

A NEONATAL FACIAL IMAGE SCORING SYSTEM (NFISS) FOR PAIN RESPONSE STUDIES

KEE-HSIN CHEN¹, SUSANNA CHANG², TZU-CHIEN HSIAO³, YUEH-CHIH CHEN¹, CHII-WANN LIN^{2,4}

¹Nursing Department, Buddhist Tzu Chi College of Technology, Hualien, Taiwan

²Institute of Biomedical Engineering, College of Medicine and College of Engineering, National Taiwan University, Taipei, Taiwan

³Department of Biomedical Engineering, I-Shou University, Kaohsiung, Taiwan

⁴Department of Electrical Engineering, College of Electrical Engineering and Computer Engineering, National Taiwan University, Taipei, Taiwan

ABSTRACT

The aim of this study was to establish a scoring system, Neonatal Facial Image Scoring System (NFISS), for investigating the pain responses in neonates during routine venipuncture. We collected 1998 facial images from 19 infants during the first phase of this study and identified twelve features of the facial image: normalized scale for NFISS (scale from 0-15), including brow bulge (0-2), vertical brow furrowing (0-1) and short distance (0-1), brow lowering (0-1), eyes close (0-1), bulging eyes (0-2), eye-eye furrowing (0-1), cheek bulge (0-1), nasal-labial furrowing (0-2), mouth open and stretch (0-1), lip purse (0-1) and taut tongue (0-1). During the second evaluation phase, with 2770 images from 31 newborn infants (average age: 2.85^o.027 days), the pain responses were graded on the NFISS scale. Two peak score values during a time course of baseline-venipuncture-recovery (3-3-10 min) were found with needle piercing and needle withdrawal. The reliability of the score analysis was evaluated using 492 randomly selected images out of the 2770 images. The results for intra-scorer (author, one week interval) and inter-scorer (author and one experienced nurse) has correlation coefficients of 0.916 and 0.826, respectively.

Biomed Eng Appl Basis Comm, 2005(April); 17: 19-26.

Keywords: Neonatal Facial Image Scoring System (NFISS), venipuncture, neonatal, pain response

1. INTRODUCTION

Due to the decline birth rate, the quality of neonatal cares has been an important issue for recent

announcement of several international pediatrics societies (2000). In these announcements, recommendations for safe and effective pain relieve measures were addressed without proper pain assessment methods. According to our clinical observations, on average, there are about 2 to 5 times painful procedures, venipuncture or heel lance for metabolic screening and serum bilirubin test, which need to be done during the first three days in hospital for a healthy neonate. Due to the short of language capabilities for subjective description and large variations in tolerance, it is difficult to establish a quantifiable and effective pain response index for these routine clinical procedures of neonatal cares. The

Received: Jan 14, 2005; Accepted: Jan 20, 2005

Correspondence: Chil-Wann Lin, Associate Professor
DIstitute of Biomedical Engineering, College of
Medicine and College of Engineering, National Taiwan
University, 6F, IBME, No. 1, Sec. 4, Roosevelt Road,
106, Taipei, Taiwan

E-mail: cwlinx@ntu.edu.tw

description of neonatal facial expression in response to pain can trace back to Charles Darwin on 1872 (JOHNSTON 1989). Based on such research, Grunau and Craig developed "Neonatal Facial Coding System (NFCS)" on 1987 (GRUNAU and CRAIG, 1987). It has been extensively applied to the neonatal pain studies since then. Consistent results have been reported by various groups that the responses of facial features, including brow bulge, eye squeeze, nasal-labial furrowing and mouth open are statistically related with pain stimulus. (GRUNAU *et al.*, 1990; HADJISTAVROPOULOS *et al.*, 1997; GRUNAU *et al.*, 1998; WOLF 1999). Further extension of the development is the multidimensional scoring systems, which combined facial expression with physiological parameters. These include premature infant pain profile (PIPP) (RUNEFORS *et al.*, 2000; STEVENS *et al.*, 1996), neonatal infant pain scale (NIPS) (LAWRENCE *et al.*, 1993), and Crying Requirement of oxygen supplementation, Increase of heart rate and blood pressure, facial Express, Sleeplessness (CRIES) (KRECHEL and BILDNER 1995).

