Inter-observer agreement in the identification of pain faces in full-term and late preterm newborns: cross-sectional study

Concordância interavaliadores na identificação de faces de dor de recém-nascidos a termo e pré-termo tardio: estudo transversal

Tatiany Marcondes Heiderich[1], Marina Carvalho de Moraes Barros[1], Ruth Guinsburg[1]

DOI 10.5935/2595-0118.20200194

ABSTRACT

BACKGROUND AND OBJECTIVES: Difficulty in neonatal assessment is a challenge for the development of pain prevention and treatment strategies. The objective of this study was to analyze the agreement among health professionals in the identification of facial pain movements in images of neonates submitted or not to a painful procedure and to evaluate the discriminatory capacity of these facial movements regarding the presence of pain.
METHODS: Cross-sectional study. Six health professionals trained in neonatal pain assessment evaluated 30 images of newborns undergoing a painful procedure and 30 images of the same newborns at rest, without pain. Each professional evaluated five facial movements that are part of the Neonatal Facial Coding System. Sensitivity, specificity, and positive and negative predictive values were determined. Agreement among professionals was assessed using the kappa coefficient.
RESULTS: The six observers correctly assessed 94±9% of the images obtained at rest as absence of pain and 88±28% of the images obtained during the painful procedure as presence of pain. Protruding forehead, narrowed eyelid cleft, deepened nasolabial furrow, and open mouth showed high sensitivity, specificity, and positive and negative predictive values in the diagnosis of pain, with values between 78-90%. The inter-observer agreement for all 60 images showed a kappa coefficient of 0.60 (95%CI 0.55-0.66).
CONCLUSION: The evaluation of the forehead, eyelid, nasolabial furrow and mouth of newborns showed high sensitivity and specificity to discriminate the presence and absence of pain in static images. The agreement between the evaluators in identifying facial movements related to the expression of pain in newborns was moderate.

Keywords: Facial expression, Newborn, Observer variation, Pain measurement.

RESUMO

JUSTIFICATIVA E OBJETIVOS: A dificuldade na avaliação da dor do recém-nascido é um desafio para o desenvolvimento de estratégias de prevenção e tratamento da dor. O objetivo deste estudo foi analisar a concordância entre profissionais de saúde na identificação de movimentos faciais de dor em imagens de recém-nascidos submetidos ou não a um procedimento doloroso e a capacidade discriminatória quanto à presença de dor desses movimentos faciais.
MÉTODOS: Estudo transversal. Seis profissionais de saúde treinados na avaliação da dor neonatal avaliaram 30 imagens de recém-nascidos submetidos a um procedimento doloroso e 30 imagens em repouso dos mesmos recém-nascidos, sem dor. Cada profissional avaliou cinco movimentos faciais que fazem parte do Sistema de Codificação Facial Neonatal. Sensibilidade, especificidade e valores preditivos positivos e negativos foram determinados. A concordância inter-avaliadores foi avaliada pelo coeficiente kappa.
RESULTADOS: Os seis observadores avaliaram corretamente 94±9% das imagens obtidas em repouso como ausência de dor e 88±28% das imagens obtidas durante o procedimento doloroso como presença de dor. Fronte saliente, fenda palpebral estreitada, sulco nasolabial aprofundado e boca aberta mostraram alta sensibilidade, especificidade e valores preditivos positivo e negativo no diagnóstico de dor, com valores entre 78 e 90%. A concordância inter-avaliadores para todas as 60 imagens mostrou um kappa 0,60 (IC95%0,55-0,66).
CONCLUSÃO: A avaliação da fronte, pálpebra, sulco nasolabial e boca de recém-nascidos mostrou alta sensibilidade e especificidade para discriminar a presença e ausência de dor em imagens estáticas. A concordância inter-avaliadores na identificação de movimentos faciais relacionados à expressão da dor em recém-nascidos foi moderada.

Descritores: Expressão facial, Medição da dor, Recém-nascido, Variações dependentes do observador.

Tatiany Marcondes Heiderich – https://orcid.org/0000-0003-2038-7719;
Marina Carvalho de Moraes Barros – https://orcid.org/0000-0001-6989-3474;
Ruth Guinsburg – https://orcid.org/0000-0003-1967-9861.

1. Universidade Federal de São Paulo, Disciplina de Pediatria Neonatal, Departamento de Pediatria, São Paulo, SP, Brasil.

