

FOUR score versus GCS in patients with traumatic brain injury in the prehospital setting

Matej Rubelli Furman (✉ matej.rubelli.furman@siol.net)

Zdravstveni dom Koper <https://orcid.org/0000-0003-0413-0606>

Mario Gorenjak

Univerza v Mariboru Medicinska fakulteta

Janez Ravnik

Univerzitetni Klinicni Center Maribor

Research article

Keywords: traumatic brain injury, coma, GCS (Glasgow Coma Scale), FOUR score (Full Outline of UnResponsiveness), mortality

Posted Date: March 19th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-18036/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: In the last few decades, different coma scoring scales have been proposed. The purpose of this study is to compare two coma scales: the GCS (Glasgow Coma Scale) and the FOUR score (Full Outline of UnResponsiveness score), aiming to examine which scale is better in predicting mortality outcome in traumatic brain injury (TBI) patients in the prehospital setting.

Methods: We evaluated the GCS and FOUR score in the prehospital setting at three different prehospital timepoints, and we reassessed the scores in surviving patients 24 hours, one month and three months after the injury. Then, we compared the mortality outcomes. We used the χ^2 method with the best cut-off point for each model; we calculated sensitivity, specificity and correct prediction of the outcome with four severity scores. The Youden index, Z score, McNemar's test and ROC curve were also assessed. $P < 0.05$ was considered statistically significant. Both scales were ranked with gain ratio.

Results: We included 75 TBI patients who were treated in a prehospital setting by a prehospital specialized medical unit. In terms of predictions of mortality, our study showed the following: 1. Twenty-four hours after the injury, the best cut-off points obtained were from the FOUR 2 and FOUR 3 models. 2. One month after the injury, the best cut-off obtained was from the GCS 2 model. 3. Three months after the injury, the best cut-offs obtained were from the GCS 2 and GCS 3 models.

Conclusions: In our study, we were able to demonstrate that the FOUR 2 and FOUR 3 models, compared to the GCS 2 and GCS 3 models, resulted in only a marginally better prediction of mortality outcome 24 hours after TBI. Statistically significant differences were found in neither the Youden index nor the area under the ROC curve after 24 hours and one and three months after the injury.

Background

The assessment of the depth of consciousness in patients with brain injuries is very important and demanding in emergency scenarios outside of the hospital. In the last few decades, several different coma score scales have been proposed. In 1974, a new scale, the GCS (Glasgow Coma Scale), was introduced, and since then, it has been widely used and is still largely used worldwide [1]. The GCS offers a good tool for assessing the depth of consciousness and coma. Many studies have suggested that this scale should be used with other neurological assessments [2–5] and should be administered as soon as possible in emergency situations. The GCS may be repeated at intervals, especially when neurological function fluctuates [6]. The purpose of the GCS is to evaluate patients with TBI and predict their chances of neurological recovery. The GCS presents several weaknesses, such as limited utility in intubated patients as well as inability to estimate brainstem reflexes. Considering these limitations, a new coma score, the Full Outline of Un Responsiveness (FOUR), has been developed [7] to overcome these shortcomings and to provide further neurological details that might lead to a better prediction of outcomes in coma patients.

Emergency departments and intensive care units (ICUs) are ideal places to perform studies in patients after TBI.

The intent of this study was to compare the GCS and FOUR score and to verify their ability to predict mortality outcome in TBI coma patients out of the hospital setting, specifically in the prehospital scenario. Until now, these two scoring scales have not been evaluated and compared in TBI patients in the prehospital setting.

Methods

The present study was approved by the National Medical Ethics Committee of the Republic of Slovenia. We performed a prospective observational cohort study over a one-and-a-half-year period, from March 2012 to September 2013, in a prehospital setting in a large community area. TBI patients were treated by emergency prehospital medical unit personnel (team formed by a physician specialized in emergency medicine, who also performed the GCS and FOUR score assessment, and two highly skilled critical care technicians). The arrival time to the nearest regional hospital was up to 15 minutes. All patients were treated according to the ATLS guidelines.

