

Comparison of interrater reliability and predictive validity of FOUR score and Glasgow Coma Scale in multi traumatic patients

Samad Shams Vahdati,MD¹; Jafar Ghobadi,MD²; MohammadReza Bazavar,MD³; Fatemeh Seifar⁴

1-Associate professor of emergency medicine, Road Traffic Injury Research Center, Emergency department, Tabriz University of medical science, Tabriz/Iran

2- Resident of emergency medicine, emergency department, Tabriz University of medical science, Tabriz/Iran

3- Assistant professor of orthopedi, department of orthopedi, Tabriz University of medical science, Tabriz/Iran

4- Medical student, neuroscience research center AND students research committee, Tabriz University of medical sciences, Tabriz/Iran

Abstract

Background: Multi traumatic injuries impose health care concern and major burden for society. The Glasgow Coma Scale (GCS) is a routine scale for assessing levels of consciousness and prognosis of traumatic patients. The Full outline of unresponsiveness (FOUR) score is a new coma scale developed to overcome the limitations of GCS. In this prospective study, we aimed to compare the predicting outcomes and inter-rater reliability of the GCS and FOUR score in a group of multi traumatic patients. 96 consecutive multi trauma patients admitted in emergency departments were enrolled in the study. GCS and FOUR score were documented on arrival to the emergency room. Their correlation with patients 'outcomes was analyzed. In terms of predictive power for in-hospital mortality, calculated mortality rate was 33.1 for FOUR score and 30.21 for GCS. Mean value of GCS and FOUR score were 14.83 and 13.68, respectively. Mortality rate was determined 9.3% and mean duration of hospitalization was 7.86 ± 8.73 days. In addition, inter-rater reliability was determined $\kappa = 0.84 \pm 0.01$ for GCS score and $\kappa = 0.86 \pm 0.01$ for FOUR score rating. Inter-rater reliability and outcome predictability for FOUR score was superior to the GCS in this study, therefore FOUR score can be considered as a viable alternative to the GCS in the emergency department by accurately predicting outcome and improving the quality of management in trauma patients.

Key Words: Interrater reliability; Prediction; Glasgow Coma Scale; FOUR score; Trauma

Corresponding author:

Fatemeh Seifar

Medical student, neuroscience research center, students research committee, Tabriz University of medical sciences, Tabriz/Iran

Email: f.seifar@gmail.com

Receive date: 2017-06-11 | Accept date: 2017-09-24 | Publish date: 2017-10-01

DOI: 10.7575/aiac.abcm.17.05.04.03





Introduction

Multiple traumas refer to severe damages affecting two or more organs. The most possible organs in these kinds of injuries are head and neck as well as the extremities, while abdominal and thoracic damages could commonly happen (1). Trauma is the fourth most common reason of mortality in the world (2). And it's been reported that more than 5 million people have died as the result of trauma in 2010 (3). Crude mortality rate for Trauma is 89% in developing countries and the youth are more likely to die compared with other age groups (4). In evaluating traumatic patients, mental component is often the most important in predicting possible mortality. Lack of an objective scale leads to ambiguous estimates of the level of consciousness, which obscures the ultimate results of any managements (5,6). Education and workshops are necessary for health care providers and physicians and must be taken into consideration (7,8). In evaluation of trauma patients, it is necessary to check vital signs and use ultrasonography as an instrument which helps physicians to evaluate patients in an easy way to find critical situations (9,10,11). Glasgow Coma Scale (GCS) is a routine component of a neurological exam following traumatic injuries. This scale was created to facilitate the assessment and grading of the status of the central nervous system and recovery evaluation in critically ill patients (12). It is also an index of injury severity, as the scores are related to either mortality or morbidity. GCS measures a trauma patient's motor, eye, and verbal responses to external stimuli and it consists of 15 points. Traditionally, mild brain injury is defined as a summed GCS score ranging from 13-15,

Moderate injury 9-12, and Severe injury 3-8 (13).