Submitted on June 18, 2020.
Accepted for publication on August 11, 2020.
Conflict of interests: none – Sponsoring sources: CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior): FAPESP scholarship (Fundação de Amparo à Pesquisa do Estado de São Paulo - Process: 2012/50157-0 - Regular Research Aid): finacial support.

Correspondence to:
Tatiany Marcondes Heiderich
Rua: Marselhesa, 630
04020-060 São Paulo, SP, Brasil.
E-mail: tatianyms@gmail.com

INTRODUCTION

The cortex of the newborn (NB), even during development, is capable of processing impulses originated in the nociceptive receptors in response to pain stimuli, resulting in the sensation of pain[1]. For NBs hospitalized in the Neonatal Intensive Care Unit (NICU), several procedures that are necessary for their treatment are painful or stressful and not always the pharmacological or non-pharmacological interventions for the treatment of pain are used[2].
The evaluation of pain in the NB is fundamental for its proper management and requires knowledge, attitudes and evaluation skills of each professional involved in the neonatal care[2,3]. However, determining the presence of pain in this population is a challenge for health professionals. Several factors are involved in the difficulty to evaluate pain, as the subjective nature of the phenomenon of pain, the impossibility of these patients to verbalize pain, the lack of precise pain indicators in the neonatal period and the non-existence of a universal standard accepted for the qualification of infant pain[2,4]. Among the various behaviors triggered by pain and expressed by the NB, facial movement seems to be the most specific[5-7], besides being the most observed by caregivers[8].
The Neonatal Facial Coding System (NFCS) pain assessment scale has been applied to evaluate facial pain expression in full-term[9,10] and preterm[6,11] NBs. The scale is based on the Facial Action Coding System (FACS)[12] approach, which was developed to codify facial movements and subsequently adapted for NBs[13]. NFCS, described originally by Grunau and Craig[10], evaluates 10 facial movements, but several studies have shown good scale reliability with the reduction in the number of facial movements evaluated[5,14,15]. Although the NFCS scale is easy to apply, since the evaluator only checks for the presence or absence of certain facial movements, there may be variability in the detection of these movements among observers, variability associated with demographic characteristics, previous personal experiences with painful phenomena and/or the evaluator's emotional state[16].
It's possible to observe a growing concern in the literature in the last years in describing ways to assess pain[17,18], investigate the perceptions of the nurses and neonatologists about the knowledge and practice of pain evaluation of NB[2,6] and describing the knowledge, attitudes and practices of the professionals in regard to their handling of pain[19-21]. Such global concern finds wide resonance in Brazil, with several studies performed in the country[22,23]. Generally, the national and international literature indicate that differences and particularities in the comprehension of the NB pain among the several health professionals that care for the NB has been making it difficult to develop strategies of prevention and treatment of pain in the clinical practice[19,24,25].
In this context, the present study had the objective of analyzing the agreement among health professionals in regard to the identification of pain face in images of NBs submitted or not to a painful procedure and evaluating the discriminatory capacity of the NFCS facial mimic scale for the presence or absence of pain in the NBs. The study aimed at specifically assessing the agreement of the health professionals in the identification of the presence of facial movements that characterize pain in pictures of NBs obtained during a painful procedure, and the identification of the absence of these facial movements in front of the same NBs while resting.

