The ability of GCS, FOUR, and APACHE II in predicting the outcome of patients with traumatic brain injury: A comparative study

Masoum Khoshfetrat¹, Mostafa Arab Yaghoubi², Bibi Mahdie Khodadadi Hosseini³, Reza Farahmandrad⁴

ABSTRACT

Introduction: Assessment of the neurological condition of patients admitted to intensive care units gives healthcare professionals the necessary knowledge to attempt to improve their condition and final outcome. Such an assessment cannot be done without a highly reliable and accurate instrument. This study aimed to compare the ability of GCS and FOUR scales to assess the level of consciousness and dysfunction in patients with traumatic brain injury. Methods: This descriptive cross-sectional study was conducted on 102 patients with traumatic brain injury who were found by convenience sampling. The condition of patients during hospitalization up until determination of the final outcome was assessed using GCS, FOUR, Karnofsky, and APACHE-II scales. In the end, the ability of these scales to predict outcome and dysfunction of these patients, and their correlation in this application, was measured using one-way ANOVA and the Pearson correlation test. Results: Of the 102 patients, 80 (78.4%) were male, and 22 (21.6%) were female. The mean age of patients was 32.2±14.8. There was a significant correlation between the results of FOUR and GCS in the assessment of the patients' consciousness level (r = 0.925), which was statistically significant (p = 0.0001). The mean scores of FOUR, GCS, and Karnofsky scales were significantly higher in survivors and in patients with neurological deficits than in non-surviving patients. The patients who died also had a significantly higher APACHE-II score than those who did not. FOUR and GCS showed a strong positive correlation in the assessment of outcomes, and both of them also exhibited a high correlation with APACHE-II in this respect. Conclusion: GCS and FOUR are both suitable scales for assessing consciousness level and outcome of patients with traumatic brain injury.

Key words: FOUR, Glasgow Coma Scale, APACHE II, Traumatic Brain Injury, Karnofsky scale, Intensive care unit

INTRODUCTION

Traumatic brain injury (TBI) is one of the world's most important causes of mortality and severe disability and imposes a devastating financial burden on communities. This issue is exacerbated by the fact that most people suffering from TBI are young. Mortality from TBI can be reduced by improving patient assessment and management systems. In modern evidence-based medicine, access to reliable assessment scales is essential for predicting mortality risk and the ultimate outcome of intensive care so as to allow caregivers to determine the severity and outcome of patient condition and to make important health care decisions accordingly. One of the scales used for such evaluations is the Glasgow Coma Scale (GCS). This scale was first designed in 1974 to assess the level of consciousness of patients with head injury. GCS can predict the primary outcome of TBI (mortality and morbidity) and help healthcare professionals devise a model for care delivery. However, some researchers have questioned the validity of this scale since it does not evaluate brainstem reflexes and cannot assess verbal responses in patients with intubation and tracheostomy. To avoid these shortcomings, researchers have developed other scales for measuring consciousness in TBI patients. One of these scales is the FOUR (Full Outline of UnResponsiveness) score, developed by researchers at Mayo Clinic (United States), which consists of four components: eye responses, motor responses, brainstem reflexes, and respiration pattern. Several studies have suggested that the FOUR scales can provide a better assessment of patients' final outcome than GCS. However, these studies have reached this conclusion by comparing these two scales with each other and have not assessed the consistency of these two scales with the results of APACHE-II or other major scales. Therefore, this study aimed to compare the ability of GCS and FOUR, versus other scales (Karnofsky, APACHE-II, etc.), to evaluate the level of consciousness and dysfunction in TBI patients.
METHODS

This descriptive cross-sectional study was performed with the authorization of the Department of Research and Technology of Zahedan University of Medical Sciences (Iran) and the approval of the Ethics Committee of the university (with Code IR.ZAUMS.REC.1396.33). Using the reports of previous studies and the sample size formula, the sample size was calculated as 102 [14]. Sampling was performed using the convenience sampling method according to the defined inclusion criteria.

