Turkish validity and reliability of the COVERS pain scale

Müjde Çalıkuşu İncekar[1] , Nehir Ulu Öğüt[2] , Birsen Mutlu[3]* , Eda Çeçen[2] , Emrah Can[2]

SUMMARY

OBJECTIVE: The objective of this study is to determine the Turkish validity and reliability of COVERS.

METHODS: This study was conducted on 41 newborns as methodological design. The scales, such as newborn information form, COVERS, preterm infant pain profile (PIPP), and neonatal infant pain scale (NIPS), were used in the study. Validity (e.g., language, content concurrent, and construct) and internal consistency and inter-rater reliability of the scale were conducted.

RESULTS: It was found that COVERS showed a high correlation with PIPP and NIPS, and the item-total correlation of COVERS was above 0.30 during and after heel lance procedure. The Cronbach's α values were 0.77 and 0.83 during and after heel lance procedure, respectively. The kappa values of the items of COVERS were between 0.38 and 0.78 during heel lance procedure.

CONCLUSIONS: It was concluded in this study that there was a moderate correlation in intraclass correlation coefficients for scores of COVERS during both diaper change and heel lance procedures. It has been concluded that the scale is valid and reliable in 27-week-old and older newborns.

KEYWORDS: Neonatal intensive care unit. Newborn. Pain. Reliability. Validity.

INTRODUCTION

Preterm and term newborns experience pain and stress due to numerous and very different reasons such as intubation, venipuncture, and nasogastric/orogastric tube insertion in Neonatal Intensive Care Units (NICUs)[1]. Newborns experience approximately 70 stressful procedures[2] or 51 painful stimuli per day in the NICU[3]. Newborns may respond to pain in an exaggerated or attenuated manner as a result of frequent painful and invasive procedures[4]. The pain experienced by the newborns can not only prevent his/her behaviors, family infant interaction, and infant's adaptation to the outside world, but also cause changes in the development of brain and senses and affect the growth negatively[1].

Failure to appropriately assess the pain of newborns may result in delayed treatment and negative consequences[5].

Emotional status that cannot be expressed verbally by infants is involved in pain assessment, thus resulting in problems about the definition and treatment of pain. Newborns are dependent on others for the definition and treatment of pain[1]. Since newborns cannot express themselves, they show pain in behavioral and physiological ways[6]. Physiological changes related to pain in newborns include the changes in heart rate, respiratory rate, blood pressure, and oxygen and carbon dioxide levels in the blood. Behavioral changes related to pain include crying, facial expressions, motor movements, and behavioral status changes[7].

For the objective assessment of pain, the American Academy of Pediatrics recommends the use of validated and reliable scales[8]. Many scales have been developed for the measurement and evaluation of pain in newborns[5,9]. It has been reported that pain scales that are not suitable for the unsuitable population are used,

so it is important to use appropriate pain scales for newborns[9]. Hand et al.[10] developed COVERS and reported that it is a valid scale for both preterm and term newborns. O'Sullivan et al.[11] proved that the scale is valid and reliable. Considering the negative consequences of acute pain in newborns, it is very important to evaluate the pain of the newborn with a valid and reliable scale in daily clinical practice in the NICU[12]. This study was conducted to determine the reliability and validity of COVERS, which has not been used yet in newborns in Turkey.

METHODS

Setting, design, and sample
This study was conducted using a methodological design in a 16-bed NICU of a public hospital between January and June 2018. The case number was determined for a moderate effect size [(Δ: 0.50), β: 0.20 (80% power: 1–β) and α: 0.05] was found as 28 newborns. Forty-one newborns who met the criteria were included in this study. The inclusion criteria for the infants were determined as follows: being 24 weeks old and older[11], receiving no analgesic treatment within the last 12 h, and having no congenital anomaly.

Newborn information form
The form included questions about gestational week, weight, length of hospital stay (in days), gender, diagnosis, respiratory support, and the type of receiving respiratory support.

