








ORIGINAL ARTICLE

Variation Among Evaluators of The Full Outline of Unresponsiveness Score and The Glasgow Coma Scale in Critically Ill Patients: A Prospective Study

Kritik Hastalarda Full Outline of Unresponsiveness Skoru ve Glasgow Koma Skalası Değerlendiricileri Arasındaki Varyasyon- Prospektif Bir Çalışma

¹Yeşim Şerife BAYRAKTAR , ¹Yasemin CEBECİ , ²Büşra PEKİNCE , ³Dilara CARI GÜNGÖR , ⁴Muslu Kazım KÖREZ 
¹Faruk ÇİÇEKÇİ , ¹Ale Bengi ÇELİK 

¹Department of Anesthesiology, Intensive Care Unit, Selçuk University School of Medicine, Konya, Türkiye
²Department of Pulmonary Diseases, Intensive Care Unit, Selçuk University School of Medicine, Konya, Türkiye
³Department of Neurology, Intensive Care Unit, Selçuk University School of Medicine, Konya, Türkiye
⁴Department of Biostatistics, Selçuk University School of Medicine, Konya, Türkiye

Correspondence

Yeşim Şerife Bayraktar
Department of Anesthesiology,
Intensive Care Unit, Selçuk University
School of Medicine, Konya, Türkiye

E-Mail: bayraktaryesim83@gmail.com

How to cite ?

Bayraktar YŞ, Cebeci Y, Pekince B, Cari G, G, Körez MK, Çiçekçi F, Çelik JB. Variation Among Evaluators of The Full Outline of Unresponsiveness Score and The Glasgow Coma Scale in Critically Ill Patients: A Prospective Study. Genel Tıp Derg. 2024;34(6):861-866

ABSTRACT

Background/Aims: This study aims to compare the reliability of the Full Outline of Unresponsiveness (FOUR) score and the Glasgow Coma Scale (GCS) when used by specialists from different medical disciplines.

Methods: This prospective observational study was conducted at Selçuk University Faculty of Medicine between December 2023 and June 2024. Eighty-two patients in the Anesthesiology and Reanimation Intensive Care Unit (ICU) were assessed by three specialists—the pulmonologist, the neurologist, and the anesthesiologist—within 24 hours of ICU admission. The scores for each patient by each specialist independently were recorded within a time interval of one hour, consecutively, as possible to minimize the likelihood of any changes in the patient's condition. The variation between evaluators of both scoring systems was analyzed using the Intraclass Correlation Coefficient (ICC). If ICC was below 0.50, the agreement was interpreted as poor.

Results: The study included 82 patients. There were no statistically significant differences in the FOUR and GCS scores assigned by the three specialists. The mortality rate among patients with low scores on both FOUR and GCS was higher than the hospital mortality rate.

Conclusions: Scales used in the ICU should be simple, reliable, and predictive. This study demonstrated that the FOUR is at minimum equivalent to the GCS in meeting these criteria.

Keywords: Full outline of unresponsiveness score, Glasgow coma scale, Intensive care unit

ÖZ

Amaç: Bu çalışma, farklı tıp disiplinlerinden uzmanlar tarafından kullanılan FOUR (Full Outline of Unresponsiveness) skoru ile GKS'nin (Glasgow Koma Skoru) güvenilirliğini karşılaştırmayı amaçlamaktadır.

Gereç ve Yöntem: Bu prospektif gözlemsel çalışma Aralık 2023-Haziran 2024 tarihleri arasında Selçuk Üniversitesi Tıp Fakültesi'nde gerçekleştirilmiştir. Anesteziyoloji ve Reanimasyon Yoğun Bakım Ünitesindeki (YBÜ) 82 hasta, YBÜ'ye kabul edildikten sonraki 24 saat içinde göğüs hastalıkları uzmanı, nörolog ve anesteziyolog olmak üzere üç uzman tarafından değerlendirildi. Her bir hasta için her bir uzman tarafından bağımsız olarak verilen skorlar, hastanın durumunda herhangi bir değişiklik olması ihtimalini en aza indirmek için mümkün olduğunca bir saatlik bir zaman aralığında ve ardışık olarak kaydedildi. Her iki skora sisteminin değerlendiricileri arasındaki varyasyon Sınıf İçi Korelasyon Katsayısı (ICC) kullanılarak analiz edildi. ICC 0,50'nin altında ise uyum zayıf olarak yorumlandı.

