PAIN Publish Ahead of Print
DOI: 10.1097/j.pain.0000000000001814

Quantification of neonatal procedural pain severity: a platform for estimating total pain burden in individual infants

Authors: Laudiano-Dray, Maria Pureza[a], Pillai Riddell, Rebecca R.[c,d], Jones, Laura[a], Iyer, Rajeshwari[a], Whitehead, Kimberley[a], Fitzgerald, Maria[a], Fabrizi, Lorenzo[a], Meek, Judith[b]

Affiliations:

[a]Department of Neuroscience, Physiology and Pharmacology, University College London, London, WC1E 6BT, United Kingdom

[b]Elizabeth Garrett Anderson Obstetric Wing, University College London Hospitals, London, WC1E 6DB, United Kingdom

[c]Department of Psychology, Faculty of Health, The O.U.C.H. Lab, York University, Toronto, Ontario, Canada,

[d]Hospital for Sick Children, Toronto, Ontario, Canada

Author Correspondence: Maria Pureza Laudiano-Dray

Address: Department of Neuroscience, Physiology and Pharmacology, Medawar Building, Room G17, University College London, Gower Street, London, WC1E 6BT

Telephone: +44 (0)20 76793533

Email: m.laudiano-dray@ucl.ac.uk

URL: https://www.ucl.ac.uk/biosciences/departments/npp

Number of text pages of the entire manuscript (including pages containing figures and tables) (17)

Number of figures (3)

Number of tables (1)

Abstract

There is increasing evidence that long-term outcomes for infants born prematurely are adversely affected by repeated exposure to noxious procedures. These interventions vary widely, for example, in the extent of damage caused and duration. NICU (neonatal intensive care unit) procedures are therefore likely to each contribute differently to the overall pain burden of individual neonates, ultimately having a different impact on their development. In order for researchers to quantify the procedural pain burden experienced by infants on NICU, we aimed to estimate the pain severity of common NICU procedures using published pain scores. We extracted pain scores over the first minute (pain reactivity) from the literature, using 59 randomized controlled trials for 15 different procedures. Hierarchical cluster analysis of average pain scores resulted in five discrete severity groups; mild (n=1), mild to moderate (n=3), moderate (n=7), severe (n=3) and very severe (n=1). The estimate of the severity of individual procedures provided new insight into infant pain reactivity which is not always directly related to the invasiveness and duration of a procedure; thus both heel lance and skin tape removal are moderately painful procedures. This estimate of procedural pain severity, based on pain reactivity scores, provides a novel platform for retrospective quantification of an individual neonate's pain burden due to NICU procedures. The addition of measures that reflect the recovery from each procedure, such as brain activity and behavioural regulation, would further improve estimates of the pain burden of neonatal intensive care.

Keywords: neonatal; preterm; infant; pain; nociception; PIPP; NFCS; procedure, pain burden; severity; NICU; pain score; reactivity; placebo; standard care

Introduction

Babies delivered preterm are subjected to repeated noxious clinically essential procedures which are likely to have adverse effects on their long term developmental outcomes [28; 62]. Recent studies have shown that infants who were born prematurely have altered brain structure [21] and functional pain and somatosensory development [63; 64] correlated with the number of painful interventions they have experienced during their neonatal care [10; 51]. Much of this important work quantifies pain exposure using the number of skin-breaking or acutely painful procedures. However, clinical procedures vary widely in nature and therefore might differ in their contribution to the overall burden of pain experienced by an individual baby. Adults do not consider clinical procedures to be equally painful [41; 49; 57]. Therefore, the possibility that infants also do not experience all acutely painful procedures equally merits further investigation. For example, during a routinely occurring heel lance, a lancet automatically releases a blade, cutting the skin of the heel to a depth of 0.65 to 2.5mm in a sharp and quick action. Venipuncture uses a needle of approximately 0.56mm in diameter to pierce the skin, subcutaneous tissue and eventually the vessel walls of the vein. The time required to complete the procedure depends on the skill of the operator. Endotracheal intubation involves introducing a 3.4-5.4 mm diameter tube via the mouth and passing it through the vocal cords with the help of a scope, to enter the windpipe (trachea). This procedure could take several attempts and the mean duration to successful intubation during neonatal resuscitation at birth could last 25-51 seconds [42]. Thus, it may not be correct to expect that the impact of 3 heel lances is the same of 3 different procedures such as endotracheal intubation, a venipuncture, and a heel lance.

The current study aims to estimate the pain severity of individual neonatal procedures in order to understand their specific contribution to the overall pain burden of hospitalized neonates. We

estimated the pain severity of NICU procedures by averaging neonatal pain scores derived from the literature and performing a hierarchical cluster analysis.

Methods

1. Identification of common clinical procedures

Epidemiological studies on NICU pain were used to identify a list of commonly occurring noxious clinical procedures experienced by newborn babies in hospital [6; 9; 11; 12; 15; 31; 50; 56]. The resulting 17 procedures were: heel lance, intramuscular injection, venipuncture, peripheral arterial puncture, lumbar puncture, nasogastric and orogastric tube insertion, nasal prongs insertion for continuous positive airway pressure (CPAP), naso/oropharyngeal suction, endotracheal suction, endotracheal intubation, chest tube insertion, chest tube removal, urethral catheterization, dressing change or adhesive tape removal, eye examination and instillation of eye drops.

2.1 Selection of publications for pain score extraction

Neonatal pain scores for each procedure were extracted from randomized controlled trials (RCTs) that tested analgesic interventions compared to a no-treatment control group. Suitable RCTs were identified in two steps. First, all reviews within the Neonatal and Pain & Anesthesia Groups (n= 850) of the Cochrane Database of Systematic Reviews were searched for the following key terms; neonate, newborn, neonatal, infant, baby and pain. Eleven reviews were selected as suitable, and searched for articles describing RCTs involving the procedures listed above [7; 8; 17; 25; 33; 44; 46; 53-55; 60].

A second literature search was conducted to find RCTs that involved the procedures that could not be found in the Cochrane Reviews. These were nasal prongs insertion for CPAP, endotracheal intubation, chest tube insertion, chest tube removal, and urethral catheterization. Searches were conducted in Ovid PubMed, and EBSCO Host using key words related to these procedures, neonatal and the names of commonly used pain scores. In total, 227 articles were found through the Cochrane Database and 488 articles were found through other database searching after duplicates were removed with 619 articles screened for inclusion.