Neonatal infant pain scale
The scale consisting of behavioral parameters was developed by Lawrence et al.[13]. It consists of six subscales such as facial expression, crying, breathing patterns, arm movement, leg movement, and state of arousal. The scores of the scale range between 0 and 7, and the score higher than 4 points signifies pain. The Cronbach's α value of neonatal infant pain scale (NIPS) was reported as 0.95, 0.87, and 0.88 before, during, and after the procedure, respectively[13]. Akdovan[14] adapted the scale into Turkish, and the Cronbach's α value of the scale was found to be between 0.83 and 0.86. In this study, the Cronbach's α values were 0.79 and 0.87 during and after heel lance procedure, respectively.

Preterm infant pain profile
The scale was developed for 28- to 36-week preterm infants by Stevens et al.[15]. It evaluates behavioral (e.g., eyebrows, eyes, nasolabial furrow, and facial movement) and physiological (e.g., heart rate and oxygen saturation) parameters of the infant. The score of the scale ranges between 0 and 21. While a score of <5 indicates no pain, a score of >10 indicates moderate to severe pain. The Cronbach's α value of preterm infant pain profile (PIPP) was found to be between 0.59 and 0.76[15]. The reliability and validity study of the scale was conducted by Akcan and Yiğit[16] in Turkey. The Cronbach's α value of the scale was between 0.68 and 0.78[16]. In this study, the Cronbach's α values were 0.72 and 0.32 during and after heel lance procedure, respectively.

COVERS pain scale
The scale, which was developed by Hand et al.[10], consists of six subscales and depends on physiological (e.g., oxygen requirement and vital signs) and behavioral (e.g., crying, expression, resting, and signaling distress) measurements (Table 1). The subscales of the COVERS scale are scored with 0, 1, and 2 points, and the scale score ranges between 0 and 12. The validity of the scale in 27- to 40-week infants was analyzed[10]. The scale was valid and reliable in newborns older than 24 weeks, and it can be applied in the clinic[11].

Forward and backward translation
The researchers translated COVERS scale into Turkish, and the translated scale was checked by 10 pediatric nursing academicians. The scale translated into Turkish was translated back into English by a linguist. The original scale and the translated English scale were checked and found similar by another linguist. The language validity was performed in this way.

Content validity
The scale was evaluated by 10 pediatric nursing academicians. Each expert was asked to evaluate the relevance level of each item about the purpose of the questionnaire with a 4-point Likert scale as "completely appropriate" to "not appropriate." The content validity index of the scale was calculated as 0.91.

Study protocol
This study was conducted in supine position during both diaper change and heel lance procedures between 10:00 a.m. and 11:00 a.m. during weekdays while the infants were awake, and the process was recorded on video. Both diaper change and heel lance procedures were applied to each infant. The scales were applied 1 min before, during, 1 min after the diaper change, and during heel lance procedure (i.e., at the first time when manual lancet penetrated) and 1 min after heel lance procedure (i.e., applying cotton). The application lasted for an average of 5 min.

Concurrent validity and construct validity
For concurrent validity, the infants were evaluated with the scales such as COVERS, PIPP, and NIPS. For construct validity, they

Table 1. Original COVERS scale.

	0	1	2
Crying	No	High pitched or visibly crying	Inconsolable or difficult to soothe
Oxygen requirement	None	<30%	>30%
	At baseline O_2	<20%	>20%
	Breathing comfortably	Change in breathing pattern	Significant change in breathing pattern
Vital signs	HR and/or BP WNL for age or at baseline	HR and/or BP <20% of baseline	HR and/or BP >20% of baseline
	No apnea or bradycardia or at baseline	in frequency of apnea and bradycardia	in frequency and severity of apnea and bradycardia
Expression	None/facial muscles relaxed	Grimace, min-mod brow bulge, eye squeeze, nasolabial furrow	Grimace/grunt, mod-max brow bulge, eye squeeze, and nasolabial furrow
Resting	Sleeping most of the time	Wakes at frequent intervals – fussy	Constantly awake (even when not disturbed)
Signaling distress	Relaxed	Arms/legs flexed or extended "time-out signals"	Flailing and arching

HR: heart rate; BP: blood pressure; WNL: within normal limit.

were assessed using the scales in painful and nonpainful procedures. According to the literature[10], nonpainful procedure was determined as the diaper change, and the painful procedure was determined as heel lance procedure.

