. However this better result corresponded with higher total morphine dosing in the infusion group. The outcomes of this study were further blurred by the fact that the trial was not blinded and that additional acetaminophen (orally or rectally) was given.

In our study the continuous (CM) and intermittent morphine administration (IM) routes were equally effective in reducing postoperative pain. Only for the 1 to 3-year old infants, the CM was somewhat more effective. This trial also showed inter-individual variability in analgesic need after major surgery, which however was not explained by surgery-related

characteristics. The neonates were least painful, and most of them were satisfied with 10 μg/kg/h. The young infants of 1 to 6 months of age were most painful. In both morphine conditions extra morphine was given. This resulted in a median dosage of 10.8 μg/kg/h (interquartile range 10.0 to 12.2) in the CM and a median dosage of 10.4 μg/kg/h (interquartile range 10.0 to 12.5in the IM group, which is not significantly different. Fifteen infants were excluded because they received additional other analgesics or sedatives. The inter-individual variability in morphine dosage after surgery remained largely unexplained.
The surgical stress score (SSS) was not significantly related to postoperative pain. This was perhaps partly due to the limited variability in SSS scores in this sample.

The most unwanted adverse effect of morphine infusion in infants is ventilatory depression, which is the reason why many physicians have been hesitant to use it (Purcell-Jones et al., 1988). The literature reports two methods to determine ventilatory depression. Firstly, by assessing clinical signs of ventilatory depression, e.g. shallow or slow breathing, apnoe, requiring oxygen; secondly, by blood gas analysis showing elevated $PaCO_2$ levels (Lynn et al., 2000; Lynn et al., 1993).
In our study only one toddler showed clinical signs of ventilatory depression and was excluded. The results of the blood gas analyses in relation to morphine plasma levels in our study will be described elsewhere.

A considerable part of our study sample required extra morphine. This suggests that pain prevention was not always achieved. For that reason, more double-blind placebo controlled randomised trials to achieve tailored pain treatment have to be developed (van Lingen, 2000).
In addition to the stressfulness of surgery, a score designed to rank the painfulness of surgical procedures could be useful in the future.

8.6 Long term consequences of neonatal pain

Repeated painful procedures and prolonged stress in preterm neonates during NICU stay, as well as their immature pain system, are reflected in a low pain threshold, hypersensitivity, and prolonged periods of windup (Anand, 2000a; Fitzgerald et al., 1989).

This knowledge initiated interest in the subsequent and long-term consequences of neonatal pain.

The possible short and long-term consequences of early pain experiences received increased attention within the past five years. However, these studies in infants showed inconsistent results. In chapter 6 the available studies on long term consequences of early pain are described as well as the reviews in this area.
As these studies and reviews contain incongruous elements, in comparing them one runs the risk of comparing apples and oranges. First, they differ considerably in time span between pain procedure(s) and the assessment of long-term effects. It is therefore important to differentiate between subsequent pain response, as, for instance, described in the study of (Johnston and Stevens, 1996), and long-term consequences, as for example in the study of Grunau in 8 to10 year old children (Grunau et al., 1998). Secondly, prior pain experiences appear to increase subsequent behavioural response to pain in healthy term babies, whereas in premature babies it appears to diminish these responses (Anand, 2000b; Grunau, 2000; Whitfield and Grunau, 2000).
Thirdly, past pain experiences vary in number and intensity. A history of intensive care stay with many daily painful procedures differs from that of 'just' a circumcision or a stressful delivery.

Furthermore, studies which evaluate the long-term consequences of pain on behavioural pain responses differ in design from those on physiological stress responses, as expressed e.g. by cortisol levels, which makes comparisons difficult. Similar designs are essential when studying pain-stress relationship (Aynsley-Green, 1996).

In Chapter 6 the influence of previous hospital experiences on postoperative pain and stress responses was examined. Previous hospital experiences comprised: number of procedures under anaesthesia, number of days receiving morphine infusion, number of days mechanical ventilation and number of days hospitalised. Multiple regression analyses were performed to estimate the correlation between hospital history and current postoperative pain and stress, controlling for patient characteristics and surgical characteristics.

Previous hospital experiences proved not predictive for these outcome variables. Several explanations for these findings were considered possible. Firstly, the long-term behavioural consequences of neonatal pain may not be clinically significant (Grunau, 2000; Whitfield and Grunau, 2000). However, we should be cautious to suggest that long-term consequenses of neonatal pain are relatively mild, because a recent animal study showed that localised inflammation during the neonatal period in rat pups permanently alters the neuronal circuits that process pain in the spinal cord (Ruda et al., 2000).
Second, the shift towards pain prevention at our hospital which might explain the limited impact of previous hospital experiences which are then less painful than we had expected. Thirdly it may be caused by the heterogeneity of the sample size or variables which are omitted in the analyses e.g. painful procedures at the NICU not recorded in the medical records.

Future research should incorporate follow-up studies of large cohorts of neonates, both term and premature. A prospective design will enable to carefully document all invasive painful procedures which take place at NICUs, including attempted procedures that are usually not documented. In addition, temperament and mother-child interaction (e.g. period parent visits NICU, period parent has bodily contact) could be assessed across time as well, in order to obtain a greater insight in the interplay of those variables with the painful experiences.

8.7 Discussion Part II

Things learned so far[1] about performing clinical research

Introduction

During the past five years a multidisciplinary team at the Sophia Children Hospital performed a double-blind randomised trial in 0-to-3-year-old infants. The team incorporated a pediatric intensivist, an anaesthesiologist, two medical psychologists, a health scientist, a developmental psychologist and a psychologist/methodologist and pediatric intensive care nurses. We entered 204 infants who were admitted to the PSICU after abdominal or thoracic surgery. The trial was the first study on pain assessment in our hospital. Therefore, some of the flaws in our study may have been caused by inexperience. On the other hand, some obstacles during the study were unavoidable for organisational reasons. This part of the discussion deals with these problems in chronological order, and suggests solutions that may be useful for future research.

8.8 Interdependent research questions

The trial was aimed at answering two research questions.

1. How reliable, valid and feasible is the multidimensional COMFORT scale to assess postoperative pain in children from 0-3 years of age?
2. What is the difference between intermittent morphine administration and continuous intravenous morphine in terms of quality and effectiveness of analgesia for postoperative pain?

For answering the second question we had to use the instruments under investigation. This dependency is not commendatory, because if both COMFORT and VAS should prove to be invalid, the second question could not be answered unequivocally. For a non-significant difference in pain scores between the conditions might then also be due to the invalidity of the pain instrument. Therefore, in this trial psychometric evaluation of the COMFORT scale was started as soon as a substantial number of infants had been included.

[1] This title was inspired by an article of the eminent Jacob Cohen (1990)

8.9 Treatment violations

Inclusion and exclusion criteria determine to which population the results of the study can be generalised. In our study, an exclusion criterion was the '*use of co-medication influencing the measured potency of morphine*', for example other analgesic or sedatives. When infants received analgesics or sedatives **before** surgery, these were stopped the evening before surgery. Despite the protocol, fifteen infants (7%) received other analgesics (APAP or diclofenac) or sedatives (midazolam or) next to morphine **after** surgery during the trial period. On the one hand, this may have been due to their 'restlessness', which could only be adequately treated with sedatives. On the other hand, we speculate that some physicians and nurses may prefer certain combinations of analgesic treatment that are in contrast with the trial medication. The only way to overcome these problems to some extent is by ongoing communication with participating professionals.
For the analyses to compare the efficacy of the two routes of morphine administration, these fifteen infants were excluded. For psychometric testing of the COMFORT scale these cases could remain included.

8.10 Omitted but relevant variables

Although new ideas may come up during any study, one should not change too much during a trial. And as Jacob Cohen would say, 'less is more' (Cohen, 1990). Not only for statistical or methodological reasons is it unwise to change too much but also because all participants have to be informed and changes may increase errors. These new ideas are best converted into new research proposals. For example, as the study progressed, nurses started to discriminate between pain and other forms of distress (e.g. anxiety, anger, sadness and noise). Sometimes these observations were recorded on the case record form. In future studies they could be included as a standard at each pain assessment, in order to obtain more insight in the relation between pain and other forms of distress.

Missing data
In our design the arterial lines were placed during induction, to avoid unnecessary distress. According to the protocol, a failing arterial line should be replaced. However, as this was distressing for the infants, nurses in the PSICU disapproved to have it replaced for research reasons only. After consultation with the nurses involved, we decided to only reinsert a line

when necessary for the treatment of the child. Another option was to perform the usual blood sampling related to the treatment of the child at the time of our blood sampling. In that way no extra venapunctures were required. In fifteen cases (7%) the arterial line failed at a certain point during the trial, which led to missing data for HR and mean arterial blood pressure, and missing blood analyses.

Morphine plasma levels were analysed in an external laboratory. As analysis for only a few samples is quite expensive, a large number of samples were saved until analysis. In 43 cases no morphine plasma levels were detected in blood presumably taken 5 minutes after the morphine loading dosage, which was impossible. We found out that the tube had been interchanged with another (which contained blood taken after surgery but prior to the loading dosage). The solution simply was found in better labelling of the tubes.

Another problem was encountered with the morphine-M3-glucuronides. In 187 of 954 samples (from 42 cases) the M3-glucuronides could not be determined due to unexplained 'disturbances'. The laboratory had no explanations for these findings as well. Therefore, it would be preferable to determine laboratory results at shorter time frames in order to detect these errors earlier on.

Inclusion of older infants

Our sample included 50% fewer infants of 6 months to 3 years than anticipated. This was caused by the following reasons. Firstly, 'rooming in' of the parents involved was not possible at the PSICU, which resulted in 8 documented refusals for older infants who otherwise would have been transferred to the normal care surgical unit. Secondly, some infants were operated on as a neonate and had therefore been included in the trial at an earlier age, which led to exclusion at the next surgery. And thirdly, in general, older children undergo fewer surgical interventions than younger ones do. The stratification in four age groups in this study, was chosen to compare physiological responses i.e. hormonal, metabolic and morphine plasma levels across age groups. However, we would have preferred to include more infants over 6 months because pain behaviour may be developmentally sensitive.

A threat for the generalizability of a study is selection bias (Keirse and Hanssens, 2000), which may result from unknown differences between refusals and participants. Because

parental refusal was not always properly recorded in our study, some selection bias may have occurred. In future the recording of the number and reasons for refusal should be better recorded.

8.11 Communication

Full participation of the PSICU nurses and staff was crucial in our study, because they performed pain assessments and blood sampling. Involving the nurses rather than researchers during the study for pain assessment, may facilitate implementation of a pain instrument at a later stage.
We organised meetings to inform nurses about the study progress, which gave them the opportunity to raise questions. Their critical comments about the trial were taken seriously and we tried to find solutions for the points raised. For instance, the attending nurses did not know how much extra morphine very painful infants were allowed to receive within a certain time frame. For that reason, we created an algorithm visualising the correct way to administer extra morphine. This was considered helpful. Next to informal consults, we had monthly meetings with representatives of all disciplines, and also used the PSICU standard monthly newsletter as a means of communication.

Many workers in the operation room, the department of anaesthesiology, and the laboratories in our hospital who had no direct benefit from this study were willing to co-operate. At the end of this study we gave all participating nurses from the PSICU a small present and we had pastry for the surgical and anaesthesiology ward to show our gratitude for their participation. Although this may seem trivial, we think that some tangible form of gratitude is deserved and appreciated.

8.12 Influence of the trial on attitudes toward pain

This kind of research will inevitably influence staff mentality with regard to pain. For instance, we trained nurses to assess pain and we gave pain medication according to a protocol. We had the impression that caregivers became more focused on the efficacy of pain treatment during the trial. This resulted in a critical attitude towards the protocol which stimulated discussion among caregivers and researchers.