Bulgular: Çalışmaya 82 hasta dahil edilmiştir. Üç uzman tarafından verilen FOUR ve GKS skorları arasında istatistiksel olarak anlamlı bir fark yoktu. Hem FOUR hem de GKS skorları düşük olan hastalar arasındaki mortalite oranı hastane mortalite oranından daha yüksekti.

Sonuç: YBÜ'de kullanılan ölçekler basit, güvenilir ve öngörücü olmalıdır. Bu çalışma, FOUR skorunun bu kriterleri karşılamada en azından GKS'ye eşdeğer olduğunu göstermiştir.

Anahtar kelimeler: Glasgow koma skalası, Tepkisizlik skorunun tam taslağı, yoğun bakım ünitesi

Introduction

The scoring systems in the intensive care unit (ICU) serve as a guide in determining consciousness levels (1, 2). Scales have been developed to enable healthcare providers to communicate to each other the general clinical condition of patients, neurological and cardiopulmonary, in a concise and standardized manner (3). Among these scales, the Glasgow Coma Scale is the most widely used (4). Nevertheless, the GCS has several drawbacks. It is limited in cases where

verbal responses may not be evaluated, especially in patients who are intubated and aphasic. It does not include the brainstem reflexes in the scale and fails to take into account the patient's respiration characteristics (1, 4). Quick shifts in mental state, along with changes in breathing and brainstem reflexes, can easily be overlooked.

The Glasgow Coma Scale was published 50 years ago by Teasdale and Jennet for the assessment of

coma and impaired consciousness sources (5). Celebrating its 50th anniversary (6), this scale is still widely used today, although alternative scales have been developed. The Full Outline of Unresponsiveness (FOUR) scoring system, a recently developed coma scale from the Mayo Clinic, assesses four aspects: eye responses, motor responses, brainstem reflexes, and respiration (7). The FOUR allows for the inclusion of brainstem reflexes, respiratory patterns, and verbal responses, which provide additional information for patient assessment, aspects not covered by the GCS (8, 9). This makes it particularly useful in the ICU, where factors like intubation, sedation, or delirium can hinder an accurate assessment of verbal responses, making the FOUR a valuable alternative (8). To ensure the correct understanding and precise application of this scale in Turkey, in 2010, a reliability trial of the Turkish version of the FOUR was performed (10). Especially for unconscious patients, these scales play a crucial

role in determining diagnosis and treatment, and their effectiveness can be somewhat influenced by the user.

This study aimed to evaluate the reliability of the FOUR by comparing it with the GCS assigned to the patients in the ICU by specialists from three different disciplines, the pulmonologist, the neurologist, and the anesthesiologist.

Material and Methods

This prospective observational study was conducted between 15th December 2023 and 15th June 2024 at Selcuk University, the Faculty of Medicine, the ICU of the departments of Anesthesiology and Reanimation. Three different specialists, a pulmonologist (Group A), a neurologist (Group B), and an anaesthesiologist (Group C), independently assessed the patient within the first 24 hours of admission to the ICU and applied

Table 1. GCS and FOUR scoring systems

Category	GCS	FOUR
Eye Response	4 = Spontaneous eye-opening	4 = Eyelids open, tracking, or blinking to command
	3 = Eye-opening to verbal command	3 = Eyelids open but not tracking
	2 = Eye opening to pain	2 = Eyelids closed, open to loud voice
	1 = No eye opening	1 = Eyelids closed, open to pain
		0 = Eyelids remain closed with pain stimuli
Motor Response	6 = Obeying commands	4 = Thumbs up, fist, or peace sign
	5 = Localizing pain	3 = Localizing pain
	4 = Withdrawal from pain	2 = Flexion response to pain
	3 = Abnormal flexion (decorticate posturing)	1 = Extension response (decerebrate posturing)
	2 = Abnormal extension (decerebrate posturing)	0 = No response to pain or generalized myoclonus status
	1 = No motor response	
Verbal Response	5 = Oriented	Brain Stem Reflexes
	4 = Confused	4 = Pupil and corneal reflexes present
	3 = Inappropriate words	3 = One pupil wide and fixed
	2 = Incomprehensible sounds	2 = Pupil or corneal reflexes absent
	1 = No verbal response	1 = Pupil and corneal reflexes absent
		0 = Absent pupil, corneal, or cough reflex
Respiration	N/A	Respiration
		4 = Regular breathing pattern
		3 = Cheyne-Stokes breathing pattern
		2 = Irregular breathing
		1 = Triggering ventilator or breathing above ventilator rate
	0 = Apnea or breathing only at ventilator rate	

