ACTA PÆDIATRICA
NURTURING THE CHILD

Acta Pædiatrica ISSN 0803-5253

REGULAR ARTICLE

# Neonatal procedural pain can be assessed by computer software that has good sensitivity and specificity to detect facial movements

Tatiany Marcondes Heiderich, Ana Teresa Figueiredo Stochero Leslie, Ruth Guinsburg (ruthgbr@netpoint.com.br)

Division of Neonatal Medicine, Department of Paediatrics, Escola Paulista de Medicina, Universidade Federal de São Paulo, São Paulo, SP, Brazil

**Keywords**
Computer software, Facial recognition, Newborn infant, Pain assessment, Procedural pain

**Correspondence**
Ruth Guinsburg, Rua Vicente Felix 77 apt 09, São Paulo, SP, Brazil CEP 01410-020.
Tel: +55-11-5084-0535 |
Fax: +55-11-5084-0535 |
Email: ruthgbr@netpoint.com.br

**Received**
17 July 2014; accepted 7 November 2014.

DOI:10.1111/apa.12861

## ABSTRACT

**Aim:** The difficulty in assessing pain during the neonatal period is one of the main obstacles for appropriate analgesia in intensive care units. The aim of this study was to develop and validate computer software to monitor neonatal facial movements of pain in real time.

**Methods:** The software was developed in the Delphi integrated development environment and provides real-time image analysis during monitoring, based on image recognition of pain-related facial actions. To validate the software performance, facial images were obtained during the monitoring of 30 neonates who were subjected to painful procedures related to daily care management. Of the 5644 images identified and analysed by the software, 360 images – 12 per infant – were randomly selected and assessed by six healthcare professionals with experience of recognising neonatal pain.

**Results:** The agreement between the examiners and the software assessment was excellent ($\kappa$ = 0.975). The software exhibited 85% sensitivity and 100% specificity in detecting neutral facial expressions in the resting state and 100% sensitivity and specificity in detecting pain during painful procedures.

**Conclusion:** It is possible to assess neonatal procedural pain using computer software that has good sensitivity and specificity to detect facial movements.

## INTRODUCTION

Newborn infants are exposed to painful experiences that might increase their short- and long-term morbidity and mortality, in addition to being associated with neurological developmental disorders (1,2). Improvements in this scenario require practical and efficacious methods of pain assessment, safer drugs, organisation of services allowing for the detection of flaws in the care provided and continuous updating and sensitisation of neonatal care providers (3).

Facial expression analysis provides valid, highly sensitive and specific information on the nature and intensity of pain, thus allowing for efficacious communication between newborn infants and neonatal healthcare providers (4). The Neonatal Facial Coding System is the scale most widely used to assess neonatal pain based on facial expressions, as defined by the presence or absence of eight facial actions (5). At the same time, the human face provides unique and individual biometric parameters. Several studies on automatic facial recognition have sought to develop systems able to identify individuals and their expressions, even when the subjects are unaware that they are being observed. Image recognition, and more particularly facial image recognition, might contribute to increasing the flexibility of the interaction between computer-based systems and human beings, allowing the former to detect the emotional states of the latter and ultimately enabling the automatic assessment of pain (6).

Few studies in the literature have reported on the development of technology to analyse the facial expressions of pain in general or of neonatal pain in particular (7–11). In 2008, Schiavenato et al. (11) developed a computer-based method to compare the distances measured between Neonatal Facial Coding System reference points before and after the application of a heel stick. That method was able to detect consistent differences in the distances between point-pairs, related to the actions of opening the mouth, drawing in the brows and closing the eyes. However, the aim of that

### Key notes

- Difficulty in assessing pain is one of the main obstacles for appropriate neonatal analgesia.
- This study reports how computer software was developed to monitor neonatal facial movements of pain in real time and tested by healthcare professionals experienced in recognising neonatal pain.
- The software exhibited 85% sensitivity and 100% specificity in detecting neutral facial expressions in the resting state and 100% sensitivity and specificity in detecting procedural pain in neonates.

 

study was to compare the facial expressions of full-term neonates in response to a painful stimulus, while controlling for possible confounding variables. Therefore, a computer-based methodology was designed to refine the measurement of facial expressions, rather than as a tool for monitoring pain.