METHODS

Cross-sectional study in which health professionals assessed presence or absence of pain in images of full-term and late preterm NBs, submitted or not to a painful procedure. The study was done in the NICU of a tertiary university hospital, in the period of June to August of 2013. Six female professionals, 3 doctors, 2 nurses and 1 physiotherapist specialized in neonatology and that worked in the center for formation of human resources for NICU were included. The inclusion criteria demanded that they worked in the NICU of the institution and had more than 10 years of expertise in neonatology and activity in NICU. There were no exclusion criteria.
Through an image bank of NBs generated in the study[26], two images of 30 NBs were selected: one captured during rest, before a painful procedure and considered as an image of without pain; and another captured during punction, in the exact moment of the invasion of the skin, considered as an image of pain.
The images were obtained through three Foscam IPCAM cameras (FOSCAM Manufacturer - China, with authorized FOSCAM BRAZIL resale), in the following configurations: high sensitive microphone, 300k pixels, digital remotepan/tilt, color, wireless, night vision, motion detect, wi-fi, MJPEG video compression. The NBs had a gestational age between 34 and 41 weeks, evaluated by the best obstetric estimate, had no congenital malformations or facial anomalies, and didn't require ventilatory support or naso/orogastric probe. The painful procedures, performed by medical indication, included capillary or venous puncture and intramuscular injection, being performed between 24 and 168 hours of life.
Each health professional was oriented to individually evaluate the 60 images in the sequence in which they were arranged, randomly, with no time restriction. For each image, observers pointed out whether it was a NB with or without pain and scored the presence or absence of 5 pain facial movements that make up the NFCS scale[9,10]. This scale was chosen because it provides valid and sensitive information regarding the nature and intensity of pain, with an inter-observer reliability of 88%, allowing for an effective communication between the NB and the people involved in their care[10]. For this study, the following facial movements on the scale were considered indicative of pain: protruding forehead, narrowed eyelid cleft, deepened nasolabial furrow, open mouth and mouth horizontally or vertically stretched. For each present movement, one point was attributed. The presence of pain, according to the evaluation of the facial mimic, was defined when three or more movements were present[9,10].
For the analyses, pain was considered present when the image was obtained during the procedure, and it was considered absent when the image was obtained during rest. Sensitivity, specificity and positive and negative predictive values were calculated for each observer individually when identifying absent or present pain for the 60 analyzed images. Inter-observer agreement on the presence or absence of pain was assessed by Fleiss' Kappa coefficient[27]. In order to identify which of the facial movements had the highest inter-observer agreement, the number of images in which there

was agreement between 4, 5 and 6 evaluators for each of the facial movements was verified for the images of absent and present pain separately. Next, the paired images of the 30 NBs, i.e., the image obtained during rest followed by the image obtained during the procedure of the same NB were evaluated. Also, an analysis was made for each observer and for each of the facial movements regarding the percentage of images in which the movements were correctly evaluated, i.e., facial movement absent at rest and facial movement present during the procedure.

It's worth highlighting that, when designing the original study[26], the authors carefully considered the possibility of administering non-pharmacological analgesia to the NBs, but this would be an important limitation of the studies' objectives. The study was approved by the Research Ethics Committee with the premises that only minor routine procedures would be done, only once, and the matter of pain would be openly and transparently discussed with the parents, as well as the possible benefits of the study, before requesting for their consent. The image bank obtained from these premises allowed for the present data analysis.

The study was approved by the Research Ethics Committee (CEP 1299/09) of the institution. The Free and Informed Consent Term (FICT) was obtained from those responsible for the NBs from whom the images were obtained and from the professionals who evaluated them.

Statistical analysis

Performed in the SPSS Statistics software, version 17.0 (IBM SPSS Statistics for Windows, USA). The Fleiss Kappa Coefficient was calculated using the "Online Kappa Calculator "[28]. The adopted significance level was p<0.05.

RESULTS

The mean age of evaluators was 42 years old (variation: 34-46), were graduated 18 years ago (variation: 12-23), with time of activity in neonatology of 15 years (variation: 11-19) and time of performance in NICU also of 15 years (variation: 11-20).

Of the 30 photographed NBs, 15 were male, with a mean gestational age of 37±1.4 weeks (variation: 35-41), being 7 late prematures and 23 full-term. The mean birth weight was 2962±593g (variation: 2115-4100), and seven NBs were underweight. Birth was vaginal in 11 and cesarean in 19. The median and Apgar bulletin variation in the first and fifth minute of life were 9 (6-9) and 9 (7-10), respectively. The images were obtained before and during the following painful procedures: blood collection by venipuncture in the back of the hand for laboratory tests (n=13), blood collection through capillary puncture in the calcaneus for neonatal screening tests (n=12) and for blood glycemia dosing (n=3), and intramuscular injections in the thigh for hepatitis B vaccine (n=2).

Each image obtained at rest and during the procedure was evaluated by the six professionals who assigned them a score of 0-5. For each image, the mean scores of the six evaluators was calculated and, subsequently, the overall mean of the images obtained in the absence and presence of the procedure. The overall mean score of the 30 images in the absence of the procedure was 0.56±0.46 (variation: 0.10-1.33), lower than the overall mean of the 30 images obtained during the procedure of 3.95±0.99 (variation: 0.50-4.67) (p<0.001).