Inclusion criteria were as follows: age of 18-65 years, hospitalization for head trauma, mechanical ventilation, and at least 24 hours of stay in the intensive care unit. Exclusion criteria were as follows: lack of consent of family members, the discovery of malignant and fatal diseases, cervical spinal cord injury, musculo-skeletal paralysis, deafness, blindness, and history of eye surgery resulting in a change in pupil size and shape.

All eligible patients were monitored using GCS, FOUR, APACHE II, and Karnofsky scales until discharge or death. For sedated patients, the infusion was stopped one hour before measurements. The scores obtained from the scales were analyzed in SPSS using descriptive statistics, one-way ANOVA, and Pearson’s correlation test.

The FOUR scale consists of four components: eye responses, motor responses, brainstem reflexes, and respiration patterns (Figure 1). The validity and reliability of GCS and FOUR for Iranian patients were assessed in a 2011 study conducted in Tehran on 155 patients with 4 nurses and anesthesiology residents (including the author) in two groups of trained and untrained staff, which observed excellent inter-group agreement. In this study, the coefficient of correlation of FOUR scores in the trained and untrained groups were 0.998 and 0.993, respectively; the kappa coefficient of intra-group agreement in these groups were 0.981 and 0.986, respectively; as well, the coefficient of correlation between FOUR and GCS scores in these groups were 0.980 and 0.925, respectively [15].

APACHE II (Table 1) consists of 3 main dimensions: a physiological dimension, which includes GCS, temperature, mean arterial blood pressure, heart rate, respiratory rate, Pao2, arterial blood pH, serum sodium, potassium, creatinine, hematocrit, and white blood cell count. The score of each item in the physiological dimension ranges from 0 to 4, giving this dimension a total score of 0-59. The second part of this scale is a score between 0 and 6, which is awarded based on the patient’s age group. The third part involves the evaluation of chronic diseases and the failures of one or more organs. The sum of scores obtained from these three parts is the APACHE II score, which ranges from 0 to 71. The higher the APACHE II score, the higher the patient’s risk of mortality [16]. Karnofsky scale is used for assessing the patient’s degree of dysfunction. This scale takes values between 0 (death) and 100 (no dysfunction), and reflect the patient’s activity level; the lower the Karnofsky score, the more intensive care the patient requires [17,18].

RESULTS

Of the 102 patients, 80 (78.4%) were male, and 22 (21.6%) were female. The mean age of patients was 32.2 ± 14.8 (Figure 2).

The mean GCS score at admission was 8.6 ± 3.5, and at the end of the study was 12.9 ± 3.5. The mean GCS score during the hospitalization period was 10.8 ± 2.9. The mean FOUR score at admission was 9.3 ± 4.5 and at the end of the study was 14.8 ± 4.8. Also, the mean FOUR score during the hospitalization was 11.7 ± 3.9. The mean Karnofsky core was 23.9 ± 5.5, with the lowest score being 10 and the highest being 40.

The mean APACHE II score at admission was 11.5 ± 6.7, with the lowest score being 1 and the highest being 29 (Figure 3).

There was a significant correlation between the FOUR and GCS scores for the level of consciousness (r=0.925), which was statistically significant (p=0.0001). The mean scores of FOUR, GCS and Karnofsky scales were significantly higher in survivors and patients with neurological deficits than in patients who died. Patients who died also had a significantly higher APACHE II score than those who did not (Table 2).

A high degree of positive correlation was observed between assessments by FOUR and GCS for patient outcome (healthy, neurological deficits, or death). Also, both of these scales showed very high negative correlation with APACHE II. The correlation between Karnofsky and APACHE II scores in healthy patients was r = -0.457; in patients with neurological deficits r = -0.481, and in deceased patients, was r = -0.796. These correlations were statistically significant in all cases except for patients with neurological deficits (Table 3).