This study was carried out against the background of the difficulties inherent in assessing pain in newborn infants and the need for methods to assess pain with the least possible human interference. The aims of the present study were to develop specific software that could recognise the facial expressions of pain in neonates, based on comparing facial images captured in real time to previously calibrated facial images acquired in the resting state. We also wanted to investigate whether that software could distinguish between neonates' facial expressions in the resting state and the expressions displayed when painful procedures were performed.

## PATIENTS AND METHODS

### Software development

The Neonatal Facial Coding System was selected as the pain indicator (4,6). On that scale, each occurrence of facial actions is given one point and the presence of pain is defined by a total score equal to, or greater than, three (12). On these grounds, software was developed using Embarcadero Delphi XE2, which is an integrated development environment for generating stand-alone graphical user interface programs (Embarcadero Technologies Inc., San Francisco, CA, USA). The software allowed us to acquire facial images in real time during monitoring sessions using three video cameras. The first camera was placed to the right, the second to the left and the third above the studied infant. The method for facial identification and establishing the main points, namely eyes, brows, nose, lips and chin, included the use of the Luxand Face Software Development Kit (Luxand Inc., Alexandria, VA, USA), which is an application for biometric identification that performs automatic mapping and detection of the face and 66 main points.

From the 66 nodal points identified, 16 main nodal points were selected, based on tests conducted with five newborn infants to identify the points that exhibited the greatest movement when the infants felt pain. Between these 16 points, 14 distances were chosen to assess similarities among images (Fig. 1). These distances were used to detect facial actions as follows: bulging brow, defined as the reduction of the distance between points P13-P14 and, or, P16-P17; narrowing of the lid slit, defined as the reduction of the distance between points P27-P28 and, or, P28-P16 and, or, P31-P32 and, or, P32-P17; deepening of the nasolabial furrow, defined as the increase of the distance between points P3-P45 and, or, P4-P46 and, or, P5-P45 and, or, P6-P46 and increased or reduced distance between P3-P5 and, or, P4-P6; open lips, defined as the increase of the distance between points P54-P55; and mouth stretching, defined as the increase of the distance between points P3-P4.

Thus, the software that was developed compared the distances between the main nodal points on the image acquired in resting state, which is the calibrated image, to those same distances on images captured during monitoring sessions. After calculating the score for similarity, the software attributed points to the detected facial actions following the Neonatal Facial Coding System to define the presence or absence of pain. Only five of the eight facial actions included in the Neonatal Facial Coding System were used for analysis because the position and movement of the chin and tongue are difficult to detect using the Luxand Software Development Kit (Luxand Inc., Alexandria, VA, USA). Therefore, the software was able to detect the following movements: brow bulging, narrowing of the lid slit, deepening of the nasolabial, open lips and mouth stretching. For each detected movement, the software attributed one point, with the total score varying from zero to five. The captured images were scored every 3-sec and classified as painless neutral faces, as in detection of up to two facial actions, or painful faces, as in detection of three or more facial actions. If infants moved their faces away from the cameras, or if the cameras were moved from their optimum positions, the captured images would not be recognised as faces and could not be scored.

### Participants and examiners for software validation

The number of images required for analysis was established following calculation of the sample size: analysis of 102 images would be necessary to find a $\kappa$ of $\geq$0.80, with an

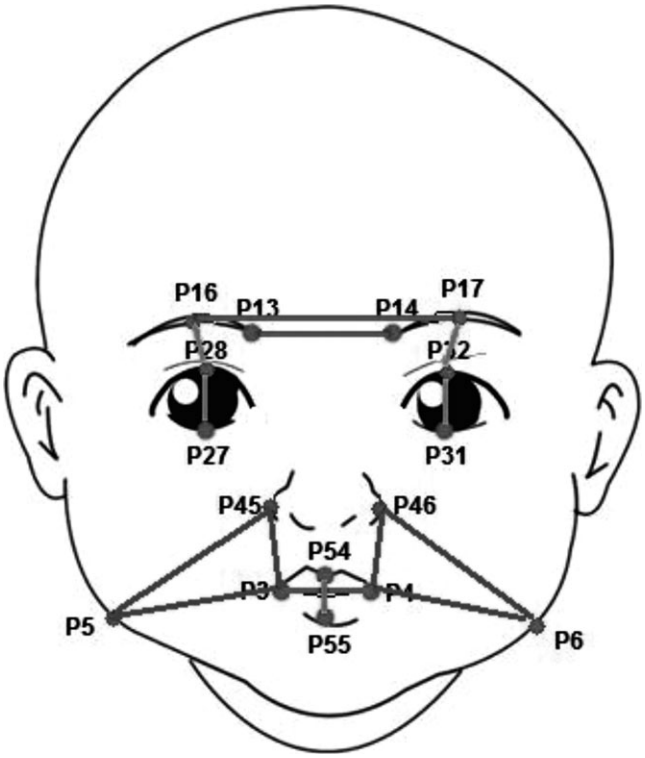


**Figure 1** Score to assess similarities among images with 14 distances measured between the 16 points (P).


 