The same analysis was performed individually for each of the facial movements. The mean score of the 6 evaluators for the presence of protruding forehead on the 30 images without pain was 0.11±0.14 (variation: 0.00-0.50) and, on the images of pain, 0.88±0.23 (variation: 0.00-1.00; p<0.001). For the presence of a narrowed eyelid cleft, the mean score for the images without and with pain was 0.11±0.15 (variation: 0.00-0.50) and 0.87±0.21 (variation: 0.33-1.00; p<0.001), respectively. As for the presence of deepened nasolabial furrow, the mean score of the images without pain and with pain was, respectively, 0.11±0.18 (variation: 0.00-0.83) and 0.78±0.28 (variation: 0.00-1.00; p<0.001). For the open mouth facial movement, the mean score for the images without pain was 0.13±0.17 (variation: 0.00-0.67) and for the images with pain was 0.81±0.30 (variation: 0.00-1.00; p<0.001). Finally, for the stretched mouth, the images without pain had a mean score of 0.09±0.14 (variation: 0.00-0.50) and those with pain of 0.59±0.20 (variation: 0.00-1.00; p<0.001).

Following that, the percentage of images correctly evaluated was verified, considering the total score of zero, 1 and 2 as an indicative of absence of pain and the total score 3, 4 and 5 as presence of pain. For this analysis, the percentage of correct evaluations by the six evaluators was calculated, considering the images obtained before the procedure, during rest, as absence of pain and, during the puncture, as presence of pain.

On average, the six evaluators evaluated correctly 94.4±9.1% (variation: 66.7-100%) of the images obtained before the procedure as absent pain and 87.8±27.6% (variation: 0.0-100%; p=0.219) of the images obtained during the painful procedure as present pain. After that, the sensitivity, specificity, positive predictive value and negative predictive value were calculated for the response of the six evaluators for the 60 images, taking into account all movements and each individual facial movement (Table 1).

Table 1. Values (variation) of sensitivity, specificity, positive predictive value and negative predictive value of the six evaluators, in the evaluation of images during rest and during the procedure for the five facial movements individually and together

Facial movements	Sensitivity (%)	Specificity (%)	Predictive positive value (%)	Predictive negative value (%)
Protruding forehead	88.9 (80.0 - 96.7)	88.9 (76.7 - 100.0)	89.7 (80.6 - 100.0)	89.4 (83.3 - 95.8)
Narrowed eyelid cleft	87.8 (80.0 - 100.0)	88.9 (70.0 - 100.0)	90.0 (76.9 - 100.0)	88.6 (83.3 - 100.0)
Deepened nasolabial furrow	78.3 (53.3 - 90.0)	88.9 (80.0 - 100.0)	88.8 (79.3 - 100.0)	81.4 (68.2 - 88.9)
Open mouth	81.1 (63.3 - 93.3)	86.7 (60.0 - 100.0)	88.5 (70.0 - 100.0)	83.3 (73.2 - 90.0)
Stretched mouth	59.4 (64.0 - 93.3)	90.6 (80.0 - 100.0)	81.3 (25.0 - 100.0)	73.2 (46.2 - 93.5)
All movements	87.8 (80.0 - 93.3)	94.4 (86.7 - 100.0)	94.5 (87.1 - 100.0)	88.7 (83.3 - 93.8)

Inter-observer agreement was verified for the images obtained during rest and during the procedure by the Fleiss Kappa coefficient, considering the facial movements together and separately (Table 2). The evaluation of inter-observer agreement for all 60 images showed a coefficient of 0.60 (variation: 0.55-0.66). In the evaluation of the images during rest, the coefficient was 0.65 (0.58-0.72) and in the images obtained during the procedure, it was 0.56 (0.48-0.64). In the separate evaluation of the facial movements, the coefficient of agreement varied from 0.59 to 0.69 for the images obtained during rest, and from 0.56 to 0.79 for those obtained during the procedure, except for the coefficient observed for the stretched mouth, which was 0.04.

Table 2. Inter-observer agreement for the images obtained during rest and during the procedure, considering the facial movements together and separately

Inter-observer agreement	Kappa Fleiss Index	Confidence interval of 95%
Total of 60 images*	0.60	0.55 - 0.66
30 images of rest#	0.65	0.58 - 0.72
Protruding forehead	0.62	0.46 - 0.78
Narrowed eyelid cleft	0.64	0.47 - 0.82
Deepened nasolabial furrow	0.62	0.46 - 0.78
Open mouth	0.59	0.42 - 0.76
Stretched mouth	0.69	0.53 - 0.86
30 images during procedure$	0.56	0.48 - 0.64
Protruding forehead	0.79	0.64 - 0.93
Narrowed eyelid cleft	0.71	0.54 - 0.88
Deepened nasolabial furrow	0.57	0.41 - 0.73
Open mouth	0.70	0.55 - 0.85
Stretched mouth	0.04	-0.07 - 0.16

*Referring to the absence of signs on the 30 images during rest and the presence of signs on the 30 images during the procedure; #referring to the absence of facial movements; $referring to the presence of facial movements.

