"THE RELIABILITY AND EFFECTIVENESS OF TOOLS FOR ASSESSING PAIN IN NEWBORNS"

Rustam Abdumutalovich Ismailov

Department of Anaesthesiology, Reanimatology, and Emergency Medical Care,

Andijan State Medical Institute, Uzbekistan

ANNOTATION:

A variety of neonatal observational tools are available for pain assessment. However, their use in clinical practice is limited. This study compares three predictive scales for neonatal pain assessment: *NFCS-R* (*Neonatal Facial Coding System - Revised*), *CHIPPS* (*Children and Infant's Postoperative Pain Scale*), and the in terms of their psychometric properties. All three tools demonstrated high relative convergent validity. Therefore, when selecting a tool, it is essential to consider its clinical utility and explore opportunities for further improvement.

Keywords: Pain, assessment, newborns, *NFCS-R* scale, *CHIPPS*

The relevance of the problem.

The need to assess pain in infants is widely recognized and considered a prerequisite for adequate pain management. Today, there are several tools available for assessing pain in newborns [1, 11, 12]. Despite the fact that the availability of these tools is a significant advantage, their use in clinical practice is limited [2, 3, 5]. One of the main issues may be that there is no "gold standard" for these tools due to the lack of systematic psychometric comparative studies [11, 12]. The *NFCS-R scale (Neonatal Facial Coding System*) [13] (Grunau and Craig, 2010) is widely used in research and can also be applied in clinical settings. The postoperative pain scale for children and infants (*CHIPPS)* [7, 8] is easily accessible and is often used not only because of its availability but also due to its high reliability, time efficiency, and simplicity in coding and assessment. Due to the limited available information,

many users are unaware that *CHIPPS* has only been validated for full-term neonates experiencing postoperative pain, but it is used for both preterm and full-term neonates with various pain conditions. The limited use of standardized assessment tools [5, 12] also implies that healthcare professionals often assess infant pain intuitively .

Perhaps comparing intuitive assessments of observers with pain scores included in observation tools could shed light on which signals are used when evaluating pain. Encouragingly, some studies have shown that the crying and facial movements of infants influenced pain assessment by clinical observers [15], as these signals directly correspond to elements of categories included in observation tools.

The goal of this article is to conduct a psychometric comparison of the effectiveness of the *CHIPPS, FLACC*, and *NFCS-R scales*, with a particular focus on their ability to differentiate pain.

Materials and Methods

In the present study, a series of video fragments depicting the facial expressions and behaviors of 44 newborns admitted to the neonatal intensive care unit of the Andijan Regional Children's Multidisciplinary Medical Center were consecutively collected. Among them, 30 were preterm (gestational age less than 37 weeks) with an average gestational age of 33.57 weeks (SD=1.48). For full-term newborns, the average gestational age was 38.69 weeks (SD=1.6). There were 27 girls (16 preterm and 11 full-term). At the time of recording, the average age of the newborns was 1.92 days (SD=2.58). Exclusion criteria from the study included newborns with neurological pathologies, hepatorenal disorders, syndromic diseases, facial malformations, and altered muscle tone.

All infants were recorded on video using an HD camera (Canon Legria HF M46) during a painful situation (venipuncture or peripheral venous catheter

placement), and the videos were processed (segmented) using TechSmith Camtasia 2019 software (TechSmith Corporation, USA). The extracted video sequences started 10 seconds before the onset of the painful or stressful situation. These situations were chosen because pediatricians and neonatologists considered them painful or stressful and similar to those described in the literature as inducing pain and stress [9]. The order in which the videos were recorded was not predetermined or randomized. However, pain and stress situations were recorded on the same day for each newborn, with an average interval of 9.5 hours. Data on the order of situations were available for 40 newborns. Of these, 17 newborns experienced stress before the painful situation, and 23 newborns experienced pain before the stressful situation. T-tests for mean differences showed that the order had no significant impact on the pain assessment using *NFCS-R* and *CHIPPS.*

The *NFCS* is a unidimensional tool for assessing pain in newborns, as it is based solely on facial movements. While the original version (*NFCS*) included ten facial movements, the reduction to five key points increased specificity without reducing sensitivity . Therefore, the authors revised *NFCS* by reducing the number of items to the following five key elements [13]: 1) brow bulge (bulging, folds, and/or vertical furrows above the eyebrows and between them); 2) eye squint (squeezing and/or bulging of the eyelids); 3) nasolabial furrow (pulling up and deepening of the nasolabial furrow, a line or wrinkle starting near the nostrils and extending downward and outward toward the corners of the mouth); 4) horizontal mouth stretch (distinct horizontal stretching at the corners of the mouth, sometimes accompanied by a taut upper lip); and 5) tight tongue (elevated, cup-shaped tongue with sharp, taut edges). Only the presence or absence of facial movements was coded with the words "present" (score 1), "absent" (score 0), or "uncertain" (score "NC"). The occurrence of each facial action (one after another) was coded within predefined time segments. Scores for individual items are summed, with possible scores ranging from 0 to 5 for each coded time segment.