agreement rate of 90%, a power of 80% and a two-tailed test with an alpha error of 5%, considering that the null hypothesis was set as $\kappa \leq 0.40$ (13).

The criteria to include newborn infants for monitoring were as follows: admission to the neonatal care unit; gestational age of 34–41 weeks as indicated by the best obstetric estimate or physical examination of the infants; postnatal age >24 and <168 h; no need for any modality of respiratory support; no need for a gastric tube; absence of congenital malformations; medical indication for capillary, venous, or arterial puncture or intramuscular or subcutaneous injections; and signed informed consent by the parents. Importantly, the painful procedures were not performed to comply with the study's aims, but were indicated based on strict clinical need.

Six healthcare professionals were selected to assess and score the infants' facial images, based on the Neonatal Facial Coding System (4,6). The examiners scored the same facial features that were analysed by the software as present or absent: brow bulging, narrowing of the lid slit, deepening of the nasolabial, open lips and mouth stretching. The inclusion criteria for the examiners were as follows: healthcare professionals specialised in neonatology and working in university centres who agreed to take part in the study as examiners.

### Image acquisition by the developed software

For the purpose of image acquisition, images of 30 newborn infants were recorded. The monitored images were acquired before, during and after the performance of painful procedures. The images acquired before the performance of painful procedures were designated as painless or resting situations; the ones captured during the performance of painful procedures were designated as painful or procedure situations, and the ones acquired after the performance of painful procedures were designated as recovery. The latter were not classified as actual painless or painful situations because it was not possible to establish the duration of pain after a venous or capillary puncture or intramuscular injection.

### Image selection

The images were acquired by the software every 3-sec during the 10-min monitoring sessions. We selected 12 facial images from each of the 30 newborn infants, and these were randomly captured every 45/50 sec, approximately one image per minute of film. Thus, a total of 360 images were selected to establish the software's accuracy to identify the presence or absence of pain in comparison with the actual situations, as well as to compare the software's judgment to that made by healthcare providers experienced in neonatal pain assessment.

For the examiners' assessment, the 12 images corresponding to each infant were randomly placed on two size A4 white paper sheets so that six images were printed per sheet using a colour laser printer. Space was left on each image for the examiners to indicate the presence or absence of the five Neonatal Facial Coding System actions that were assessed in this study.

### Description of results and statistical analysis

First, the agreement among the six examiners was assessed, and then, the agreement between the software and the six examiners with regard to the pain diagnosis in the 360 selected images was assessed by generalised *Kappa*. Subsequently, the concordance was assessed between each actual situation, that is pain present or absent, and the corresponding assessments. These comprised the assessments made by the consensus of the six examiners and by the software for each of the 60 images acquired during the painless phase and the 30 images acquired during the painful phase of the procedure.

To establish whether the assessment of the number of actions detected by the software and the examiners was homogeneous relative to both the presence and absence of pain, Bland–Altman plots were used (14). For this purpose, the average of the scores attributed by the examiners to each of the 60 images corresponding to the painless situation was calculated, with scores varying from zero to five facial actions. This result was added to the score calculated by the software and then divided by two, with the final value being plotted on the *x*-axis of a graph. The value to be plotted in the *y*-axis, for each one of the 60 facial images on rest, was calculated by subtracting the software score from the examiners' average score. The same procedure was applied to the 30 images corresponding to the painful situation and the 192 images acquired during the stage of recovery.

Statistical analysis was performed using SPSS 17.0 (SPSS Inc., Chicago IL, USA) The significance level was established as $p \leq 0.05$. The Bland–Altman plots were analysed using MedCalc software, version 12.7.3.0® (MedCalc Software bvba, Ostend, Belgium).

