ORIGINAL ARTICLE

Predictive Validity and Inter-Rater Reliability of the Persian Version of Full Outline of Unresponsiveness Among Unconscious Patients with Traumatic Brain Injury in an Intensive Care Unit

Somayeh Momenyan[1] · Sey.Mojtaba Mousavi[2] · Tahmineh Dadkhahtehrani[3] ·
Fatemeh Sarvi[4] · Reza Heidarifar[5] · Faezeh Kabiri[6] · Erfan Mohebi[2] ·
Mohammad Koohbor[2]

Abstract

Introduction The Glasgow Coma Scale (GCS) has some limitations when evaluating the unconscious patient. This study aims to validate the Persian version of the FOUR (Full Outline of Unresponsiveness) score as a proposed substitute.
Methods Two nurses, two nursing students, and two physicians scored the prepared Persian version of the FOUR and GCS in 84 patients with acute brain injury. The inter-rater agreement for the FOUR and the GCS scores was evaluated by the weighted kappa (κ_w). The outcome prediction power of the scales was assessed by the area under the curve (AUC) in the ROC curve.
Results The inter-rater agreement of the FOUR was excellent ($\kappa_w = 0.923$, 95 % CI, 0.874–0.971) and comparable with the one of the GCS ($\kappa_w = 0.938$, 95 % CI, 0.889–0.987). The area under the curve (AUC) for predicting in-hospital mortality (modified Rankin Scale: 6) was 0.835 for the FOUR (95 % CI, 0.739–0.907) and 0.772 for the GCS (95 % CI, 0.668–0.856) ($P = 0.01$). AUC for predicting poor outcome (modified Rankin Scale: 3–6) for the total FOUR score was 0.983 (95 % CI, 0.928–0.999), which is comparable with 0.987 for the total GCS score (95 % CI, 0.934–1.000).
Conclusions The researchers conclude that the Persian version of the FOUR score is a reliable and valid scale to assess unconscious patients with traumatic brain injury and can be substituted for the GCS.

Keywords Full Outline of Unresponsiveness score · Acute brain injury · Coma scale · Outcome predictor · Glasgow Coma Scale · Persian · Iran

Introduction

There is no objective scale to assess comatose patients, and it just depends on the clinical skill of physicians [1]. There are some scales to evaluate these patients. One of the most commonly used scales is the Glasgow Coma Scale (GCS) [2, 3], but it has some important limitations such as impossibility to assess the verbal component among intubated comatose patients. In this condition, some physicians record the lowest possible score for this component. In addition, the GCS does not have a clinical index for brainstem reflexes. Trained personnel prefer to apply the GCS, although interpretation of intermediate scores on the GCS remains difficult for emergency physicians. In general, the GCS cannot recognize precisely the clinical changes of comatose patients [4]. Some efforts have been made to improve the GCS, and many scoring systems were developed to be substituted for the GCS [5–7]. The Full Outline of Unresponsiveness (FOUR) score was developed by Wijdicks et al. to evaluate the consciousness in comatose patients [1, 8]. This scale has

✉ Somayeh Momenyan
somy.momenyan@gmail.com

1 Epidemiology and Biostatistics Department, Qom University of Medical Sciences, Qom, Iran

2 Qom University of Medical Sciences, Qom, Iran

3 Academic Member of Midwifery, Midwifery Department, Faculty of Nursing and Midwifery, Isfahan University of Medical Sciences, Isfahan, Iran

4 Department of Epidemiology and Biostatistics, School of Public Health, Hamadan University of Medical Science, Hamadan, Iran