The following inclusion criteria were used: (i) the study reported either a validated behavioural or bio-behavioural neonatal pain assessment tool (e.g. Neonatal Facial Coding System or Premature Infant Pain Profile) or a behavioural pain indicator (e.g. crying time (cry) or proportion of time spent in facial action or grimace (grimace)). Studies were included if they quoted a mean and standard deviation or median and range score. (ii) The score was measured intra-procedurally (e.g. at insertion of device or skin break) or up to one minute after the procedure; (iii) the subjects were preterm neonates up to 44 weeks postmenstrual age or full term neonates up to 30 postnatal days at time of study. All the neonates whose scores were included had been assigned to the placebo arm of randomized controlled trials. They had therefore received (a) inactive ingredients topically, orally, or intravenously (sterile water, saline or inert lotions), (b) standard care (including nesting, therapeutic touch and odours which have been shown to have no effect on pain behaviours [47]), (c) no interventions, either immediately before and/or during the painful procedure, or (d) no continuous analgesia or sedation. The responses of the infants in the intervention arms of the RCTs were not included in this study.

Based on consensus discussion with the research team, there were minor exceptions to the aforementioned inclusion criteria. For eye examination, all nine studies reviewed involved the use of topical anaesthetic eye drops for all subjects, as this is a standard of care. Of these studies, two were considered suitable for inclusion because the neonates in the control arms did not receive any other active intervention prior to and during the eye examination [37; 52]. In addition, only one of the nine eye examination studies measured the pain response to the instillation of topical anaesthetic eye drops separately and although the control group received a pacifier 1.5 minutes before the procedure [40], this study was included. For reported pain scores from nasal prongs insertion, one of the two studies included had 11 (30%) neonates in the CPAP group who were being weaned from mechanical ventilation prior to the procedure, but it was unclear whether they were still receiving continuous analgesia in the hours leading to the application of CPAP [45]. Standard clinical practice in most NICUs is to wean the administration of analgesics and sedatives prior to discontinuing invasive mechanical ventilation, therefore this study was included. In the same study, non-randomized patients were assigned to receive CPAP (versus heated, humidified high-flow nasal cannulae) based on clinical need, however all the other criteria for a controlled observation of pain behaviour were met [45].

Of the 274 full text articles reviewed, 215 were excluded (Figure 1). 59 eligible neonatal studies were included, with 3 of 59 studies reporting pain scores for two procedures, resulting in the data extraction for 15 of the previously described 17 procedures. There were no eligible studies for chest tube insertion and chest tube removal that separately reported the pain scores of our selected age group. Supplementary Table 1 (available at http://links.lww.com/PAIN/A950) describes the study participants, control arm or standard care provided and pain scores extracted for all included studies.

2.2 Selection of pain assessment tool to be extracted per study

We found a variety of pain assessment tools in the literature. Unidimensional scores record only one attribute such as behaviour (DAN and NFCS), or crying time (cry) or grimace. Unidimensional physiological scores (e.g. heart rate or oxygen saturation) are not thought to be specific for pain reactivity. Multidimensional scores, also referred to as composite scores, record both physiological and behavioural measures (e.g. NIPS), and may also include gestational age (e.g. PIPP). Reviews recommend the use of pain assessment tools which have been validated and are multidimensional over unidimensional tools [22].

We selected one pain assessment tool per study based on a hierarchy selection process which we devised for this purpose (see Supplementary Figure 1, available at http://links.lww.com/PAIN/A950). If more than one pain measure was used, precedence was given to validated pain assessment tools over behavioural indicators such as grimacing, cry duration, cry time. If two or more validated neonatal pain assessment tools were reported per study, the most frequently used tool used for that particular procedure was selected in order to reduce the variability of extracted pain scores per procedure. If there were an equal number of validated pain assessment tools eligible for extraction for a particular procedure (i.e. 2 Premature Infant Pain Profile and 2 Neonatal Facial Coding System) then a hierarchy of pain tools for extraction was designed based on its psychometric properties [22]. Thus, validated multidimensional behavioural scales were prioritized over validated unidimensional behavioural scales.

2.3 Extraction of pain scores

We selected the 'pain reactivity' score that was taken closest to the procedural event. The reported scores could be from a single epoch, proportion of time or a cumulative score. This procedural event is defined as the point of skin break or insertion of device. In some studies the score was applied 'during' the procedure. For example, if study observations were segmented into phases or time epochs that described the duration of the intra-procedure, then the score that was reported for that phase or epoch was extracted. If two scores were reported, for example, first 30 seconds and first minute from the procedure, then the score during the first 30 seconds was extracted. If duration of the intra-procedure was reported per minute, then the first minute (sometimes described as '0' minute) was extracted. If the duration of observations was not immediately clear but the mean duration of the procedure was within 1 minute, the score was extracted. If two observers reported a pain score each using the selected pain assessment tool per included study, the mean score was calculated.

2.4. Transformation, pooling and ranking of neonatal pain reactivity scores

For each eligible study, the mean and standard deviation or median and range pain scores were extracted. In studies where the median was reported, the estimated mean and standard deviation were calculated according to Hozo's method [30]. The extracted mean and standard deviation scores were normalized to a 0-100% scale by calculating the proportion of the pain score from the maximum possible score of the scale used. For example, a Premature Infant Pain Profile score of 7 out of a range of 0-21 was transformed into a normalized score of 33.33%. We then pooled the normalized pain scores per procedure by calculating the average and standard

deviation, weighted for pooled sample size. The weighted average pain scores per procedure were ranked in order and displayed using a forest plot.

3. Statistical analysis

Statistical analyses were performed using SPSS (IBM Corp, Version 22). We employed a hierarchical cluster analysis using the between-groups linkage method based on squared euclidean distances [13]. The optimum number of clusters was determined using the agglomeration schedule coefficients as a measure of within cluster variation (Elbow method).