### Ethical considerations

When we designed the study, we carefully considered the possibility of giving some nonpharmacologic analgesia to the patients, but this would have severely limited the validation of the software. Therefore, after discussing the issue with the Institutional Review Board, it was decided that: only routine minor procedures would be studied; the procedures would only be carried out once and that each patient would only be studied for one procedure. Most important of all, the Institutional Review Board recommended that the harm to the baby, such as feeling pain, and the possible benefit of the study, which was to develop a tool that would maybe help health professionals to better assess neonatal pain presence and treat this pain, would be thoroughly discussed with each family prior to them providing written consent. The Research Ethics Committee of the Federal University of São Paulo approved the study.

## RESULTS

For the purpose of software validation, images of 30 neonates born at the University Hospital from June to

**Table 1** General characteristics of 30 newborns

| | Values | Variation |
|---|---|---|
| Gestational Age (weeks; mean ± SD) | 37 ± 1 | 35–41 |
| Birthweight (grams; mean ± SD) | 2962 ± 593 | 2120–4100 |
| Males | 15 (50%) | |
| Caesarean delivery | 63% | |
| 1st minute Apgar (median) | 9 | 6–9 |
| 5th minute Apgar (median) | 9 | 7–10 |
| Postnatal days (mean ± SD) | 2 ± 1 | 1–6 |
| Procedures | | |
| Intramuscular injection | 2 (7%) | |
| Heel lancing | 3 (10%) | |
| Venipuncture | 25 (83%) | |

SD = Standard deviation.

August 2013 were recorded. The infants' characteristics are shown in Table 1.

The films lasted 10 min each and allowed for the acquisition of a total of 8457 images, with an average of 282 images per infant. From that total, the software identified 5644 (67%) images as corresponding to the face of the neonates, with an average of 188 images per infant. The software analysed these images and classified them according to the facial expressions as neutral or painful. From this group of images, 12 images corresponding to each infant were selected and a total of 360 were printed in colour. Of these, 138 (38.3%) were acquired during the resting state, 30 (3.8%) during the painful procedure when the skin was broken and 192 (53.3%) during recovery.

Six examiners were chosen to analyse the 360 selected images, and the examiners' characteristics are described in Table 2.

The generalised *Kappa* test was used to establish the concordance between the software and the human examiners' assessments relative to the presence or absence of pain in the 360 selected images. The concordance among the six examiners was assessed first, which resulted in $\kappa = 0.986$ ($p < 0.001$). Next, the concordance between the assessments made by the software and the examiners was tested, resulting in generalised $\kappa = 0.975$ ($p < 0.001$).

The examiners' assessments were subjected to consensus analysis, according to whether each image was classified as the presence or absence of pain based on the majority opinion, while the 22 inconclusive images were counted as presence of pain. As a result, 198 images were classified as absence of pain and 162 as presence of pain. The comparison between the examiners' consensus and the software's assessment showed that they agreed on the presence of pain in 149 images and to absence of pain in 83 images, resulting in overall agreement on 232 (64%) images. The examiners disagreed with the software when it came to 128 images: in the case of 115 images, the examiners said they showed absence of pain when the software concluded that they showed pain and the examiners identified pain in 13 images that the software had not identified as painful. Of these 128 disagreements, 57 corresponded to the resting period, three to the time during which the procedure was performed and 68 to the recovery period after painful procedures.

The concordance of the software and examiners with the actual painless or painful situation was assessed by three images from each newborn infant: two images of the resting state and one image obtained during the painful procedure. Table 3 shows that the software classified 85% of the images during the resting period as absence of pain, while the examiners' assessments of those images varied from 77% to 98%. When it came to the period when the painful procedures were performed, the software identified the presence of pain in 100% of those images, while the examiners' assessment of those images varied from 77% to 93%. Table 4 shows the Neonatal Facial Coding System scores relating to the disagreements between the software and the examiners.

To analyse potential differences in the modified Neonatal Facial Coding System scores given by the software and by human examiners, we compared the number of facial actions detected by them using Bland–Altman plots. This compared 60 images during the resting state, 30 images during the painful procedures and 192 images during the recovery period.

Figure 2A shows that the average scores recorded by the software and by the examiners were <3 for almost all of the images at rest. The difference between the examiners' and the software's assessments relative to the resting state varied from −1 to +1.2, indicating homogeneity between the assessments.