8.13 Success of blinding

During the trial, nurses were prejudiced that intermittent morphine would result in more pain, in particular shortly before the next bolus injection. To check for adequate blinding, a question was added to the case record form, asking at each pain assessment the attending nurse's opinion on the morphine condition. The closed question was: What treatment do you think this child receives? The possible answers were:
1) continuous, 2) intermittent or, 3) no idea.
For infants in the continuous morphine condition, 43% of the opinions were correct, 21 % opted for intermittent, and in 35% one had no idea. For infants in the intermittent morphine condition 22% of the opinions were correct, 41% opted for continuous, and in 37% one had no idea. This confirmed that blinding was successful because the nurse's prediction of morphine condition was worse than when they should have merely guessed.

8. 14 Recommendations for future research

- Efficient and effective communication at all levels is a priority during the entire study. Take comments seriously and keep others well informed.
- Good monitoring during the trial is essential but requires a researcher who has enough time to do this job.
- A pilot study may unravel problems or errors which may be solved before the actual trial start.
- Some problems can not be foreseen or prevented. Cope with them and accept them.

8.15 Future and current pain-related research in the Sophia Children's Hospital

This trial did not stand on its own but is the first of a number of pain studies which will be performed in the Sophia Children's Hospital.
A double-blind randomised trial was completed in March 2000, comparing the efficacy of rectally and orally administered paracetamol in 40 infants after craniofacial surgery.
Currently, a double-blind randomised trial is carried out comparing the safety and efficacy of diclofenac and paracetamol after (adeno) tonsillectomy in 100 infants aged 3 to 12 years. In this trial both the COMFORT scale and POCIS are used to compare the usefulness of both instruments.

In January 2001 a double-blind randomised trial will be started to compare the efficacy of morphine with that of morphine/ paracetamol in 108 infants aged 0 to 1 year. In this study the COMFORT, POCIS, and VAS will be used for pain assessment.
In 1999 we started with implementing of the COMFORT scale in several surgical wards. Currently, we collect COMFORT scores before and after extra analgesia in order to estimate the sensitivity to change of the COMFORT scale. Furthermore, we use the COMFORT in combination with the VAS, and determined cutoff points for both. These guide current pain treatment in 0 to 3-year-old infants in the Pediatric Surgical Intensive Care Unit. To facilitate implementation of the COMFORT scale, a Dutch manual has been written which will be extended when new results related to its psychometric properties are available (see appendix).
Our research group will continue to develop research protocols related to pain assessment and pain treatment in order to minimise **unheard pain**.

8.16 References

Anand KJ. Effects of perinatal pain and stress. Prog Brain Res 2000a;122:117-129.

Anand KJ. Pain, plasticity, and premature birth: a prescription for permanent suffering? Nat Med 2000b;6:971-973.

Aynsley-Green A. Pain and stress in infancy and childhood--where to now? Paediatr Anaesth 1996;6:167-172.

Berde CB, Lehn BM, Yee JD, Sethna NF and Russo D. Patient-controlled analgesia in children and adolescents: a randomized, prospective comparison with intramuscular administration of morphine for postoperative analgesia. J Pediatr 1991;118:460-466.

Bray RJ. Postoperative analgesia provided by morphine infusion in children. Anaesthesia 1983;38:1075-1078.

Buchholz M, Karl HW, Pomietto M and Lynn AM. Pain scores in infants: a modified infant pain scale versus visual analogue. J Pain Symptom Manage 1998;15:117-124.

Cohen J. Things I have learned (so far). American Psychologist 1990;45:1304-1312.

Fitzgerald M, Millard C and McIntosh N. Cutaneous hypersensitivity following peripheral tissue damage in newborn infants and its reversal with topical anaesthesia. Pain 1989;39:31-36.

Grunau RVE. Long-term consequences of pain in human neonates. In: KJS Anand, BJ Stevens and PJ McGrath (Eds.). Pain in neonates, 2nd revised and enlarged edition, Vol. 10, Elsevier, 2000. pp. 55-76.

Grunau RVE, Whitfield MF and Petrie JH. Children's judgements about pain at age 8-10 years: do extremely low birthweight (< 1000g) children differ from full birthweight peers? J Child Psychol Psychiatr 1998;39:587-594.

Hendrickson M, Myre L, Johnson DG, Matlak ME, Black RE and Sullivan JJ. Postoperative analgesia in children: A prospective study of intermittent intramuscular injection versus continuous intravenous infusion of morphine. J Pediatr Surg 1990;25:185-191.

Johnston CC and Stevens BJ. Experience in a neonatal intensive care unit affects pain response. Pediatrics 1996;98:925-930.

Kart T, Christrup LL and Rasmussen M. Recommended use of morphine in neonates, infants and children based on a literature review: Part 2--Clinical use. Paediatr Anaesth 1997;7:93-101.

Keirse MJ and Hanssens M. Control of error in randomized clinical trials. Eur J Obstet Gynecol Reprod Biol 2000;92:67-74.

Lawrence J, Alcock D, McGrath PJ, Kay J, Brock MacMurray S and Dulberg C. The development of a tool to assess neonatal pain. Neonatal Netw 1993;12:59-66.

Lynn AM, Nespeca M, Bratton SL and Shen DD. Intravenous morphine in postoperative infants: intermittent bolus dosing versus targeted continuous infusions. Pain 2000;88:89-95.

Lynn AM, Nespeca M, Opheim KE and Slattery JT. Respiratory effects of intravenous morphine infusions in neonates, infants, and children after cardiac surgery. Anesth Analg 1993;77:695-701.

McGrath PJ, Johnson G, Goodman JT, Schillinger J, Dunn J and Chapman JA. CHEOPS: A behavioral scale for rating postoperative pain in children. In: HL Fields, R Dubner and F Cervero (Eds.). Advances in Pain Research and Therapy, Vol. 9, Raven Press, New York, 1985. pp. 395-402.

Purcell-Jones G, Dormon F and Sumner E. Paediatric anaesthetists' perceptions of neonatal and infant pain. Pain 1988;33:181-187.

Ruda MA, Ling QD, Hohmann AG, Peng YB and Tachibana T. Altered nociceptive neuronal circuits after neonatal peripheral inflammation. Science 2000;289:628-631.

Scott J and Huskisson EC. Graphic representation of pain. Pain 1976;2:175-184.

Taddio A, Nulman I, Koren BS, Stevens BJ and Koren G. A revised measure of acute pain in infants. J Pain Symptom Manage 1995;10:456-463.

Tarbell SE, Cohen TI and Marsh JL. The Toddler-preschooler postoperative pain scale: an observational scale for measuring postoperative pain in children aged 1-5. Preliminary report. Pain 1992;50:273-280.

Truog R and Anand KJ. Management of pain in the postoperative neonate. Clin Perinatol 1989;16:61-78.

van Lingen R. Pain assessment and analgesia in the newborn: an integrated approach, Erasmus University Rotterdam, Rotterdam, 2000, pp. 1-177.

Whitfield MF and Grunau RE. Behavior, pain perception, and the extremely low-birth weight survivor. Clin Perinatol 2000;27:363-379.

Summary / Samenvatting

Summary

Pain management, including pain assessment, is in particular important for infants who are hospitalised for a long period at a very young age and may undergo multiple painful procedures without adequate pain management. The past fifteen years there has been increased awareness that young infants experience pain. However, pain treatment was not immediately changed due to the fear for side effects of opioids and the lack of valid ways to assess pain. In order to improve pain assessment methods and analgesic treatment in young infants the current study was set up.

The specific aim of this study, which was sponsored by NWO (grant nr. 940-31-031), was to answer the following research questions:

- How reliable, valid, and feasible is the multidimensional COMFORT scale to assess postoperative pain in infants and toddlers 0-3 years of age?
- What is the difference between intermittent morphine administration and continuous intravenous morphine in terms of quality and effectiveness of analgesia for postoperative pain?

During data collection a third research question came up,

- Are the present postoperative pain and stress response related to past experiences with pain?

Within a prospective double-blind randomised trial at the Pediatric Surgical Intensive Care of the Sophia Children's Hospital, 204 infants were studied to answer the above questions. The infants were stratified into four age groups and randomised to intermittent or continuous morphine administration after major abdominal or thoracic surgery. Postoperative pain was assessed by means of two pain measurement instruments: the COMFORT scale and Visual Analogue Scale (VAS) every three hours for the first 36 hours postoperative. The COMFORT scale was chosen because it is a multidimensional instrument comprising both behavioural and physiological indicators of pain, which had been developed for the intensive care environment to assess distress /comfort in ventilated children. The VAS was chosen to determine concurrent validity of the COMFORT scale and as a criterion for extra pain medication.

Chapter 2 evaluates the psychometric properties of the COMFORT scale. Interrater reliability of the COMFORT items proved to be acceptable for all items. The repeated postoperative COMFORT scale scores and VAS pain scores from an initial sample of 158 neonates and infants aged 0-3-years were used to perform structural equation modelling (SEM). SEM was used to determine the psychometric qualities of the COMFORT scale.

SEM analyses showed that the internal structure of the COMFORT data was best represented by three latent variables: COMFORT 'behaviour' with loadings from the behavioural items and separate latent variables for 'Heart rate baseline' (HR) and 'Mean arterial blood pressure baseline' (MAP). Factor loadings of the items were invariant across time, indicating stability of the structure.
We found the COMFORT 'behaviour' to be a reliable and valid instrument to assess postoperative pain in neonates and toddlers.

Chapter 3 reviews the pediatric pain literature in the English language focusing on those studies that reported quantitative information on reliability and/or validity of the observation Visual Analogue Scale (VAS_{obs}). In this application an observer, e.g. a nurse, uses the VAS to rate the intensity of the pain experienced by another person. In addition, we looked for evidence on optimal cutoff points on the VAS_{obs} to discriminate between different pain states.
The psychometric properties of the VAS_{obs} were promising. Further work needs to be done on intraobserver reliability, sensitivity to change, and optimal cutoff points. We argued that the VAS_{obs} is a helpful tool next to a validated pain instrument because high behavioural pain ratings may in some children express fear, anxiety, or other forms of distress, which can be distinguished with the VAS_{obs}. A consideration is to not use it until substantial experience with pain and pain assessment in children of different ages has been gained, and adequate interobserver reliability has been proven. While most pain instruments are based on detailed behavioural observations, the global rating on the VAS_{obs} may account for additional knowledge on individual variations in pain sensitivity, idiosyncratic behaviours, and situational influences.

Although the SEM analysis had shown a limited association between the behavioural and physiological items of the COMFORT, the question was if this was partly caused by the format of the items, which reduces the original heart rate (HR) and mean arterial pressure (MAP) scores to five response categories. Therefore the six HR and six MAP values at each assessment were used to determine the relationship between COMFORT 'behaviour' and the mean and standard deviation of HR and MAP at repeated measurements. A second aim was to identify determinants predicting the level of correlation between COMFORT 'behaviour' and the physiological scores. The within-subject correlations, using the repeated measures of the postoperative pain assessments, were moderate and ranged from

0.37, to 0.49 for COMFORT 'behaviour' with the mean and variability of heart rate and mean arterial pressure.
The wide range in behaviour-physiology correlations reflected large individual differences. Neonates had lower behaviour-physiology correlations than other age groups. Pain characteristics significantly predicted the COMFORT 'behaviour'-HR/MAP correlations, suggesting that the behavioural-physiological correlations increase with increasing pain. The behaviour-physiology correlations were not greatly affected by physical condition. These findings suggest that physiological pain indicators are not really valid for postoperative pain.