Results

Neonatal participants with a gestational age range of 23-43 completed weeks provided a total of 2281 pain scores that were extracted from fifty nine studies (Table 1). Preterm born participants who were studied at postmenstrual age of less than 36 weeks and 6 days contributed 41% (n=936) of the scores while term age participants who were up to one month old at time of study provided 47% (n=1081) of the scores. 12% (n=264) of scores were from mixed ages (preterm and term participants). Pooling of pain scores was possible for 15 of 17 chosen procedures, as there were no suitable studies for chest tube insertion and chest tube removal. The largest pooled samples per procedure were for heel lance (n=895), venipuncture (n=450) and intramuscular injection (n=264). There were almost an equal number of preterm and term participants for heel lance whereas over 71% were term participants for venipuncture. Preterm participants were represented in all but two procedures, namely intramuscular injection and peripheral arterial puncture, whereas term participants were represented in 8 of 15 procedures. A small number of

study participants had an endotracheal tube immediately prior to or during the procedure, namely those that received endotracheal suction (n=90), nasal prongs insertion for CPAP (n=21) and endotracheal intubation (n=20).

Supplementary Table 1 (available at http://links.lww.com/PAIN/A950) lists all included studies and description of the control participants and extracted pain scores. Three studies provided scores for two different procedures, (heel lance and venipuncture [39; 43]; heel lance and naso/oropharyngeal suction [5]). Data extraction for 7 procedures came from single studies with small sample sizes (range of n=15-52). The most commonly extracted pain scores were the Premature Infant Pain Profile (30 studies), Neonatal Facial Coding System (16 studies) and Neonatal Infant Pain Scale (8 studies). Supplementary Table 2 (available at http://links.lww.com/PAIN/A950) lists the pain tools extracted per procedure.

Figure 2 shows a forest plot summarizing the ranking of mean pain scores of each procedure, arranged from highest to lowest on a normalized 0-100% scale. Lumbar puncture, peripheral arterial puncture, endotracheal intubation and intramuscular injection were the highest ranking procedures with average pain scores greater than 70%. Instillation of anaesthetic eye drops scored less than 30%. The threshold for providing pain relief is marked at 30% on a 0-100% scale based on the self-report of children and adolescents [18; 29].

Hierarchical cluster analysis of the data revealed that each procedure fell into one of five severity groups: mild (n=1), mild to moderate (n=3), moderate (n=7), severe (n=3) and very severe (n=1), shown as five bands in Figure 2. The number of clusters was validated by the elbow method by plotting the agglomeration schedule coefficients against number of clusters shown in Figure 3.

At stage 9, the agglomeration schedule coefficients begin to ‘elbow’ and at this point the data clusters into five groups.

The analysis shows that the mild and very severe groups contained only one procedure, eye drops instillation and lumbar puncture respectively, and were derived from limited studies with a small number of subjects. The three middle groups, mild to moderate (n=3), moderate (n=7), severe (n=3) contained the remaining 13 procedures, with the most common severity group being moderate. The moderate group contained procedures of very different duration and invasiveness, such as tape removal, heel lance and endotracheal suction.

Discussion

Infants in intensive care experience numerous clinically required, tissue-damaging procedures. However due to their inability to self-report, researchers and clinicians have been unable to quantify the total pain experience or burden of individual babies, using a method that takes not only the number, but also the severity of each noxious procedure, into account. In order to estimate the pain severity of different NICU procedures, we have used published pain scores derived from neonatal pain assessment tools during the first minute after these procedures (pain reactivity). The results show that NICU procedures fall into five groups of pain severity and that the majority of procedures are in three middle severity groups: mild to moderate, moderate, and severe. The two outside groups, very severe and mild, each contain one procedure represented by single studies with a small number of subjects making their definition less reliable. Seven procedures were within the moderate range, namely eye examination, endotracheal and naso/oropharyngeal suctioning, venipuncture, tape removal, nasogastric tube insertion and heel

lance which are some of the most frequently performed in the NICU [6; 9; 11; 12; 15; 31; 50; 56].

It is interesting that the procedures are not predictably ranked by duration or tissue invasiveness. They also do not correspond directly with clinicians' estimates [2; 3; 20; 32; 48; 56] which may be formed from a composite of observations of infant behavioural responses, personal experience, adult perception [41; 49; 57] and self-report by verbal patients [4; 26]. One reason for this could be that we confined the measurement of pain severity to the 'reactivity score', or the score up to the first minute after the noxious event. However it is important to consider the known limitations of neonatal pain assessment scores [58]. Facial changes and movements are assessed crudely in neonatal tools, and many groups working with older infants use greater temporal resolution in their assessment of facial behaviours [19]. The limitation of behavioural scores as measures of pain has been highlighted by their poor correlation with brain cortical pain responses [58], particularly in situations where infants have high levels of physiological stress [36]. Evidence from fMRI studies suggest that the distribution of the cortical BOLD response to a sharp pinprick is similar in newborns and adults [27] suggesting that the basic brain connections required to experience pain are present, but does not validate the underlying assumption that an observational infant pain score can be described in terms of adult pain experience. Indeed neonatal pain behaviour reflects the lack of inhibitory controls in immature pain pathways [24] and the differing pattern of cortical activity following an adult and a neonatal heel lance, highlights their functional differences in brain processing [23]. Nevertheless, to date, pain scores remain the only quantitative measure of responses to the range of NICU procedures that clinicians have available and provide the only method of differentiating the severity of individual routine procedures.

Our study also shows wide variability of reported pain scores per procedure. We have used normalized data from different behavioural pain scores and assumed that they are directly comparable and can be pooled. Although this is not always the case [61], a recent study reported good comparability between the most commonly used pain tools in both research and clinical settings [38]. In clinical practice and throughout the literature it is recognized that the magnitude and response time of procedural pain scores vary depending on factors such as gestational age, postnatal age, sleep state, time since last procedure [35], caregiving practice and operator skill. Therefore, it is not surprising that the pooled scores have wide ranges, and interesting that clear clusters of pain severity have emerged.

A limitation of the study is the assumption that pain scores for babies born both prematurely and at term can be directly combined. Both the Premature Infant Pain Profile and Neonatal Facial Coding System scores included adjustments for prematurity and these were the most frequently represented scoring systems in the study. However babies born preterm have been shown to have longer latencies for their facial response, which vary with postmenstrual age [59], and smaller PIPP scores as already discussed [35]. The results include an assumption that the responses to all procedures develop in parallel over time.