FOUR: Full Outline of Unresponsiveness score, GCS: Glasgow Coma Scale, N/A: Non-applicable

the GCS and FOUR scores (Table 1) in the neurological examination. Patients for whom all three specialists were able to rate consecutively within a one-hour time interval were included. The scoring was conducted at the earliest opportunity following admission, with each patient being evaluated only once. The raters were unaware of the other ratings or their outcomes. Demographic data of the patients, APACHE II scores, ICU length of stay, and patient's outcome as survivor or nonsurvivor were collected. Diagnosis of the patients as ischemic or hemorrhagic stroke, pulmonary and cardiac arrest, sepsis-septic shock, and others were also recorded. The study included patients between 18 and 65 years of age who were treated in the ICU for 24 hours or more and who were not sedated. The study excluded patients who were younger than 18 years or older than 65 years, those who received sedation, those who died within the first 24 hours, and those who stayed in intensive care for less than 24 hours.

Statistical Analysis

All statistical analyses were performed using R version 4.2.1 (www.r-project.org). Repeated measures analysis of variance was used to test whether there was a significant difference between the scores given by the experts for each scoring system. In addition, the agreement of the scores given by the experts both in each scoring system and between the two scoring systems was evaluated with the Intraclass correlation coefficient (ICC) and 95% confidence interval. If the ICC was below 0.50, the agreement was interpreted as poor, between 0.50 - 0.75 as moderate, between 0.75 - 0.90 as good, and above 0.90 as excellent according to the levels determined by Koo and Li(11).

Results

A total of 82 patients, 50 males (61%) and 32 females (39%), were included in the study. The mean age of the patients was 62.27 ± 19.87 years (age range: 18-91 years), the mean length of ICU stay was 24.19 (range: 2-145 days) days and the mean length of hospitalization was 34.18 (range: 4-160 days) days. The mean APACHE II score was 11.11 ± 6.79 (range: 3-39). Of the 82 patients, 47 were discharged (57.3%), 35 were died in hospital (42.7%). The mortality rate was 42.6%. Twenty-two patients (26.8%) were intubated or tracheotomized and were using mechanical ventilation, and 60 (73.1%) were breathing spontaneously. The diagnoses of the patients are indicated in Table 2.

Table 2. Diagnosis of the patients in the study

Diagnosis	n (%)
Ischemic and hemorrhagic stroke	7 (8.5)
Pulmonary and cardiac arrest	3 (3.6)
Multitraumas	18 (21.9)
Sepsis	32 (39)
Others	22 (26.8)
Total	82

Patients with low FOUR and GCS scores were found a mortality rate higher than the hospital mortality. For FOUR scores, there were found to be no statistically significant differences between groups A, B, and C ($p=0.743$). Nor were there any statistically significant differences in GCS scores, among groups A, B, and C ($p=0.927$). Reliability results for overall patients are compared in Table 3, for survivors and non-survivors in Table 4, and for female and male patients in Table 5.

Table 3. Reliability results for all patients

	FOUR	GCS	ICC ¹ [95% CI]
Group A	13.04 ± 3.79	12.11 ± 3.81	0.959 [0.846 – 0.983]
Group B	13.05 ± 3.76	12.12 ± 3.75	0.959 [0.840 – 0.983]
Group C	13.06 ± 3.75	12.11 ± 3.77	0.958 [0.828 – 0.983]
p-value	0.743	0.927	
ICC ² [95% CI]	0.999 [0.998 – 0.999]	0.981 [0.972 – 0.987]	

The data were expressed as mean±standard deviation, a p-value was obtained by Repeated Measures analysis of variance (ANOVA). ICC1 indicates overall agreement among specialists, and ICC2 indicates the agreement between scoring systems.