5 Medicine, Quran and Hadith Research Center, Bagiyatallah University of Medical Sciences, Tehran, Iran

6 Qom Branch Azad University, Qom, Iran

Table 1 Definition of the FOUR score and the GCS

FOUR score	GCS
Eye response	Eye response
4 = eyelids open or opened, tracking, or blinking to command	4 = eyes open spontaneously
3 = eyelids open but not tracking	3 = eye opening to verbal command
2 = eyelids closed but open to loud voice	2 = eye opening to pain
1 = eyelids closed but open to pain	1 = no eye opening
0 = eyelids remain closed with pain	Motor response
Motor response	6 = obeys commands
4 = thumbs-up, fist, or peace sign	5 = localizing pain
3 = localizing to pain	4 = withdrawal from pain
2 = flexion response to pain	3 = flexion response to pain
1 = extension response to pain	2 = extension response to pain
0 = no response to pain or generalized myoclonus status	1 = no motor response
Brainstem reflexes	Verbal response
4 = pupil and corneal reflexes present	5 = oriented
3 = one pupil wide and fixed	4 = confused
2 = pupil or corneal reflexes absent	3 = inappropriate words
1 = pupil and corneal reflexes absent	2 = incomprehensible sounds
0 = absent pupil, corneal, and cough reflex	1 = no verbal response
Respiration	
4 = not intubated, regular breathing pattern	
3 = not intubated, Cheyne–Stokes breathing pattern	
2 = not intubated, irregular breathing	
1 = breathes above ventilator rate	
0 = breathes at ventilator rate or apnea	

four components: eye, motor, brainstem, and respiration. The score of each component ranges 0–4 (Table 1). Brainstem reflexes and respiratory components provide an assessment rather than verbal responding [9]. Some studies have translated the FOUR score into different languages and have assessed its validity and reliability for different populations [10–17], but there is no study conducted to validate the Persian version. This study aims to assess the predictive validity and inter-rater reliability of the Persian version of the FOUR score in unconscious patients with traumatic brain injury in an intensive care unit.

Methods

Development of the Persian Version of the FOUR Score

Translation of the FOUR score into Persian followed standardized forward–backward procedure. First, two professional expert translators carried out the forward translation from English into Persian independently. Then, the first consensus meeting was held to compare the two

Persian versions and discuss the accuracy of the statements to reach an agreement on a fully comprehensible and accurate Persian translation consistent with the original English text. Next, another expert translator who did not have access to the original English scale carried out the back-translation of the Persian version. After that, in the second consensus meeting, the back-translated version was compared with the original scale to develop the final Persian version. Finally, the Persian version of the FOUR score was validated.

Inter-Rater Reliability and Predictive Validity

To assess inter-rater reliability of the FOUR score, the GCS was applied as a standard scale for comparison. To compare the results of the FOUR score with the GCS, three different types of raters scored the Persian version of the FOUR score and the GCS including two ICU physicians, two ICU head nurses, and two senior students of nursing. Each of the nurses and physicians had at least 2 years of clinical experience in an intensive care unit (ICU). Prior to the study, the raters were instructed to apply the FOUR score and the GCS accurately. Subsequently, a trial session was performed on a few patients to ensure they understood

the procedure perfectly. The raters were provided with written instructions and scoring sheets to be used during the examination of all patients. Six categories of pairwise ratings were analyzed including (1) physician–nurse, (2) physician–student, (3) nurse–student, (4) physician–physician, (5) nurse–nurse, and (6) student–student. Each pair of raters scored 14 patients using both the FOUR score and the GCS, which resulted in 84 patients. To reduce the bias, the order of examining and scoring of each patient was randomly set. The raters of each pair completed their scorings within a period of 1 h without awareness of the other's scores.

To assess the predictive validity of the FOUR score, the outcome was assessed at discharge from the hospital using the modified Rankin Scale (mRS) by one of the raters of each pair who was randomly selected. Then, the results were compared with the ones of the FOUR score as a common standard scale. The rating of mRS scale was done according to 7 points as follows: 0 = no symptoms, 1 = no significant disability, 2 = slight disability, 3 = moderate disability, 4 = moderately severe disability, 5 = severe disability, and 6 = dead [18]. In this study, mRS score 0–2 was considered as good outcome and score 3–6 as poor outcome.