Nowadays, GCS score is widely used as a prognostic indicator of traumatic patients, while in recent studies the verbal component of GCS score has shown to be inaccurate in up to 10% of patients due to injury or intubation (14). In addition, studies, in which GCS inter-observer agreement was examined, validation of its inter-observer reliability was questioned in patients with GCS score < 13, averaging in some cases more than one point to five-point scales of the GCS (15,16). Moreover, it has been criticized for not accurately evaluating the severity of consciousness according to the altered inspiration patterns which was detected in intubated patients or abnormal brain stem reflexes (17). Because of these limitations of the GCS, many scoring systems have emerged and been compared with GCS. The FOUR (Full Outline of Unresponsiveness) score is a new coma scale that consists of 4 components eye, motor, brainstem, and respiration (18). Recent studies on the full outline of unresponsiveness (FOUR) Scale suggested that it can overcome the limitations of GCS (15,17). In this study, we determined the prognostic value of the GCS and FOUR scores in a group of multi traumatic patients. We also aimed to identify the validity based on the inter-rater reliability of each score. We investigated whether the FOUR score is better than GCS in predicting mortality in patients with multi-traumatic injuries.

Method

This prospective analytic study was conducted between the March 2013 and October 2013 in emergency department of Imam Reza and Sina hospitals in Tabriz, Iran. This study was approved by the ethical committee of Tabriz





University of medical sciences; written consent was obtained from all the participants or a legal next relative. Our study population included all multi-traumatic patients admitted to the emergency departments. We excluded the patients under the age of 18, patients with ESI (triage system) score of 1, 4 or 5 and Injury Severity Score (ISS) of 15 or more. In addition, patients who were unwilling to participate in the practice were omitted from the study. Prior to the study, an instruction about the FOUR score and the GCS was given to the physicians and nurses in the emergency departments. They were asked to grade their patients, according to both GCS and FOUR score scales. At first hours after the admission, patients were graded according to their clinical condition (shock index and shock stage) by the attending physician. The FOUR score and the GCS were prospectively studied at the first possible time after admittance. The patients were separately rated by both of the emergency staff to evaluate the inter-rater agreement. The expected chance of death was calculated by means of formulas of each severity scoring system. $GCS \leq 5$ and $FOUR \text{ score} \leq 6$ were determined as the cutoff points of each scaling system. Then all patients were followed up according to the duration of hospitalization and the incidence of death. We provided a check list for each patient which included the following information; age, sex, type of the trauma (including car accident, falling from the height and vehicle roll-over ...), emergency transferring device, shock index (heart rate/systolic blood pressure) greater than 0.8 identified as compensated shock, shock stage (according to the volume of blood loss) was defined in three stages; Stage I (compensated),

Stage II (decompensated), and Stage III (irreversible), FOUR score and GCS which were rated by the physician and the triage nurse, the prognosis of each patient and the duration of hospitalization. Data were collected and analyzed using SPSS software version 15.0. Continuous variables were analyzed using means and median. The Spearman test was used to determine possible correlation between the calculated FOUR score and the GCS score. Kappa qualification was determined to get the inter-rater reliability. Pearson's Chi-square Test was used to find associations between the prognosis and mean hospital duration. Moreover, the correlation between the clinical condition and prognosis was determined by Pearson's Chi-square Test. Then Kruskal-Wallis test was used to evaluate any correlation between the GCS and Four scores with prognosis. The mortality rate was calculated according to both GCS and FOUR score. Mortality prevalence was determined using denominator of the total number of included multi trauma patients. The P - value of < 0.05 was defined as meaningful.