Table 3 presents the values of agreement between 4, 5 and 6 observers for each facial movement on the images obtained during the painful procedure and during rest. For the images of rest, in the absence of pain, all observers agreed with the absence of the five facial movements evaluated in 53.3 to 66.7% of the images. Also for the images of rest, 80% agreement was obtained by at least five evaluators for the protruding forehead, deepened nasolabial furrow and stretched mouth.

For the movements of narrowed eyelid cleft and open mouth, 80% agreement was obtained by four or more evaluators. Regarding the images obtained during the procedure (presence of pain), the agreement between the six evaluators varied from 3.3 to 73.3%. 80% agreement was obtained by at least five evaluators for the protruding forehead and narrowed eyelid cleft movements.

For the deepened nasolabial furrow and open mouth, 80% of agreement was obtained by four or more evaluators. Stretched mouth did not reach 80% of agreement when four or more evaluators were considered.

Table 4 presents the results of inter-observer agreement in the analysis of the paired images of the 30 NBs with the image obtained at rest, followed by the image obtained during the procedure, regarding the evaluation of facial movements as absent in

Table 3. Agreement between 4, 5 and 6 evaluators for the different facial movements in the images of newborns without pain (rest) and with pain (painful procedure)

Inter-observer agreement	Protruding forehead		Narrowed eyelid cleft		Deepened nasolabial furrow		Open mouth		Stretched mouth	
	n	%	n	%	n	%	n	%	n	%
Absence of pain										
6 evaluators	16	53.3	19	63.3	18	60.0	16	53.3	20	66.7
5 evaluators	9	30.0	3	10.0	8	26.7	7	23.3	4	13.3
4 evaluators	4	13.3	7	23.3	2	6.7	5	16.7	5	16.7
Presence of pain										
6 evaluators	22	73.3	21	70.0	12	40.0	17	56.7	1	3.3
5 evaluators	3	10.0	3	10.0	10	33.3	6	20.0	4	13.3
4 evaluators	2	6.7	1	3.3	3	10.0	2	6.7	14	46.7

Table 4. Inter-observer agreement on the analysis of paired images of the 30 newborns regarding the evaluation of facial movements as absent in the images obtained without pain at rest and present in the images obtained with pain during the procedure

Inter-observer agreement	Protruding forehead		Narrowed eyelid cleft		Deepened nasolabial furrow		Open mouth		Stretched mouth	
	n (%)	% accum	n (%)	% accum	n (%)	% accum	n (%)	% accum	n (%)	% accum
6 evaluators	12 (40.0)	40.0	13 (43.3)	43.3	6 (20.0)	20.0	9 (30.0)	30.0	-	0
5 evaluators	8 (26.7)	66.7	5 (16.7)	60.0	10 (33.3)	53.3	6 (20.0)	50.0	3 (10.0)	10.0
4 evaluators	5 (16.7)	83.4	7 (23.3)	83.3	7 (23.3)	76.6	7 (23.3)	73.3	12 (40.0)	50.0
3 evaluators	2 (6.7)	90.1	2 (6.7)	90.0	2 (6.7)	83.3	2 (6.7)	80.0	7 (23.3)	73.3
2 evaluators	2 (6.7)	96.8	1 (3.3)	93.3	1 (3.3)	86.6	2 (6.7)	86.7	4 (13.3)	86.6
1 evaluators	-		1 (3.3)	96.6	1 (3.3)	89.9	1 (3.3)	89.9	3 (10.0)	96.6
Total	29 (96.8)		29 (96.6)		27 (89.9)		27 (89.9)		29 (96.6)	

% accum: accumulated percentage.

the images obtained during rest, without pain, and present in the images obtained during the procedure, with pain. The protruding forehead was scored as absent in the image without pain and present in the image with pain by six evaluators in 40.0% of the NB photos, the eyelid cleft in 43.3%, the nasolabial furrow in 20% and the open mouth in 30%. The absence and presence of stretched mouth on the images with and without pain, assessed in pairs for the same NB, was not correctly identified by the six observers in none of the NB.