Inter-rater reliability

The diaper change and heel lance procedures were performed by a nurse included in this study. Another nurse involved in this study recorded the application process and the monitor showing heart rate and oxygen saturation of the infant via a video recorder. The nurse making the video recording in this study and the academician nurse included in this study watched the videos of the infants independently and evaluated the infants according to COVERS, PIPP, and NIPS.

Ethics statement

Ethics committee approval permission (IRB: 2017,12,4,05,021) was obtained from the hospital. Permissions to use scales were obtained by authors via e-mail. The written informed consent was obtained from the parents.

Statistical analysis

The IBM SPSS Statistics 22 (IBM SPSS, Turkey) was used for the statistical analyses. Normality assessment of the variables was made by using the Shapiro-Wilk test. The statistical significance level was set at 0.05. The internal consistency of COVERS at each time point was established using mean inter-item correlations, corrected item-total correlations, and the Cronbach's α reliability coefficients. Inter-rater reliability was established using kappa measure of agreement for categorical data and intraclass correlation coefficients (ICCs) for the continuous data. While the Spearman's rho correlation coefficient was used for concurrent validity, the Wilcoxon signed rank test was used for construct validity.

RESULTS

The characteristics of the infants were shown in Table 2.

Concurrent validity

In <37 weeks (n=28), a statistically significant correlation was determined between COVERS and PIPP scores during (p=0.768, p<0.001) and after (p=0.617, p<0.001) heel lance procedure, and between COVERS and NIPS scores during (p=0.785, p<0.001) and after (p=0.800, p<0.001) heel lance procedure.

In >37 weeks (n=13), a statistically significant correlation was determined between COVERS and PIPP scores during (p=0.854, p<0.001) and after (p=0.869, p<0.001) heel lance procedure,

Table 2. Characteristics of newborns (n=41).

Identify			Min–max	M±SD
Gestation weeks			27–41	33.98±4.05
Weight (g)			780–4,200	2,368.41±922.14
Number of days			1–75	9.34±15.62
			n	%
Gender		Female	18	43.9
		Male	23	56.1
Diagnose		Premature	26	63.4
		Respiratory distress syndrome	8	19.5
		Sepsis	4	9.8
		Hyperbilirubinemia	2	4.9
		Hypoglycemia	1	2.4
Receiving respiration support		Yes	12	29.3
		No	29	70.7
If yes		CPAP	5	41.7
		Incubator inside oxygen	4	33.3
		Hood oxygen	3	25

M: mean; SD: standard deviation; CPAP: continuous positive airway pressure.

and between COVERS and NIPS scores during (p=0.823, p<0.001) and after (p=0.951, p<0.001) heel lance procedure.

Construct validity
Table 3 shows the distribution of the COVERS mean scores of the newborns before, during, and after the diaper change and heel lance procedures.

Internal consistency
The item-total correlation values of COVERS were 0.32–0.82 and 0.39–0.86 during and after heel lance procedure. The Cronbach's α values of COVERS were 0.77 and 0.83 during and after heel lance procedure.

Inter-rater reliability
Table 4 shows the kappa results during diaper change and heel lance procedures. ICC values obtained for COVERS total score were 0.741 (95%CI, 0.514–0.862) during diaper change procedure and 0.579 (95%CI, 0.211–0.776) during heel lance procedure.