Like *NFCS-R,* the *CHIPPS* scale relies on behavioral signals for pain assessment. However, its five categories cover several behavioral aspects: 1) crying; 2) facial expression; 3) body posture; 4) leg posture; and 5) motor agitation. The newborns were observed for 15 seconds, and based on behavioral definitions, each item was rated as "0" (no pain), "1" (potential pain), or "2" (obvious pain) during this time frame. Scores for individual items were summed, so the final score for the 15-second time segment ranged from 0 to 10. A total score of 4 or higher indicates the need for analgesic treatment [8].

Data were analyzed using Microsoft Excel 2016. For better comparability, we calculated the average score for each item, meaning that the total score for the tool was divided by the number of items. As a measure of relative validity, Spearman's correlation coefficients between *NFCS-R* and *CHIPPS* were selected.

Results. Since the intraclass correlation coefficients (ICC) were high for both tools across all newborns, the scores were averaged across all rating scales for further analysis (Table 1).

Table 1

Intraclass Correlation Coefficients (ICC) for NFCS-R and CHIPPS Scales in Painful Situations for Preterm and Full-Term Newborns

№	Pain Assessment Scale	Pain Rating	Preterm Newborns	Full-Term Newborns	All Newborns
1	NFCS-R	Pain Assessment	0.980 (0.963-0.990)	0.944 (0.864-0.980)	0.975 (0.958-0.986)
2	CHIPPS	Pain Assessment	0.970 (0.945-0.985)	0.954 (0.889-0.984)	0.968 (0.947-0.981)

Table 2

Intraclass Correlation and Convergent Validity for *NFCS-R and CHIPPS* Scales in Painful Situations for Preterm and Full-Term Newborns

№	Pain Assessment Scale	ICC for Pain Situations	Preterm Newborns	Full-Term Newborns	All Newborns	Internal Consistency (α)

1	*NFCS-R*	Pain Rating	0.980 (0.963-0.990)	0.944 (0.864-0.980)	0.975 (0.958-0.986)	All Newborns: α = 0.936
						Preterm: α = 0.943
						Full-Term: α = 0.880
2	*CHIPPS*	Pain Rating	0.970 (0.945-0.985)	0.954 (0.889-0.984)	0.968 (0.947-0.981)	All Newborns: α = 0.83
						Preterm: α = 0.82
						Full-Term: α = 0.85

Convergent Validity (Effect Size):

The correlation between the *NFCS-R* and *CHIPPS* scales showed a large effect size for the associations between the two tools in assessing pain in preterm newborns. The effect size was also large in the full-term group, further confirming the convergent validity of these two pain assessment tools.

Table 3. Results of Repeated Measures ANOVA for NFCS-R and CHIPPS Scales

Scale	Source of Variation	*Df*	*F*	*p*	η^2	*d*	*Er*
NFCS-R	Pain Situation	1	10.886	0.002	0.214	0.48	40
	Gestational Age	1	1.167	0.286	0.028		
	Pain Situation × Gestational Age	1	1.264	0.268	0.031		
CHIPPS	Pain Situation	1	13.161	0.001	0.239	0.52	42
	Gestational Age	1	1.372	0.248	0.032		
	Pain Situation × Gestational Age	1	1.416	0.241	0.033		

Note: Df – Degrees of freedom, *F* – F-ratio (variation coefficient*), p* – p-value, η^2 – Eta squared (effect size), *d* – Effect size (Cohen's d*), Er* – Within-group error

Key Findings: Both the *NFCS-R* and *CHIPPS* scales showed a significant main effect for the pain situation (*p = 0.002 for NFCS-R and p = 0.001 for CHIPPS),* indicating that pain significantly affected the assessments in both scales. The gestational age and pain situation × gestational age interaction did not show

significant effects ($p > 0.05$), suggesting that gestational age did not significantly influence the pain assessments when using either scale.

Discussion. Despite limited preparation time, both scales demonstrated a high degree of agreement on the pain assessment tool items (*CHIPPS* range: 0.918 - 0.981, *NFCS* - R: 0.910 - 0.980). The inter-rater reliability for *NFCS-R* was comparable to that established in the literature [14], while the inter-rater reliability for *CHIPPS* was higher [4, 7], presumably due to the extensive time spent in advance clarifying the coding criteria. Both tools achieved good internal consistency (*NFCS-R*: 0.94; *CHIPPS*: 0.83), suggesting they are homogeneous. The validity scores for CHIPPS in our study were slightly lower compared to existing data on *CHIPPS* [7] and *NFCS-R.* Since *CHIPPS* has not yet been evaluated for premature newborns, it is promising that the validity assessment for premature newborns in *CHIPPS* is similar to the evaluation for all newborns (0.82; 0.83). The relationship between the two pain assessment tools, i.e., relative validity, was significantly high in terms of effect size and resembled the results of other studies [7, 10]. Perfect associations between both pain assessment tools are unlikely, as *NFCS-R* measures facial expression, while *CHIPPS* includes additional pain behaviors. However, the relationships between the two tools were higher than the associations between each tool and clinicians' intuitive assessments *(NFCS* 0.55–0.66; *CHIPPS* 0.53–0.55). Although these latter associations are still strong, intuitive judgments often arise implicitly, tend to be biased, and may vary widely. Therefore, the use of an objective assessment tool is strongly recommended.