Chapter 5 represents a study comparing the efficacy of 10 µg/kg/h morphine continuous IV infusion (CM) with 30 µg/kg morphine (IM) every three hours after major abdominal or thoracic surgery, in a subset of 181 infants aged 0 to 3 years. Efficacy was assessed with the COMFORT 'behaviour' and VAS every three hours in the first 24 hours after surgery. Random regression modelling was used to simultaneously estimate the effect of the route of morphine administration, actual morphine dosage (protocol dosage plus extra morphine when required), age group, Surgical Stress Score and the time-varying covariate mechanical ventilation on COMFORT 'behaviour' and VAS pain. Overall, continuous and intermittent morphine administration were equally effective in reducing postoperative pain. Only for the 1 to 3-year old infants, the CM was somewhat more effective. Only one toddler showed clinical signs of ventilatory depression and was excluded. Actual morphine dosage and age group predicted the level of repeated pain assessments. The greatest differences in pain response and actual morphine dosage were shown between neonates and infants aged 1 to 6 months, with lower pain response in neonates. Surgical stress score and mechanical ventilation were not related to postoperative pain or morphine dosages, leaving the inter-individual differences in pain response and morphine requirement largely unexplained.

Chapter 6 examines long-term consequences of early pain experiences. This topic originated from two sources. Firstly, from our study sample which included many children with major multiple congenital anomalies and also prematurely born infants, who had been long hospitalised with adverse events. Secondly, from the available literature on short and long-term consequences of early pain experiences, showing inconsistent results. This chapter also describes the influence of previous hospital experiences on postoperative pain

and stress responses in our study. We demonstrated that in our sample previous hospital experiences were not substantially predictive of postoperative pain response or stress response. This was against our expectations based on our personal clinical experience and literature findings. Several explanations for these findings are available. First, after major surgery, pre-emptive analgesics are given as a rule in our hospital. This shift from pain treatment toward pain prevention, resulting in low pain scores, might explain the limited impact of previous hospital experiences. On the other hand, short painful procedures which in the past were performed without analgesia had not been documented. Secondly, the heterogeneity of our sample with regard to prior hospital experiences, and diagnosis may have influenced the results. We concluded that long-term consequences of early pain experiences seem not as detrimental as expected. However, we should be cautious and continue patient related research. Future research designs should incorporate follow-up studies of large cohorts of neonates, both term and premature.

The many debates concerning the IAPS definition of pain concern the standard of observational pain assessment in preverbal infants. There was a need for a valid way to assess pain in preverbal infants, either as a 'gold standard' or 'silver' one. This need resulted in the development of approximately thirty instruments in the past fifteen years. In Chapter 7 we critically review the developments with regard to pain assessment instruments of the past five years. The various instruments use mainly the same indicators. The additional value of physiological indicators for postoperative pain assessment is limited because they are not specific for pain and influenced by many other factors such as blood loss, fluid intake, body temperature, and medical interventions. Most instruments have been used for research purposes only and the feasibility and the clinical utility of the instruments are neglected areas. It seems unnecessary to develop new pain instruments unless new indicators for pain are added.

Chapter 8 is divided in two parts. The first part summarised the findings from the previous chapters. The results were discussed and future directives were given.
The second part of Chapter 8 summarises our experiences during the trial. The issues were related to the design of the study, such as interdependent research questions, omitted variables, and practical matters related to data collection such as treatment violations and missing data. Some practical recommendations for future research were given. Finally, current and future pain studies were outlined, showing that the study in this thesis is part of a whole line of pain research in the Sophia Children's Hospital.

Samenvatting

Een goed pijnbestrijdingbeleid omvat zowel pijnbeoordeling als pijnbehandeling. Dit is met name van belang voor kinderen die op zeer jonge leeftijd lange tijd in het ziekenhuis liggen en mogelijk verschillende pijnlijke procedures ondergaan. In de afgelopen vijftien jaar is het inzicht gegroeid dat pasgeborenen en jonge kinderen ook pijn ervaren. De pijnbestrijding werd echter niet meteen overal aangepast omdat men bang was voor bijwerkingen van opiaten en omdat gevalideerde methoden om pijn bij deze kinderen te beoordelen ontbraken. Het onderzoek dat hier gerapporteerd wordt, is opgezet om pijnbeoordelingsmethoden en de pijnbehandeling voor jonge kinderen te verbeteren. Het specifieke doel van dit onderzoek, dat werd mogelijk gemaakt door een subsidie van NWO (nr. 940-31-031), was het beantwoorden van de volgende onderzoeksvragen:

- Hoe betrouwbaar, valide en praktisch in gebruik is de multidimensionale COMFORT schaal voor het beoordelen van postoperatieve pijn in baby's en kleuters in de leeftijd van 0-3 jaar?
- Wat is het verschil in effectiviteit tussen intermitterende en continue intraveneuze toediening van morfine op postoperatieve pijn?

Tijdens het verzamelen van de gegevens kwam een derde onderzoeksvraag naar voren:

- Is er een verband tussen huidige postoperatieve pijn en stressreactie enerzijds, en pijnervaringen in het verleden anderzijds?

In het kader van een prospectief, dubbelblind, gerandomiseerd onderzoek op de kinderchirurgische intensive care afdeling van het Sophia Kinderziekenhuis werden de gegevens van 204 jonge kinderen verzameld en geanalyseerd. De kinderen werden in 4 leeftijdsgroepen ingedeeld en werden na een grote buik- of borstoperatie gerandomiseerd naar intermitterende of continue morfine toediening. De postoperatieve pijn werd in de eerste 36 uur na de operatie elke drie uur beoordeeld door verpleegkundigen met behulp van twee pijnmeetinstrumenten: de COMFORT schaal en de Visueel Analoge Schaal (VAS). De COMFORT schaal werd gekozen omdat dit een multidimensioneel instrument is met zowel gedrags- als fysiologische pijnindicatoren, en is oorspronkelijk ontwikkeld om op intensive care afdelingen de mate van distress /comfort van kinderen aan de beademing te beoordelen.

In hoofdstuk 2 worden de psychometrische eigenschappen van de COMFORT schaal geanalyseerd. De tussenbeoordelaarsbetrouwbaarheid van de COMFORT items bleek voldoende tot goed te zijn voor alle items. Dit toonde aan dat het mogelijk was verpleegkundigen te leren op een betrouwbare wijze pijngedrag te observeren.

De postoperatieve COMFORT en VAS pijn scores van de eerste 158 pasgeborenen en jonge kinderen in de leeftijdsgroep van 0-3 jaar werden gebruikt voor het modelleren van structurele vergelijkingen (SEM). Met behulp van SEM werden de psychometrische eigenschappen van de COMFORT schaal geanalyseerd. Het bleek dat de interne structuur

van de COMFORT gegevens het best werd weergegeven door drie latente variabelen: de COMFORT 'gedrag' met factorladingen op de gedragsitems (Alertheid, Kalmte, Ademhalingsreactie/Huilen, Lichaamsbeweging, Spierspanning en Gelaatsspanning) en aparte latente variabelen voor 'Hartslag' en 'Gemiddelde arteriële bloeddruk'. De factorladingen van de items waren constant door de tijd, hetgeen wijst op stabiliteit van de structuur.
Ofschoon 'Hartslag' en 'Gemiddelde arteriële bloeddruk' door de tijd heen stabiel waren, bleken ze zwak gerelateerd aan de VAS en COMFORT 'gedrag'. We kwamen tot de conclusie dat de optelling van de scores van de gedragsitems, de zogenaamde COMFORT 'gedrag' schaal, voldoende betrouwbaar en valide is om postoperatieve pijn bij pasgeborenen en kleuters te beoordelen.

Hoofdstuk 3 geeft een overzicht van de Engelstalige literatuur op het gebied van pijn bij kinderen, met de nadruk op onderzoeken die de betrouwbaarheid en/of de validiteit van de observationele Visueel Analoge Schaal (VAS_{obs}) toetsen met onderzoeksgegevens. In deze toepassing maakt een observator, bijvoorbeeld een verpleegkundige, een inschatting van de intensiteit van de pijn die een ander ervaart. Bovendien zochten we naar onderzoeksresultaten betreffende optimale afkappunten op de VAS_{obs} die het mogelijk maken gradaties van pijn te onderscheiden.
De psychometrische eigenschappen van de VAS_{obs} bleken veelbelovend. Terwijl de meeste pijnmeetinstrumenten zijn gebaseerd op gedetailleerde gedragsobservaties, kan de globale inschatting die met de VAS_{obs} wordt gedaan aanvullende kennis opleveren over individuele variaties in pijngevoeligheid en gedrag, en situationele invloeden.
Nader onderzoek behoeven de binnenbeoordelaarsbetrouwbaarheid, de gevoeligheid voor verandering, en de optimale afkappunten. Afkappunten beogen de grens tussen 'pijn' versus 'geen pijn' te representeren. We stelden dat de VAS_{obs} van nut kan zijn in aanvulling op een ander, gevalideerd pijninstrument, aangezien hoge gedragsscores bij sommige kinderen angst of andere vormen van distress uitdrukken, die met behulp van de VAS_{obs} kunnen worden onderkend. Het lijkt raadzaam de VAS_{obs} niet eerder te gebruiken voordat er voldoende ervaring is opgedaan met pijn en pijnbeoordeling bij kinderen van verschillende leeftijden, en voldoende tussenbeoordelaarsbetrouwbaarheid is aangetoond.

Omdat de statistische analyse een lage samenhang had aangetoond tussen de gedragsitems en de fysiologische items van de COMFORT schaal, rees de vraag of dit mede kwam door het 'format' van de items, waarin de oorspronkelijke hartslag en bloeddrukwaarden zijn teruggebracht tot vijf antwoordcategorieën. Om die reden werden in hoofdstuk 4 de bij elke beoordeling geregistreerde zes hartslag- en bloeddrukwaarden gebruikt om de relatie vast te stellen tussen COMFORT 'gedrag' en de gemiddelde en standaarddeviatie van hartslag en bloeddruk. Een tweede doel was het identificeren van determinanten die de mate van samenhang tussen COMFORT 'gedrag' en de fysiologische items zouden kunnen voorspellen.

De 'binnen-subject' correlaties, op basis van de herhaalde metingen van de postoperatieve pijnschattingen, waren matig, uiteenlopend van 0,37 tot 0,49 voor COMFORT 'gedrag' met de gemiddelde en standaarddeviatie van hartslag en bloeddruk.
De brede spreiding in gedrag-fysiologie correlaties weerspiegelde de grote individuele verschillen. De correlaties bij pasgeborenen waren lager dan bij de andere leeftijdsgroepen. Als mogelijke oorzaak hiervoor werden genoemd de lage pijnscores en lage morfinedoseringen bij pasgeborenen. Dit suggereert dat naarmate de pijn toeneemt de samenhang tussen het gedrag en de fysiologie maten (hartslag en bloeddruk) groter wordt. De gedrag-fysiologie correlaties werden niet in hoge mate beïnvloed door factoren gerelateerd aan chirurgische stress of ziektebeeld. Deze bevindingen wijzen er op dat fysiologische pijnindicatoren niet echt valide zijn voor postoperatieve pijn. Echter, voor situaties waarin gedrag niet te beoordelen is (b.v. beademde zuigelingen die spierverslappers krijgen) of minder betrouwbaar is (b.v. bij zeer zieke kinderen) kunnen ze misschien wel valide zijn. Dit dient nog nader te worden onderzocht.