Participants

A total of 87 patients admitted to the intensive care unit (ICU) in Shahid Beheshti Hospital of Qom, Iran, were enrolled from March to December 2013. They were evaluated within 7 days from admission to ICU. The inclusion criteria were an age > 18 years and unconsciousness due to an acute traumatic brain injury. Exclusion criteria were treatment with neuromuscular junction blockers and sedatives and interval longer than 1 h between assessment and pairwise scoring of the raters. An informed written consent was obtained from the patient's legal surrogate. The Ethics Committee of Qom University of Medical Sciences approved this study.

Statistical Analysis

The inter-rater reliability of the FOUR score and the GCS was assessed using the weighted Cohen's kappa (κ_{w}) for the total score as well as the score of each item. The κ_{w} coefficients of 0.4 or less were considered as poor agreements, and values greater than 0.8 were considered as excellent agreements between the raters [19]. Internal consistency was assessed by Cronbach's α, and concurrent validity was done via calculating the Spearman's correlation coefficients between the FOUR score and the GCS. Predictive validity was assessed by receiver operating characteristic (ROC) curve. This curve shows the power of

the FOUR score and the GCS to predict the mortality or poor outcome at discharge from the hospital. The sensitivity and specificity were calculated for both scales. Logistic regression was used to show the odds ratios of the FOUR score and the GCS in predicting the mortality or poor outcome at discharge. The mean ratings of two raters for each patient were calculated for the ROC curve and regression analysis. SPSS V20 and MedCalc 14 were applied to analyze the data. The level of statistical significance was set at $P < 0.05$.

Results

Eighty-seven unconscious patients with acute traumatic brain injury were included in our study, but three of them were excluded because no pairwise rating occurred within a time interval of 1 h. Eventually, statistical analysis was performed on 84 patients. Their mean age was 42.6 ± 11.7 years (25–70 years) and 63 (74.1 %) were men. Sixty-one patients were intubated and mechanically ventilated at the time of scoring (71.8 %); thus, score one was recorded for the GCS verbal subscore.

In total, 168 ratings were performed for 84 patients by the FOUR score and the GCS. The frequency of the total score for each scale and its subscales is illustrated in Fig. 1. Cronbach's α showed a high degree of internal consistency for the GCS (α = 0.82) as well as the FOUR score (α = 0.93).

Spearman's correlation coefficient was high ($r = 0.95$, $P < 0.001$) between the total scores of the scales.

The inter-rater agreement for each pair of raters was excellent both for the total FOUR score (k_w 0.923, 95 % CI, 0.874–0.971) and for the total GCS score (k_w 0.838, 95 % CI, 0.889–0.987). Kappa values for all pairs of raters and for each subscale of the FOUR score and the GCS are shown in Table 2. The inter-rater agreement of both scales for each pair of raters was excellent independent of the level of expertise and experiences (Table 3).

Sixteen patients (18.8 %) died at hospital (mRS = 6), and 40 patients (47.1 %) had poor outcomes (mRS = 3–6) at hospital discharge. The area under the curve (AUC) in the ROC curve was estimated to compare the scales in the prediction power of in-hospital mortality and poor outcome at discharge (Fig. 2).

AUC values in prediction of in-hospital mortality were significantly different between the FOUR score (AUC = 0.835; 95 % CI, 0.739–0.907) and the GCS (AUC = 0.772; 95 % CI, 0.668–0.856) ($P = 0.01$).