Gestational age did not affect the pain assessment with either of these tools. Additionally, the relationship between gestational age and the situation was not significant. This is particularly encouraging because it suggests that *CHIPPS* can be used not only for term but also for preterm newborns.

Regarding practicality, both tools are equally short, and obtaining the final score is relatively simple.

Our results show that both tools can be easily improved. For *NFCS-R,* it is necessary to create training opportunities that are accessible and time-efficient. For *CHIPPS*, further research is needed with preterm and term newborns suffering from various painful conditions, and there is a need for guidelines that provide clinicians with clear definitions of the elements and their coding, as well as answers to questions related to the coding process. Due to the significant consequences of pain assessment, additional research is undoubtedly needed to gather information on what implicit decision-making strategies healthcare professionals use to assess pain in newborns.

Thus, the psychometric results are promising for both tools, especially *CHIPPS,* as our results show that it can also be used to assess acute pain in preterm newborns. Both tools can be improved in terms of clinical applicability. The pain assessment scales through observations for newborns studied in this research, *NFCS-R* and *CHIPPS*, are comparable in terms of their reliability and validity. Although both could potentially be used quite easily in everyday clinical practice, both have different shortcomings that may hinder their current clinical application.

List of references used.

1. Киреев С.С. Боль и стресс у новорожденных //Вестник новых медицинских технологий – 2016 – Т.23.-№4 – С.328-342

2. Сатвалдиева Э.А., Расулева Н.Р., Кдырниязова Р.Д., Мамадалиева З.А. Оценка адекватности концепции «предупреждающей аналгезии» в педиатрической анестезиологии //Материалы I съезда ассоциации врачей экстренной медицинской помощи. – Ташкент, 2009 – С.512-514.

3. Агзамходжаев Т.С., Юсупов А.С., Файзиев О.Я. и др. Послеоперационное обезболивание промедолом после абдоминальных операций у детей. //Вестник экстренной медицины.-2013.-№3.-С.250-251

4. Alves M.M., Carvalho P.R., Wagner M.B. et all. Cross -validation of the Children's and Infants' Postoperative Pain Scale in Brazilian children. //Pain Pract.-2018.-№8.-P.171 -176.

5. Andersen R. D. Do you see my pain? Aspects of pain assessmentin hospitalized preverbal children.-2018. Retrieved from https://openarchive.ki.se/xmlui/bitstream/handle/10616/46289/Thesis_Randi_Dovland_Andersen.pdf?sequence=1&isAllowed=y

6. Blauer T., Gerstmann D. A simulatneous comparison of three neonatal pain scales during common NICU procedures. //Clin J Pain.-2008.-№14.-P.39-47.

7. Buettner W., Finke W. Analysis of behavioural and physiological parameters for the assessment of postoperative analgesic demand in newborns, infants and young children: a comprehensive report on seven consecutive studies. //Pediatr Anesth.-2010.-№10.-P.303 –318.

8. Buettner W., Finke W., Hilleke M. et all. Entwicklung eines Fremdbeobachtungsbogens zur Beurteilung des postoperativen Schmerzes bei Säuglingen. //Anasthesiol Intensivmed Notfallmed Schmerzther 2008.-№33.-P.353 –361.

9. Carbajal R., Rousset A., Danan C. et al. Epidemiology and treatment of painful procedures in neonates in intensive care units. //JAMA.-2010.-№300.-P.60 –70.

10. Craig K.D., Hadjistavropoulos H.D., Grunau R.V. A comparison of two measures of facial activity during pain in the newborn child. //J. Pediatr Psychol.-1998.-№19.-P.305 –318.

11. Duhn L.J., Medves J.M. A systematic integrative review of infant pain assessment tools. //Adv Neonatal Care.-2010.-№4.-P.126-140.

12. Franck L.S., Ridout D., Howard R. et all. A comparison of pain measures in newborn infants after cardiac surgery. //Pain.-2011.-№152.-P. 1758 –1765.

13. Grunau R.E., Craig, K.D. Neonatal Facial Coding System: Revised: Training Manual. Neonatal facial and cry responses to invasive and non -invasive procedures. //Pain.-2011.-№42.-P.295 –305.

14. Grunau R.E., Oberlander T., Holsti L., Whitfield M.F. Bedside application of the Neonatal Facial Coding System in pain assessment of premature neonates. //Pain.-2010.-№76.-P.277 -286.

15. Hadjistavropoulos H.D., Craig K.D., Grunau R.V., Johnston, C.C. Judging pain in newborns: facial and cry determinants. //J Pediatr Psychol.-2014.-№19.-P.485 – 491.