Our inclusion criteria were minor or moderate-to-severe TBI patients with altered mental status and/or coma due to road traffic or motor vehicle incidents, falls, impacts, assaults and violence, etc. either in polytrauma or isolated head injury situations. We did not include patients under the age of 18, patients who required CPR or patients who died before arriving at the hospital.

Our study aimed to compare the GCS and FOUR score in TBI coma patients in a prehospital setting, evaluating the best prediction ability in terms of mortality outcome. We evaluated the GCS and FOUR score in the prehospital setting at three different time points: immediately at first contact with the patient at the scene, after the management of the patient by the prehospital medical unit, and during the patient handover by ambulance staff at the hospital. We also compared the GCS and FOUR score data in patients who survived 24 hours (A) and one (B) and three months (C) after the injury in regard to mortality outcomes (Fig. 1).

The sensitivity, specificity, correct prediction and Youden index [13] were obtained with a two-by-two table. We calculated sensitivity (true positive) by choosing survivors with equal or more points according to the selected cut-off point. In addition, we calculated specificity (true negative), where we chose non-survivors with fewer points according to the selected cut-off point. The outcome with four severity scores assessed the predictive power of each cut-off point scoring system (the higher the index, the better the prediction of the cut-off point). The percentages of correct predictions of mortality were obtained according to these cut-off points. For each score, receiver operating characteristic (ROC) curves were obtained [14–17]. The greater the area under the ROC curve, the better the scoring system. Data were analysed with IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp. The outcome prediction data were analysed with McNemar's test. The comparisons of the areas under the ROC curves and the analyses of the

differences of the Youden index were performed using the method described by Hanley and McNeil [14–17]. $P < 0.05$ was considered statistically significant.

Results

We included 75 patients with TBI (52 men, 23 women) in our study. Fifty-one of the 75 patients (68%) presented an isolated moderate TBI, 5 of the 75 patients (6.7%) were polytraumatized patients with severe TBI, and 19 of the 75 patients (25.3%) presented with other injuries and minor head injury. The mean age of the patients was 53.6 years.

Within 24 hours after the injury, 3 patients (4%) died; one month after the injury, we registered 12 deceased patients (16%); and three months after the injury, we recorded a cumulative total of 14 deceased patients (18.7% of the total number of studied patients). The average score of the FOUR 1 model (first contact by prehospital unit ambulance with patients at the scene, before medical treatment) was 12.6/16. The GCS 1 model at the same time was 10.9/15. After the first contact at the scene, we assessed the FOUR 1 score and the GCS 1 score: the prehospital medical unit performed head-to-toe immobilization and fixation, they placed an IV line and gave IV therapy, and oxygen was administered. Among all patients, 12 of 75 (16%) were ET intubated (with appropriate sedation, analgesia and muscular relaxation), mechanically ventilated and monitored with end-tidal CO_2 .

The best cut-off points for predicting mortality within 24 hours after the injury were 4 for the FOUR 2 and FOUR 3 models and 5 for the GCS 2 and GCS 3 models. Table 1 provides data regarding sensitivity, specificity and correct predictions of mortality outcome 24 hours after the injury for the best cut-off points for the Youden index. We observed no statistically significant differences between the FOUR score and the GCS for assessment 2 and 3.

Table 1

Sensitivity, specificity, ROC area, and correct prediction of mortality outcome 24 hours after injury for selected cut-off points in the FOUR and GCS models based on the best Youden index.