Both GCS and FOUR were as capable as APACHE II in predicting the outcomes. The highest agreement between the scales was in the assessment of mortality rates, which was slightly higher between GCS and APACHE II.
Table 1: The APACHE II scale

| Physiologic variable            | High abnormal range | Low abnormal range |
|--------------------------------|---------------------|--------------------|
|                                | +4  | +3  | +2  | +1  | 0   | +1  | +2  | +3  | +4  |
| Rectal temp (°C)               | ≥41 | 39-40.9 | 38.5-38.9 | 36-38.4 | 34-35.9 | 32-33.9 | 30-31.9 | ≤29.9 |
| Mean Arterial Pressure (mmhg)  | ≥160 | 159-130 | 129-110 | -   | 109-70 | -   | 69-50 | -   | ≤49  |
| Heart Rate                     | ≥180 | 179-140 | 139-110 | 109-70 | 69-55 | 54-40 | ≤39  |
| Oxygenation                     |      |      |      |      |      |      |      |      |      |
| a) Fio2 ≥ 0.5 record A-aDo2    | ≥500  | 499-350 | 349-200 | -   | <200  | Pao2  | -   | Pao2  | Pao2  |
| b) Fio2 < 0.5 record Pao2      | ≥7.7  | 7.69-7.6 | -   | 7.59-7.5 | 7.49-7.33 | -   | 7.32-7.25 | 7.24-7.15 | <7.15 |
| Arterial pH                    | ≥52  | 51.9-41 | -   | 40.9-32 | 31.9-22 | -   | 21.9-18 | 17.9-15 | <15  |
| HCO₃ (mEq/l)                   | ≥7   | 6.9-6  | -   | 5.9-5.5 | 5.4-3.5 | 3.4-3 | 2.9-2.5 | -   | <2.5  |
| K (mEq/l)                      | ≥180 | 179-160 | 159-155 | 154-150 | 149-130 | -   | 129-120 | 119-111 | <110  |
| Na (mEq/l)                     | ≥3.5 | 3.4-2  | 1.9-1.5 | -   | 1.4-0.6 | -   | <0.6  | -   | -     |
| S.Creat (mqm/dl)               | ≥60  | -   | 59.9-50 | 49.9-46 | 45.9-30 | -   | 29.9-20 | -   | <20   |
| Hematocrit (%)                 | ≥40  | -   | 39.9-20 | 19.9-15 | 14.9-3 | -   | 2.9-1  | -   | <1    |
| WBC (10³ /cc)                  |      |      |      |      |      |      |      |      |      |
| Age-score                      |      |      |      |      |      |      |      |      |      |
| < 44− 0                        | 15−0 | 10−5 | 5−10 | 10−5 |     |     |     |     |      |
| 45−54− 2                       | 14−1 | 9−6 | 4−11 | 11−6 |     |     |     |     |      |
| 55−64− 3                       | 13−2 | 8−7 | 3−12 | 12−7 |     |     |     |     |      |
| 65−74− 5                       | 12−3 | 7−8 |     |     |     |     |     |     |      |
| ≥ 75− 6                       | 11−4 | 6−9 |     |     |     |     |     |     |      |
Figure 1: The FOUR scale.

Table 2: Comparison of mean scores of the FOUR, GCS, Karnofsky, and APACHE II scales using one-way ANOVA

| Variable       | Outcome at discharge | P-value |
|----------------|----------------------|---------|
|                | Survival n = 78      | Neurological deficit n = 14 | Death n = 10 |
| GCS            | 11.6 ± 1.9           | 9.4 ± 1.3 | 7.7 ± 4.7 | 0.0001 |
| FOUR           | 12.9 ± 2.4           | 11.4 ± 1.2 | 8.8 ± 6.7 | 0.0001 |
| Karnofsky      | 24.8 ± 5.5           | 21.4 ± 3.6 | 20 ± 4.7  | 0.005  |
| APACHE II      | 9 ± 4.7              | 17.9 ± 6.1 | 22.1 ± 5  | 0.0001 |
Figure 2: Age frequency of patients.