Results

A total of 96 patients, who matched the criteria, were included in this study. Among them, 74 (77.1%) were male and 22 (22.9%) were female. Their mean age was determined 31.86 ± 16.88 years. Mean duration of hospitalization was evaluated 7.86 ± 8.73 days. Mean GCS and FOUR score were 14.83 ± 2.65 and 13.68 ± 2.28 , respectively. In terms of trauma type, 32.3% of patients had a car accident injury, 13.5% were with motorcycle accident injury, 15.6 % of them had fallen from a height and 5.2% were with vehicle rollover. 28.1% of patients were pedestrians, 13.5%





were struck by passenger car and 13.5 % had a motorcycle collision. In 9.4% of the patient's, other reasons which weren't categorized in this summary accounted for the injury. This shows that the major number of multi traumatic patients was car and motorcycle passengers. In terms of transferring device type, the majority, i.e. 78.1% of the patients were transferred to the emergency department by emergency medical services (EMS). 40.6% were dispatched from other medical centers and 37.5% were transferred during a routine emergency call. Clinical condition was examined according to the shock indexes and shock staging. Out of 96 patients, 11.2% had compensated shock (shock index < 0.8) and in the rest of them (88.8 %) shock index was more than 0.8. According to shock staging results, 91.4% of patients had normal condition and were not categorized in any of shocking stages. 5.4% were graded in stage I and 3.2% were in stage II or III. The mortality rate was 9.3 % for multi trauma patients; among them, 2 (2.0 %) expired in the emergency department and 7 (7.3%) expired in admitting departments. A significant relationship was evaluated between shock stage and shock index (Pearson's Chi-square Test, 0.001). So for the patients at the shock index < 0.8, shock stage was normal in 97.4% and the rest of them (2.6%) were in stage I. Among the patients with shock index >0.8, 33.3% were in stage I, 33.3% at stage II and 33.3% were categorized at stage III. Our results revealed that there are significant relationships between the shock stage and shock index with patients' prognosis. Among the patients with shock index < 0.8; 60% were discharged from the department, 29% were hospitalized and 10.2% expired; while at shock index > 0.8; 70%

of patients were hospitalized or discharged and 10% of them expired in the emergency department (X^2 test, 0.012). Similarly, among the patients with normal shock index 62.4% were discharged, 29.4% were hospitalized and 8.3% expired. At the stage I patients, 20% were discharged, 60% were hospitalized and 20% expired and also for the Stage II and stage III patients, 66.7% were hospitalized and 33.3% did not survive. Spearman's correlation coefficient between GCS and FOUR scores was high (0.981 and $p=0.0001$). In addition, inter-rater reliability was determined 0.84 ± 0.01 for GCS score and 0.86 ± 0.01 for FOUR score rating. The GCS has been shown to correlate with morbidity and mortality, the evaluated scores of discharged patients were significantly higher than the hospitalized or expired patients ($p= 0.001$). FOUR score was significantly lower in expired patients (Kruskal-Wallis, 0.001) compared with admitted or discharged patients. Among the patients admitted from other departments, the mean hospital duration was 15.2 days for the expired patients and 6.63 days for the rest; this means that there would be a relationship between the hospital duration and mortality rate (P value = 0.04).

Discussion

Trauma imposes a substantial burden on national economies, costing billions of dollars for health care system each year. According to the annual reports of police and government officials, road traffic accidents are the second most common cause of death after cardiovascular events among civilians in Iran (1). Multiple traumas constitutes the highest proportion of mortality in Iranian young men (under 40 years) (2). According to the results of our study men are more susceptible to





traumatic accidents than women in Iran. This was similar to the results of Roudsari et al. in Iran (3).

Car and motorcycle crash were responsible for the majority of the patients in our study accounting for 45.8% of cases, and among them, car passengers were the major group of victims. The findings of this study didn't match those reported by Fazel et al. in which, the most common injury mechanism was motorcycle events on city roads (4). Prognostic factors play a significant role in clinical management of traumatic patients. It is important to identify patients at risk of early mortality who cannot benefit from an effective treatment. Ideally, Prediction models' outcome should be reproducible and accurate to support clinical judgement.

Classification of trauma severity was almost similar between the FOUR score and the GCS categories, according to the correlation coefficient of 0.98. FOUR score have been reported to get the highest inter-rater reliability the in most studies as supported by the present study, in which k was 0.86 ± 0.01 for FOUR scale compared with $k=0.84 \pm 0.01$ for GCS. Wijdicks et al. compared the reliability of FOUR score according to coordination between physicians and nurses in 2005 and 2011, which was estimated $k=0.82$ (5). In addition, the data reported in Fischer's study supported our results, in which inter-rater coordination for GCS and FOUR score was $k=0.71$ and $k=0.82 \pm 1$, respectively (6).