Pain is a multifaceted experience with multiple physiological responses, but only a subset of these are accessible when studying infants. We have attempted to derive a measure of pain severity from the most commonly used pain assessment tools, while acknowledging their limitations. We have categorized the severity of different NICU procedures so that future studies can ascribe a pain burden associated with a hospital admission for neonatal care retrospectively. However, by using pain reactivity scores alone, the emphasis may be more upon the brief ‘hurt’ of a noxious event, that might act as a ‘pain switch’ to activate later, more important responses

[16], rather than measure the total impact of that event upon the infant brain and physiological homeostasis. Our estimation of pain severity may be improved by including later phases or recovery components of the response [1; 34], both in brain and behaviour, but these aspects cannot currently be extracted from the literature in a consistent manner. This estimation of pain severity is not intended to replace bedside pain assessment practice. Rather, it allows researchers to better quantify pain burden from retrospective chart review beyond simple counts of acutely painful procedures. Given procedures are more reliably recorded than pain scores, this information will allow for better chart review quantification of an infant's NICU pain burden. However, it is important to recognize that acute procedures form only one component of the NICU pain experience. Future work needs to quantify other factors such as episodes of mechanical ventilation and surgery.

The clinical implications of this analysis are limited by the small numbers in the studies of the least and most painful procedures, and by the pooling of data for babies of different gestational and postnatal ages. The study was not intended to change clinical practice, but rather to be considered as a research tool. However it is not surprising that lumbar puncture has the highest pain score, and most guidelines recommend using analgesia for this procedure [14]. Our results re-inforce the need to provide appropriate analgesia to infants undergoing NICU procedures, as we have shown that the response to each procedure is subject to a wide range of individual variability. The use of pain relief intervention would modify our results. The study is not relevant for guidelines for managing post-operative pain.

In conclusion, we have shown that common noxious NICU procedures can be grouped according to their pain reactivity score. Our findings provide some preliminary insights but in view of the limitations, we require further elucidation of factors in addition to procedural pain to calculate

the total pain burden, particularly for babies who have undergone surgery. This analysis provides a first step in the process of quantifying this pain burden for individual babies, which is essential if we are to measure the adverse impact of repeated painful procedures and introduce preventative strategies. We suggest that the addition of bio-behavioural and cortical 'recovery' activity measures will refine this concept in the future.

Acknowledgements

The authors have no conflicts of interest to declare. This work was funded by the Medical Research Council UK (MR/M006468/1, MR/L019248/1 and MR/S003207/1). The work was also supported by an International Association for the Study of Pain's Collaborative Research Grant and a grant from the Canadian Institutes of Health Research (CHRP 538853-19).

(1) Contributions that need acknowledging but do not justify authorship:

Costa, Stefano[c]; Rumero, Carla[c]

References

[1] Ahola Kohut S, Pillai Riddell R. Does the Neonatal Facial Coding System differentiate between infants experiencing pain-related and non-pain-related distress? J Pain 2009;10(2):214-220.

[2] Akuma AO, Jordan S. Pain management in neonates: a survey of nurses and doctors. J Adv Nurs 2012;68(6):1288-1301.

[3] Anand KJS, Eriksson M, Boyle EM, Avila-Alvarez A, Andersen RD, Sarafidis K, Polkki T, Matos C, Lago P, Papadouri T, Attard-Montalto S, Ilmoja ML, Simons S, Tameliene R, van Overmeire B, Berger A, Dobrzanska A, Schroth M, Bergqvist L, Courtois E, Rousseau J, Carbajal R, Consortium EswgotN. Assessment of continuous pain in newborns admitted to NICUs in 18 European countries. Acta Paediatr 2017;106(8):1248-1259.

[4] Andersen RD, Nakstad B, Jylli L, Campbell-Yeo M, Anderzen-Carlsson A. The Complexities of Nurses' Pain Assessment in Hospitalized Preverbal Children. Pain Manag Nurs 2019.

[5] Axelin A, Salantera S, Kirjavainen J, Lehtonen L. Oral glucose and parental holding preferable to opioid in pain management in preterm infants. Clin J Pain 2009;25(2):138-145.

[6] Barker DP, Rutter N. Exposure to invasive procedures in neonatal intensive care unit admissions. Arch Dis Child Fetal Neonatal Ed 1995;72(1):F47-48.

[7] Beirne PV, Hennessy S, Cadogan SL, Shiely F, Fitzgerald T, MacLeod F. Needle size for vaccination procedures in children and adolescents. Cochrane Database Syst Rev 2015(6):CD010720.

[8] Brady-Fryer B, Wiebe N, Lander JA. Pain relief for neonatal circumcision. Cochrane Database Syst Rev 2004(4):CD004217.

[9] Britto CD, Rao Pn S, Nesargi S, Nair S, Rao S, Thilagavathy T, Ramesh A, Bhat S. PAIN--perception and assessment of painful procedures in the NICU. J Trop Pediatr 2014;60(6):422-427.

[10] Brummelte S, Grunau RE, Chau V, Poskitt KJ, Brant R, Vinall J, Gover A, Synnes AR, Miller SP. Procedural pain and brain development in premature newborns. Ann Neurol 2012;71(3):385-396.

[11] Carbajal R, Rousset A, Danan C, Coquery S, Nolent P, Ducrocq S, Saizou C, Lapillonne A, Granier M, Durand P, Lenclen R, Coursol A, Hubert P, de Saint Blanquat L, Boelle PY, Annequin D, Cimerman P, Anand KJ, Breart G. Epidemiology and treatment of painful procedures in neonates in intensive care units. JAMA 2008;300(1):60-70.

[12] Cignacco E, Hamers J, van Lingen RA, Stoffel L, Buchi S, Muller R, Schutz N, Zimmermann L, Nelle M. Neonatal procedural pain exposure and pain management in ventilated preterm infants during the first 14 days of life. Swiss Med Wkly 2009;139(15-16):226-232.

[13] Clatworthy J, Buick D, Hankins M, Weinman J, Horne R. The use and reporting of cluster analysis in health psychology: a review. Br J Health Psychol 2005;10(Pt 3):329-358.