The correct identification of the absence and presence of pain in 80% of the NBs in the images analyzed in pairs was observed by four evaluators for the facial movements of protruding forehead and narrowed eyelid cleft. The correct identification of the absence and presence of pain in 80% of the NBs on the images analyzed in pairs was observed by three evaluators for the facial movements of deepened nasolabial furrow and open mouth.

DISCUSSION

In the present study, health professionals evaluated images of NBs obtained during rest and during a painful procedure, assigning one point to the presence of each of the five facial movements: protruding forehead, narrowed eyelid cleft, deepened nasolabial furrow, open mouth and stretched mouth, according to the NFCS scale[9,10]. The score attributed to the obtained images during the painful procedure was higher when compared to the attributed to the images during rest, indicating a good discriminatory capacity of the scale in differentiating, in NBs, the presence or absence of facial movements of pain. The evaluation of NB photos by health professionals showed values of sensitivity, specificity and positive and negative predictive value between 88 and 94%, showing a good discriminatory power of the observation of facial movements in order to identify the presence or absence of pain in this age group.

According to the literature, the ability to assess pain through facial mimicry has made it one of the main indicators of pain in pre-verbal patients[5], which is confirmed by the present study in the difference between the mean scores obtained in the evaluation of facial movements together and individually on the images obtained at rest and during the painful procedure. Authors[29] obtained similar results when comparing NFCS with the Children's and Infant's Postoperative Pain Scale (CHIPPS). These authors report that for NFCS the internal consistency was high (α=0.936 for all NBs, α=0.943 for premature and α=0.880 for full-term NBs) and conclude that NFCS is a good tool to assess the presence and absence of pain.

Among the NFCS parameters used for the assessment of NB photos during rest and during painful procedures, the protruding forehead, narrowed eyelid cleft, deep nasolabial furrows and open mouth showed sensitivity, specificity and positive and negative predictive value between 78 and 90%.

The stretched mouth presented the worst performance, with sensitivity below 60%. Such results were reinforced when the images were analyzed in pairs, comparing the image of the same NB before and during the painful procedure. In this evaluation, the absence and presence of stretched mouth, without and with pain, respectively, on the images for the same NB were not correctly identified by the six evaluators in none of the 30 patients. A similar result was pointed out by the pioneer study[10]. The authors evaluated 140 healthy NBs regarding facial expression for a period of 60 seconds, followed by a period during which the capillary puncture was performed in the calcaneus. The study showed that the forehead movement, with protrusion of the eyebrow, closing of the eyelids, deepening of the nasolabial furrow and opening of the lips were observed in 99% of the NBs during puncture; while the opened or vertically stretched mouth was observed in only 70 and 43%, respectively.

In none of the children the horizontally stretched mouth or the crimped lips were observed and, because of that, these two facial movements were not included in the further studies that used facial movements for the evaluation of pain in NBs.

Similarly, the authors[22] observed during exam collection, using the NFCS scale to assess NB pain, that some facial manifestations were more frequent, such as open mouth (96.1%), followed by protruding forehead (88.4%) and narrowed eyelid cleft (76.9%).

Not using the 10 facial movements proposed by NFCS to evaluate pain was also addressed by Peters et al.[14] and other researchers[6], who suggested assessing only the three movements most commonly observed in the context of neonatal pain: protruding forehead, narrowed eyelid cleft and deepened nasolabial furrow. The study[30] examined two coding systems, the NFCS and the Modified Behavior Pain Scale (MBPS) and observed the factorial structure of these scales. The authors confirmed that the internal consistency of the NFCS scale with three items (squeezed eyes, vertically stretched mouth and horizontally stretched mouth) is similar to the NFCS scale of seven items (α between 0.75 and 0.87 for the 3 items). In addition to that, the authors[30] stressed that shorter versions of the scale increase its potential for clinical use.

In the analysis of agreement among the health professionals regarding the evaluation of NB images during rest and the painful procedure, moderate agreement was observed when all 60 images were analyzed together or separately, with 89.4% agreement for images at rest and 83.4% for those obtained during the procedure. The inter-observer agreement in the evaluation of the images obtained during the procedure was similar to that pointed out by other authors, as in the study[5], in which agreement of 86% was reported in the response of three professionals, a nurse, an occupational therapist and a social worker in the evaluation of acute pain in 40 premature infants using NFCS. The evaluation was performed during the procedures of removal of the blanket involving the patient, heel friction with a cotton swab, heel puncture with a lancet, heel compression with cotton for hemostasis and in the recovery phase after procedure. Study[11] analyzed 56 full-term and preterm NBs photographed during heel puncture using NFCS.