Table 3: Comparison of correlation between FOUR and GCS, Karnofsky and APACHE II scores for the assessment of patient outcomes using Pearson correlation test

| Outcome at discharge | Compared variable          | r     | P-value |
|----------------------|----------------------------|-------|---------|
| Survival             | Healthy                    | 0.826 | 0.0001  |
|                      | neurological deficit       | 0.746 | 0.002   |
| Death                | GCS-FOUR                   | 0.871 | 0.001   |
| Survival             | Healthy                    | -0.457| 0.0001  |
|                      | neurological deficit       | -0.481| 0.082   |
| Death                | Karnofsky-APACHE II        | -0.796| 0.006   |
| Survival             | Healthy                    | -0.737| 0.000   |
|                      | neurological deficit       | -0.7   | 0.01    |
| Death                | GCS-APACHE II              | -0.911| 0.000   |
| Survival             | Healthy                    | -0.699| 0.0001  |
|                      | neurological deficit       | -0.646| 0.00    |
| Death                | FOUR-APACHE II             | -0.891| 0.000   |
DISCUSSION
Access to reliable and uniform scales for assessing a patient’s level of consciousness can help caregivers provide a better assessment of patient condition. The results of this study showed that GCS and FOUR are perfectly consistent in the assessment of patients’ level of consciousness. Most of the previous studies have reported the better ability of the FOUR scale in assessing consciousness and have suggested that GCS is unable to measure brainstem reflexes, is inconsistent in measuring verbal reflexes in confused patients without consciousness disorder, and is unable to give reliable results for patients with intubation and tracheostomy. However, the results of this study showed no significant difference between the initial score obtained from these two scales. There was also no significant difference between the two scales in the assessment of consciousness and outcome. In other words, GCS was as capable as FOUR in predicting patient outcomes. Considering the ease of use of GCS and the fact that it can be used by all healthcare team members and does not need any special measurement equipment, GCS can still serve as a good scale for measuring consciousness and outcome.

Contrary to this finding, Keykha et al. reported that the FOUR scale is a better measure of consciousness level and should be used instead of GCS when measuring consciousness in ICU patients. This difference in the results can be attributed to the fact that the above study was performed on patients receiving continuous sedation infusion, whereas sedation of the patients in this study was stopped one hour before measurements. Stopping the sedation infusion before consciousness measurements have the benefit of ensuring that the patient’s level of consciousness is not physically reduced. To avoid high levels of reduced consciousness, it is best to use on-off infusions, because sedative drugs may accumulate in the fat tissue, and their sedative effects can remain for as long as 24 hours after stopping infusion. In cases where continuous sedation is absolutely needed for pain control or tolerance of mechanical ventilation, although the scores of both scales would drop, the FOUR scale appears to be more suitable for consciousness assessment. In a study by Jalali, it was reported that the FOUR scale was better in assessing consciousness and predicting mortality than GCS, but the low score of neither scale was associated with mortality because deceased patients had a higher mean score in both
In terms of the scales’ ability to measure consciousness, these results are consistent with previous findings but are inconsistent with the findings of the present study. In terms of predicting mortality, these results are inconsistent with all other reports because all previous studies have shown that patients with a low admission score, in either scale, have a mortality outcome. Unlike the present study, Phuping et al. recognized the FOUR scale as a reliable predictor of recovery and mortality and reported that most of the recoveries occurred in patients who earned a FOUR score of more than 10. They argued that since the lower scores of the FOUR scale could provide a better assessment of patient condition (compared to GCS), this scale is recommended for measuring changes in the level of consciousness in neurosurgery and similar patients. It should be noted that there are etiological differences between these two studies, as subjects of that study included patients with tumors, aneurysms, hydrocephalus, and brain infections. Yet, this study was conducted only on patients with head trauma. Consistent with the results of the present study, Bruno et al. reported that GCS and FOUR are well-correlated in predicting mortality and recovery. But in intensive care units, these scales must not only measure prognosis and consciousness but also meet other requirements, one of which is the distinction between coma, vegetative state and locked-in syndrome (LIS). According to the GCS scoring criteria in the eye response dimension, open eyes earn a complete score. Meanwhile, vegetative patients cannot close their eyes or follow a target, and LIS patients only have vertical eye movements. In the eye response component of the FOUR scale, a complete score is awarded only when the patient can blink on-demand or follow a target, which helps make a distinction between the aforementioned states. Despite all these shortcomings of GCS, Hosseini et al. reported that this scale was even better than APACHE II in predicting patient mortality. Cho et al. stated that although GCS is suitable for assessing early mortality, it is not a good replacement for APACHE, as APACHE II is a better predictor of late mortality.