The sum of sensitivity and specificity was maximized to predict in-hospital mortality at a total score of 6 for both the FOUR (sensitivity = 100 %; specificity = 62 %) and the GCS (sensitivity = 100 %; specificity = 61 %).
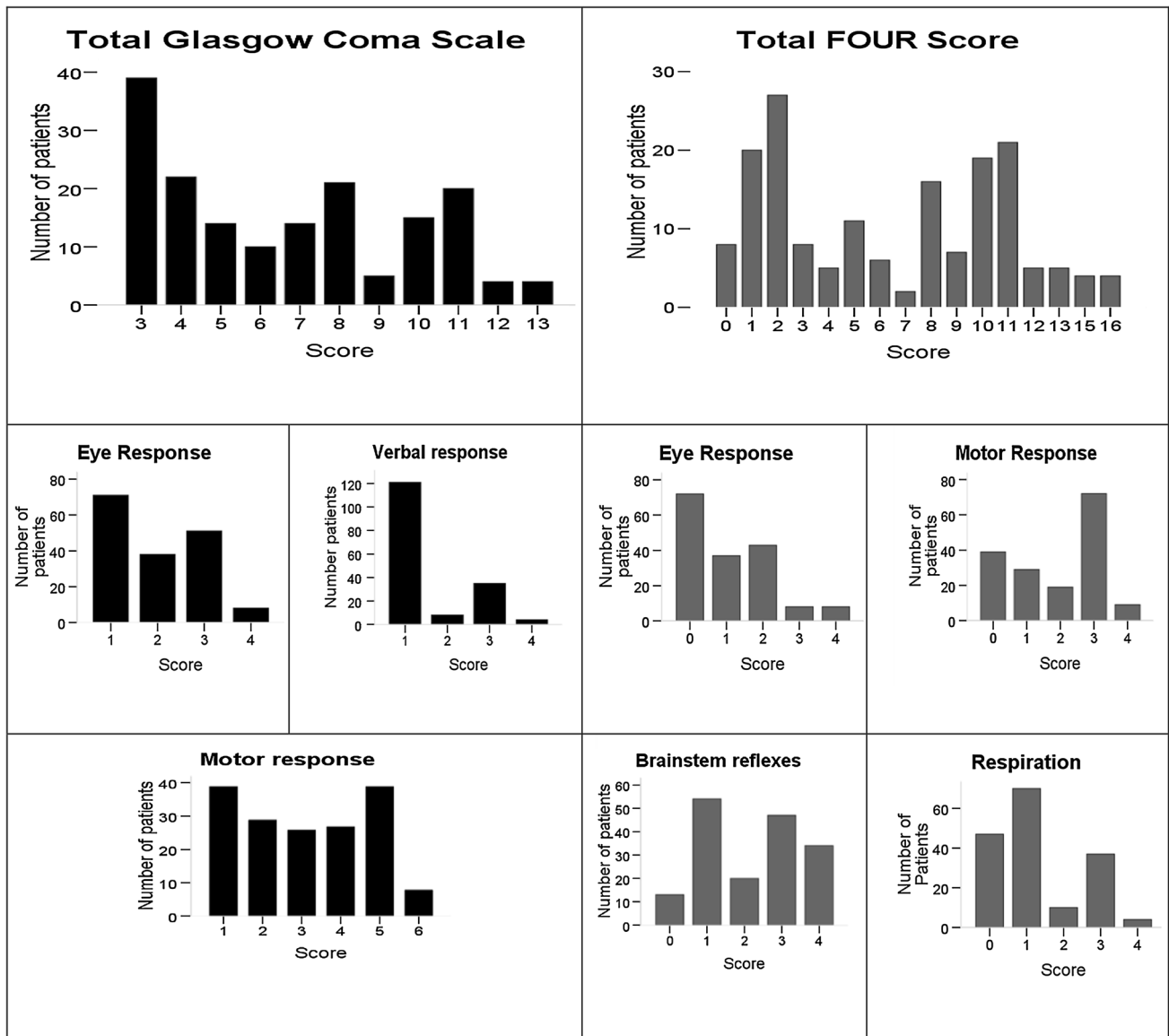
Fig. 1 Frequency of scores of the GCS and the FOUR scores for 168 raters

**Table 2** Inter-rater agreement for the GCS and the FOUR scores by κ_w

Pair of raters	No. of patients	The FOUR score					GCS			
		Eye	Motor	Brainstem	Respiration	Total	Eye	Verbal	Motor	Total
P/N	14	0.972	0.901	0.855	0.882	0.923	0.967	1.000	0.863	0.938
N/S	14	1.000	0.864	0.864	0.940	0.909	1.000	1.000	0.939	0.967
P/S	14	1.000	0.885	0.862	0.868	0.909	0.905	1.000	0.908	0.937
P/P	14	0.872	0.892	0.982	0.972	0.901	0.857	0.972	0.896	0.970
N/N	14	1.000	0.932	0.894	0.952	0.866	0.955	1.000	0.931	0.832
S/S	14	0.902	0.885	0.862	0.818	0.809	0.905	0.806	0.908	0.837
Overall	84	0.961	0.897	0.894	0.954	0.923	0.971	0.989	0.961	0.838