Scale	Best cut-off point	Sensitivity (%)	Specificity (%)	Correct prediction of outcome (%)	Youden index	ROC area
FOUR 2	4	86	100	87	0.86 ± 0.04	0.88 ± 0.04
FOUR 3	4	86	100	87	0.86 ± 0.04	0.88 ± 0.04
GCS 2	5	79	100	80	0.79 ± 0.05	0.90 ± 0.05
GCS 3	5	79	100	80	0.79 ± 0.05	0.90 ± 0.05
2: After initial management and intervention of the patient; 3: During the patient handover by ambulance staff at the hospital; ROC: Receiver operating characteristic; FOUR: Full Outline of UnResponsiveness; GCS: Glasgow Coma Scale						

The two scores were also compared by drawing ROC curves to avoid fixing arbitrary cut-off points. The comparison of the GCS score and the FOUR score showed only a minimal significant difference in correct prediction of mortality in surviving patients 24 hours after the injury, particularly comparing GCS 1 to FOUR 2 and GCS 1 to FOUR 3. These results confirmed that the FOUR 2 model and the FOUR 3 model showed slightly better predictive power in terms of patient mortality outcome (Table 1). The best cut-offs for the Youden index after 24 hours were 0.86 for the FOUR 2 model and 0.86 for the FOUR 3 model. The area under the ROC curve (area \pm standard error) obtained after 24 hours was 0.884 ± 0.042 for FOUR 2 and the same for FOUR 3. We conclude that no differences in Youden index or area under ROC curve 24 hours after the injury were found.

In addition to 24-hour follow-up outcomes, we also evaluated GCS and FOUR scores in surviving patients one and three months after the injury.

Analysis performed one month after the injury showed that the best cut-off point was 5 for the GCS 2 model, with the highest mortality outcome prediction value among all combined model scales. The cut-off Youden index was 0.90 for the GCS 2 model. The ROC curve was 0.927 ± 0.031 for the GCS 2 model. No differences in Youden index or area under the ROC curve were found.

Data obtained from the analysis performed three months after the injury showed that among all combined model scales, the best models were the GCS 2 (cut-off 7) and GCS 3 (cut-off 9) models. The Youden index values were 0.73 for the GCS 2 model and 0.73 for the GCS 3 model. The ROC was 0.916 ± 0.032 for the GCS 2 model and 0.917 ± 0.032 for the GCS 3 model. No significant differences in Youden index or area under ROC curve were found.

According to McNemar's test, we found a statistically significant difference in correct prediction after 24 hours only between models GCS 1 and FOUR 2 and between models GCS1 and FOUR 3, but no statistically significant differences were found in Youden index or area under ROC curve after 24 hours. Moreover, no statistically significant differences were found one month or three months after the injury.

Discussion

Patients commonly, for one clinical reason or another, have some degree of impaired consciousness. Our efforts are directed to evaluate the depth of consciousness to properly manage these patients. One of our duties is also to predict mortality outcome in patients with impaired consciousness. In our study, we evaluated consciousness and depth of coma in patients with brain injury in the prehospital setting. We focused specifically on comparing two coma scales: the GCS (Glasgow Coma Scale) and the FOUR score (Full Outline of UnResponsiveness).

Since 1974 [1], when the GCS was introduced, the GCS has been widely used in the prehospital setting. In 2005, the FOUR score [7] was proposed to reduce some limitations of GCS assessment. Currently, the FOUR score is mostly employed in intensive and neurological care units.

The advantages of the FOUR score have been assessed by Wijdicks et al. [7], especially in neurologically critically ill patients who are intubated. When we compared the GCS and FOUR scores, we noticed a key difference: the verbal response is not an intrinsic part of the FOUR score; taking this into consideration, the FOUR score is fully useful in intubated patients [7]. Intubation is a common procedure in the field, in the emergency department, in the ICU and in the prehospital setting. The FOUR score tests essential brainstem reflexes and provides information about stages of brainstem injury that is unavailable with the GCS. The FOUR score can distinguish a locked-in syndrome and a possible vegetative state [7] and includes signs suggesting uncal herniation [7]. The evaluation of the respiratory patterns in the FOUR score may also add information about the presence of a respiratory drive [7]. Studies also show that in-hospital mortality between scales was higher for the lowest total FOUR scores when compared with the GCS scores [7]. Conclusions obtained in previous studies have shown that the FOUR score is an accurate predictor of outcome mortality in TBI patients, has some advantages over the GCS [8] and can be performed in a variety of ICU contexts. After all, the FOUR score is easily taught, simple to administer, and provides essential neurologic information that allows for an accurate assessment of patients with altered consciousness with excellent interrater agreement among medical intensivists [9]. The FOUR score might be a better prognostic tool of ICU mortality than the GCS, most likely because it integrates brainstem reflexes and respiration [10]. Other studies have shown the predictive value of the FOUR score on admission of patients after moderate and severe TBI. These studies shown also that the predictive ability for the primary outcome of mortality 2 weeks after injury was no better than that of the GCS score [11]. For nontraumatic comatose patients, different parameters as predictors of mortality in the prehospital environment were also studied [12].