Abbreviations: CI: Confidence intervals, FOUR: Full Outline of Unresponsiveness score, GCS: Glasgow Coma Scale, ICC: Intraclass correlation coefficients

Discussion

The present prospective observational study compared the agreement between the evaluators of GCS and FOUR scores. The evaluators were composed of three specialists in different fields, studying as fellows in the intensive care unit, one pulmonologist, one neurologist, and one anesthesiologist. The scores for each patient were recorded as close together within a time interval of one hour as possible to minimize the likelihood of any changes in the patient's condition. In the primary outcome, agreement between evaluators, there was no significant difference between the GCS and the FOUR score, nor between the evaluations

Table 4. Reliability results for non-survivors and survivors

	FOUR	GCS	ICC ¹ [95% CI]
Non-survivors (n=35)			
Group A	11.771 ± 4.264	11.000 ± 4.221	0.965 [0.906 – 0.985]
Group B	11.800 ± 4.234	11.000 ± 4.165	0.967 [0.900 – 0.986]
Group C	11.800 ± 4.220	11.000 ± 4.172	0.966 [0.900 – 0.985]
p-value	0.885	0.885	
ICC ² [95% CI]	0.999 [0.999 – 0.999]	0.999 [0.998 – 0.999]	
Survivors (n=47)			
Group A	13.979 ± 3.138	12.936 ± 3.273	0.943 [0.671 – 0.980]
Group B	13.979 ± 3.103	12.957 ± 3.223	0.942 [0.686 – 0.979]
Group C	14.000 ± 3.071	12.936 ± 3.246	0.941 [0.638 – 0.979]
p-value	0.613	0.999	
ICC ² [95% CI]	0.998 [0.996 – 0.998]	0.998 [0.997 – 0.998]	

The data were expressed as mean±standard deviation, a p-value was obtained by Repeated Measures analysis of variance (ANOVA). ICC1 indicates overall agreement among specialists, and ICC2 indicates the agreement between scoring systems.

Abbreviations: CI: Confidence intervals, FOUR: Full Outline of Unresponsiveness score, GCS: Glasgow Coma Scale, ICC: Intraclass correlation coefficients

Table 5. Reliability results for male and female patients

	FOUR	GCS	ICC ¹ [95% CI]
Male (n=50)			
Group A	13.480 ± 3.632	12.560 ± 3.753	0.961 [0.820 – 0.985]
Group B	13.500 ± 3.598	12.600 ± 3.636	0.959 [0.828 – 0.984]
Group C	13.520 ± 3.558	12.540 ± 3.764	0.958 [0.784 – 0.984]
p-value	0.372	0.175	
ICC ² [95% CI]	0.999 [0.999 – 0.999]	0.999 [0.999 – 0.999]	
Female (n=32)			
Group A	12.344 ± 4.004	11.406 ± 3.842	0.953 [0.844 – 0.981]
Group B	12.344 ± 3.964	11.375 ± 3.883	0.957 [0.828 – 0.984]
Group C	12.344 ± 3.972	11.438 ± 3.741	0.956 [0.848 – 0.983]
p-value	0.999	0.723	
ICC ² [95% CI]	0.998 [0.997 – 0.999]	0.997 [0.996 – 0.998]	

The data were expressed as mean±standard deviation, a p-value was obtained by Repeated Measures analysis of variance (ANOVA). ICC1 indicates overall agreement among specialists, and ICC2 indicates the agreement between scoring systems.

Abbreviations: CI: Confidence intervals, FOUR: Full Outline of Unresponsiveness score, GCS: Glasgow Coma Scale, ICC: Intraclass correlation coefficients

of the anaesthesiologist, the neurologist, and the pulmonologist.