The authors observed an 89% inter-observer agreement in response to the pain evaluation during the painful procedure. Similarly, another study[15] found a 94% agreement between two evaluators when analyzing the behavioral response to pain of 36 full-term infants and 31 healthy premature infants submitted to capillary puncture through four facial actions: protruding forehead, narrowed eyelid cleft, deep nasolabial furrow and open mouth.

It's worth noting that the present study presented the limitation of a small number of evaluated professionals. In addition, the analysis of the presence of pain in photos of NBs at rest and submitted to painful procedures should also be considered because, since this is the evaluation of a static 2D image, it is not possible to analyze all facial movements suggested by NFCS, such as tense tongue and chin tremor.

Studies done at bedside are needed to confirm the present findings. However, even with a reduced number of evaluators, the use of printed images allowed all professionals to visualize the same faces of NBs, which made it possible to compare the observation made by the professionals under equal conditions. Further studies should be performed in order to analyze the dynamic of the health professional's view when evaluating NB pain. Moreover, new researches are necessary to comprehend how the presence of devices in front of the NBs' faces interfere with the assessment and, in consequence, with the management of pain, since the fixation of the nasal probe and orotracheal tube, as well as the eye protection used when the NB is under photo therapy, may hinder the adequate visualization of the facial points and interfere in the assessment of pain in the NB.

CONCLUSION

The assessment of facial movements of the forehead, eyelid cleft, nasolabial furrow and the mouth of NBs presented high sensibility and specificity for determining the presence and absence of pain in images of NBs submitted or not to painful procedures.

The inter-observer agreement on the identification of facial movements related to pain expression in NBs was moderate. Of the analyzed facial movements, the protruding forehead, the narrowed eyelid cleft, the deep nasolabial furrow, and the open mouth showed high sensitivity in the identification of the presence of NB pain.

ACKNOWLEDGMENTS

Thanks to CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Coordination for the Improvement of Higher Education Personnel) for the scholarship and to FAPESP (Fundação *de Amparo à Pesquisa do Estado de São Paulo* - Foundation for the Support of Research in the State of São Paulo - Process: 2012/50157-0 - Regular Research Aid), for the financial support.