N Nurse; *P* physician; *S* nursing student

Table 3 Receiver operating characteristic curve analyses in predicting mortality (mRS = 6) and poor outcome (mRS = 3–6) at discharge for the GCS and the FOUR scores and their subscales

Variable	Modified Rankin score = 6 AUC (95 % CI)	Modified Rankin score 3–6 AUC (95 % CI)
Total FOUR score	0.835 (0.739–0.907)	0.983 (0.928–0.999)
Eye	0.75 (0.650–0.842)	0.961 (0.895–0.991)
Motor	0.763 (0.658–0.848)	0.961 (0.895–0.991)
Brainstem reflexes	0.873 (0.783–0.935)	0.971 (0.909–0.995)
Respiration	0.796 (0.695–0.876)	0.905 (0.822–0.958)
Total GCS	0.772 (0.668–0.856)	0.987 (0.934–1.000)
Eye	0.741 (0.635–0.830)	0.961 (0.895–0.991)
Motor	0.777 (0.674–0.860)	0.968 (0.905–0.994)
Verbal	0.674 (0.564–0.772)	0.767 (0.662–0.852)

Fig. 2 Receiver operating characteristic curve for the GCS and the FOUR scores for poor outcome (mRS = 3–6) and mortality (mRS = 6) at discharge
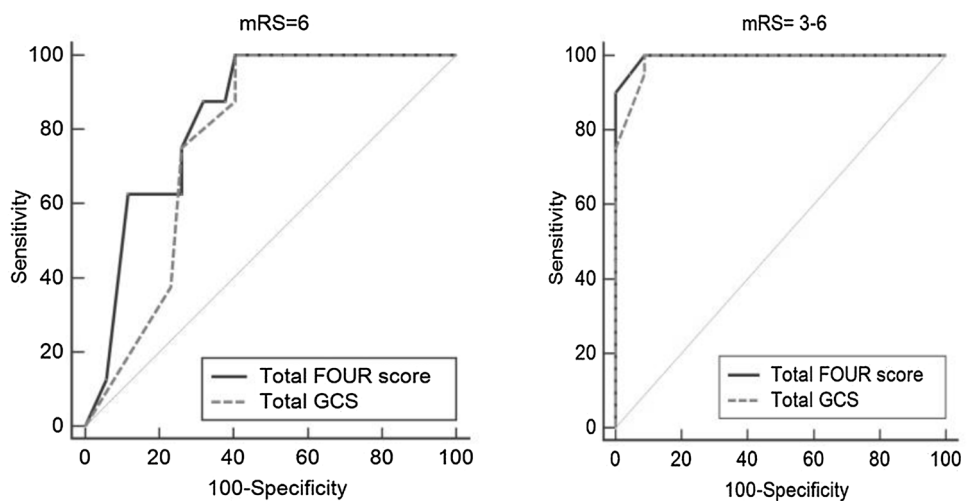
In prediction of poor outcome, AUC values were not significantly different between the total FOUR score: 0.983 (95 % CI, 0.928–0.999) and the total GCS: 0.987 (95 % CI, 0.934–1.000).

The sum of sensitivity and specificity was maximized to predict poor outcome at a total score of 6 for both the FOUR score (sensitivity = 100 %, specificity = 91.1 %) and the GCS (sensitivity = 100 %, specificity = 95 %).

Table 4 shows the results of logistic regression between the total score and patient outcome for the two scales.