Hoofdstuk 5 rapporteert het studieonderdeel waarin de effectiviteit van continue morfine infuus van 10 µg/kg/h morfine (CM) wordt vergeleken met die van intraveneuze bolusinjecties van 30 µg/kg morfine (IM) à drie uur, na een grote buik- of borstoperatie. Voor dit doel werd, bij 181 kinderen in de leeftijd van 0 tot 3 jaar, gedurende de eerste 24 uur na de ingreep elke drie uur de COMFORT 'gedrag' score en de VAS score bepaald. Met random regressie analyse werd het effect van de wijze van morfinetoediening en de hoeveelheid toegediende morfine op COMFORT 'gedrag' en VAS pijn geschat, rekening houdend met de leeftijdsgroep, de 'Chirurgische Stress Score' en het wel of niet kunstmatig beademd zijn. Globaal bleken CM en IM even effectief postoperatieve pijn te verminderen, hoewel voor de leeftijdsgroep van 1 tot 3 jaar, CM enigszins effectiever bleek te zijn. De morfinedosering en de leeftijdsgroep waren beide voorspelbaar voor de mate van postoperatieve pijn. De grootste verschillen in pijnreactie en daadwerkelijke morfinedosering kwamen voor tussen pasgeborenen en zuigelingen van 1 tot 6 maanden, met lagere pijnscores bij pasgeborenen. Voor de meeste pasgeborenen volstond de dosering van 10 µg/kg/h volgens het protocol. 'Chirurgische stress score' en kunstmatige beademing waren niet gerelateerd aan postoperatieve pijn of morfinedoseringen, hetgeen de interindividuële verschillen in pijnrespons en morfinebehoefte grotendeels onverklaard laat. Slechts één kind vertoonde klinische tekenen van ademhalingsdepressie.

Hoofdstuk 6 gaat in op de lange termijn gevolgen van vroege pijnervaringen. Dit onderwerp vond zijn oorsprong in twee bronnen. Ten eerste onze steekproef, waarin veel kinderen zaten die vanwege ernstige meervoudige aangeboren afwijkingen of vroeggeboorte reeds meerdere keren geopereerd waren of pijnlijke ingrepen moesten ondergaan. Ten tweede, de beschikbare literatuur over de gevolgen op de korte en lange termijn van vroege pijnervaringen. Deze literatuur liet overigens tegenstrijdige bevindingen zien.

Dit hoofdstuk beschrijft de invloed van vroegere ziekenhuiservaringen op postoperatieve pijn en stressresponses in onze patiëntengroep. In onze patiëntengroep bleken vroegere ziekenhuiservaringen zoals totale ziekenhuisopnameduur, aantal ingrepen onder algehele narcose en aantal dagen beademing, niet substantieel voorspellend voor de huidige postoperatieve pijn- of stressresponses. Dit hadden wij niet verwacht gezien onze eigen klinische ervaring en de bevindingen uit de literatuur. Verschillende verklaringen zijn mogelijk. Ten eerste, na een ingrijpende operatie worden in ons ziekenhuis in de regel preventief pijnstillende middelen gegeven. Deze verschuiving van pijnbehandeling naar pijnpreventie, die lage pijnscores ten gevolge had, zou de beperkte invloed van vroegere ziekenhuiservaringen kunnen verklaren. Daarentegen, korte pijnlijke procedures die in het verleden zonder pijnstilling werden uitgevoerd, waren niet gedocumenteerd. Ten tweede, de heterogeniteit van onze patiëntengroep met betrekking tot vroegere ziekenhuiservaringen en diagnosis kan de resultaten hebben beïnvloed. Ter illustratie, zowel kinderen die een acute operatie ondergingen als zij die sinds kort een maligniteit bleken te hebben, hadden geen ziekenhuisgeschiedenis, maar maakten wel een mogelijk pijnlijke en stressvolle periode door tijdens het huidige verblijf in het ziekenhuis.
We concludeerden dat de latere gevolgen van vroege pijnervaringen niet zo nadelig leken als was verwacht. Het verdient echter aanbeveling prospectief onderzoek op te zetten in de vorm van 'follow-up studies' van grote cohorten pasgeborenen, zowel voldragen als prematuur. Dit om de late gevolgen van vroege pijnervaringen beter in kaart te brengen.

Er was behoefte aan een goed gevalideerde methode om pijn te beoordelen bij jonge kinderen. In de afgelopen vijftien jaar heeft dit geleid tot de ontwikkeling van zo'n dertig instrumenten. Hoofdstuk 7 geeft een kritisch overzicht van de ontwikkelingen op het gebied van pijnmeetinstrumenten in de afgelopen vijf jaar. Er is inhoudelijk veel overlap tussen de verschillende pijnmeetinstrumenten, aangezien deze veelal gebruik maken van dezelfde indicatoren voor pijn. De toegevoegde waarde van fysiologische indicatoren, zoals hartslag en bloeddruk, voor met name postoperatieve pijnbeoordeling is beperkt, omdat deze niet specifiek zijn voor pijn. Daarnaast worden zij beïnvloed door vele andere factoren zoals bloedverlies, vochtinname, lichaamstemperatuur, en medisch handelen.
De meeste pijnmeetinstrumenten zijn alleen gebruikt voor onderzoek bij acute pijn; het nut in de dagelijkse praktijk is een nog onderbelicht terrein. Andere gebieden waarop onderzoek ontbreekt is het bepalen van zogeheten afkappunten, en de vraag hoe een onderscheid te maken tussen pijn en andere vormen van 'distress', zoals boosheid, angst of stress bij kinderen die nog niet kunnen praten. Bovendien zouden de bestaande pijnmeetinstrumenten moeten worden getest in verschillende settings, zoals de dagbehandelingskliniek en verkoeverkamer, en bij meer chronische pijn, zoals bijvoorbeeld necrotiserende enterocolitis bij pasgeborenen.
Het lijkt niet zinvol om nieuwe pijninstrumenten te ontwikkelen tenzij nieuwe indicatoren worden toegevoegd.

Hoofdstuk 8 bestaat uit twee delen. Het eerste bevat een discussie van de belangrijkste bevindingen van het onderzoek. Het tweede deel recapituleert enkele positieve en negatieve ervaringen gedurende het onderzoek met aansluitend aanbevelingen voor toekomstig onderzoek, en de huidige stand van zaken met betrekking tot het onderzoek op het gebied van pijnbestrijding in het Sophia Kinderziekenhuis.

Dankwoord

Lieve Pieter, eindelijk meer tijd om 'de cabrio's' af te maken en een wat minder neurotische partner (?). Lieve Sam en Nicky, bedankt voor jullie bijdrage aan dit boekje en ik beloof zoiets nooit meer te doen!
Beste Ko, je hebt me vooral in de laatste fase enorm geholpen met het redigeren van een aantal hoofdstukken. Jouw bescheiden, doch scherpe observaties hebben me enorm geholpen.
Lieve Margo, bedankt voor jouw hulp bij de lay-out maar ook voor jouw positieve kijk op het geheel.
Prof.dr. Passchier, beste Jan, bedankt voor jouw vakkundige begeleiding en vertrouwen in de goede afloop.
Mw. Dr. de Boer, beste Josien, als co-promotor heb jij mij enorm geholpen met jouw enthousiaste begeleiding en het meedenken en bijschaven van de verschillende artikelen.
Beste Dick, jouw enorme energie werkt (soms te) aanstekelijk. Bedankt voor jouw hulp op allerlei gebied. De samenwerking gaat gelukkig door in een volgend project.
Dr. Duivenvoorden, beste Hugo, vanaf 1991 heb ik een zeer uitgebreide postdoc opleiding Methoden en Technieken van jou gekregen. Maar daarnaast ook veel gezelligheid en steun ondervonden. Bedankt voor jouw jarenlange begeleiding, die hopelijk nog niet afgelopen is.
Prof. Koot, beste Hans, niemand kan zo goed mijn ideeën verwoorden als jij. Je hebt de kwaliteit van de artikelen op een hoger plan gebracht. Bedankt en hopelijk blijft de samenwerking op de een of andere wijze doorgaan.
Some other people were an inspiration to me during this research. Dr. Anand, dear Sunny, you inspired me through your important studies and your gift to describe complicated matters in a clear and understandable way.
Dr. Ron Barr, your original articles were inspiring and I hope to meet you in person one of these days.
Dr. Bouwmeester, Beste Nancy, jij bent een van de pijlers geweest waarop het onderzoek steunde. Ik kon altijd bij jou terecht voor vakinhoudelijke vragen. Bedankt hiervoor.
De verpleging van de kinderchirurgische intensive care onder aanvoering van Heleen Verwijs heeft enorm bijgedragen aan het welslagen van dit project. Niet alleen deden zij alle pijnmetingen en bloedafnames maar ook keken zij kritisch mee en wezen ons op onvolkomenheden binnen het onderzoek.
De overige leden van de kleine commissie, prof. dr. Huijer Abu Saad en prof. dr. van den Anker wil ik hartelijk bedanken voor hun tijd en moeite om het proefschrift te lezen en te beoordelen. Tevens wil ik prof. dr. Hazebroek en prof. dr. Zuurmond bedanken voor hun bereidheid zitting te nemen in de grote commissie.
Pappa, ik kan het niet uitstaan dat jij er niet bij bent om trots te zitten zijn.

Curriculum Vitae

Monique van Dijk werd op 20 december 1958 geboren te Rotterdam. In 1977 behaalde ze het VWO-diploma aan het Montessori Lyceum in dezelfde stad. In de periode 1978 tot en met 1984 behaalde ze resp. het A en B- diploma verpleegkundige. Van 1984 tot en met 1990 was ze werkzaam als verpleegkundige in verschillende ziekenhuizen en in de wijkverpleging. Van 1987 tot en met 1993 studeerde ze Psychologie te Leiden. In 1993 legde zij het doctoraal- examen af met als afstudeerrichting Methoden en Technieken. Van 1993 tot en met 1995 was zij werkzaam als onderzoeker bij het Helen Dowling Instituut bij een onderzoek naar 'aanpassing aan kanker'. Van 1995 tot en met 1999 was ze met subsidie van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) verbonden aan de afdeling Medische Psychologie en Psychotherapie van de Erasmus Universiteit Rotterdam voor het onderzoek dat beschreven is in dit proefschrift. Dit onderzoek vond plaats op de afdeling Intensive Care Kinderheelkunde, in samenwerking met de afdeling Kinder- en Jeugdpsychiatrie en de afdeling Anesthesiologie, in het Sophia Kinderziekenhuis.
Sinds juni 2000 is ze verbonden aan de afdeling Kinderheelkunde als postdoc. Zij is momenteel betrokken bij een door NWO gesubsidieerd onderzoek naar pijnmeting bij beademde pasgeborenen binnen de neonatale Intensive Care van het Sophia Kinderziekenhuis.

De COMFORT schaal

Handleiding versie 1.0

Josien de Boer
Hans Koot
Monique van Dijk

Pijnkenniscentrum Rotterdam

Sophia Kinderziekenhuis Rotterdam

De COMFORT schaal

Handleiding versie 1.0

J.B. de Boer[1], H.M. Koot[2], M. van Dijk[3].

[1] Afdeling Medische Psychologie en Psychotherapie, Erasmus Universiteit Rotterdam
[2] Afdeling Kinder- en Jeugdpsychiatrie, Sophia Kinderziekenhuis Rotterdam
[3] Afdeling Kinderheelkunde, Sophia Kinderziekenhuis Rotterdam

ISBN 90-73235-42-1

1 Handleiding bij de COMFORT schaal, versie 1.0

Voorwoord

Dat kinderen van alle leeftijden, dus ook zeer jonge kinderen, pijn kunnen ervaren is overtuigend aangetoond in verschillende onderzoeken (Anand and Hickey 1987; Anand et al. 1987). Om iets aan de pijn te kunnen doen dienen we te weten wanneer deze optreedt en in welke mate. Voor het vaststellen van pijn en de evaluatie van de pijnbestrijding zijn betrouwbare en valide pijnmeetinstrumenten onontbeerlijk.