During repeated clinical observations, we found that there are some detailed facial expressions, which are not listed in the NFCS and yet also strongly related to pain responses. Other system (e.g. NIPS) can only provide 1 or 0 index, which is difficult for quantitative comparison. In order to provide a system for clinical studies of neonatal pain responses, we reported a refined facial imaging system based on above mentioned NFCS and named as "Neonatal Facial Image Scoring System" (NFISS). It was then applied to assess the pain responses of routine venipuncture procedure for its feasibility. The preliminary results of 31 infants in the second phase experiment confirmed NFISS is applicable for neonatal pain evaluation.

2. MATERIALS AND METHODS

2.1 Subjects

Criteria for the infant selection in the experiment were as following: Gestation Age 37-42 weeks, single birth, birth weight above 2500 gram and in general health. Those who had the following situations were excluded from experiment: Apgar score less than 7 at 5 minutes, congenital diseases, subjected to paralytic, analgesic or sedatives drugs within 4 hours before experiment.

2.2 Experimental Setting

The experiments were carried out in the baby room of National Taiwan University Hospital (NTUH), which has 2,000 newborn infants in average annually. To avoid unnecessary interference from other babies,

an isolated, quiet study room with constant temperature and light conditions was prepared for this study.

2.3 Apparatus and Measures

The physiological parameter, heart rate, and facial images of infants were video recorded for off-line analysis. The EKG monitor (GE : Marquette, USA) was operated with alarm-off to minimize environmental interference. A digital video recorder (SONY DCR-PC110, Japan), which provided a close-up view of the infant's facial images, was mounted on the bedside to continuously monitor the facial images of infants on a monitor (JVC TM-920u, Japan) and simultaneously recorded on a videocassette recorder (JVC BR9060U, Japan). The video images were then transferred to a PC via a frame grabber (PCI-1408, National Instruments, Austin USA) and then analyzed by using an imaging analysis package (Vision Builder and LabVIEW 6.1, National Instruments, Austin USA) with interactive analytical tools. The identified facial features of recorded images were processed off-line to extract the pattern for matching scores. The necessary processing steps include: extract color plane to HSL-Luminance, learn the template with shift and rotation, match the pattern with rotation invariant and set the minimum score of 700 as threshold (0-1000 for degree of similarity). Each individual feature is counted as an independent event to assign the final score for summation.

2.4 Procedures

The study procedures in this research included two phases:

Phase I -pilot study

Facial images of 19 infants were continuously recorded on a video recorder during venipuncture for 16 minutes. It was then digitized off-line by above mentioned image acquisition system. Serial images of the facial features were stored for further analysis. There were 1998 frames of images in total during this phase of experiment.

Qualitative content analysis based on static images was performed to establish and categorize the general facial features during pain and calm situations. Facial feature under calmness was determined before venipuncture stimulus and recorded as baseline imaging. Some features were decided not to include for further studies due to the limitations of static image, e.g., quivering. Five areas and twelve features of neonatal facial images were identified and compared to NFCS as in Table 1. It was then used in the second phase of study to verify its applicability.

Table 1 Comparison of NFISS and NFCS

Facial features and characteristics		NFISS	NFCS
brow	brow bulge	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	vertical brow furrowing	<input checked="" type="checkbox"/>	<input type="checkbox"/>
	shorting brow distance	<input checked="" type="checkbox"/>	<input type="checkbox"/>
	brow lowering	<input checked="" type="checkbox"/>	<input type="checkbox"/>
eyes	Eyes close (eye squeeze)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	bulging of the fatty pads around the infant's eyes (between the brow and eyes)	<input checked="" type="checkbox"/>	<input type="checkbox"/>
	eye-eye furrowing (between eyes)	<input checked="" type="checkbox"/>	<input type="checkbox"/>
cheek	cheek bulge	<input checked="" type="checkbox"/>	<input type="checkbox"/>
nasal-labial	nasal-labial furrowing	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
mouth	mouth open*	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	lip purse	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	stretch mouth/horizontal	<input type="checkbox"/>	<input checked="" type="checkbox"/>
	stretch mouth/vertical	<input type="checkbox"/>	<input checked="" type="checkbox"/>
taut	taut tongue	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
chin	chin quiver	<input type="checkbox"/>	<input checked="" type="checkbox"/>