The GCS has several disadvantages, such as the lack of ability to evaluate the verbal responses in intubated and aphasic patients, and the omission of the brain stem reflexes and the respiratory patterns from its evaluation criteria (2, 12). Thus, quick shifts in the mental status caused by alterations in the respiration patterns and the brainstem reflexes might go undetected. While assessing the patient's state of consciousness, these shortcomings of the Glasgow Coma Scale can lead to mistakes regarding the level of the

coma recovery, and the brain death (10). Unlike the GCS, the FOUR has respiratory parameters. Using this parameter, healthcare professionals can determine if intubated or tracheotomized patients on mechanical ventilation are either apnoeic or breathing at the ventilator frequency. Wijdicks et al. found that the FOUR has advantages over the GCS. One key benefit is that it includes the brainstem reflexes, providing more information about the patient's progress and helping guide urgent interventions for intubated patients (13). Additionally, the FOUR scoring system predicts mortality more accurately. Patients with the low FOUR score have a higher mortality rate compared to those with

the low GCS (13). Finally, the FOUR score shows better observer compliance. However, the drawback is that it does not assess all the behavioral criteria needed to diagnose the minimally conscious state (MCS). MCS is characterized by the patients demonstrating inconsistent but noticeable minimal signs of consciousness, such as responding to noxious stimuli, eye fixation or tracking, reproducible movements to commands, or nonfunctional verbalization (14). The FOUR scale includes an assessment of eye tracking, which helps distinguish between the vegetative state and the MCS patients. However, it is important to note that both acute and chronic patients might only exhibit visual fixation, a factor that is not evaluated by the FOUR scale (15).

Temiz et al. studied 47 patients in the neurosurgical ICU, and the agreement between the practitioners in terms of FOUR and GCS evaluations was found to be high. Wijdicks et al. found that incorporating the brainstem and the respiratory parameters made the FOUR scale a more accurate predictor of mortality compared to the GCS (13). However, other studies have shown that the two scoring systems produce similar outcomes. Eken et al. concluded that the FOUR was not superior to the GCS in patients with altered consciousness arriving at the emergency department (16). Studies comparing GCS and FOUR have shown high inter-observer agreement and correlation (10, 17). In our study, both scoring systems showed similar results between the pulmonologist, the neurologist, and the anesthesiologist. We conducted a similar study in two different ICUs with two specialists from different fields in 2019 and no significant difference was found between the two evaluators. We extended our previous study to evaluate the difference in scoring between internal and surgical departments and to obtain more effective results with a larger number of evaluators. Given the advantages of the FOUR scale, the comparable results between the three specialists suggest that it may be a more beneficial tool for patient assessment. Jalali et al. studied the agreement between users in terms of scoring in a specific group of patients with traumatic brain injury and found that the FOUR score provided more neurological information than the GCS (7). The absence of a clear superiority between the two scores in our findings may be attributed to the selection of a nonspecific patient population and the relatively limited sample size.

Our study has several limitations. A key limitation is that

it was conducted at a single center and within a single intensive care unit, which means its generalizability to other ICUs has not been established. Additionally, the sample size may have been relatively small. We consider that future research involving multiple centers, more ICUs, and larger sample sizes would provide a more robust evaluation.

Conclusion

While the GCS has been the standard scoring system for several decades, various scales are currently in use. These scoring systems must be simple, reliable, and predictive. This study demonstrated that the FOUR is, at minimum, equivalent to the GCS in meeting these criteria.

Informed Consent

Informed consent was obtained from all of the participants included in the study.

Ethical Approval

The protocol of this study was approved by the Clinical Research Ethics Committee of Konya Selçuk University (No: 2023/541, 21 November 2023).

Author Contribution

Conceptualization, YŞB, and JBÇ.; Methodology, YŞB, JBÇ, FÇ.; Formal analysis, MKK.; Investigation, DCG, BP, and YC.; Resources, DCG, BP and YC.; Writing—original draft preparation, YŞB, DCG, BP and YC.; Writing—review and editing, YŞB, JBÇ and FÇ. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Conflicts of Interest

The authors have no conflicts of interest to declare.

Main Points

The Glasgow Coma Scale, while widely used for assessing consciousness, has limitations such as excluding brainstem reflexes and respiratory patterns, which the recently developed FOUR scale addresses.

The FOUR score was found to be at least as reliable as the GCS when used by specialists from three different

fields—pulmonology, neurology, and anesthesiology—in a study conducted in ICU patients, with no significant differences between evaluators.