With the FOUR score, each 1-point increase in total score was associated with an estimated 33 % reduction in odds of experiencing in-hospital mortality under the unadjusted model (OR = 0.67, 95 % CI, 0.54–0.85) and 85 % reduction in odds of poor outcome (OR = 0.15, 95 % CI, 0.04–0.6). These relations remained after adjusting for age and sex.

With the GCS total score, each 1-point increase in total score was associated with an estimated 40 % reduction in odds of in-hospital mortality (OR = 0.6, 95 % CI, 0.43–0.83) and estimated 80 % reduction in odds of poor outcome under the unadjusted model (OR = 0.2, 95 % CI, 0.04–0.4). These relations remained after adjusting for age and sex.

Table 4 Odds ratios, confidence intervals, and the percent of cases correctly classified for the GCS and the FOUR scores for poor outcome (mRS = 3–6) and mortality (mRS = 6) at discharge

Variable	Modified Rankin score = 6		Modified Rankin score 3–6	
	OR (95 % CI)	Percentage correctly classified	OR (95 % CI)	Percentage correctly classified
Total FOUR score	0.67 (0.54–0.85)	78.8	0.15 (0.04–0.6)	95.3
Total GCS	0.6 (0.43–0.83)	81.2	0.2 (0.04–0.4)	92.9

Discussion

The present study assessed the predictive validity and inter-rater reliability of the Persian version of the FOUR score among unconscious patients with traumatic injuries in an intensive care unit by comparing it with the GCS as standard scale.

The FOUR score is simply applied and includes the minimal necessities in impaired consciousness and distinguishes specifically certain unconscious states. It has been developed to overcome the limitations of the GCS, which is unable to assess the verbal score in intubated patients and test brainstem reflexes.

The results showed that the inter-rater agreement was excellent for the FOUR score ($\kappa_w = 0.923$) and comparable with the GCS ($\kappa_w = 0.838$). This finding is consistent with the results of the developers of the scale($\kappa_w = 0.82$ for both scales) [1], the French version ($\kappa_w = 0.86$ for the FOUR score and $\kappa_w = 0.85$ for the GCS) [10], the Spanish version ($\kappa_w = 0.93$ for the FOUR score and $\kappa_w = 0.96$ for the GCS) [12], and also the Italian version of the scale ($\kappa_w = 0.953$ for the FOUR score and $\kappa_w = 0.943$ for the GCS) [11].

It is interesting that our findings about inter-rater agreement are similar to some studies with different raters and various levels of experience. For instance, in the present study, scorings were performed by two nurses, two physicians, and two nursing students and in the Italian version were performed by neurologists and neurology residents with clinical expertise [11]; however, both studies found similar results of inter-rater agreement.

We found that the inter-rater agreement was excellent for all pairs, even for the student–student pair who was less experienced than the nurses and physicians. This finding is consistent with the study of Eelco F. M. Wijdicks that showed good inter-rater agreements in nurse–physician pair for the GCS ($\kappa_w = 0.77$) and the FOUR score ($\kappa_w = 0.75$) [1], but it is slightly at variance with the finding of a study on the Italian version that involved highly, moderately, and less experienced raters and showed that performances of the FOUR and GCS were comparable only among the highly and moderately experienced raters. The difference between our findings and those of the Italian version may be due to various patients [20]. Considering that standard instruction is required to apply a scale accurately [21], the excellent inter-rater agreement of the present study may result from a standard and perfect instruction of applying the new scale before scoring for the raters. The difference between our findings and those of the French version may be related to various approaches and quality of instruction.

In contrast, in a study by Michael Fischer, physician-nurse pair agreement (neurologist–ICU staff) was 0.56 with the GCS and 0.66 with the FOUR score [14]. Albeit, in their study, the agreement in the pairs of neurologist–neurologist and nurse–nurse was also less than ours with both scales, it can rationalize the difference between the findings of two studies.