[14] Committee On F, Newborn, Section On A, Pain M. Prevention and Management of Procedural Pain in the Neonate: An Update. Pediatrics 2016;137(2):e20154271.

[15] Cruz MD, Fernandes AM, Oliveira CR. Epidemiology of painful procedures performed in neonates: A systematic review of observational studies. Eur J Pain 2016;20(4):489-498.

[16] Davis KD, Kucyi A, Moayedi M. The pain switch: an "ouch" detector. Pain 2015;156(11):2164-2166.

[17] Dempsey E, McCreery K. Local anaesthetic eye drops for prevention of pain in preterm infants undergoing screening for retinopathy of prematurity. Cochrane Database Syst Rev 2011(9):CD007645.

[18] Demyttenaere S, Finley GA, Johnston CC, McGrath PJ. Pain treatment thresholds in children after major surgery. Clin J Pain 2001;17(2):173-177.

[19] DiLorenzo MG, Pillai Riddell R, Flora DB, Craig KD. Infant Clinical Pain Assessment: Core Behavioral Cues. J Pain 2018;19(9):1024-1032.

[20] Dodds E. Neonatal procedural pain: a survey of nursing staff. Paediatr Nurs 2003;15(5):18-21.

[21] Duerden EG, Grunau RE, Guo T, Foong J, Pearson A, Au-Young S, Lavoie R, Chakravarty MM, Chau V, Synnes A, Miller SP. Early Procedural Pain Is Associated with Regionally-Specific Alterations in Thalamic Development in Preterm Neonates. J Neurosci 2018;38(4):878-886.

[22] Duhn LJ, Medves JM. A systematic integrative review of infant pain assessment tools. Adv Neonatal Care 2004;4(3):126-140.

[23] Fabrizi L, Verriotis M, Williams G, Lee A, Meek J, Olhede S, Fitzgerald M. Encoding of mechanical nociception differs in the adult and infant brain. Sci Rep 2016;6:28642.

[24] Fitzgerald M. What do we really know about newborn infant pain? Experimental Physiology 2015;100(12):1451-1457.

[25] Foster JP, Taylor C, Spence K. Topical anaesthesia for needle-related pain in newborn infants. Cochrane Database Syst Rev 2017;2:CD010331.

[26] Gibbins S, Stevens B, Dionne K, Yamada J, Pillai Riddell R, McGrath P, Asztalos E, O'Brien K, Beyene J, McNamara P, Johnston C. Perceptions of health professionals on pain in extremely low gestational age infants. Qual Health Res 2015;25(6):763-774.

[27] Goksan S, Hartley C, Emery F, Cockrill N, Poorun R, Moultrie F, Rogers R, Campbell J, Sanders M, Adams E, Clare S, Jenkinson M, Tracey I, Slater R. fMRI reveals neural activity overlap between adult and infant pain. Elife 2015;4.

[28] Grunau RE. Neonatal pain in very preterm infants: long-term effects on brain, neurodevelopment and pain reactivity. Rambam Maimonides Med J 2013;4(4):e0025.

[29] Hirschfeld G, Zernikow B. Cut points for mild, moderate, and severe pain on the VAS for children and adolescents: what can be learned from 10 million ANOVAs? Pain 2013;154(12):2626-2632.

[30] Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. BMC Med Res Methodol 2005;5:13.

[31] Jeong IS, Park SM, Lee JM, Choi YJ, Lee J. The frequency of painful procedures in neonatal intensive care units in South Korea. Int J Nurs Pract 2014;20(4):398-407.

[32] Jeong IS, Park SM, Lee JM, Choi YJ, Lee J. Perceptions on pain management among Korean nurses in neonatal intensive care units. Asian Nurs Res (Korean Soc Nurs Sci) 2014;8(4):261-266.

[33] Johnston C, Campbell-Yeo M, Disher T, Benoit B, Fernandes A, Streiner D, Inglis D, Zee R. Skin-to-skin care for procedural pain in neonates. Cochrane Database Syst Rev 2017;2:CD008435.

[34] Johnston CC, Filion F, Campbell-Yeo M, Goulet C, Bell L, McNaughton K, Byron J, Aita M, Finley GA, Walker C-D. Kangaroo mother care diminishes pain from heel lance in very preterm neonates: A crossover trial. BMC Pediatrics 2008;8(1):13.

[35] Johnston CC, Stevens BJ, Franck LS, Jack A, Stremler R, Platt R. Factors explaining lack of response to heel stick in preterm newborns. J Obstet Gynecol Neonatal Nurs 1999;28(6):587-594.

[36] Jones L, Fabrizi L, Laudiano-Dray M, Whitehead K, Meek J, Verriotis M, Fitzgerald M. Nociceptive Cortical Activity Is Dissociated from Nociceptive Behavior in Newborn Human Infants under Stress. Curr Biol 2017;27(24):3846-3851 e3843.

[37] Kabatas EU, Dursun A, Beken S, Dilli D, Zenciroglu A, Okumus N. Efficacy of Single Dose Oral Paracetamol in Reducing Pain During Examination for Retinopathy of Prematurity: A Blinded Randomized Controlled Trial. Indian J Pediatr 2016;83(1):22-26.

[38] Kappesser J, Kamper-Fuhrmann E, de Laffolie J, Faas D, Ehrhardt H, Franck LS, Hermann C. Pain-specific Reactions or Indicators of a General Stress Response?: Investigating the Discriminant Validity of 5 Well-established Neonatal Pain Assessment Tools. Clin J Pain 2019;35(2):101-110.

[39] Larsson BA, Tannfeldt G, Lagercrantz H, Olsson GL. Venipuncture is more effective and less painful than heel lancing for blood tests in neonates. Pediatrics 1998;101(5):882-886.

[40] Mitchell A, Stevens B, Mungan N, Johnson W, Lobert S, Boss B. Analgesic effects of oral sucrose and pacifier during eye examinations for retinopathy of prematurity. Pain Manag Nurs 2004;5(4):160-168.

[41] Morrison RS, Ahronheim JC, Morrison GR, Darling E, Baskin SA, Morris J, Choi C, Meier DE. Pain and discomfort associated with common hospital procedures and experiences. J Pain Symptom Manage 1998;15(2):91-101.