DISCUSSION
In a systematic review, it was stated that the validity and reliability of a scale must necessarily have construct validity,

internal consistency, and inter-rater reliability[17]. In this study, content validity, concurrent validity, and construct validity were tested for the validity of the COVERS scale. For the reliability of the scale, internal consistency and inter-rater reliability were tested.

For correlation values, it was reported that values between 0.70 and 0.89 showed a high correlation and values between 0.90 and 1.00 showed a very high correlation[18]. It was determined in a study conducted by Hand et al.[10] that while COVERS showed a high degree of correlation with PIPP (r=0.84) in preterm infants, it showed a very high degree of correlation with NIPS (r=0.95) in full-term infants. In this study, COVERS was found to be highly correlated with PIPP and NIPS. According to the results of this study, COVERS was found to have concurrent validity. Construct validity is defined as the degree to which a test measures what it is supposed to measure[19]. It was found in this study that there was a significant difference between the mean scores of COVERS in nonpainful and painful procedures. The results of this study are consistent with the literature[10]. In this study, it was determined that COVERS, which was adapted to Turkish, had construct validity.

In the literature, it has been reported that 0.30 and above is accepted as the optimum for corrected item-total correlation, and it becomes perfect as it approaches to 1[20]. In this study,

Table 3. Distribution of COVERS pain mean scores before, during, and after the diaper change and heel lance procedure (n=41) determined by using Wilcoxon signed rank test.

COVERS	Before	During	After
Diaper change	0.636	3.415	1.439
Heel lance	1.341	5.122	1.829
p	<0.001	0.003	0.439

Table 4. COVERS kappa measure of agreement analysis results of two observers during diaper change and heel lance procedure (n=41).

COVERS	During diaper change				During heel lance procedure			
	Kappa	SE	95%CI		Kappa	SE	95%CI	
			Lower	Upper			Lower	Upper
Crying	0.681	0.134	0.418	0.943	0.789	0.115	0.564	1.000
Oxygen requirement	0.448	0.152	0.151	0.746	0.522	0.132	0.264	0.779
Vital signs	0.394	0.139	0.122	0.666	0.414	0.142	0.135	0.693
Expression	0.561	0.121	0.324	0.798	0.474	0.114	0.251	0.698
Resting	0.345	10.150	0.051	0.639	0.433	0.117	0.204	0.661
Signaling distress	0.473	0.155	0.169	0.777	0.381	0.176	0.037	0.726

SE: standard error; CI: confidence interval.

since item-total correlation of COVERS was above 0.30 during and after heel lance procedure and showed that the items were appropriate, no item was excluded from the scale. It was reported in the study by O'Sullivan et al.[11] that corrected item-total correlation of COVERS was 0.19–0.68 during heel lance procedure and only the score of the item "oxygen requirement" was below 0.30. In contrast to the literature[11], the score of this item was found to be above 0.30 in this study.

The Cronbach's α value between 0.70 and 0.95 is reported to be an acceptable value[21,22]. In the literature[11], the internal consistency of COVERS was 0.78. The results of this study were found to be compatible with the literature. For kappa statistic values, it is reported that <0.20 is weak agreement, 0.20–0.40 is acceptable agreement, 0.40–0.60 is moderate agreement, 0.60–0.80 is good agreement, and 0.80–1.00 is perfect agreement[23]. In the literature, the kappa values of COVERS are reported to be between 0.29 and 0.78 at baseline and between 0.22 and 0.67 at heel lance. Acceptable agreement of COVERS was seen in "vital signs, expression, and signaling distress" items at baseline and in "vital signs, expression, resting, and signaling distress" items at heel lance[11]. In this study, acceptable agreement of COVERS was found in "vital signs and resting"

items during diaper change and in "signaling distress" item during heel lance procedure. For ICC, it is reported that <0.5 is weak, 0.5–0.75 is moderate, 0.75–0.90 is good, and >0.90 is perfect[24]. In the study by O'Sullivan et al.[11], ICCs for scores of COVERS were 0.82 (95%CI, 0.72–0.88) at baseline and 0.80 (95%CI, 0.69–0.87) during heel lance. In this study, it was found that there was a moderate correlation in ICCs for scores of COVERS during both diaper change and heel lance procedures.