CONCLUSIONS
The results of this study showed that GCS and FOUR are both very good scales for measuring the consciousness level and outcome of TBI patients and that neither scale can be considered preferable over the other in measuring the consciousness level. Since many healthcare professionals with varying degrees of expertise participate in the treatment process, the instruments used to evaluate patients’ condition should be able to provide a uniform assessment of consciousness level and outcomes when applied by different people. Therefore, it is best to choose the right instrument according to the type of department and the skill of healthcare professionals who are going to employ it.

LIMITATIONS
The limitations of this study included the absence of any measurement during sedation and the small size of the sample. The authors recommend conducting a similar study with a greater number of subjects to confirm the findings.

COMPETING INTERESTS
The authors declare that there is no conflict of interest regarding the publication of this paper.

AUTHORS’ CONTRIBUTIONS
Masoum Khoshetrat: Participation in implementation of research and supervision over data collection and revising this paper. Mostafa Arab Yaghoubi: Participation in implementation of research and data collection. Bibi Mahdie Khodadadi Hosseini: Compilation of the paper. Reza Farahmandrad: Analyzing the data and editing this paper carefully. All authors reviewed, commented and approved the final version.

FUNDING
The present study was financially supported by Zahedan University of Medical Sciences, Zahedan, Iran.

ACKNOWLEDGMENT
This article has been adopted from the Thesis written by Resident of Anesthesiology and Critical Care, approved and funded by Research and Technology Deputy of Zahedan University of Medical Sciences. The researcher would like to express own gratitude to the Research and Technology Deputy of the University and the staff and ICU ward at the Khatam-ol-Anbia Hospital of Zahedan.

REFERENCES
1. Jalali R, Rezaei M. A comparison of the Glasgow Coma Scale score with full outline of unresponsiveness scale to predict patients’ traumatic brain injury outcomes in intensive care units. Critical care research and practice. 2014;2014:1–4. Available from: 10.1155/2014/289803.
2. Gururaj G. Epidemiology of traumatic brain injuries: indian scenario. Neurological Research. 2002;24(1):24–8. PMID: 12078375. Available from: 10.1179/016164102101199503.
3. Celso B, Tepas J, Langland-Orban B, Pracht E, Papa L, Lottenberg L. A systematic review and meta-analysis comparing outcomes of severely injured patients treated in trauma centers following the establishment of trauma systems. The Journal of Trauma. 2006;60(2):371–8. PMID: 16508498. Available from: 10.1097/01.ta.0000197916.99629.eb.
4. Ayazoglu TA. A comparison of APACHE II and APACHE IV scoring systems in predicting outcome in patients admitted with stroke to an intensive care unit. Anaesthesia, Pain & Intensive Care. 2011;19(20):7–12.

5. Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. Lancet. 1974;2(7872):81–4. PMID: 4136544. Available from: 10.1016/0140-6736(74)91639-0.

6. Gill MR, Reiley DG, Green SM. Interrater reliability of Glasgow Coma Scale scores in the emergency department. Annals of Emergency Medicine. 2004;43(2):215–23. PMID: 14747811. Available from: 10.1016/S0196-0644(03)00814-X.

7. Khajeh A, Fayyazi A, Miri-Allabad G, Askari H, Noori N, Khajeh B. Comparison between the ability of Glasgow coma scale and full outline of Unresponsiveness score to predict the mortality and discharge rate of pediatric intensive care unit patients. Iranian Journal of Pediatrics. 2014;24(5):603–8. PMID: 25793069.

8. Keykha A, Askari H, Navidian A, Hosseini BM. Ability of Glasgow coma Scale and Full Outline of Unresponsiveness Score in Measuring Level of Consciousness and Outcome in Patients Receiving Sedation Under Mechanical Ventilation. Journal of Critical Care Nursing. 2017;10(1):e10160. Available from: 10.5812/ccnr.10160.

9. Fischer M, Ruegg S, Czapinski A, Strohmeier M, Lehmann A, Tschan 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(2):R64. PMID: 20398274. Available from: 10.1186/cc8963.