* In NFISS, including mouth stretch

Phase II -formal study

Subjects were randomly selected from the baby room of National Taiwan University Hospital by investigator. Informed consent was obtained in written form prior to enrollment. Subjects were required to receive a venipuncture for newborn metabolic screening, hemotocrit and serum bilirubin at third day

of birth and all of them breast-fed 30-80 minutes before the observation. The whole procedure was conducted in a quiet and isolated room inside baby room. Infants were placed EKG leads over left, right chest wall and abdominal wall, as well as remained partly swaddled in the bassinet. All of subjects rested for 15 minutes after preparation. The same investigator carried out all procedures. The standard protocol for

venipuncture included: holding infant's palm, swabbing the puncture site with 75% alcohol to disinfect the skin around 1 inch, venipuncture with a 25-gauge 5/8" needle, collecting blood in absorbent newborn screen cards (at least five areas), and covering the puncture site with bandage. After the procedure, investigator was required to remain in the study room for 10 minutes to ensure the infant did not experience any adverse reactions to the venipuncture. Only one trial of venipuncture was performed in all infants. After these procedures, video images were digitized for further analysis. There were 2770 images from 31 newborn infants (average age: $2.85 \pm .027$ days) during the second phase of this study. The pain responses were assessed by physiological and behavioral parameters. Heart rate and facial expression were recorded on the video continuously during a 3 minutes baseline, period of venipuncture and for a recovery period of 10 minutes after last contact without soothing or calming interactions.

3. DATA ANALYSIS

The data were analyzed with SPSS statistics package (version 8.0 for window), using descriptive analysis (count, percentage, mean, 95% C.I.) and graphs. Kappa Statistic revealed a substantial agreement between two scorers for the reliability of NFISS.

4. RESULTS

4.1 Demographic Data

Subjects were recruited in the baby room of NTUH by investigator during Dec. 2000-Mar. 2001 for phase II study. Those who met the screening criteria when transferred into baby room will be randomly selected for this study by coin-toss procedure (heads up). Due to the constraints of available resources, this preliminary result can only allow for enrollment of 31 infants (20 males, 11 females). With informed consents from parents, 96.7% family members in this study agreed to participate in the evaluation processes. Mean gestation age was 38.52 weeks (S.D. = 1.1216, range 37-41); Birth weight was 3280.71 gram (S.D. = 376.11, range 2648-4000). Methods of delivery were normal spontaneous vaginal, 17 (54.8%), and Caesarian section, 14 (45.2%).

4.2 NFISS

Qualitative content analysis was performed on the recorded image frames from pilot experiment to identify the pain related facial changes and limitations

of static images. NFISS divided the facial image into five areas, including brow, eyes, cheek, nasal-labial and mouth in order to characterize detail features of pain responses. There were totally twelve features that can be identified, ¹brow bulge; ²vertical brow furrowing; ³shorting distance; ⁴brow lowering; ⁵eyes close; ⁶bulging eyes; ⁷eye-eye furrowing; ⁸cheek bulge; ⁹nasal-labial furrowing; ¹⁰mouth open and stretch; ¹¹lip purse; ¹²taut tongue. A score from 0 to 2 was assigned to each identified feature of facial image with higher value for more intensive pain responses. This would result in the lowest score of 0 to highest score of 15 for graded pain responses. As shown in Fig. 1, three typical facial images of neonate in different degree of pains will have score value of 0, 7 and 13 according to NFISS, respectively.

4.3 NFISS Score Changes during Venipuncture

By using established NFISS, the resultant measurement of pain responses from 31 subjects during venipuncture were shown in Fig. 2 with 95% confidence interval. The baseline score was below 2 during the pilot phase. The score increased significantly right after the needle piercing and subdued to lower value during the blood sampling period. It reached another peak value when needle was withdrawn from the infant. These two peak values indicated the intensive pain during this invasive manipulation.

4.4 The HR Changes during Venipuncture

Physiological monitoring of the heart rate (beat per minute) was displayed on an EKG monitor and film recorded with facial images simultaneously. During the period of venipuncture, the heart rate reached two peak values in two particular time points as indicated in the Fig. 3 with 95% confidence interval. One peak was the first 10 seconds after the needle pierce into skin, and the other is the moment when the needle was removed.



Fig. 1 Three typical facial images of neonate during calm and pain situations, from left to right, the NFISS score value is (a) 0, (b) 7 and (c) 13, respectively. (*Images are taken with parent's informed consent.)