Our results show that the AUC values from ROC curves are analogous and excellent to predict the poor outcome for both scales; but the AUC value in predicting in-hospital mortality was significantly different between the scales, as it was better with FOUR score. This finding is not consistent with the results of the Italian version, which indicates that both scoring systems are excellent outcome predictors of in-hospital mortality and less accurate response in patients with a poor outcome [11]. In addition, they reported that the scales were comparable in prediction power of in-hospital mortality; but the prediction power of the FOUR score was lower than the GCS in poor outcome. The difference between our findings and those of the Italian version may result from various patients and the settings of sampling.

In our study, among the patients with a poor outcome (mRS $\geq$ 3), the odds ratio for the FOUR score is rather lower than that for the GCS. The lower odds ratios have been associated with a positive predictive value for a higher chance of a positive outcome with increased total score values [1, 13]. The proportion of cases correctly classified for both poor outcome and in-hospital mortality was analogous for both the GCS and the FOUR scores. This is consistent with the result of Cohen [14].

Conclusion

The present study shows that the reliability of the Persian version of the FOUR score as well as its prediction power for poor outcome (predictive validity) is comparable to those of the GCS; moreover, it is superior to the GCS due to its higher prediction power for in-hospital mortality as well as its ability to assess the brainstem reflexes. Therefore, the Persian version of the FOUR score is a simple-to-use, easy-to-teach, and reliable scale for all practitioners, even less-experienced ones such as nurse students. Also, it can be a proper communicating tool among various members of a treatment team that can be applied reliably to assess patients with impaired consciousness and patients with traumatic brain injury in intensive care units if a standard instruction is performed for them.

We conclude that the Persian version of the FOUR score could be a good substitution for the GCS among unconscious patients. Further studies are recommended in various patients and settings.

Appendix

امـتیـاز	شـاخص
	پـاسخ چشمی
4	پـلک هـا بـاز است یـا بـاز مـی شونـد و بـاز مـی مـانـند و پـلک مـی زنـد
3	پـلک هـا بـاز مـی شونـد امـا مـسیر درخواست شده را بـا چشم دنـبـال نـمیتوانـند بـکنند
2	پـلک هـا بـسته امـا بـا صدای بـلند بـاز مـی شونـد
1	پـلک هـا بـسته و بـا تحریک دردنـاک بـاز مـی شونـد.
0	پـلک هـا بـسته و بـا تحریک دردنـاک بـاز نمی شونـد.
	پـاسخ حرکتی
4	مشت کردن، نشانـه صلح را انجام مـیدهد
3	نشان دادن محل درد
2	چرخش غیر طبیعی بـه سمت داخل بـا تحریک دردنـاک
1	چرخش غیر طبیعی بـه سمت خارج بـا تحریک دردنـاک
0	بـدون پـاسخ بـه تحریک دردنـاک
	رفـلکس هـا ساقـه مغز
4	وجود رفـلکس مـردمک و قـرنـیه
3	یک مـردمک گشاد و ثـابت
2	عدم وجود رفـلکس مـردمک یـا قـرنـیه
1	عدم وجود رفـلکس مـردمک و قـرنـیه
0	عدم وجود رفـلکس مـردمک و قـرنـیه و سرفـه
	پـاسخ هـای تـنفسی
4	الـگوی تـنفسی مـنظم، ایـنتوبـه نیست
3	الـگوی تـنفسی شایـن استوک ایـنتوبـه نیست
2	تـنفس نـامـنظم ، ایـنتوبـه نیست
1	تـعداد تـنفس بیش از مقـدار تحت ونـتیلاتـور
0	تـعداد تـنفس تحت ونـتیلاتـور یـا آپـنه

References

1. Wijdicks EF, Bamlet WR, Maramattom BV, Manno EM, McClelland RL. Validation of a new coma scale: the FOUR score. Ann Neurol. 2005;58(4):585–93.

2. Laureys S, Perrin F, Brédart S. Self-consciousness in non-communicative patients. Conscious Cognit. 2007;16(3):722–41.

3. Teasdale G, Jennett B. Assessment of coma and impaired consciousness: a practical scale. The Lancet. 1974;304(7872):81–4.