In de afgelopen jaren is een groot aantal pijnmeetinstrumenten ontwikkeld, die verschillen in doelgroep, inhoud, lengte en in de mate waarin hun psychometrische eigenschappen bekend zijn. Eén ervan is de COMFORT schaal, ontwikkeld in de Verenigde Staten door Celeste Marx, Bruce Ambuel en Kim Hamlett (Ambuel et al. 1990; Ambuel et al. 1992). Het is oorspronkelijk een observatie-instrument voor het meten van 'distress' bij beademde kinderen van 0-18 jaar die opgenomen zijn op een intensive care (IC).
De afgelopen jaren is de COMFORT schaal diverse malen toegepast in onderzoek naar de doeltreffendheid van de sedatie (El-Khatib et al. 1994; Marx et al.1994; Reed et al.1996). Recentelijk is de COMFORT schaal gebruikt in een onderzoek naar pijnlijke procedures op een afdeling neonatologie, zoals intubaties of uitzuigen (Blauer and Gerstmann 1998).

In deze handleiding wordt verslag gedaan van onze bevindingen met de COMFORT schaal. De COMFORT schaal werd gebruikt in een placebo-gecontroleerde studie naar de effectiviteit van continue versus intermitterende toediening van morfine bij kinderen van 0 tot 3 jaar die een zware borst- of buikoperatie moesten ondergaan. Voor dit onderzoek zijn 39 verpleegkundigen en 2 anesthesisten getraind in het gebruik van de COMFORT schaal. Na afloop van het onderzoek zijn de overige verpleegkundigen en de artsen van de afdeling intensive care kinderchirurgie van het Sophia Kinderziekenhuis geschoold in het gebruik van de COMFORT schaal. De ervaringen van de artsen en verpleegkundigen met de schaal alsmede de resultaten van het onderzoek zijn een belangrijke bron van informatie geweest voor deze handleiding.

In hoofdstuk 1 wordt de COMFORT schaal besproken. In hoofdstuk 2 worden de resultaten besproken van Amerikaans en Nederlands onderzoek naar de psychometrische eigenschappen van de schaal. Tenslotte worden in hoofdstuk 3 een aantal conclusies en aanbevelingen gepresenteerd.
In de bijlagen 1 en 2 worden respectievelijk de scoringsprocedure en de interpretatie van de items en antwoordcategorieën uitgebreid toegelicht.
Bijlage 3 bevat de Nederlandse versie 2 van de COMFORT schaal.
Het testen en valideren van de COMFORT schaal werd mede mogelijk gemaakt door een subsidie van N.W.O. (projectnr. 940-31-031). De auteurs zijn de verpleegkundigen en

artsen van de afdeling Intensive Care Kinderchirurgie van het Sophia Kinderziekenhuis zeer erkentelijk voor hun bijdrage en danken dr. Hugo Duivenvoorden voor zijn methodologische en statistische adviezen en drs. Nancy Bouwmeester, Prof. dr. Dick Tibboel, en Prof. dr. Jan Passchier voor hun aandeel in het onderzoek.

Voorwaarden voor gebruik

Alvorens de Nederlandse versie van de COMFORT schaal te mogen gebruiken dienen gebruikers een training van twee dagdelen te volgen. Deze wordt tegen betaling twee keer per jaar door het Pijnkenniscentrum Rotterdam thema 'Pijn op de kinderleeftijd' georganiseerd.
In publicaties dienen gebruikers te verwijzen naar deze handleiding. Gebruikers dienen schriftelijke toestemming te hebben gekregen van de Nederlandse auteurs.
Aangezien de COMFORT nog verder gevalideerd dient te worden, wordt gebruikers verzocht de door hen verzamelde COMFORT gegevens ter beschikking te stellen voor psychometrisch onderzoek.
Alvorens de COMFORT schaal in onderzoek te gebruiken dient een tussenbeoordelaarsbetrouwbaarheid (uitgedrukt als een gewogen Cohen's kappa) van minimaal 0.60 te worden behaald op basis van 10 COMFORT observaties.

Verdere informatie

Informatie over de training in het gebruik van de COMFORT schaal en het verkrijgen van materialen is verkrijgbaar bij: mw. Kitty Oosterbaan

Adres: Sophia Kinderziekenhuis Rotterdam
Intensive Care Kinderchirurgie
Postbus 2060
3000 CB Rotterdam
Telefoon: 010-4636203
Fax: 010-4636288
E-mail: oosterman@chis.azr.nl

Hoofdstuk 1 De COMFORT schaal

1.1 Beschrijving van de COMFORT schaal

De oorspronkelijke Amerikaanse versie van de COMFORT schaal bevat 8 items; 6 gedragskenmerken en twee fysiologische kenmerken. In overleg met de auteurs zijn de schaal en de handleiding vertaald in het Nederlands. Om het gebruik van de schaal niet te beperken tot beademde kinderen werd een equivalent van het item 'ademhalingsreactie' geconstrueerd voor niet-beademde kinderen, het item 'Huilen'.

De Nederlandse versie van de COMFORT schaal bestaat derhalve uit negen items: zeven gedragskenmerken: Alertheid, Kalmte/Agitatie, Ademhalingsreactie, Huilen, Lichaamsbeweging, Spierspanning, en Gelaatsspanning, en twee fysiologische kenmerken: Bloeddruk en Hartslag. Elk kenmerk wordt op een 5-puntsschaal beoordeeld, die oploopt van 'geen distress' naar 'ernstige distress'.

Voorafgaand aan de observaties worden de baseline van de hartslag en de gemiddelde arteriële bloeddruk (MAP) bepaald. Het kind wordt vervolgens gedurende twee minuten geobserveerd. De bloeddruk en hartslag worden geobserveerd van de monitor, de spierspanning wordt bepaald door een arm of been van het kind te bewegen aan het einde van de observatieperiode, en de overige gedragskenmerken worden door middel van observatie van het kind beoordeeld. Na afloop van de twee minuten wordt elk kenmerk op de betreffende 5-puntsschaal gescoord.

Voor een kind dat beademd wordt, wordt het item 'Ademhalingsreactie' gescoord en niet het item 'Huilen'. Voor een kind dat niet beademd wordt, wordt het item 'Huilen' ingevuld, het item 'Ademhalingsreactie' wordt overgeslagen. Voor elk kind worden dus 8 items gescoord.

De uiteindelijke COMFORT score is het totaal van de scores op deze acht items, en kan aldus uiteenlopen van 8 tot 40. Hoe hoger de score, des te meer distress.

De schaal mag gebruikt worden bij al dan niet kunstmatig beademde pasgeborenen, kinderen en adolescenten. De COMFORT is oorspronkelijk niet bedoeld voor het meten van distress bij premature pasgeborenen of bij kinderen met gestoorde neurologische functies of beperkingen van het bewegingsapparaat. De toepasbaarheid bij deze groepen alsmede bij kinderen die niet op een afdeling voor intensive care worden verzorgd, wordt onderzocht.

Hoofdstuk 2. Psychometrisch onderzoek

2.1 Psychometrisch onderzoek naar de Amerikaanse versie

De betrouwbaarheid en validiteit van de Amerikaanse versie van de COMFORT schaal zijn gedocumenteerd in twee onderzoeken (Ambuel et al. 1992; Marx et al. 1994). De interne consistentie van de schaal (8 items) was hoog (Cronbach's alpha =0.90). In beide studies was de overeenstemming tussen twee beoordelaars over de totale COMFORT score goed (Pearson correlatiecoëfficiënten respectievelijk 0.82 en 0.84). De overeenstemming tussen de beoordelaars op de acht afzonderlijke items, uitgedrukt in Pearson correlatiecoëfficiënten, varieerde van 0.51 voor zowel Bloeddruk als Gelaatsspanning tot 0.73 voor Alertheid.

Steun voor de klinische validiteit van de COMFORT schaal werd gevonden in een hoge correlatie (r=0.75) tussen de totale COMFORT score en de beoordeling van een ervaren verpleegkundige over de mate van distress op een visueel analoge schaal (VAS). Verder bleken de gemiddelde scores van de, volgens de artsen van de IC-afdeling, optimaal, onvoldoende en overmatig gesedeerde kinderen significant van elkaar te verschillen. Er was goede overeenstemming tussen de COMFORT score en het oordeel van de IC-arts over de doeltreffendheid van de sedatie van het kind (r=0.66).

Blauer (1998) heeft de gevoeligheid vergeleken van drie meetinstrumenten, waaronder de COMFORT schaal, voor veranderingen in pijngedrag tijdens pijnlijke procedures (intubaties, uitzuigen) op een neonatale IC-afdeling. Het bleek goed mogelijk om met de COMFORT veranderingen in pijngedrag waar te nemen.

2.2 Psychometrisch onderzoek naar de Nederlandse versie

2.2.1 Beschrijving onderzoek en onderzoeksgroep

De COMFORT schaal werd afgenomen in het kader van een placebo-gecontroleerde studie waarin de effectiviteit en tolerantie van continue versus intermitterende toediening van morfine werden vergeleken. Een belangrijke vraagstelling was in hoeverre de COMFORT schaal geschikt was om postoperatieve pijn bij kinderen van 0-3 jaar te meten. Voor een uitgebreidere bespreking van deze studie wordt verwezen naar het originele onderzoeksverslag in het tijdschrift *Pain* (Dijk van et al. 2000).

In dit onderzoeksdeel werden de gegevens van de eerste 158 kinderen van 0 tot 3 jaar geëvalueerd. Verderop in dit hoofdstuk worden de gegevens van de uiteindelijke steekproef van 204 kinderen besproken.

De COMFORT schaal werd volgens een tevoren vastgesteld protocol 13 keer afgenomen: één keer voor de ingreep (baseline meting), en twaalf keer na de ingreep voorafgaand aan elk 3-uurs controle. De laatste meting was 36 uur na de operatie. Na het invullen van de COMFORT schaal, gaf de verpleegkundige op een VAS (lijn van 10 centimeter met de ankerpunten 'geen pijn' en 'ernstige pijn') aan hoeveel pijn het kind op dat moment had (Huskisson 1974). De verpleegkundige observeerde het kind steeds voordat de verpleegkundige handelingen werden uitgevoerd.

2.2.2 Gebruikersvriendelijkheid

In de praktijk blijkt dat verpleegkundigen goed kunnen werken met de COMFORT schaal. Dit blijkt zowel uit de geringe hoeveelheid ontbrekende gegevens in het onderzoek als uit een kleinschalig onderzoek onder de verpleging naar de gebruikersvriendelijkheid (Koole et al. 1998). De verpleegkundigen gaven aan dat de gedragskenmerken goed aansluiten bij hun klinische observaties. Verder meenden zij dat de regelmatige observaties hen meer alert maakte op de aanwezigheid van pijn waardoor ook de pijnbestrijding verbeterde.

2.2.3. Tussenbeoordelaarsbetrouwbaarheid

Vóór en tijdens het onderzoek zijn verpleegkundigen van de afdeling in twee- vier- of zestallen getraind in het gebruik van de COMFORT schaal. De training bestond uit een uitleg van de verschillende items en antwoordcategorieën, het toepassen van de COMFORT schaal bij een aantal videofragmenten, en bij een aantal baby's en kinderen op zaal. Daarna werden de deelnemers verzocht minimaal 10 keer samen te scoren met een verpleegkundige die al lange tijd ervaring had met het scoren met de COMFORT schaal. Voor de tussenbeoordelaarsbetrouwbaarheid werd gesteld dat de gewogen Cohen's kappa groter dan 0.60 moest zijn (Spitzer et al. 1967) voordat een verpleegkundige in het onderzoek mee kon doen.
In totaal hebben 39 verpleegkundigen en 2 anesthesisten de training gevolgd en ieder 10 gepaarde observatiescores ingeleverd. De overeenstemming met de ervaren verpleegkundige bleek zeer goed te zijn voor de afzonderlijke items, met uitzondering van 'ademhalingsreactie' (0.54) (zie tabel 1). Deze relatief lage kappa zou te wijten kunnen zijn aan het feit dat het scoren van dit item afhankelijk is van de interpretatie van de uitslag van de monitor. Niet alle verpleegkundigen die de training volgden, werkten dagelijks met de beademingsapparatuur zodat het scoren van dit item beïnvloed kan zijn door de mate van ervaring met deze apparatuur.
De mediane kappa was 0.70.