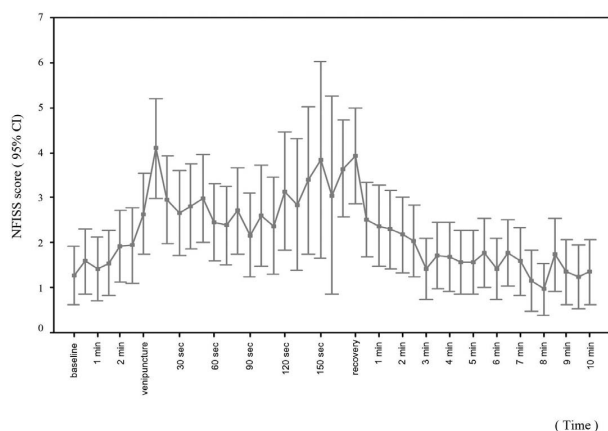


Fig. 2 The profile of NFISS score value during venipuncture has two peak responses, which correspond to the strong pain response of needle piercing and needle withdrawing.

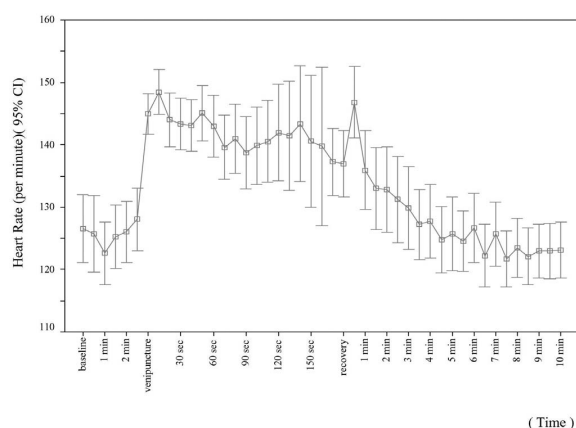


Fig. 3 The HR variation during venipuncture.

4.5 Reliability

For the intra-scorer reliability of established NFISS, the investigator repeated the scoring procedure after two weeks with randomly selected images from one tenth of previous recorded images. This comparison resulted in 0.916 correlation coefficient value during the phase II study, indicated the highly consistence between two subsequently measurements. While for the inter-scorer reliability, the NFISS color atlas was shown and explained to an experienced nurse in the baby room. Then she was asked to put down the score with randomly selected images from one tenth of previous recorded images. This score was then compared to the result from investigator. Apparently, this approach could not account for the background difference of investigators. However, after reviewing

the color atlas, this comparison results in 0.826 during the phase II study.

5. DISCUSSION

On average, there will be 2-5 times of venipuncture or heel lance for metabolic screening and serum bilirubin test during the first three days in the hospital. However, pain assessment is one of the most difficult challenges for health professions and researchers who work with infants. It is thus important to develop a low cost and simple system for routine clinical evaluation of neonatal pain responses and effective managements. Images of facial expression contain rich information for direct pain assessments. Derived from NFCS, our NFISS identified five areas and twelve features of static neonatal facial image for quantitative pain responses. Some of these identified features have not been reported in previous literatures yet. These were then assigned to a numerical index for graded pain responses. This results in the lowest score of 0 to highest score of 15 for graded pain responses. When applied to measure the pain responses from 31 subjects during routine venipuncture procedure, the index value of NFISS not only gives a quantitative index of pains but also shows coincident with needle piercing and withdrawing actions. However, current version of NFISS is not capable of measuring fast changing facial images, especially for chin quivering and mouth stretching during intensive cry, and fully automation yet. It would be possible to enhance NFISS for these aspects in the future.

Nociceptive stimulus can induce physiological and behavior changes in neonate. Heart rate, respiratory rate, oxygen saturation, sweating, cranial pressure, vagal tone, cortisol, and catecholamine have all been monitored as physiological parameters in pain responses (OWENS M.E. and TODT E.H. 1984; STEVENS and JOHNSTON 1994; PORTER *et al.*, 1988; JOHNSTON 1989; ANAND *et al.*, 1987). The behavior changes include voice volume (crying, weeping, moaning), facial expressions (brow bulge, eye squeeze, nasal-labial furrowing, mouth open, lip purse, stretch mouth, taut tongue, and chin quiver etc), posture (withdrawing, quivering, stiff, hand holding), alertness, feeding pattern, and interacting modes (MARSHALL 1989). In the early development of fetus, they start to feel nociception or pain stimulus through skin pain receptors at around week 20 (ANAND *et al.*, 1987; HUMPHREY 1964; STEVENS and JOHNSTON 1993), and the thalamocortical connections for signal conduction are fully integrated at about week 24 (ANAND *et al.*, 1987; ANAND *et al.*, 1989; FRIENDRICH *et al.*, 1995; WOLF 1999). The neurotransmitter, substance P, exists abundance at