It has been reported in the literature that PIPP and NIPS are frequently used to evaluate acute pain of preterm and term newborns[9]. The COVERS scale includes the parameters of both the NIPS and PIPP scale. However, the criteria used for scoring on the COVERS scale include newborns with a wider week[10]. Although crying is a parameter of behavioral responses to pain, it is unlikely that an intubated infant will have a high-pitched crying[10]. Compared with NIPS[13], the COVERS scale included visible crying within the behavioral parameter[10]. Compared with the PIPP scale, the COVERS scale has brought a new perspective to oxygen requirements. Since the oxygen requirement of the infant is not always an indicator of pain, the COVERS scale focuses on the change of oxygen demand rather than its value[10].

CONCLUSIONS

It was concluded that based on content validity, concurrent validity, and construct validity analyses of COVERS adapted into Turkish, it is a valid scale, and based on internal consistency and inter-rater reliability analyses, it is a reliable scale. COVERS can be used in all newborns with a gestational age of ≥27 weeks. In future studies, it may be suggested to adapt the COVERS scale to different cultures and to investigate the validity and reliability of the COVERS scale on infants in the postneonatal period.

ACKNOWLEDGMENTS

AUTHORS' CONTRIBUTIONS

MCI: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Software, Supervision, Validation, Writing – original draft, Writing – review & editing. **NUÖ:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Resources, Software, Validation, Writing – original draft. **BM:** Conceptualization, Funding acquisition, Methodology, Validation, Writing – original draft, Writing – review & editing. **EÇ:** Conceptualization, Data curation, Funding acquisition, Methodology, Resources, Software, Validation, Writing – original draft. **EC:** Conceptualization, Data curation, Funding acquisition, Writing – original draft.

REFERENCES

1. Witt N, Coynor S, Edwards C, Bradshaw H. A guide to pain assessment and management in the neonate. Curr Emerg Hosp Med Rep. 2016;4:1-10. https://doi.org/10.1007/s40138-016-0089-y

2. Weber A, Harrison TM. Reducing toxic stress in the neonatal intensive care unit to improve infant outcomes. Nurs Outlook. 2019;67(2):169-89. https://doi.org/10.1016/j.outlook.2018.11.002

3. Field T. Preterm newborn pain research review. Infant Behav Dev. 2017;49:141-50. https://doi.org/10.1016/j.infbeh.2017.09.002

4. Agarwal A, Gupta NK, Chaudhary R, Upadhyay A, Dubey A. Repeated painful stimuli may affect pain response in neonates of more than 34 weeks post-conceptional age: a case control study. Sri Lanka Journal of Child Health. 2019;48(3):251-5. http://doi.org/10.4038/sljch.v48i3.8761

5. Chen Y, Tong Y, Xue Z, Cheng Y, Li XY. Evaluation of the reliability and validity of the behavioral indicators of infant pain scale in Chinese neonates. Pain Manag Nurs. 2020;21(5):456-61. http://doi.org/10.1016/j.pmn.2020.01.001

6. Faye PM, De Jonckheere J, Logier R, Kuissi E, Jeanne M, Rakza T, et al. Newborn infant pain assessment using heart rate variability analysis. Clin J Pain. 2010;26(9):777-82. http://doi.org/10.1097/ajp.0b013e3181ed1058

7. Melo GM, Lélis LPA, Moura AF, Cardoso MVLML, Silva VM. Pain assessment scales in newborns: integrative review. Rev Paul Pediatr. 2014;32(4):395-402. http://doi.org/10.1016/j.rpped.2014.04.007

8. Committee on Fetus and Newborn and Section on Anesthesiology and Pain Medicine, Keels E, Sethna N, Watterberg KL, et al. Prevention and management of procedural pain in the neonate: an update. Pediatrics. 2016;137(2):e20154271. http://doi.org/10.1542/peds.2015-4271