[42] O'Donnell CP, Kamlin CO, Davis PG, Morley CJ. Endotracheal intubation attempts during neonatal resuscitation: success rates, duration, and adverse effects. Pediatrics 2006;117(1):e16-21.

[43] Ogawa S, Ogihara T, Fujiwara E, Ito K, Nakano M, Nakayama S, Hachiya T, Fujimoto N, Abe H, Ban S, Ikeda E, Tamai H. Venepuncture is preferable to heel lance for blood sampling in term neonates. Arch Dis Child Fetal Neonatal Ed 2005;90(5):F432-436.

[44] Ohlsson A, Shah PS. Paracetamol (acetaminophen) for prevention or treatment of pain in newborns. Cochrane Database Syst Rev 2016;10:CD011219.

[45] Osman M, Elsharkawy A, Abdel-Hady H. Assessment of pain during application of nasal-continuous positive airway pressure and heated, humidified high-flow nasal cannulae in preterm infants. J Perinatol 2015;35(4):263-267.

[46] Pillai Riddell RR, Racine NM, Gennis HG, Turcotte K, Uman LS, Horton RE, Ahola Kohut S, Hillgrove Stuart J, Stevens B, Lisi DM. Non-pharmacological management of infant and young child procedural pain. Cochrane Database Syst Rev 2015(12):CD006275.

[47] Pillai Riddell RR, Racine NM, Turcotte K, Uman LS, Horton RE, Din Osmun L, Ahola Kohut S, Hillgrove Stuart J, Stevens B, Gerwitz-Stern A. Non-pharmacological management of infant and young child procedural pain. Cochrane Database Syst Rev 2011(10):CD006275.

[48] Porter FL, Wolf CM, Gold J, Lotsoff D, Miller JP. Pain and pain management in newborn infants: a survey of physicians and nurses. Pediatrics 1997;100(4):626-632.

[49] Puntillo KA, Max A, Timsit JF, Vignoud L, Chanques G, Robleda G, Roche-Campo F, Mancebo J, Divatia JV, Soares M, Ionescu DC, Grintescu IM, Vasiliu IL, Maggiore SM, Rusinova K, Owczuk R, Egerod I, Papathanassoglou ED, Kyranou M, Joynt GM, Burghi

G, Freebairn RC, Ho KM, Kaarlola A, Gerritsen RT, Kesecioglu J, Sulaj MM, Norrenberg M, Benoit DD, Seha MS, Hennein A, Periera FJ, Benbenishty JS, Abroug F, Aquilina A, Monte JR, An Y, Azoulay E. Determinants of procedural pain intensity in the intensive care unit. The Europain(R) study. Am J Respir Crit Care Med 2014;189(1):39-47.

[50] Roofthooft DW, Simons SH, Anand KJ, Tibboel D, van Dijk M. Eight years later, are we still hurting newborn infants? Neonatology 2014;105(3):218-226.

[51] Schneider J, Duerden EG, Guo T, Ng K, Hagmann P, Bickle Graz M, Grunau RE, Chakravarty MM, Huppi PS, Truttmann AC, Miller SP. Procedural pain and oral glucose in preterm neonates: brain development and sex-specific effects. Pain 2018;159(3):515-525.

[52] Seifi F, Peirovifar A, Gharehbaghi MM. Comparing the Efficacy of Oral Sucrose and Acetaminophen in Pain Relief for Ophthalmologic Screening of Retinopathy of Prematurity. American Journal of Medical Sciences and Medicine 2013;1(2):24-27.

[53] Shah PS, Herbozo C, Aliwalas LL, Shah VS. Breastfeeding or breast milk for procedural pain in neonates. Cochrane Database Syst Rev 2012;12:CD004950.

[54] Shah PS, Shah VS. Propofol for procedural sedation/anaesthesia in neonates. Cochrane Database Syst Rev 2011(3):CD007248.

[55] Shah VS, Ohlsson A. Venepuncture versus heel lance for blood sampling in term neonates. Cochrane Database Syst Rev 2011(10):CD001452.

[56] Simons SH, van Dijk M, Anand KS, Roofthooft D, van Lingen RA, Tibboel D. Do we still hurt newborn babies? A prospective study of procedural pain and analgesia in neonates. Arch Pediatr Adolesc Med 2003;157(11):1058-1064.

[57] Singer AJ, Richman PB, Kowalska A, Thode HC, Jr. Comparison of patient and practitioner assessments of pain from commonly performed emergency department procedures. Ann Emerg Med 1999;33(6):652-658.

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

[59] Slater R, Cantarella A, Yoxen J, Patten D, Potts H, Meek J, Fitzgerald M. Latency to facial expression change following noxious stimulation in infants is dependent on postmenstrual age. Pain 2009;146(1-2):177-182.

[60] Stevens B, Yamada J, Ohlsson A, Haliburton S, Shorkey A. Sucrose for analgesia in newborn infants undergoing painful procedures. Cochrane Database Syst Rev 2016;7:CD001069.

[61] Uyan ZS, Bilgen H, Topuzoglu A, Akman I, Ozek E. Comparison of three neonatal pain scales during minor painful procedures. J Matern Fetal Neonatal Med 2008;21(5):305-308.

[62] Walker SM. Long-term effects of neonatal pain. Semin Fetal Neonatal Med 2019;24(4):101005.

[63] Walker SM, Franck LS, Fitzgerald M, Myles J, Stocks J, Marlow N. Long-term impact of neonatal intensive care and surgery on somatosensory perception in children born extremely preterm. Pain 2009;141(1-2):79-87.

[64] Walker SM, Melbourne A, O'Reilly H, Beckmann J, Eaton-Rosen Z, Ourselin S, Marlow N. Somatosensory function and pain in extremely preterm young adults from the UK EPICure cohort: sex-dependent differences and impact of neonatal surgery. Br J Anaesth 2018;121(3):623-635.