about week 16 and has even lower threshold than adult (hypersensitive) (FITZGERALD *et al.*, 1988). Even though previous study on pre-mature infants has shown that the short of pain sensation may be due to un-matured myelinated sheath fiber, recent publications indicate that full-month neonates have fully integrated pain sensations as in children or adults (ANAND *et al.*, 1987; STEVENS and JOHNSTON 1993; JOHNSTON and STEVENS 1990; STEVENS and JOHNSTON 1994). Due to the short of language capabilities for subjective description and large variations in tolerance, it is difficult to establish a quantifiable and effective pain response index for clinical procedures of neonatal cares. However, the consistence of facial features in neonatal pain responses has been studied and adopted for pain studies especially in qualitative aspects, e.g. Neonatal Facial Coding System (NFCS) (GRUNAU & CRAIG, 1987 ; GRUNAU *et al.*, 1990). Even though, it is difficult to register the sequence of imaging for fully automatic analysis, it seems that the other behavior parameters all subject to more interferences, e.g. body movements, which lead to larger variations in evaluations. We have refined the typical features of neonatal facial expression into five areas and twelve features for scoring and named as NFISS. The comparison between these two systems is listed in Table 1. Both systems have some similar facial features, which can be easily identified during neonatal pain responses, including brow bulge, eyes close, nasal-labial furrowing, mouth open, lip purse, and taut tongue. However, several refined features are identified with repeated efforts during pilot experiment. All these features are located in the rich facial expression areas, primarily brow, eyes and mouth. For brow and eyes area, NFISS covers seven features comparing to two in NFCS. For mouth area, only two features are identified in NFISS. Because mouth open and stretch either vertically or horizontally can also be seen in non-pain related facial expression, NFISS groups these three features into one feature. The other omitted feature in NFISS is chin quivering. This is due to the limitation of NFISS. When applying NFISS for clinical evaluations, it never reaches full scale of 15 even under extreme painful responses (e.g., crying, red face, sweating, and screaming). This might be due to above mentioned limitation of static image analysis, especially for chin quivering and mouth stretching during intensive cry, which would require image registration for automation. It would be possible to enhance NFISS for this aspect in the future.

Even though, all the infants are randomly selected by coin-toss prior to the evaluation, the demographic data does indicate slightly bias in the gender appearance. Part of the reason can be attributed to the small sample size in this experiment. However, it is also interesting to know that the gender ratio is around

109:100 (compared to normal at about 105:100, male: female) in a recent report from Minister of Interior of Taiwan Government. As shown in Fig. 2 and 3, during the period of venipuncture, both heart rate and NFISS scores reached peak values in two distinguished time points. One was the first 10 seconds after the needle injected into skin, and the other was the moment when the needle was removed. The time course of NFISS is coincident with physiological recording of heart rate and respiratory rate changes during experimental manipulations. To our knowledge, no study has been found to compare such a correlation between behavioral and physiological changes for neonatal pain responses. There are findings of heart rate increasing due to venipuncture, which are in accordance with previous studies from different groups (JOHNSTON 1989; ANAND *et al.*, 1987; STEVENS and JOHNSTON 1994; OWENS M.E. and TODT E.H. 1984). However, the changes of heart rate during nociceptive stimulus can be increased or decreased depend on the status of body activities. Therefore, it is not a suitable evaluating parameter for pain studies along (BLASS and WATT 1999). Similar reason holds true for the crying behavior of neonate (RUNEFORS *et al.*, 2000).

One of the future applications for NFISS would be the effective measures for pain management in neonates. These include analgesia by chemical administration or simply by different types of solutions, e.g. water, milks, or sugar. Further investigations by using NFISS in these aspects can contribute to both the quality of neonatal cares and its activation mechanisms.

6. CONCLUSIONS

Pain assessment is one of the most difficult challenges for health professions and researchers who work with infants. Based on the Neonatal Facial Coding System (NFCS), we have redefined the typical features of neonatal facial expression into five areas and twelve features for scoring and named as Neonatal Facial Image Scoring System (NFISS). The result indicates that NFISS score correlated very well with physiological parameter of heart rate. NFISS can thus be functioned as a non-invasive and effective assessment tool for clinical workers. In addition, it can be a stepping stone for development of pain management plans and pain relief interventions in the future.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the support of the National Taiwan University Hospital for this work. They also wish to thank for the parents of all the infants who participated in the study.