According to our study, the clinical condition of the patients at first admission hours was relatively capable of predicting mortality in multi-traumatic patients. In line with other studies Sock Index (SI) is an accurate

predictor of mortality in multi traumatic patients (12,13). Prognosis of patients with SI lower than or equal to 0.8 was hopefully better than those with SI greater than 0.8. The $SI < 0.8$ was associated with greater survival rate in the emergency department, while it was in controversy to survive in other departments, in which higher mortality rate attributed to lower SI. This difference in survival is probably due to a combination of factors, including organ dysfunction, time wasting management in other departments, and lack of appropriate diagnostic and therapeutic facilities for the care of these patients in the admitted departments. Our study results also demonstrated an association between shock stages and higher mortality rate, reflecting that patients at decompensated and irreversible shock stage were more susceptible to early death compared to those in compensated stage and this was directly related to possible intracranial hemorrhage.

Similar to previous studies (14,15), GCS and FOUR scores were both predictive of mortality in trauma patients admitted to our emergency; however, our data indicated that FOUR score is a more effective indicator for in-hospital mortality and discharge outcome. Their efficacy in mortality prediction in hospital was 33.1 for FOUR and 30.21 for GCS. In the current study, the calculated mortality rate was lower than other studies, like Eken (16), Gorji (17), Sadaka (18), and Akavipat (19) study, which can be explained with the following reasons. The study population in the Sadaka study included identified Traumatic Brain Injury (TBI) patients admitted to the Neuro-ICU, so the studied sample included all neurologically ill patients, which was associated with a greater





mortality rate compared to all multi traumatic patients in our study (18). The inconsistency with Gorji et al. Akavipat et al. and Eken et al. (16,17,19) may be related to different courses of cut off points in the research plan and difference in injury severity. Our results were in contrast with Akavipat et al. study to determine the value of the two scoring systems in reliability and feasibility; the results revealed that FOUR score is not superior to the GCS in neurologically ill disease. This may be related to the use of small study population in this study (20).

Finally, while FOUR score has been validated among adult population, recent comparative studies of two scoring systems in

the pediatric population revealed that FOUR score provides no significant advantage over GCS in evaluation of children (21-23).

Conclusion

FOUR score appears to be a valuable tool with a great interrater agreement that is comparable to GCS in predicting in-hospital outcome. It is performable in intubated patients with identifying nonverbal signs of consciousness by assessing visual pursuit. Therefore, according to the results of this study, FOUR score can be considered as a viable alternative to the GCS in the emergency department by accurately predicting survivability and improving the quality of management in traumatic patients and their families.

References

1. Montazeri A. Road-traffic-related mortality in Iran: a descriptive study. *Public Health* 2004 Mar;118(2):110-3.
2. Zangoeei DH, Shafahi Y, Zangoeei DM. Fatal accident distribution by age, gender and head injury, and death probability at accident scene in Mashhad, Iran, 2006-2009. *Int J Inj Contr Saf Promot* 2013;20(2):121-33.
3. Roudsari BS, Sharzei K, Zargar M. Sex and age distribution in transport-related injuries in Tehran. *Accid Anal Prev* 2004 May;36(3):391-8.
4. Fazel MR, Fakharian E, Mahdian M, Mohammadzadeh M, Salehfard L, Ramezani M. Demographic Profiles of Adult Trauma During a 5 Year Period (2007-2011) in Kashan, IR Iran. *Arch Trauma Res* 2012;1(2):63-6.
5. Wijdicks EF, Rabinstein AA, Bamlet WR, Mandrekar JN. FOUR score and Glasgow Coma Scale in predicting outcome of comatose patients: a pooled analysis. *Neurology* 2011 Jul 5;77(1):84-5.
6. Fischer M, Ruegg S, Czaplinski 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. *Crit Care* 2010;14(2):R64.
7. Amiri H, Gholipour C, Mokhtarpour M, Vahdati SS, Aghdam YH, Bakhshayeshi M. Two-day primary trauma care workshop: early and late evaluation of knowledge and practice. *European Journal of Emergency Medicine*. 2013 Apr 1;20(2):130-2.
8. Amiri H, Vahdati SS. Two-Day Primary Trauma Care Workshop-Is it Beneficial?. *Turkish Journal of Emergency Medicine*. 2009;9(1):008-11.
9. Tajoddini S, Vahdati SS. Ultrasonographic diagnosis of abdominal free fluid: accuracy comparison of emergency physicians and radiologists. *European Journal of Trauma and Emergency Surgery*. 2013 Feb 1;39(1):9-13.
10. Haghighi SH, Adimi I, Vahdati SS, Khiavi RS. Ultrasonographic diagnosis of suspected hemopneumothorax in trauma patients. *Trauma monthly*. 2014 Sep;19(4).