Table 1. Number of included studies and neonatal subjects per procedure classified according to gestational age and mean postmenstrual age (preterm, mixed (term and preterm), term)

Procedure	Number of studies	Number of neonates	Number (%) of preterm infants	Number (%) of mixed age infants	Number (%) of term infants
Endotracheal intubation	1	20	20 (100)		
Endotracheal suction	3	90	56 (62)	34 (38)	
Eye drops instillation	1	15	15 (100)		
Eye examination	2	95	39 (41)	56 (59)	
Heel lance	26	895	444 (50)		451 (50)
Intramuscular injection	4	264			264 (100)
Lumbar puncture	1	30		30 (100)	
Nasal prongs insertion for CPAP	2	67	67 (100)		
Nasogastric tube insertion	3	122	73 (60)	49 (40)	
Naso/Oro Pharyngeal suction	1	20	20 (100)		
Orogastric tube insertion	1	52	52 (100)		
Peripheral arterial puncture	1	45			45 (100)
Tape removal	1	15	15 (100)		
Urethral catheterization	3	101	59 (58)	42 (42)	
Venipuncture	12	450	20 (4)	109 (24)	321 (71)

Total studies included exceeds n = 59 as three articles provided data for more than one procedure

Legends

Figure 1. PRISMA 2009 Flow Diagram of Neonatal Studies

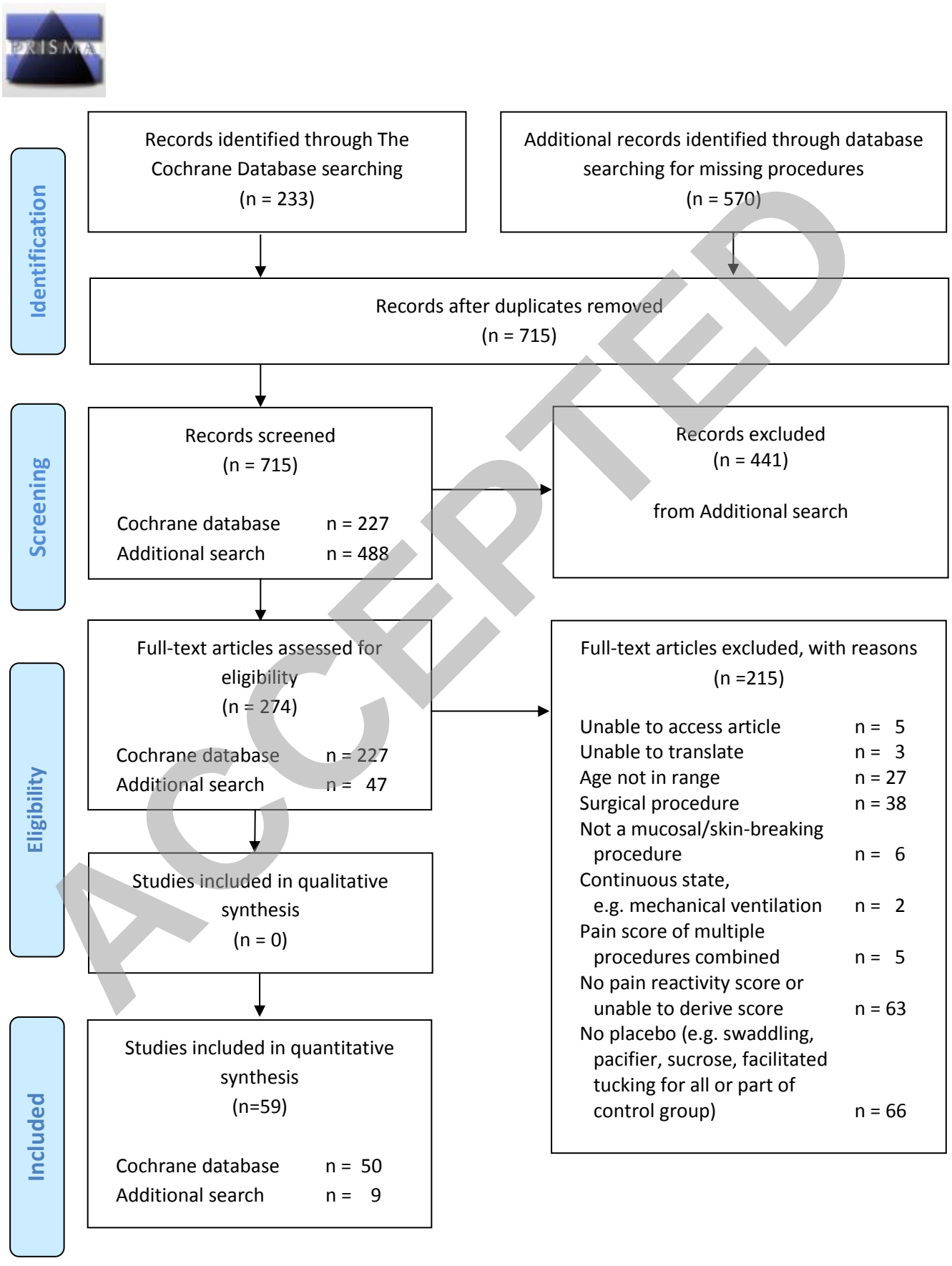
Figure 2. Forest plot of mean normalized pain scoresForest plot showing 15 procedures ranked by mean pain scores normalized to a 0-100% scale. Five clusters (shaded areas) of severity were derived from hierarchical cluster analysis. ◊ =weighted average, data points 5-13 according to pooled neonatal sample size, smallest (5) = 1-100, value increasing by +100. Line bars = weighted standard deviation.