REFERENCES

1. Marchant A. 'Neonates do not feel pain': a critical review of the evidence. Biosci Horizons. 2014;7:1-9.
2. Cong X, McGrath JM, Cusson RM, Zhang D. Pain assessment and measurement in neonates: an updated review. Adv Neonatal Care. 2013;13(6):379-95.
3. Efe E, Dikmen Ş, Altaş N, Boneval C. Turkish pediatric surgical nurses' knowledge and attitudes regarding pain assessment and nonpharmacological and environmental methods in newborns' pain relief. Pain Manag Nurs. 2013;14(4):343-50.
4. Laudiano-Dray MP, Pillai Riddell R, Jones L, Iyer J, Whitehead K, Fitzgeral M, et al. Quantification of neonatal procedural pain severity: a platform for estimating total pain burden in individual infants. Pain. 2020;161(6):1270-7.
5. 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(3):277-86.
6. Stevens B, Johnston C, Petryshen P, Taddio A. Premature Infant Pain Profile: development and initial validation. Clin J Pain. 1996;12(1):13-22.
7. Silva TP, Silva LJ. Escalas de avaliação da dor utilizadas no recém-nascido: revisão sistemática. Acta Med Port. 2010;23:437-54.
8. Hadjistavropoolos HD, Craig KD, Grunau RV, Johnston CC. Judging pain in newborns: facial and cry determinants. J Pediatr Pysichol. 1994;19(4):485-91.
9. Grunau R, Johnston C, Craig K. Neonatal facial and cry responses to invasive and non-invasive procedures. Pain. 1990;42(3):295-305.
10. Grunau RV, Craig KD. Pain expression in neonates: facial action and cry. Pain. 1987;28(3):395-410.
11. Craig KD, Whitfield MF, Grunau RV, Hadjistavropoulos HD. Pain in the preterm neonate: behavioural and physiological indices. Pain. 1993;52(3):287-99.
12. Chen Z, Ansari R, Wilkie D. Learning pain from action unit combinations: a weakly supervised approach via multiple instance learning. IEEE Trans Affect Comput. 2019;10.11091/TAFFC.2019.2949313.
13. Ahola Kohut S, Pillai Riddell R, Flora DB, Oster H. A longitudinal analysis of the development of infant facial expressions in response to acute pain: immediate and regulatory expressions. Pain. 2012;153(12):2458-65.
14. Peters JW, Koot HM, Grunau RE, Boer J, van Druenen MJ, Tibboel D, ret al. Neonatal Facial Coding System for assessing postoperative pain in infants: item reduction is valid and feasible. Clin J Pain. 2003;19(6):353-63.
15. Rushforth JA, Levene MI. Behavioural response to pain in healthy neonates. Arch Dis Child Fetal. 1994;70(3):F174-6.
16. Silva YP, Gomez RS, Máximo TA, Silva AC. Avaliação da dor em neonatologia. Rev Bras Anestesiol. 2007;57(5):565-74.
17. Caetano EA, Lemos NR, Cordeiro S, Pereira FM, Moreira DS, Buchhorn SM. O recém-nascido com dor: atuação da equipe de enfermagem. Esc Anna Nery. 2013;17(3):439-45.
18. Bueno M, Stevens B, Barwick MA, Riahi S, Li SA, Lanese A, et al. A cluster randomized clinical trial to evaluate the effectiveness of the Implementation of Infant Pain Practice Change (ImPaC) Resource to improve pain practices in hospitalized infants: a study protocol. Trials. 2020;21(1):16.
19. Balice-Bourgois C, Zumstein-Shaha M, Vanoni F, Jaques C, Newman CJ, Simonetti GD. A systematic review of clinical practice guidelines for acute procedural pain on neonates. Clin J Pain. 2020;36(5):390-8.
20. Maciel HIA, Costa MF, Costa ACL, Marcatto JO, Manzo BF, Bueno M. Medidas farmacológicas e não farmacológicas de controle e tratamento da dor em recém-nascidos. Rev Bras Ter Intensiva. 2019;31(1):21-6.
21. Hills E, Rosenberg J, Banfield N, Harding C. A multidisciplinary approach to the implementation of non-pharmacological strategies to manage infant pain. Infant. 2020;16(2):78-81.
22. Pinheiro IO, Lima FE, Magalhães FJ, Farias LM, Sherlock MS. Pain evaluation in newborns using the Neonatal Facial Activity Coding scale during blood gases analysis. Rev Dor. 2015;16(3):176-80.
23. Cardoso TA, Rocha RSB, Cunha KC. Influência da utilização da glicose 0,25% na avaliação da dor neonatal em prematuros tardios submetidos a fisioterapia respiratória. Rev Para Med. 2014;28(3):43-8.
24. Hu J, Ruan H, Li Q, Gifford W, Zhou Y, Yu L, Harrison D. Barriers and facilitators to effective procedural pain treatments for pediatric patients in the Chinese context: a qualitative descriptive study. J Pediatr Nurs. 2020;54:78-85.
25. Martins SW, Enumo SRF, Paula KMP. Manejo da dor neonatal: influência de fatores psicológicos e organizacionais. Estud Psicol. 2016;33(4):633-44.
26. Heiderich TM, Leslie AT, Giomsburg R. Neonatal procedural pain can be assessed by computer software that has good sensitivity and specificity to detect facial movements. Acta Paediatr. 2015;104(2):e63-9.
27. Hale CA, Fleiss JL. Interval estimation under two study designs for kappa with binary classifications. Biometrics. 1993;49(2):523-34.
28. Randolph JJ. Online Kappa Calculator [Computer software]. Advances in Data Analysis and Classification, http://justus.randolph.name/kappa (2008, accessed 2 May 2019).
29. Kappesser J, de Laffolie J, Faas D, Ehrhardt H, Herman C. Comparison of two neonatal pain assessment tools (Children and Infant's Postoperative Pain Scale and the Neonatal Facial Coding System-Revised) and their relations to clinicians' intuitive pain estimates. Eur J Pain. 2019;23(4):708-18.
30. DiLorenzo MG, Pillai Riddell R, Flora DB, Craig KD. Infant Clinical Pain Assessment: core behavioral cues. J Pain. 2018;19(9):1024-32.