9. Olsson E, Ahl H, Bengtsson K, Vejayaram DN, Norman E, Bruschettini M, et al. The use and reporting of neonatal pain scales: a systematic review of randomized trials. Pain. 2020;162(2):353-60. http://doi.org/10.1097/j.pain.0000000000002046

10. Hand IL, Noble L, Geiss D, Wozniak L, Hall C. COVERS neonatal pain scale: development and validation. Int J Pediatr. 2010;2010:496719. http://doi.org/10.1155/2010/496719

11. O'Sullivan AT, Rowley S, Ellis S, Faasse K, Petrie KJ. The validity and clinical utility of the COVERS scale and pain assessment tool for assessing pain in neonates admitted to an intensive care unit. Clin J Pain. 2016;32(1):51-7. http://doi.org/10.1097/AJP.0000000000000228

12. Dionysakopoulou C, Giannakopoulou M, Lianou L, Bozas E, Zannikos K, Matziou V. Validation of Greek versions of the neonatal infant pain scale and premature infant pain profile in neonatal intensive care unit. Pain Manag Nurs. 2018;19(3):313-9. http://doi.org/10.1016/j.pmn.2017.05.008

13. Lawrence J, Alcock D, McGrath P, Kay J, MacMurray SB, Dulberg C. The development of a tool to assess neonatal pain. Neonatal Netw. 1993;12(6):59-66. PMID: 8413140

14. Akdovan T. Assessment of pain in healthy newborns, ınvestigation of the effect of the sucking and holding method [thesis]. Istanbul (Turkey): Marmara University Institute of Health Sciences, 1999.

15. Stevens B, Johnston C, Petryshen P, Taddio A. Premature infant pain profile: development and initial validation. Clin J Pain. 1996;12(1):13-22. http://doi.org/10.1097/00002508-199603000-00004

16. Akcan E, Yiğit R. Prematüre bebek ağrı profili: türkçe geçerlilik ve güvenirliliği. F Ü Sağ Bil Tıp Derg. 2015;29(3):97-102. Available from: http://tip.fusabil.org/pdf/pdf_FUSABIL_1076.pdf#:~:text=Bulgular%3A%20Ara%C5%9Ft%C4%B1rma%20grubundaki%20premat%C3%BCre%20bebeklerin,g%C3%BCvenirlilik%200.85%E2%80%930.99%20aras%C4%B1nda%20bulundu

17. Giordano V, Edobor J, Deindl P, Wildner B, Goeral K, Steinbauer P, et al. Pain and sedation scales for neonatal and pediatric patients in a preverbal stage of development: a systematic review. JAMA Pediatr. 2019;173(12):1186-97. http://doi.org/10.1001/jamapediatrics.2019.3351

18. Zady MF. Correlation and simple least squares regression. WestgardQC [cited on Aug 30, 2019]. Available from: https://www.westgard.com/lesson42.htm

19. Brown JD. What is construct validity? JALT Testing & Evaluation SIG. 2000;4(2):8-12. Available from: https://hosted.jalt.org/test/PDF/Brown8.pdf

20. Cristobal E, Flavian C, Guinaliu M. Perceived e-service quality (PeSQ): measurement validation and effects on consumer satisfaction and web site loyalty. Managing Service Quality: An International Journal. 2007;17(3):317-40. https://doi.org/10.1108/09604520710744326

21. Bland JM, Altman DG. Statistics notes: Cronbach's alpha. BMJ. 1997;314:572. https://doi.org/10.1136/bmj.314.7080.572

22. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ. 2011;2:53-5. https://doi.org/10.5116/ijme.4dfb.8dfd

23. Collett D. Modelling survival data in medical research. Boca Raton: CRC Press, 2015.

24. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15(2):155-63. https://doi.org/10.1016/j.jcm.2016.02.012