REFERENCES

1. AMONYMOUS, (2000): 'Prevention and management of pain and stress in the neonate', *Pediatrics*, 105, pp. 454-461
2. ANAND, K. J. S., PHIL, D., and CARR, D. (1989): 'The neuroanatomy, neurophysiology, and neurochemistry of pain, stress, and analgesia in newborns and children', *Pediatric Clinics of North America*, 36, pp. 795-821
3. ANAND, K. J. S., PHIL, D., and HICKEY, P. R. (1987): 'Pain and its effects in the human neonate and fetus', *New England Journal of Medicine*, 31, pp. 1321-1329
4. BLASS, E. M., and WATT, L. B. (1999): 'Suckling- and sucrose-induced analgesia in human newborns', *Pain*, 83, pp. 611-623
5. FITZGERALD, M., MILLARD, C., and MCINTOSH, N. (1988): 'Hyperalgesia in premature infant', *Lancet*, 1, pp. 292
6. FRIENDRICH, J. B., GALLAGHER, D., and KIMURA, R. E. (1995): 'Where does it hurt? An interdisciplinary approach to improving the quality of pain assessment and management in the neonatal intensive care unit', *Nursing Clinics of North America*, 30, pp. 143-159
7. GRUNAN, R. V. E., and CRAIG, K. D. (1987): 'Pain expression in neonates: facial action and cry', *Pain*, 28, pp. 395-410
8. GRUNAU, R. V. E., JOHNSTON, C. C., and CRAIG, K. D. (1990): 'Neonatal facial and cry response to invasive and non - invasive procedures', *Pain*, 42, pp. 295-305
9. GRUNAU, R. V. E., OBERLANDER, T., HOLSTI, L., and WHITFIELD, M. F. (1998): 'Bedside application of the Neonatal Facial Coding System in pain assessment of premature neonates', *Pain*, 76, pp. 277-286
10. HADJISTAVROPOULOS, H. D., CRAIG, K. D., GRUNAU, R. V. E., and WHITFIELD, M. F. (1997): 'Judging pain in infants: behavioural, contextual, and developmental determinants', *Pain*, 73, pp. 319-324
11. HUMPHREY, T. (1964): 'Some correlations between the appearance of human fatal reflexes and the development of the nervous system', *Progress in Brain Research*, 4, pp. 93-135
12. JOHNSTON, C. C. (1989): 'Pain assessment and management in infants', *Pediatrician*, 16, pp. 16-23
13. JOHNSTON, C. C., and STEVENS, B. J. (1990): 'Pain assessment in newborns', *The Journal of Perinatal and Neonatal Nursing*, 4, pp. 41-52
14. KRECHEL, S. W., and BILDNER, J. (1995): 'CRIES: a new neonatal postoperative pain measurement score. Initial testing of validity and reliability', *Paediatric Anaesthesia*, 5, pp. 53-61
15. LAWRENCE, J., ALCOCK, D., MCGRATH, P., KAY, J., MACMURRAY, S. B., and DULBERG, C. (1993): 'The development of a tool to assess neonatal pain', *Neonatal Network*, 12, pp. 59-66
16. MARSHALL, R. E. (1989): 'Neonatal pain associated with caregiving procedures', *Pediatric Clinics of North America*, 36, pp. 885-903
17. OWENS M.E., and TODT E.H. (1984): 'Pain in infancy: neonatal reaction to a heel lance', *Pain*, 20, pp. 77-86
18. PORTER, F. L., PORGER, S. E., and MARSHALL, R. E. (1988): 'Newborn pain cries and vagal tone: parallel changes in response to circumcision', *Child Development*, 59, pp. 495-505
19. RUNEFORS, P., ARNBJONSSON, E., ELANDER, G., and MICHELSSON, K. (2000): 'Newborn infants' cry after heel-prick: analysis with sound spectrogram', *Acta Paediatr*, 89, pp. 68-72
20. STEVENS, B., JOHNSTON, C., PETRYSHEN, P., and TODDIO, A. (1996): 'Premature infant pain profile: development and initial validation', *The Clinical Journal of Pain*, 12, pp. 13-22
21. STEVENS, B. J., and JOHNSTON, C. C. (1993): 'Pain in the infant: theoretical and conceptual issues', *Maternal-Child Nursing Journal*, 21, pp. 3-14
22. STEVENS, B. J., and JOHNSTON, C. C. (1994): 'Physiological responses of premature infants to a painful stimulus', *Nursing Research*, 43, pp. 226-231
23. WOLF, A. D. (1999): 'Pain, nociception and the developing infant', *Pediatric Anesthesia*, 9, pp. 7-17.