Analyses performed between GCS and FOUR score scales in the hospital environment have demonstrated that GCS was missing key essential elements of a comprehensive neurological examination for comatose patients [18]. In the same study, it was confirmed that the FOUR score maintained simplicity and, at the same time, provided far better information [18]. Other previous studies have demonstrated that GCS and FOUR scores show comparable results in the assessment of patients with traumatic brain injury. These data show that there were excellent statistical correlations between the two scoring systems. Additionally, the FOUR score provides better details regarding the neurological status of the patient. The results can be considered clinically relevant because of the strong statistical association obtained as well as the agreement in the literature [19]. Overall, there are currently multiple scores used to determine the prognosis of patients in intensive care units. However, a scoring system should be simple, reliable, and predictive of morbidity and mortality.

Due to the different categories of scores, the FOUR score is more effective in evaluating patients who are unconscious and dependent on mechanical ventilation. Prospective studies with larger cohorts of patients treated in various intensive care units for longer durations are needed to evaluate whether the application of these scales influences functional and cognitive outcomes [20].

In addition, further comparative neurological outcome studies show also that the mortality of patients admitted to the ICU was significantly higher, when the GCS or the FOUR score was used. Discrimination was fair for both scores, but the FOUR score was superior to the GCS. Calibration was better for the FOUR score than for the GCS in ICU. The sensitivity, specificity, positive predictive value, negative predictive value, and accuracy were also better for the FOUR score than for the GCS score. A good correlation was observed between the two scores [21].

A comprehensive overview of the relationship between a patient's FOUR score and outcome is still lacking. A recent study on the FOUR score overall showed that the FOUR score has a close relationship to in-hospital mortality and poor functional outcome in patients with impaired consciousness. This research also claims that there was insufficient evidence to determine whether performance was modified in different groups, and there was some suggestion that the assessment of brainstem reflexes and respiratory patterns made less contribution than eye and motor scores [22].

In the present study, our data showed only a marginal statistically significant difference in terms of correct prediction of mortality outcome 24 hours after the injury, specifically between the GCS 1 and FOUR 2 model scales ($P = 0.039$) and between the GCS 1 and FOUR 3 model scales ($P = 0.039$). We found no statistically significant differences in the Youden index or area under the ROC curve after 24 hours, no statistically significant differences one month after the injury, and no statistically significant differences three months after the injury.

We think that we must gain a better understanding of the anatomical and pathophysiological pathways that are not evidenced by certain GCS and FOUR score numbers. The next level of research should be focused on the comparison between the obtained GCS and FOUR score data measurements and the anatomical substrate changes revealed by a diagnostic tool such as CT scan of the head and fMRI of the

brain. With these data, we could obtain the accurate sub-anatomical and clinical information needed to perform a specific invasive therapy to lead to a far better outcome for the patient.

Conclusions

The present study involved a comparison between the GCS and FOUR score in TBI patients in out-of-hospital scenarios at different follow-up moments. We introduced and compared different models for the prediction of morality outcome 24 hours after the injury and re-evaluated the predictive ability one and three months after the injury. Our data reveal that the FOUR 2 score and the FOUR 3 score resulted in only a marginal statistically significant prediction of short-term 24-hour mortality compared to the predictive abilities of the GCS 2 and GCS 3 models performed at the same time.