Although both the GCS and the FOUR scores showed high agreement between evaluators, the FOUR score's inclusion of brainstem reflexes and respiratory parameters makes it a potentially more informative tool, particularly in ICU settings where intubation and other factors complicate verbal assessments.

References

1. Bayraktar YS, Sahinoglu M, Cicekci F, Kara I, Karabagli H, Duman A, Celik JB. Comparison of Glasgow coma scale and full outline of unresponsiveness (Four) score: a prospective study. *Turk Neurosurg*. 2019;29(2):285-8.
2. McNett M, Amato S, Gianakis A, Grimm D, Phillipbar SA, Belle J, Moran C. The FOUR score and GCS as predictors of outcome after traumatic brain injury. *Neurocritical care*. 2014;21:52-7.
3. Kramer AA, Wijdicks EF, Snively VL, Dunivan JR, Naranjo LLS, Bible S, et al. A multicenter prospective study of interobserver agreement using the Full Outline of Unresponsiveness score coma scale in the intensive care unit. *Critical care medicine*. 2012;40(9):2671-6.
4. KARABIYIK L. Intensive care scoring systems. *J Med Surg Intensive Care Med*. 2010;9.
5. Teasdale G, Jennett B. Assessment of coma and impaired consciousness: a practical scale. *The Lancet*. 1974;304(7872):81-4.
6. Manley GT, Maas AI. The Glasgow Coma Scale at 50: looking back and forward. *The Lancet*. 2024;404(10454):734-5.
7. Jalali R, Rezaei M. A comparison of the Glasgow Coma Scale score with a full outline of unresponsiveness scale to predict patients' traumatic brain injury outcomes in intensive care units. *Critical care research and practice*. 2014;2014(1):289803.
8. Fischer M, Rüegg S, Czaplinski A, Strohmeier M, Lehmann A, Tschann F, et al. Inter-rater reliability of the Full Outline of UnResponsiveness score and the Glasgow Coma Scale in critically ill patients: a prospective observational study. *Critical care*. 2010;14:1-9.
9. Temiz NC, Kose G, Tehli O, Acikel C, Hatipoglu S. A comparison between the effectiveness of full outline of unresponsiveness and Glasgow coma score at neurosurgical intensive care unit patients. *Turk Neurosurg*. 2018;28(2):248-50.
10. Örken DN, Sağduyu AK, Şırın H, Işıkara CT, Gökçe M, Süttaş N, et al. Yeni Bir Koma Değerlendirme Ölçeği Olan FOUR Skorun Türkçe çevirisinin Güvenirlik çalışması. *Medical Journal of Trakya University/Trakya Üniversitesi Tıp Fakültesi Dergisi*. 2010;27(1).
11. Koo TK, Li MY. A guideline for selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine*. 2016;15(2):155-63.
12. Okasha AS, Fayed AM, Saleh AS. The FOUR score predicts mortality, endotracheal intubation, and ICU length of stay after traumatic brain injury. *Neurocritical care*. 2014;21:496-504.
13. Wijdicks EF, Bamlet WR, Maramattom BV, Manno EM, McClelland RL. Validation of a new coma scale: the FOUR score. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*. 2005;58(4):585-93.
14. Giacino JT, Ashwal S, Childs N, Cranford R, Jennett B, Katz DI, et al. The minimally conscious state: definition and diagnostic criteria. *Neurology*. 2002;58(3):349-53.
15. 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? *Annals of neurology*. 2006;60(6).
16. Eken C, Kartal M, Bacanlı A, Eray O. Comparison of the Full Outline of Unresponsiveness Score Coma Scale and the Glasgow Coma Scale in an emergency setting population. *European journal of emergency medicine*. 2009;16(1):29-36.
17. Şahin AS, Şahin M, Öztürk NK, Kızılateş E, Karlı B. Comparison Of GCS And FOUR Scores Used In The Evaluation Of Neurological Status In Intensive Care Units Yoğun Bakım Ünitelerinde Nörolojik Durumun Değerlendirilmesinde Kullanılan GKS ve FOUR Skorlarının Karşılaştırılması. *Journal of Contemporary Medicine*. 2015;5(3).