Figure 2B corresponds to the assessment of the 30 facial images acquired during the painful procedures and shows that the average scores attributed by the software and by the

**Table 2** General characteristics of examiners

| Examiners | Gender | Age (years) | Profession | Time since graduation (years) | Time working in NICU (years) |
|---|---|---|---|---|---|
| 1 | Female | 41 | MD | 17 | 14 |
| 2 | Female | 46 | MD | 23 | 20 |
| 3 | Female | 38 | RN | 16 | 16 |
| 4 | Female | 43 | RN | 21 | 20 |
| 5 | Female | 43 | MD | 19 | 14 |
| 6 | Female | 34 | PT | 12 | 11 |

MD = Medical doctor; RN = Registered nurse; PT = Physical therapist.


 

**Table 3** Agreement between the software and each examiners assessment with the real situation (reality) regarding the absence (neonate at rest before the invasive procedure) or the presence of pain (neonate receiving the invasive procedure)

| Agreement | Rest – painless 60 images, % | Painful procedure 30 images, % |
|---|---|---|
| Reality × Software | 85.0 | 100.0 |
| Reality × Examiner 1 | 91.7 | 93.3 |
| Reality × Examiner 2 | 76.7 | 90.0 |
| Reality × Examiner 3 | 93.3 | 83.3 |
| Reality × Examiner 4 | 98.3 | 76.7 |
| Reality × Examiner 5 | 83.3 | 90.0 |
| Reality × Examiner 6 | 81.7 | 90.0 |

examiners were equal to, or greater than, three for 27 images, indicating the presence of pain. The difference between the examiners' and the software's assessments varied from −1.0 to +1.0, with 29 points lying within two standard deviations on the *y*-axis.

Figure 2C shows that during the recovery period, the average number of facial actions detected on 192 images varied from zero to five. The difference between the examiners' and the software's assessments varied from −2 to +1.0, and 187 of the 192 points analysed were placed within two standard deviations on the *y*-axis.

## DISCUSSION

Despite the large number of scales available to measure pain in newborn infants, and acute pain in particular, better strategies are needed for pain assessment in neonatal care units (15). Changes in facial expression are known to represent one of the basic behavioural responses to painful stimuli (16). Some neonatal facial movements are considered pain specific, including lowering and drawing together of the brow, narrowing the lid slit and, or, squeezing the eyes shut, and deepening the nasolabial furrow and, or, cheek raising (4,17). Isolated or combined facial actions occurring as a response to pain have been almost universally included in the various instruments used to assess neonatal pain in clinical practice and scientific research (18). Within that context, facial actions have been given particular value because they represent the full expression of the experience of pain, including also its affective and emotional dimensions (19).

Routine bedside assessment of pain in critically ill neonates is considered crucial for the proper management of pain (20), but critically ill neonates are subjected to several potentially painful procedures, which are still poorly managed (2,21–23). Franck and Bruce observed that inappropriate assessment of neonatal pain is an integral component of the undertreatment of pain in neonatal care units, while it is not yet known whether defects in pain recognition by healthcare professionals can be corrected or compensated for or if the motivation to treat pain can be improved (24). According to these authors, humans might perhaps be inherently flawed in the assessment of pain in others, and if that were indeed the case, then human-mediated pain assessment should be replaced by technology.

The situation just depicted defines the context of the present study, which sought to replace the human with a computer eye in the recognition of the facial expressions of pain in newborn infants. The study conducted to validate the software's performance found agreement between the assessment performed by healthcare professionals of the still images acquired by the software and the decoding of the facial actions performed by the software itself relative to the actual painless and painful situations. Analysis of the heterogeneity between the number of facial actions detected by the software, and the average number of facial actions detected by the examiners, showed homogeneity in all three conditions, namely before the procedure, during the procedure and during the recovery period.

Therefore, the software developed in the present study was successful with regard to the intended goals of acquisition of facial images, decoding of facial images relevant to pain assessment, and recoding of the results as the presence or absence of pain in an almost instantaneous manner. The responses detected by the software were consistent with the actual painful or painless situations to which the infants were exposed, and they exhibited expressive concordance and homogeneity with the facial actions detected by healthcare professionals with experience in the management of neonatal pain, who assessed those very same images, although they were printed in assessment notebooks. It is impossible to compare the system developed in the present study to other tools, due to the small