11. Ala AR, Pouraghaei M, Vahdati SS, Taghizadieh A, Moharamzadeh P, Arjmandi H. Diagnostic Accuracy of focused assessment with sonography for trauma in the emergency department. *Trauma Monthly*. 2016 Jun;21(4).
12. King RW, Plewa MC, Buderer NM, Knotts FB. Shock index as a marker for significant injury in trauma patients. *Acad Emerg Med* 1996 Nov;3(11):1041-5.
13. Pandit V, Rhee P, Hashmi A, Kulvatunyou N, Tang A, Khalil M, et al. Shock index predicts mortality in geriatric trauma patients: an analysis of the National Trauma Data Bank. *J Trauma Acute Care Surg* 2014 Apr;76(4):1111-5.
14. Czaikowski BL, Liang H, Stewart CT. A pediatric FOUR score coma scale: interrater reliability and predictive validity. *J Neurosci Nurs* 2014 Apr;46(2):79-87.
15. Cohen J. Interrater reliability and predictive validity of the FOUR score coma scale in a pediatric population. *J Neurosci Nurs* 2009 Oct;41(5):261-7.
16. Eken C, Kartal M, Bacanli A, Eray O. Comparison of the Full Outline of Unresponsiveness Score Coma Scale and the Glasgow Coma Scale in an emergency setting population. *Eur J Emerg Med* 2009 Feb;16(1):29-36.
17. Gorji MA, Hoseini SH, Gholipur A, Mohammadpur RA. A comparison of the diagnostic power of the Full Outline of Unresponsiveness scale and the Glasgow coma scale in the discharge outcome prediction of patients with traumatic brain injury admitted to the intensive care unit. *Saudi J Anaesth* 2014 Apr;8(2):193-7.
18. Sadaka F, Patel D, Lakshmanan R. The FOUR score predicts outcome in patients after traumatic brain injury. *Neurocrit Care* 2012 Feb;16(1):95-101.
19. Akavipat P, Sookplung P, Kaewsingha P, Maunsaiyat P. Prediction of discharge outcome with the full outline of unresponsiveness (FOUR) score in neurosurgical patients. *Acta Med Okayama* 2011 Jun;65(3):205-10.
20. Akavipat P. Endorsement of the FOUR score for consciousness assessment in neurosurgical patients. *Neurol Med Chir (Tokyo)* 2009 Dec;49(12):565-71.
21. Czaikowski BL, Liang H, Stewart CT. A pediatric FOUR score coma scale: interrater reliability and predictive validity. *J Neurosci Nurs* 2014 Apr;46(2):79-87.
22. Kochar GS, Gulati S, Lodha R, Pandey R. Full outline of unresponsiveness score versus Glasgow Coma Scale in children with nontraumatic impairment of consciousness. *J Child Neurol* 2014 Oct;29(10):1299-304.
23. Buyukcam F, Kaya U, Karakilic ME, Cavus UY, Turan SF, Odabas O. Predicting the outcome in children with head trauma: comparison of FOUR score and Glasgow Coma Scale. *Ulus Travma Acil Cerrahi Derg* 2012 Nov;18(6):469-73.