4. Balestreri M, Czosnyka M, Chatfield D, Steiner L, Schmidt E, Smielewski P, et al. Predictive value of glasgow coma scale after brain trauma: change in trend over the past ten years. J Neurol Neurosurg Psychiatry. 2004;75(1):161–2.

5. Knaus WA, Zimmerman JE, Wagner DP, Draper EA, Lawrence DE. APACHE-acute physiology and chronic health evaluation: a physiologically based classification system. Crit Care Med. 1981;9(8):591–7.

6. Stanczak DE, White JG III, Gouview WD, Moehle KA, Daniel M, Novack T, et al. Assessment of level of consciousness following severe neurological insult: a comparison of the psychometric qualities of the glasgow coma scale and the comprehensive level of consciousness scale. J Neurosurg. 1984;60(5):955–60.

7. Gill M, Martens K, Lynch EL, Salih A, Green SM. Interrater reliability of 3 simplified neurologic scales applied to adults presenting to the emergency department with altered levels of consciousness. Ann Emerg Med. 2007;49(4):403–7.

8. Wolf CA, Wijdicks EF, Bamlet WR, McClelland RL. Further validation of the FOUR score coma scale by intensive care nurses. Pro Mayo Clinic 2007;82:435–8.

9. Wijdicks EF, Varelas PN, Gronseth GS, Greer DM. Evidence-based guideline update: determining brain death in adults report of the quality standards subcommittee of the American Academy of Neurology. Neurology. 2010;74(23):1911–8.

10. Weiss N, Mutlu G, Essardy F, Nacabal C, Sauves C, Bally C, et al. The French version of the FOUR score: a new coma score. Rev Neurol. 2009;165(10):796–802.

11. Marcati E, Ricci S, Casalena A, Toni D, Carolei A, Sacco S. Validation of the Italian version of a new coma scale: the FOUR score. Intern Emerg Med. 2012;7(2):145–52.

12. Idrovo L, Fuentes B, Medina J, Gabaldón L, Ruiz-Ares G, Abenza MJ, et al. Validation of the FOUR Score (Spanish Version) in acute stroke: an interobserver variability study. Eur Neurol. 2010;63(6):364–9.

13. Stead LG, Wijdicks EF, Bhagra A, Kashyap R, Bellolio MF, Nash DL, et al. Validation of a new coma scale, the FOUR score, in the emergency department. Neurocrit Care. 2009;10(1):50–4.

14. Cohen J. Interrater reliability and predictive validity of the FOUR score coma scale in a pediatric population. J Neurosci Nurs. 2009;41(5):261–7.

15. Bruno M-A, Ledoux D, Lambermont B, Damas F, Schnakers C, Vanhaudenhuyse A, et al. Comparison of the full outline of unresponsiveness and glasgow liege scale/glasgow coma scale in an intensive care unit population. Neurocrit Care. 2011;15(3):447–53.

16. Akavipat P, Sookplung P, Kaewsingha P, Maunsaiyat P. Prediction of discharge outcome with the full outline of unresponsiveness (FOUR) score in neurosurgical patients. Acta MedOkayama. 2011;65:205–10.

17. Fischer M, Rüegg S, Czaplinski A, Strohmeier M, Lehmann A, Tschan F, et al. Research inter-rater reliability of the full outline of unresponsiveness score and the glasgow coma scale in critically ill patients: a prospective observational study. 2010.

18. Rankin J. Cerebral vascular accidents in patients over the age of 60. II. Prognosis. Scott Med J. 1957;2(5):200–15.

19. Van Swieten J, Koudstaal P, Visser M, Schouten H, Van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. Stroke. 1988;19(5):604–7.

20. Schnakers C, Giacino J, Kalmar K, Piret S, Lopez E, Boly M, et al. Does the FOUR score correctly diagnose the vegetative and minimally conscious states? Ann Neurol. 2006;60(6):744–5.

21. Reith FC, Brennan PM, Maas AI, Teasdale GM. Lack of standardization in the use of the glasgow coma scale: results of international surveys. J Neurotrauma. 2016;33(1):89–94.