**Table 4** Disagreement among examiners and the software regarding Neonatal Facial Coding System score during the pain procedure (real situation = pain presence)

| Neonate # | Software | Examiner 1 | Examiner 2 | Examiner 3 | Examiner 4 | Examiner 5 | Examiner 6 | Mean of 6 examiners |
|---|---|---|---|---|---|---|---|---|
| 2 | 3 | 4 | 2 | 1 | 0 | 0 | 3 | 1.7 |
| 3 | 3 | 5 | 4 | 3 | 2 | 4 | 3 | 3.5 |
| 4 | 4 | 4 | 4 | 2 | 0 | 4 | 1 | 2.5 |
| 11 | 4 | 4 | 5 | 4 | 2 | 4 | 5 | 4.0 |
| 12 | 3 | 1 | 2 | 2 | 1 | 2 | 2 | 1.7 |
| 14 | 3 | 0 | 0 | 1 | 1 | 0 | 1 | 0.5 |
| 16 | 3 | 5 | 5 | 2 | 0 | 3 | 3 | 3.0 |

Grey squares indicate disagreement between the software assessment (presence of pain) and the examiners assessment (absence of pain).

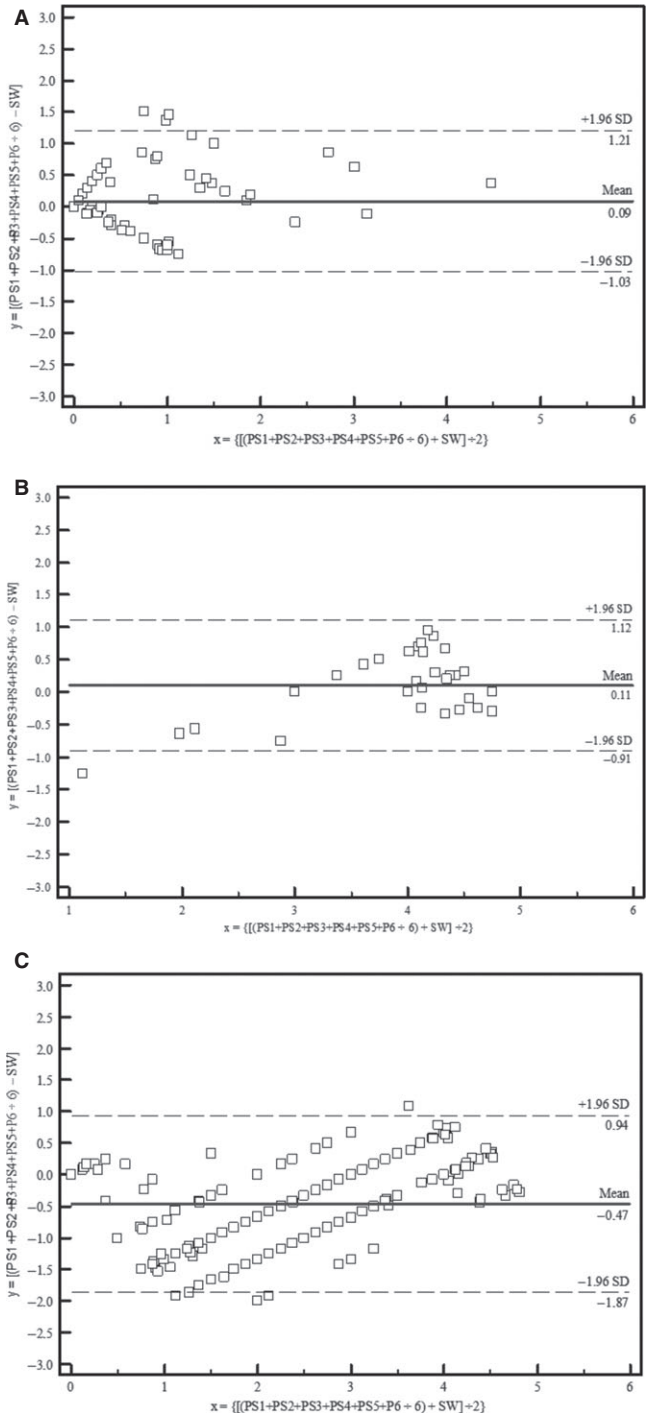


**Figure 2** Bland–Altman graph to analyse the heterogeneity of facial movements score assessed by the software and by the examiners. A: Images analysed during baseline, with infant at rest. B: Images analysed during the invasive procedure. C: Images analysed during recovery, after the invasive procedure. *x* and *y* = *x*-Axis and *y*-axis; PS = Examiner; SW = Software; SD = Standard deviation. The arithmetical mean of the NFCS score attributed by the software and by the average of the 6 examiners was plotted in *x*-axis for each image in the three studied situations (rest, procedure and recovery). The value in the *y*-axis corresponds to the subtraction of software score from the six examiners' average score.

number of studies in the literature that have reported on the development of systems for facial recognition of pain in general (7–11) and the fact that only one study has been conducted in newborn infants (11).