Tabel 1. Tussenbeoordelaarsbetrouwbaarheid

COMFORT item	Gewogen Kappa	Aantal gepaarde observaties
Alertheid	0.74	302
Kalmte/Agitatie	0.69	302
Ademhalingsreactie	0.54	131
Huilen	0.70	170
Lichaamsbeweging	0.70	302
Spierspanning	0.66	302
Gelaatsspanning	0.63	296
Bloeddruk	0.93	232
Hartslag	0.93	290

2.2.4 Interne consistentie

Voor het bepalen van de interne consistentie zijn de Cronbach's alpha's berekend voor zowel de hele schaal als voor de twee fysiologische en zes gedragsitems afzonderlijk. Dit is gedaan voor 3 meetmomenten: de baselinemeting, en 6 en 12 uur postoperatief. Uit tabel 2 blijkt dat de interne consistentie van de gehele schaal goed is (>0.70), met uitzondering van de postoperatieve meting na 12 uur. Deze lage Cronbach's alpha blijkt veroorzaakt te worden door de items bloeddruk en hartslag. De interne consistentie van de gedragsitems blijkt hoog (≥0.80), en die van de fysiologische items bijzonder laag, met uitzondering van de preoperatieve baseline meting.
Het item 'huilen' bleek in vergelijking met het item 'ademhalingsreactie' de interne consistentie niet nadelig te beïnvloeden.

Tabel 2. Interne consistentie

Schaal	Aantal items	Cronbach's alpha					
		Preoperatief	n	6 u postop	n	12 u postop	n
Totaal	8 (incl. Huilen)	-		0.81	108	0.80	113
	8 (incl. Ademhalingsreactie	-		0.79	74	0.62	68
Gedrag	6 (incl. Huilen)	0.83	162	0.92	113	0.87	118
	6 (incl. Ademhalingsreactie	0.90	24	0.89	79	0.84	71
Fysiologie[1)]	2 (Hartslag en Bloeddruk)	-		0.29	187	0.29	181

[1)] hartslag en bloeddruk werden preoperatief gemeten om als uitgangswaarde te gebruiken bij postoperatieve COMFORT metingen ; Postop = postoperatief; n=aantal kinderen

2.2.5 Congruente validiteit

Een indruk van de congruente validiteit is verkregen door de correlaties te berekenen tussen de VAS voor pijn en de items van de COMFORT schaal (zie tabel 3).
De correlatie van de VAS voor pijn met de gedragsitems van de COMFORT schaal is hoog op alle drie metingen. De samenhang met de bloeddruk is matig (range 0.20 tot 0.34), en met de hartslag laag (range 0.11 tot 0.21).

Tabel 3 Correlaties VAS 'pijn' met COMFORT items 3, 6 en 12 uur postoperatief

	VAS 3 u postop	VAS 6 u postop	VAS 12 u postop
Alertheid	0.66	0.59	0.60
Kalmte/agitatie	0.79	0.75	0.79
Huilen	0.66	0.70	0.72
Ademhalingsreactie	0.50	0.62	0.41
Lichaamsbeweging	0.64	0.57	0.60
Gelaatsspanning	0.71	0.70	0.70
Spierspanning	0.63	0.67	0.54
Bloeddruk	0.34	0.29	0.20
Hartslag	0.14	0.21	0.11
Steekproefgrootte	196	192	189

2.2.6 Stabiliteit van de factorstructuur

Bij het ontwikkelen van de COMFORT schaal is ervan uitgegaan dat alle items eenzelfde construct (distress of pijn) meten. De correlaties tussen de gedrags- en fysiologische items gaven echter aan dat er mogelijk sprake is van twee constructen.
Met behulp van LISREL 8.0 (Jöreskog and Sörbom 1993) is gekeken naar de structuur en de stabiliteit van de factorstructuur van de COMFORT over de tijd (3, 6, en 9 uur postoperatief) (zie figuur 1.). De beste 'fit' van het model werd gevonden bij 3 latente variabelen: één voor de 6 gedragsitems, één voor hartslag, en één voor bloeddruk. Dit geeft aan dat de gedragsitems een afzonderlijk construct weerspiegelen. De stabiliteitscoëfficiënten van de latente variabelen hartslag en bloeddruk waren zeer hoog (0.89), die van de gedragsitems hoog (0.58 en 0.59).

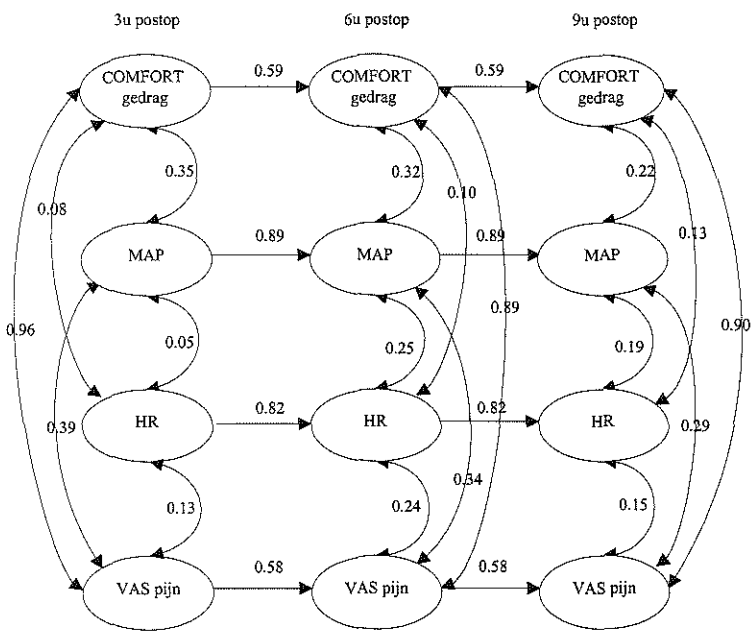


Fig 1. Structureel deel van LISREL model, ovale cirkels representeren latente variabelen, rechte lijnen geven stabiliteitscoëfficiënten weer, gebogen lijnen symboliseren de samenhang tussen de latente variabelen. MAP = gemiddelde arteriële bloeddruk, HR = hartslag. (*from van Dijk et al, Pain 2000*).

Tabel 4 Gemiddelden en standaarddeviaties (SD) voor COMFORT items en VAS pijn preoperatief, en na 6 en 12 uur postoperatief

		Preoperatief	Postoperatief	
COMFORT			6 u postop.	12 u postop.
Alertheid	Mean	3.2	2.3	2.3
	SD	1.3	1.0	1.0
	n	*186*	*194*	*189*
Kalmte/Agitatie	Mean	1.6	2.1	2.0
	SD	1.0	1.0	1.0
	n	*186*	*194*	*189*
Ademhalingsreactie	Mean	2.1	1.8	1.8
	SD	1.0	0.8	0.7
	n	*25*	*80*	*73*
Huilen	Mean	1.4	2.2	1.9
	SD	0.9	1.2	1.1
	n	*164*	*114*	*118*
Lichaamsbeweging	Mean	3.0	2.5	2.6
	SD	1.2	1.0	0.9
	n	*186*	*192*	*189*
Spierspanning	Mean	3.0	3.2	3.2
	SD	0.6	0.7	0.7
	n	*184*	*193*	*188*
Gelaatsspanning	Mean	2.6	2.7	2.6
	SD	0.8	0.9	0.7
	n	*185*	*192*	*188*
Bloeddruk baseline[1]	Mean	-	2.6	2.6
	SD	-	1.4	1.4
	n		*187*	*181*
COMFORT totaal[2]	Mean	-	20.5	19.9
	SD	-	5.8	5.1
	n		182	178
COMFORT gedrag (6	Mean	14.9	14.9	14.6
items)	SD	4.4	4.9	4.9
	n	183	192	186
Hartslag baseline[1]	Mean	-	3.0	2.8
	SD	-	1.4	1.4
	n		*195*	*189*
VAS pijn	Mean	0.5	2.5	2.0
	SD	1.4	2.1	1.8
	n	*182*	*195*	*193*

1) hartslag en bloeddruk werden preoperatief gemeten om als uitgangswaarde te dienen bij postoperatieve COMFORT metingen

2) COMFORT totaal is gebaseerd op 8 items (met item 'ademhalingsreactie' voor beademde kinderen en item 'huilen' voor niet beademde kinderen)

2.2.7 Normen/gemiddelde/mediane waarden

In tabel 4 zijn de gemiddelden en standaarddeviaties van de COMFORT items opgenomen voor de preoperatieve meting, en 6 en 12 uur postoperatief.

Hoofdstuk 3 Conclusies en aanbevelingen

3.1 Conclusies

In een onderzoek naar postoperatieve pijn bij kinderen van 0 tot 3 jaar blijkt dat het mogelijk is om een grote groep verpleegkundigen betrouwbaar met de COMFORT schaal te laten scoren.
De gehele COMFORT schaal heeft een goede interne consistentie, die nog verbetert wanneer de fysiologische items niet in de totaalscore worden opgenomen.
Het nieuwe item huilen blijkt voldoende betrouwbaar en valide om het item ademhalingsreactie te vervangen bij niet-beademde kinderen.

De gedragsitems van de COMFORT schaal blijken hoog te correleren met de VAS voor pijn. De conclusie lijkt gerechtvaardigd dat ze geschikt zijn om postoperatieve pijn te meten bij kinderen van 0 tot 3 jaar. De fysiologische items lijken minder geschikt voor dit doel. Dit kan liggen aan de antwoordcategorieën van bloeddruk en hartslag, die gebaseerd zijn op een baselinemeting. Deze baselinemeting zou idealiter de gemiddelde waarde van de afgelopen 24 uur moeten zijn. Echter, in het onderzoek was in veel gevallen maar één baselinemeting beschikbaar vanwege het ontbreken van een arteriële lijn. Bovendien waren de baselinewaarden voor een aantal kinderen bijzonder hoog (vanwege hun ziekte) zodat de postoperatieve waarden zelden boven de baselinewaarden uitkwamen. Een andere mogelijkheid is dat hartslag en bloeddruk minder geschikt zijn om pijn te meten in een postoperatieve situatie. Aanvullend onderzoek zal ingaan op de bruikbaarheid van de fysiologische items in een postoperatieve situatie.

3.2 Aanbevelingen

De gedragsitems van de COMFORT zijn geschikt voor het meten van postoperatieve pijn bij kinderen van 0-3 jaar met diverse aandoeningen. De gedragsitems mogen opgeteld worden om een totaalscore te bepalen.
De fysiologische items hebben een lage interne consistentie en blijken matig samen te hangen met de VAS voor pijn. De bruikbaarheid van deze items dient verder onderzocht te worden.
Het nieuwe item 'huilen' doet psychometrisch niet onder voor het item 'ademhalingsreactie' en kan derhalve voortaan gebruikt worden bij niet-beademde kinderen.

3.3 Verder onderzoek

De bruikbaarheid van de fysiologische variabelen hartslag en bloeddruk voor het meten van postoperatieve pijn wordt momenteel nader onderzocht. Daarnaast wordt in een vervolgonderzoek nagegaan bij welke COMFORT score er (extra) pijnmedicatie moet worden gegeven.