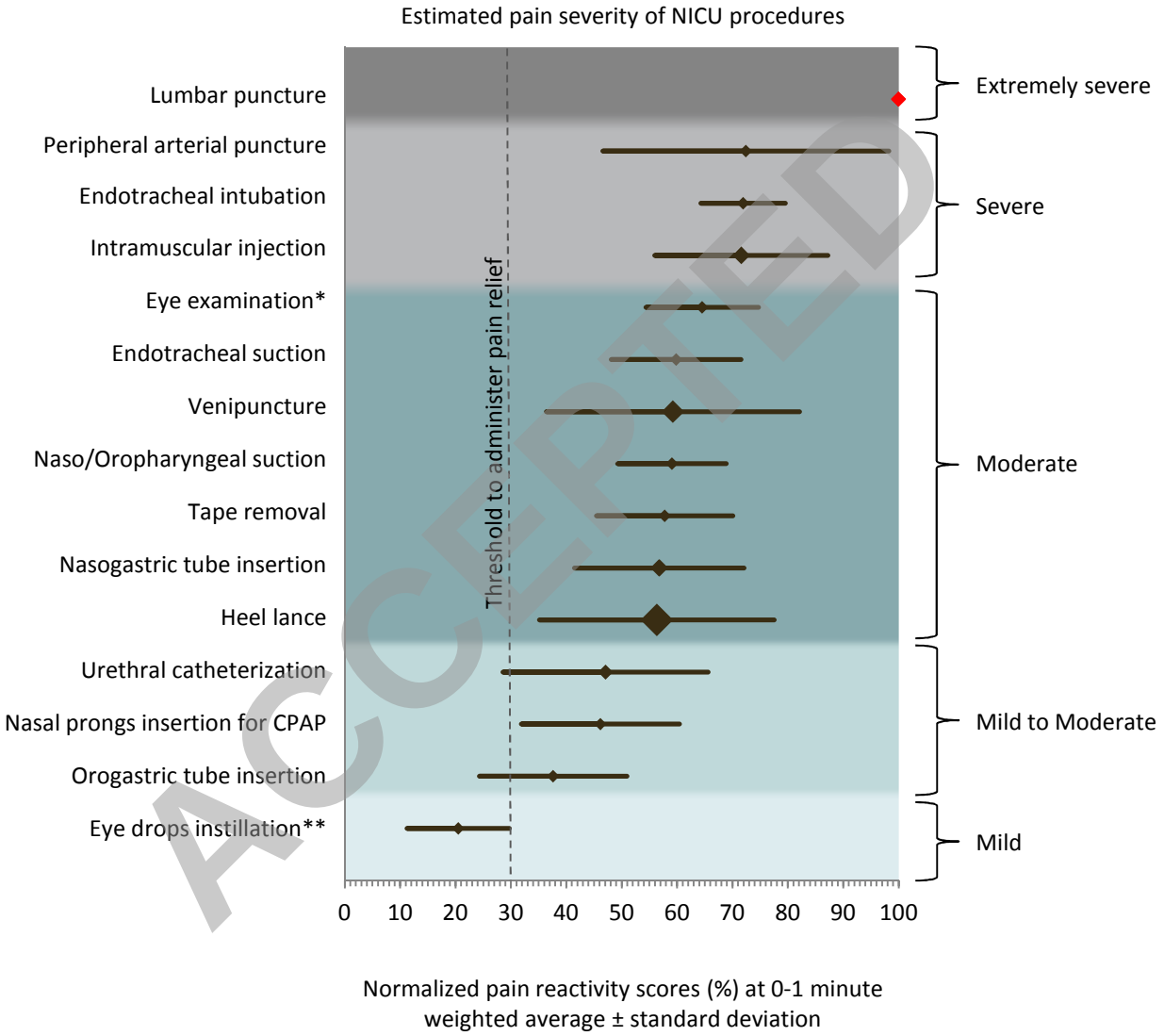
*all control groups in the neonatal eye examination studies received topical anaesthetic ophthalmic drops

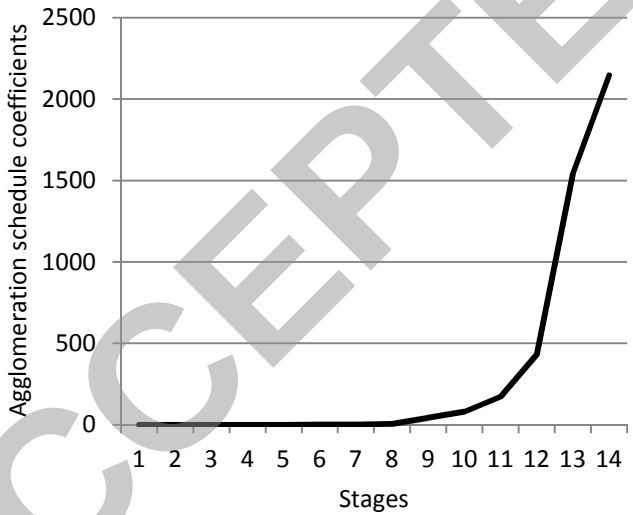
**control group received a pacifier before topical anaesthetic ophthalmic drops instillation

Figure 3. Agglomeration schedule coefficients

'Stages' represent sequential steps of the hierarchical cluster analysis


Identification
Records identified through The Cochrane Database searching (n = 233)
Additional records identified through database searching for missing procedures (n = 570)
Records after duplicates removed (n = 715)
Screening
Records screened (n = 715)
Cochrane database n = 227
Additional search n = 488
Records excluded (n = 441)
from Additional search
Eligibility
Full-text articles assessed for eligibility (n = 274)
Cochrane database n = 227
Additional search n = 47
Full-text articles excluded, with reasons (n =215)
Unable to access article n = 5
Unable to translate n = 3
Age not in range n = 27
Surgical procedure n = 38
Not a mucosal/skin-breaking procedure n = 6
Continuous state, e.g. mechanical ventilation n = 2
Pain score of multiple procedures combined n = 5
No pain reactivity score or unable to derive score n = 63
No placebo (e.g. swaddling, pacifier, sucrose, facilitated tucking for all or part of control group) n = 66
Studies included in qualitative synthesis (n = 0)
Included
Studies included in quantitative synthesis (n=59)
Cochrane database n = 50
Additional search n = 9


Estimated pain severity of NICU procedures
Lumbar puncture
Peripheral arterial puncture
Endotracheal intubation
Intramuscular injection
Eye examination*
Endotracheal suction
Venipuncture
Naso/Oropharyngeal suction
Tape removal
Nasogastric tube insertion
Heel lance
Urethral catheterization
Nasal prongs insertion for CPAP
Orogastric tube insertion
Eye drops instillation**
Threshold to administer pain relief
Extremely severe
Severe
Moderate
Mild to Moderate
Mild
0
10
20
30
40
50
60
70
80
90
100
Normalized pain reactivity scores (%) at 0-1 minute
weighted average ± standard deviation


Agglomeration schedule coefficients
2500
2000
1500
1000
500
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14
Stages