There are challenges that must be overcome to improve the system described here. First, the Luxand Face Software Development Kit (Luxand Inc., Alexandria, VA, USA) is unable to detect points corresponding to the lower area of the face, such as chin movements, or specific points on the tongue. In addition, the need for pain monitoring tools to exhibit preferentially high sensitivity or high specificity should be discussed. On the one hand, false-positive results might indicate the presence of pain and consequently the need to start analgesic measures in infants who might not actually be feeling pain. On the other hand, high specificity would improve the accuracy of diagnosis, but might lead to missing some patients actually feeling pain. As a function of the possible deleterious effects of the use of opiates during the neonatal period (25,26), prioritisation of specificity over sensitivity seems to be the most prudent approach.

One further problem that might have hindered the precision of the image analysis performed by the software is related to the position of the video cameras during the stages of calibration and monitoring. Once the images were calibrated, the users were no longer able to move the video cameras, as during the calibration stage the software identified the distances between the facial nodal points. Any change to the positions of the video cameras would have resulted in alterations to the distances measured on the images that were captured during the monitoring sessions relative to the calibrated images.

The software developed in the present study accurately detected the absence of pain in the resting state and the presence of pain during the performance of painful procedures in full-term and late premature infants exposed to acute pain. One of the goals for the software we developed was to minimise errors associated with individual behavioural differences, as the images of the facial expressions of pain of each infant were compared to his or her own facial images in the resting state. This goal represented a full methodological innovation, as other software previously developed for image analysis has compared adult faces to images stored in a database (27–30), which might lead to error due to anatomical differences among individuals. In addition, the software developed in the present study allowed users to select the time and duration of the monitoring sessions. This particular software application opens new paths for research into the phenomena of the habituation and amplification of pain-related facial movements in response to long-lasting and repeated painful stimuli.

It should be noted that the ideal tool would indicate when pain intensity deserves treatment or treatment adjustments, but due to the current state of art in neonatal pain evaluation, the software developed is able to detect pain presence, but not intensity. The use of an electronic eye, less dependent on humans, could help to integrate pain assessment with pain treatment in the context of neonatal care. The sensitivity and specificity of the software devel-

oped in the present study, to detect pain in newborn infants during the performance of painful procedures and the absence of pain during the resting state, were excellent. In addition, the number of facial actions detected by the software and by the human examiners was homogeneous.

**ACKNOWLEDGEMENTS**

This study was fully supported by a 2012/50157-0 – Regular grant from FAPESP (State of São Paulo Research Support Foundation).

**FUNDING**

TMH received a Grant from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES. The research was funded by Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP; Process number 2012/50157-0 – Regular.

**CONFLICT OF INTERESTS**

There are no conflict of interests to declare.