De COMFORT schaal wordt momenteel in een aantal onderzoeken toegepast binnen het Sophia Kinderziekenhuis Rotterdam:

1 In een onderzoek naar het meten van pijn bij diep verstandelijk gehandicapte kinderen wordt onderzocht of de COMFORT schaal betrouwbaar en valide is voor deze groep (NWO project 940-31-044 in samenwerking met de afdeling Verplegingswetenschap van de Universiteit van Maastricht).
2 In een onderzoek naar de validiteit van de NFCS (Neonatal Facial action Coding System) bij een aantal kinderen uit het hierboven beschreven postoperatieve pijnonderzoek wordt ook de COMFORT schaal geëvalueerd (SSWO-project 247).
3 In een onderzoek naar de effectiviteit en tolerantie van intraveneuze versus orale toediening van midazolam bij prematuren wordt de COMFORT schaal gebruikt om de mate van sedatie te bepalen.
4 De COMFORT schaal wordt gebruikt in een onderzoek naar de effectiviteit van rectale versus orale acetaminophen na craniofaciale chirurgie bij jonge kinderen.
5 De COMFORT schaal wordt gebruikt in een dubbelblind gerandomiseerd onderzoek waarin diclofenac en acetaminophen worden vergeleken na (adeno)-tonsillectiomie bij kinderen van 3 tot 12 jaar. De COMFORT schaal zal hierbij vergeleken worden met andere pijninstrumenten.

De resultaten van de bovenstaande onderzoeken zullen in volgende versies van de handleiding worden opgenomen. De COMFORT schaal werd medio 2000 ingevoerd als vast onderdeel bij de reguliere controles op verschillende kinderchirurgische afdelingen in het Sophia Kinderziekenhuis, en voor en na het geven van (extra) analgetica op deze afdelingen.

Literatuur

Ambuel B, Hamlett KW and Marx CM, COMFORT scale manual, 1990.

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

Anand KJS and Hickey PR. Pain and its effects in the human neonate and fetus. N Engl J Med 1987;317:1321-1329.

Anand KJS, Sippell WG and Aynsley-Green A. Randomised trial of fentanyl anaesthesia in preterm babies undergoing surgery: effects on the stress response. Lancet 1987;1:243-248.

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

Dijk van M, Boer de JB, Koot HM, Passchier J, Tibboel D and Duivenvoorden HJ. The reliability, stability and validity of the COMFORT scale as a postoperative pain instrument in 0 to 3-year-old infants. Pain 2000;84:367-377.

El-Khatib MF, Chatburn RL, Potts DL, Blumer JL and Smith PG. Mechanical ventilators optimized for pediatric use decrease work of breathing and oxygen consumption during pressure-support ventilation. Crit Care Med 1994;22:1942-1948.

Huskisson EC. Measurement of pain. Lancet 1974;2:1127-1131.

Jöreskog KG and Sörbom D. LISREL 8: Structural equation modeling with the SIMPLIS command language, Scientific Software International, Chicago, 1993.

Koole A, Verwijs HM, Booy J and Dijk van M, Het meten van postoperatieve pijn bij kinderen van 0 tot 3 jaar op een chirurgische intensive care: ervaringen van de verpleegkundigen. 6e Landelijke bijeenkomst Kinderpijngroepen, Rotterdam, 1998.

Marx CM, Smith PG, Lowrie LH, Hamlett KW, Ambuel B, Yamashita TS and Blumer JL. Optimal sedation of mechanically ventilated pediatric critical care patients. Crit Care Med 1994;22:163-170.

Reed MD, Yasmashita TS, Marx CM, Myers CM and Blumer JL. A pharmacokinetically based propofol dosing strategy for sedation of the critically ill, mechanically ventilated pediatric patient. Crit Care Med 1996;24:1473-1481.

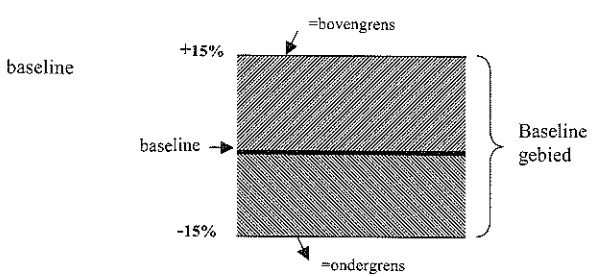
Spitzer RL, Cohen J, Fleiss JL and Endicott J. Quantification of agreement in psychiatric diagnosis. A new approach. Arch Gen Psychiatry 1967;17:83-87.

Bijlage 1. Uitgebreide beschrijving van de scoringsprocedure

Materiaal: - scoreformulier, pen
- horloge met secondewijzer of wekkertje
- rekenmachine om de baseline waarden van de hartslag en bloeddruk te berekenen

1. **De beoordelaar noteert op het scoringsformulier de datum, de naam van het kind, en de naam van de beoordelaar**. Daarnaast wordt ingevuld of het kind een arteriële lijn heeft en of het beademd wordt. Eventueel worden ook de geboortedatum van het kind of het statusnummer genoteerd.

2. **De beoordelaar berekent aan de hand van de beschikbare gegevens de baseline en boven- en ondergrens voor de hartslag en bloeddruk.** Voor de baseline waarde wordt het gemiddelde van de 12 registraties tijdens de voorafgaande 24 uur berekend, ook al zijn die waarden opgetreden na sedatie. Wanneer er geen 12 registraties beschikbaar zijn, wordt dit op het formulier genoteerd en worden zoveel mogelijk, maar minimaal 2, registraties gebruikt.
Vervolgens worden vóór het begin van de observatie de waarden berekend die 15 procent boven en onder deze baselinewaarde liggen, om snelle vaststelling van de variabiliteit mogelijk te maken. Deze boven- en ondergrens van de baselinewaarde worden op het scoreformulier bij BASELINE genoteerd. Ter illustratie een voorbeeld:
Stel, de baseline hartslag is 120. De bovengrens wordt dan 120 + (15% van 120), dus 120 + 18 =138.De ondergrens wordt dan 120 –18 = 102. In het baselinegebied vallen alle geobserveerde hartslagen tussen de 102 en 138.



3. **Het kind wordt tijdelijk blootgelegd,** zodat de beentjes, de romp en de armpjes goed zichtbaar zijn. **De observator kiest een zodanige positie dat hij/zij goed zicht heeft op zowel de monitor als het gezicht en het lichaam van het kind.**

4. De beoordelaar **noteert de aanvangstijd** en begint de observatie van 2 minuten. De beoordelaar maakt een snelle inschatting van beweging, lichaamshouding, gezichtsexpressie, respons op prikkels uit de omgeving, enz.

5. **Elke 20 seconden na het begin van de observatie, worden snel de waarden van de bloeddruk en hartslag in de kantlijn genoteerd.** De eerste observatie is na 20 seconden, de volgende observaties zijn na 40, 60, 80, 100 en 120 seconden. In totaal worden er dus voor beiden zes waarden genoteerd. Wanneer er twee observatoren zijn, geeft één van hen aan op welk moment de observatie wordt gedaan.

6. **Ongeveer 10 seconden voor het einde van de observatieperiode beoordeelt de observator de spierspanning** gebaseerd op de reactie van de patiënt op het snel en langzaam buigen en strekken van een arm of been die niet verbonden is aan instrumenten (bijvoorbeeld elleboog of knie zonder intraveneuze lijn, verband, of arterielijn). Een pols of enkel kan worden gebruikt als geen knie of elleboog beschikbaar is. Om dit aan te kondigen zegt de observator tegen het kind (zonder het bij de naam te noemen): "ik ga even aan je beentje/armpje voelen". (N.B. Twee beoordelaars die vlak na elkaar de spierspanning voelen kunnen een volstrekt andere spierspanning bij het kind waarnemen!).

7. **De beoordelaar** verlaat de patiënt en **noteert de beoordelingen voor elk item en de eindtijd.**

Vuistregels voor de scoring:

* Het meest extreme (distressed) gedrag dat geobserveerd is tijdens de observatieperiode wordt gescoord, ook al heeft het maar heel kort plaatsgevonden
* Voor de hartslag en bloeddruk wordt geturfd hoe vaak de geobserveerde waarden binnen of buiten het baselinegebied valt.
* Wanneer er een item niet gescoord kan worden (bijvoorbeeld door het ontbreken van een arteriële lijn) of wanneer het vertoonde gedrag niets te maken heeft met distress (bijvoorbeeld bij een vrolijk met de benen spartelend kind) wordt niet gescoord. In de kantlijn wordt genoteerd waarom er niet gescoord is.
* Opmerkingen die van belang zijn voor de interpretatie van de score worden onder aan het formulier genoteerd (bijvoorbeeld: kind huilt waarschijnlijk ook omdat de ouders net weg zijn).

Bijlage 2. Uitgebreide toelichting van de COMFORT items

Alertheid

Alertheid wordt beoordeeld aan de hand van de reactie van het kind op prikkels uit de omgeving zoals reacties op geluid (geruis van monitors, intercoms, mensen die praten, piepers, enz.), beweging, licht, enz. De observator mag niets doen dat de aandacht van het kind kan afleiden en er mag geen interventie (tube verwisselen, hielprikje etc) plaatsvinden.

1 Diep in slaap. De toestand van minste respons op de omgeving. De ogen van het kind zijn gesloten, ademhaling is diep en regelmatig, en het kind toont minimale reactie op veranderingen in de omgeving.

2 Licht in slaap. Het kind heeft de ogen het grootste deel van de observatieperiode gesloten, maar reageert toch nog wat op de omgeving zoals blijkt uit kleine bewegingen, bewegingen in het gelaat, niet-geslaagde pogingen om de ogen te openen, enz.

3 Slaperig. Het kind sluit vaak zijn ogen, of doet moeizame pogingen om de ogen te openen, en reageert minder op de omgeving.

4 Wakker en alert. Het kind reageert op de omgeving, maar niet overdreven. De ogen van het kind blijven het grootste deel van de tijd open, of gaan makkelijk open als reactie op prikkels uit de omgeving.

5 Wakker en hyper-alert. Het kind is overdreven waakzaam, wellicht met wijdopen ogen, reageert snel op kleine veranderingen in stimulatie uit de omgeving, en reageert overdreven op veranderingen in de omgeving.

Kalmte / Agitatie

Bij Kalmte / Agitatie wordt de mate van emotionele opwinding en angst bij het kind beoordeeld. Tekenen van angst kunnen zijn: het terugtrekken met armpjes tegen het lijf bij toenadering; wijdopen ogen; verhoogde hartslag; huilen. Bij kinderen jonger dan 1 jaar beoordeelt men de mate van onrust.

1 Kalm. De patiënt lijkt helder en rustig; er is geen teken van onrust of angst.

2 Licht angstig. Het kind is niet helemaal kalm; hij/zij vertoont lichte onrust of angst.

3 Angstig. Het kind lijkt wat onrustig en angstig, maar kan zich beheersen.

4 Zeer angstig. De patiënt lijkt zeer onrustig; angst is zichtbaar maar de patiënt kan zich nog enigszins beheersen.

5 Paniekerig. Uit de hele manier van doen van het kind spreekt onmiddellijke en ernstige onrust/ angst met verlies van beheersing.

Ademhalingsreactie

Bij Ademhalingsreactie wordt de reactie van het kind via mond en ademhaling op een endotracheale tube en intermitterende beademing beoordeeld. Er wordt zowel gelet op het ademhalingspatroon van het kind als op gegevens van de machine.

1 Geen hoesten en geen spontane ademhaling. Alleen door de ventilator opgewekte ademhaling is zichtbaar. Er is geen ademhalingsbeweging zichtbaar tussen ademhalingen teweeggebracht door de ventilator. Er vindt geen beweging van de mond of beweging van de borstwand plaats behalve wanneer die teweeggebracht is door de ventilator.