10. Murray GD, Teasdale GM, Braakman R, Cohadon F, Dearden M, Iannotti F. The European Brain Injury Consortium survey of head injuries. Acta Neurochirurgica. 1999;141(3):223–36. PMID: 10214478. Available from: 10.1007/s007010050292.

11. Born JD, Hans P, Dexters G, Kalangu K, Lenelle J, Milbourn G. (Practical assessment of brain dysfunction in severe head trauma [author's transl]). Neuro-Chirurgie. 1982;28(1):7–7. PMID: 7110496.

12. Iyer VN, Mandrekar JN, Danielsen RD, Zubkov AV, Elmer JL, Wyjdicks EF, editors. Validity of the FOUR score coma scale in the medical intensive care unit. Mayo Clinic Proceedings. Elsevier; 2009.

13. Wolf CA, Wyjdicks EF, Bamlet WR, McClelland RL, editors. Further validation of the FOUR score coma scale by intensive care nurses. Mayo Clinic Proceedings. Elsevier; 2007.

14. Bruno MA, Ledoux D, Lamberton B, Damas F, Schnakers C, Vanhaudenhuyse A. Comparison of the Full Outline of UnResponsiveness and Glasgow Liege Scale/Glasgow Coma Scale in an intensive care unit population. Neurocritical Care. 2011;15(3):447–53. PMID: 21526394. Available from: 10.1007/s12028-011-9547-2.

15. Tadrisi SD, Bahari N, EBADI A, Madani SJ. Validity and reliability of coma scale (Four Score) in adult patient hospitalized in critical care units. Critical Care Nursing. 2012;25(1):95–102.

16. Wagner DP, Draper EA. Acute physiology and chronic health evaluation (APACHE II) and Medicare reimbursement. Health Care Financing Review. 1984;1984:91–105. PMID: 10311080.

17. Mor V, Laliberte L, Morris JN, Wiemann M. The Karnofsky Performance Status Scale. An examination of its reliability and validity in a research setting. Cancer. 1984;53(9):2002–7. PMID: 6704925. Available from: 10.1002/1097-0142(19840501)53:9<2002::AID-CNCR2820530933>3.0.CO;2-W.

18. Chambless LB, Kistka HM, Parker SL, Hassam-Malani L, McGirt MJ, Thompson RC. The relative value of postoperative versus preoperative Karnofsky Performance Scale scores as a predictor of survival after surgical resection of glioblastoma multiforme. Journal of Neuro-Oncology. 2015;121(2):359–64. PMID: 25344883. Available from: 10.1007/s11060-014-1640-x.

19. Bakhtiar S, Amiri F, Kazemnegad A, Baksheshdeh H, Moma-madalha Z, Termechi G, et al. Comparison GCS and FOUR score for predicting neurological outcome in patients with impairment of consciousness. 2014;3(2):85–94.

20. Marino PL. Marino's the ICU Book. Lippincott Williams & Wilkins; 2014.

21. Akavipat P, Sookplung P, Kaewsingha P, Maunsaiyat P. Prediction of discharge outcome with the full outline of unresponsiveness (FOUR) score in neurosurgical patients. Intensive Care Medicine. 2009;35(3):447–53. PMID: 21526394. Available from: 10.1007/s007010050292.

22. Wijdicks EF, Kolmen E, O'Brien PC. Measurement of impaired consciousness in the neurological intensive care unit: a new test. Journal of Neurology, Neurosurgery, and Psychiatry. 1998;61(1):117–9. PMID: 9436740. Available from: 10.1136/jnnp.61.1.117.

23. Hosseini M, Ramazani J. Comparison of acute physiology and chronic health evaluation II and Glasgow coma scale in predicting the outcomes of post anesthesia care unit patients. Saudi Journal of Anaesthesia. 2015;9(2):136–41. PMID: 25829900. Available from: 10.4103/1658-354X.152839.

24. Cho DY, Wang YC. Comparison of the APACHE III, APACHE II and Glasgow coma scale in acute head injury for prediction of mortality and functional outcome. Intensive Care Medicine. 1997;23(1):77–84. PMID: 9037644. Available from: 10.1007/s001340050294.