## References

1. Grunau RE. Neonatal pain in very preterm infants: long-term effects on brain, neurodevelopment and pain reactivity. *Rambam Maimonides Med J* 2013; 4: e0025.
2. 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: 1058–64.
3. Aymar CLG, Coutinho SB. Fatores relacionados ao uso de analgesia sistêmica em neonatologia [Factors associated with the use of systemic analgesia in neonatology]. *Rev Bras Ter Int* 2008; 20: 405–10.
4. Grunau RE, Oberlander T, Holsti L, Whitfield MF. Bedside application of the neonatal facial coding system in pain assessment of premature neonates. *Pain* 1998; 76: 277–86.
5. Grunau RVE, Craig KD. Pain expression in neonates: facial action and cry. *Pain* 1897; 28: 395–410.
6. Zhao W, Chellappa RRA, Phillips PJ, Rosenfeld A. Face recognition: a literature survey. *ACM Comp Surv* 2003; 35: 399–458.
7. Ashraf AB, Lucey S, Cohn J, Chen T, Ambadar Z, Prkachin K, et al. The painful face – pain expression recognition using active appearance models. Proceedings of the 9th international conference on Multimodal interfaces; 2007 Nov 12–15; Nagoya, Aichi, Japan. 9–14.
8. Deyo K, Prkachin KM, Mercer SR. Development of sensitivity to facial expressions of pain. *Pain* 2004; 107: 16–21.
9. Lucey S, Matthews I, Hu C, Ambadar Z, de La Torre F, Cohn JF. AAM derived face representations for robust facial action recognition. Proceedings of the 7th IEEE International Conference on Automatic Face and Gesture Recognition; 2006 April 2–6; Southampton, UK. 155–60.
10. Prkachin KM. Assessing pain by facial expression: facial expression as nexus. *Pain Res Manag* 2009; 14: 53–8.
11. Schiavenato M, Byers JF, Scovanner P, McMahon JM, Xia Y, Lu N, et al. Neonatal pain facial expression: evaluating the primal face of pain. *Pain* 2008; 138: 460–71.
12. Grunau RVE, Johnston CC, Craig KD. Neonatal facial and cry responses to invasive and non-invasive procedures. *Pain* 1990; 42: 295–305.
13. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 2005; 85: 257–68.
14. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 10: 307–10.
15. Van Dijk M, Tibboel D. Update on pain assessment in sick neonates and infants. *Pediatr Clin North Am* 2012; 59: 1167–81.
16. Craig KD. The facial display of pain. In: Finley GA, McGrath PJ, editors. *Measurement of pain in infants and children*. Seattle, WA: IASP, 1998: 103–22.
17. Prkachin KM, Berzins S, Mercer SR. Encoding and decoding of pain expressions: a judgement study. *Pain* 1994; 58: 253–9.
18. Stevens B, McGrath P, Gibbins S, Beyene J, Breau L, Camfield C, et al. Determining behavioural and physiological responses to pain in infants at risk for neurological impairment. *Pain* 2007; 127: 94–102.
19. Hummel P, Van Dijk M. Pain assessment: current status and challenges. *Semin Fetal Neonatal Med* 2006; 11: 237–45.
20. American Academy of Pediatrics. Committee on Fetus and Newborn, Section on Surgery, and Section on Anesthesiology and Pain Medicine. Prevention and management of pain in the neonate. *Pediatrics* 2006; 118: 2231–41.
21. Carbajal R, Rousset A, Danan C, Coquery S, Nolent P, Ducrocq S, et al. Epidemiology and treatment of painful procedures in neonates in intensive care units. *JAMA* 2008; 300: 60–70.
22. Cignacco E, Hamers J, Van Lingen RA, Stoffel L, Büchi S, Müller R, et al. Neonatal procedural pain exposure and pain management in ventilated preterm infants during the first 14 days of life. *Swiss Med Wkly* 2009; 139: 226–32.
23. Lago P, Garetti E, Boccuzzo G, Merazzi D, Pirelli A, Pieragostini L, et al. Procedural pain in neonates: the estate of the art in the implementation of national guidelines in Italy. *Pediatr Anesth* 2013; 23: 407–8.
24. Franck LS, Bruce E. Putting pain assessment into practice: why is it so painful? *Pain Res Manag* 2009; 14: 13–20.
25. Anand KJ, Hall RW, Desai N, Shephard B, Bergqvist LL, Young TE, et al. Effects of morphine analgesia in ventilated preterm neonates: primary outcomes from the NEOPAIN randomized trial. *Lancet* 2004; 363: 1673–82.
26. Bellù R, de Waal KA, Zanini R. Opioids for neonates receiving mechanical ventilation. *Cochrane Database Syst Rev* 2008; (1): CD004212.
27. Bartlett M, Littlewort G, Lainscsek C, Fasel I, Movellan J. Machine learning methods for fully automatic recognition of facial expressions and facial actions. In IEEE International Conference on Systems, Man & Cybernetics, The Hague, the Netherlands; 2004 Oct 10–13; Istanbul, Turkey. 592–97.
28. Dantas GFL. Federação Nacional dos Policiais Federais [homepage on the Internet]. Tribuna Livre: Os Novos Passaportes Biométricos 2008 [Free tribune: the new biometric passports 2008] [cited 2008 Sept 2]. Available at: http://www.fenapef.org.br/htm/com_tribuna_exibe.cfm?Id=562.
29. Kanade T, Cohn JF, Tian Y. Comprehensive database for facial expression analysis. Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition; 2000 March 28–30; Grenoble, France. 46–533.
30. Pamplona Segundo M, Queirolo CC, Bellon ORP, Silva L. Detecção de características para o reconhecimento de faces 3D com expressão facial [Trait detection for 3-D face and facial expression recognition]. *Rev Eletr Inic Cient* 2008; 8: 2.