2 Spontane ademhaling met weinig of geen reactie op de beademing. De patiënt ademt in een regelmatig, normaal tempo tegelijk met de ventilator. De mond of borstwand beweegt niet tegen de beweging van de ventilator in. Aan het brandende lampje is te zien dat het kind zelf de ademhaling inzet.

3 Af en toe hoesten of verzet tegen de ventilator. Het kind beweegt af en toe de mond of borstwand tegen het patroon van de ventilator in. De patiënt ademt af en toe niet tegelijk met de ventilator. Drukverhoging is zichtbaar.

4 Ademt actief tegen de ventilator in of hoest regelmatig. Het kind beweegt vaak de mond of borstwand tegen het patroon van de ventilator in, hoest regelmatig, of ademt vaak niet tegelijk met de ventilator. 'Neusvleugelen' is zichtbaar.

5 Vecht tegen de ventilator; hoesten of zich verslikken. Het kind maakt actief mondbewegingen of bewegingen tegen het patroon van de ventilator in, hoest, verslikt zich en/of kokhalst op een manier die de beademing kan bemoeilijken. Het kind ligt zichtbaar tegen te ademen.

Huilen

Huilen wordt geobserveerd bij kinderen die niet beademd worden. Er wordt gelet op het geluid en het gedrag van het kind, niet zozeer op de aanwezigheid van tranen.

1 Rustige ademhaling, geen huilgeluiden

2 Af en toe snikken of kreunen (nasnikken). Het kan gaan om het snikken van een kind dat op het punt staat om in huilen uit te barsten, maar ook om het snikken of kreunen na een huilbui. Zowel het snikken als kreunen treedt met tussenpozen op, het is niet continu.

3 Jengelen of dreinen (monotoon geluid). Het kind huilt (nog) niet, wel maakt het continu of met kleine tussenpozen vrij monotone huilgeluiden.

4 Huilen. Het kind huilt.

5 Schreeuwen of krijsen. Het kind huilt hard of met uithalen. Het geluid is vaak hard en doordringend.

Lichaamsbeweging

Bij Lichaamsbeweging worden de frequentie en intensiteit van bewegingen van het lichaam beoordeeld. Het gaat hierbij om bewegingen die horen bij 'distress'.

1 Geen beweging. Het kind beweegt helemaal niet.

2 Incidentele kleine bewegingen. Het kind maakt drie of minder kleine bewegingen met de vingers of voeten, of zeer kleine hoofdbewegingen.

3 Regelmatige kleine bewegingen. Het kind maakt meer dan drie kleine bewegingen met de vingers of voeten, of zeer kleine hoofdbewegingen.

4 Heftige bewegingen met armen en benen. Het kind maakt bewegingen van grotere omvang, snelheid of sterkte met de handen, armen, of benen. Het hoofd kan daarbij licht bewegen. De bewegingen zijn zo heftig dat de beademingsslangen los kunnen gaan.

5 Heftige bewegingen ook met romp en hoofd. Het kind maakt bewegingen van grotere omvang, snelheid, of sterkte met hoofd en romp, zoals rollen met het hoofd, de rug buigen of de nek buigen. Armen en benen kunnen ook bewegen. De bewegingen zijn zo heftig dat ze auto-extubatie kunnen veroorzaken.

Spierspanning

Spierspanning wordt gemeten in relatie tot de normale spierspanning bij een patiënt die wakker en alert is. Beoordeeld wordt de reactie van de patiënt op het snel en langzaam buigen en strekken van een arm of been zonder intraveneuze lijn, verband, arterielijn, of andere fysieke beperking, bijvoorbeeld een spalk. Wanneer de armen en benen niet beschikbaar zijn kan worden volstaan met het optillen van de pols of enkel. Deze beoordeling is de enige die actieve interventie van de beoordelaar vereist en wordt uitgevoerd vlak voor het einde van de observatieperiode van twee minuten.

1 Spieren volledig ontspannen; geen spierspanning. Spierspanning is afwezig; er is geen weerstand tegen beweging.

2 Verminderde spierspanning. Het kind toont minder weerstand tegen beweging dan normaal, maar spierspanning is niet geheel afwezig.

3 Normale spierspanning. Weerstand tegen beweging is normaal.

4 Toegenomen spierspanning en buiging van vingers en tenen. Het kind toont duidelijk grotere weerstand dan normaal, maar het gewricht is niet stijf. Vingers en tenen kunnen in een gebogen positie worden gehouden.

5 Extreme spierstijfheid en buiging van vingers en tenen. Spierstijfheid is de overheersende toestand van het kind tijdens de observatieperiode. Dit is zelfs zichtbaar zonder bewegen van een extremiteit, onder andere aan gebalde vuistjes en gespreide teentjes.

Gelaatsspanning

Gelaatsspanning meet de tonus en spanning van de gelaatsspieren. De standaard-vergelijking is aan een kind dat wakker en alert is.

1 Gelaatsspieren volkomen ontspannen. Het kind toont geen spanning in de gelaatsspieren, en normaal sluiten van mond en ogen is afwezig. De mond kan er slap uitzien en het kind kan kwijlen.

2 Normale spanning van het gelaat. Het kind toont geen spierspanning, met normaal gesloten mond en ogen.

3 Spanning duidelijk in sommige gelaatsspieren. Lichte of infrequente spanning in sommige gelaatsspieren. Dit houdt geen aanhoudende spanning in van spiergroepen zoals de wenkbrauwen, het voorhoofd of de mond.

4 Spanning duidelijk in alle gelaatsspieren. Het kind laat opvallende, aanhoudende spanning zien van spiergroepen in het gelaat zoals de wenkbrauwen, het voorhoofd, de mond, kin of wangen.

5 Gelaatsspieren verwrongen en in een grimas. Het gezicht van het kind vertoont een grimas met een uitdrukking die de indruk geeft van huilen, ongemak, en onrust/angst. Dit is vaak te zien aan het extreem rimpelen van de wenkbrauwen en het voorhoofd, en een verwrongen mond.

Hartslag

Bij Hartslag wordt het aantal keren dat de geobserveerde waarde buiten de baselinewaarden valt, beoordeeld. De hartslag wordt berekend als de minimale normale waarde vastgelegd over de voorafgaande 24 uur (baselinewaarde). Het baselinegebied wordt begrenst door de waarde die 15% boven de baselinewaarde ligt (= bovengrens) en de waarde die 15% beneden de baselinewaarde ligt (= ondergrens). De beoordelaar observeert 6 keer gedurende de observatieperiode van twee minuten de hartslag, en noteert elke waarde in de kantlijn Beoordelingen worden gebaseerd op het aantal aflezingen boven of onder de grenswaarden.

1 Hartslag is lager dan de ondergrens van de baselinewaarde (1-6 keer).

2 Hartslag valt steeds binnen de grenswaarden (6 keer).

3 Incidentele verhogingen (1, 2 of 3 waarden boven de bovengrens).

4 Regelmatige verhogingen (4 of 5 waarden boven de bovengrens).

5 Hartslag is steeds hoger dan de bovengrens van baselinewaarde (6 keer).

Bloeddruk

Bij Bloeddruk wordt het aantal keren dat de geobserveerde waarde buiten de grenswaarden valt, beoordeeld. Het gaat hierbij om de gemiddelde arteriële bloeddruk (MAP). De bloeddruk wordt berekend als de minimale normale waarde vastgelegd over de voorafgaande 24 uur (baselinewaarde). Het baselinegebied wordt begrensd door de waarde die 15% boven de baselinewaarde ligt (=bovengrens) en de waarde die 15% beneden de baselinewaarde ligt (=ondergrens). De beoordelaar observeert 6 keer gedurende de observatieperiode van twee minuten de bloeddruk, en noteert elke waarde in de kantlijn. Beoordelingen worden gebaseerd op het aantal aflezingen boven of onder de grenswaarden.

1 Bloeddruk is lager dan de ondergrens van de baselinewaarde (1-6 keer).

2 Bloeddruk valt steeds binnen de grenswaarden (6 keer).

3 Incidentele verhogingen (1, 2 of 3 waarden boven de bovengrens).

4 Regelmatige verhogingen (4 of 5 waarden boven de bovengrens).

5 Bloeddruk is steeds hoger dan de bovengrens van de baselinewaarde (6 keer).

Bijlage 3 **COMFORT © schaal**

ALERTHEID

Diep in slaap (ogen dicht, geen reactie op omgeving)	1_____
Licht in slaap (ogen grotendeels gesloten, af en toe reactie)	2_____
Slaperig (kind sluit vaak zijn ogen, reageert minder op omgeving)	3_____
Wakker en alert (kind reageert op omgeving)	4_____
Wakker en hyper-alert (overdreven reactie op veranderingen)	5_____

KALMTE / AGITATIE

Kalm (kind lijkt helder en rustig)	1_____
Licht angstig (kind toont lichte onrust)	2_____
Angstig (kind lijkt onrustig maar kan zich beheersen)	3_____
Zeer angstig (kind lijkt zeer onrustig, kan zich nog net beheersen)	4_____
Paniekerig (ernstige onrust met verlies van beheersing)	5_____

ADEMHALINGSREACTIE

Geen hoesten en geen spontane ademhaling	1_____
Spontane ademhaling met weinig of geen reactie op de beademing	2_____
Af en toe hoesten of verzet tegen de ventilator	3_____
Ademt actief tegen de ventilator in of hoest regelmatig	4_____
Vecht tegen de ventilator; hoesten, verslikken, tegenademen	5_____

HUILEN

Geen huilgeluiden	1_____
Af en toe snikken of kreunen (nasnikken)	2_____
Jengelen of dreinen (monotoon geluid)	3_____
Huilen	4_____
Schreeuwen of krijsen	5_____

LICHAAMSBEWEGING

Geen beweging	1_____
Incidentele (3 of minder) kleine bewegingen	2_____
Frequente (meer dan 3) kleine bewegingen	3_____
Heftige bewegingen met armen en benen	4_____
Heftige bewegingen ook met romp en hoofd	5_____

Naam kind: _____ **Arteriële lijn** **ja/nee**
Datum: _____ **Beademing** **ja/nee**
Tijd: _____ **Observator:** _____

SPIERSPANNING

Spieren volledig ontspannen; geen spierspanning	1_____
Verminderde spierspanning; minder weerstand dan normaal	2_____
Normale spierspanning	3_____
Toegenomen spierspanning en buiging van vingers en tenen	4_____
Extreme spierstijfheid en buiging van vingers en tenen	5_____

GELAATSSPANNING

Gezichtsspieren volkomen ontspannen	1_____
Normale spanning van het gelaat	2_____
Spanning duidelijk in sommige gelaatsspieren (niet aanhoudend)	3_____
Spanning duidelijk in alle gelaatsspieren (aanhoudend)	4_____
Gelaatsspieren verwrongen en in een grimas	5_____

HARTSLAG BASELINE_____

Hartslag is verlaagd (1-6 waarden onder de ondergrens)	1_____
Hartslag valt steeds binnen de grenswaarden	2_____
Incidentele verhogingen (1-3 waarden boven de bovengrens)	3_____
Regelmatige verhogingen (4 of 5 waarden boven de bovengrens)	4_____
Hartslag is steeds verhoogd (6 waarden boven de bovengrens)	5_____

BLOEDDRUK BASELINE _____

Bloeddruk is verlaagd (1-6 waarden onder de ondergrens)	1_____
Bloeddruk valt steeds binnen de grenswaarden	2_____
Incidentele verhogingen (1-3 waarden boven de bovengrens)	3_____
Regelmatige verhogingen (4 of 5 waarden boven de bovengrens)	4_____
Bloeddruk is steeds verhoogd (6 waarden boven de bovengrens)	5_____

TOTAAL _____

Bijzonderheden:_____

_____

_____