The results of our research confirm that there are no practical or clinical differences between the GCS and FOUR score scales in terms of predicting morality outcome 24 hours, one month, and three months after injury. Due to the different assessment categories, the FOUR score is more effective in evaluating patients who are unconscious and dependent on mechanical ventilation.

We believe that the FOUR score has promising mortality outcome predictive potential and could be regularly performed in the prehospital setting, especially in intubated patients with brain injury. To be routinely utilized, the FOUR score needs a more detailed investigation on a larger group of selected patients. Until then, the GCS will continue to be the cornerstone diagnostic coma scale in everyday clinical practice. We suggest that accurate studies on a wider scale of different types of patients be carried out before we can unanimously conclude that the FOUR score should replace the GCS. Until supporting data are obtained or other prognostic scales that are far better are introduced and validated, we admit that the GCS will remain the most performed coma score tool in the prehospital setting in everyday clinical practice.

Abbreviations

FOUR - Full Outline of Unresponsiveness

GCS - Glasgow Coma Scale

TBI - Traumatic Brain Injury

ICU - Intensive Care Unit

ROC - Receiver Operating Characteristic

IV - Intravenous Lines

ET - Endotracheal Tubes

ATLS - Advanced Trauma Life Support

CT - Computed Tomography

fMRI - functional Magnetic Resonance Imaging

Declarations

Ethics approval and consent to participate

The present study was approved by the National Medical Ethics Committee of the Republic of Slovenia. We recruit only patients who gave signed consent to participate. Informed signed consent was obtained from the patients' or patient's guardian/caregivers.

Consent for publication

Informed signed consent and explanation that data may be published in the study, was obtained.

Availability of data and material

The datasets collected and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

Not applicable

Authors' contributions

MRF contributed to the acquisition and interpretation of the data. MG performed the analysis of the data. JR revised the study. All authors read and approved the final manuscript.

Acknowledgements

Not applicable

References

1. Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. *Lancet* 1974;2(7872): 81-4. [PubMed](#) [TheLancet](#) [ScienceDirect](#)

2. Levy DE, Bates D, Caronna JJ, Cartlidge NE, Knill-Jones RP, Lapinski RH, Singer BH, Shaw DA, Plum F: Prognosis in nontraumatic coma. *Ann Intern Med* 1981; 94:293–301. [PubMed](#) [ACP](#)
3. Jensen M, Neunzig HP, Emskotter T: Outcome prediction in comatose patients: significance of reflex eye movement analysis. *J Neurol Neurosurg Psychiatry* 1987;50:389–92. [PubMed](#)
4. Starmark JE, Holmgren E, Stalhammar D: Current reporting of responsiveness in acute cerebral disorders: a survey of the neurosurgical literature. *J Neurosurg* 1988;69:692–8. [JNS](#) [PubMed](#)
5. Edgren E, Hedstrand U, Nordin M, Rydin E, Ronquist G: Prediction of outcome after cardiac arrest. *Crit Care Med* 1987;15:820–5. [PubMed](#) [CCM](#)
6. Jennett B, Teasdale G, Braakman R, Minderhoud J, Heiden J, Kurze T: Prognosis of patients with severe head injury. *Neuro- surgery* 1979;4:283–9. [PubMed](#)
7. 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. [PubMed](#)
8. 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. doi: 10.1007/s12028-011-9617-5. [PubMed](#)
9. Iyer VN, Mandrekar JN, Danielson RD, Zubkov AY, Elmer JL, Wijdicks EF. Validity of the FOUR Score Coma Scale in the Medical Intensive Care Unit. *Mayo Clin Proc*. 2009;84(8):694–701. [PubMed](#)
10. Wijdicks EF, Kramer AA, Rohs T Jr, Hanna S, Sadaka F, O'Brien J, Bible S, Dickess SM, Foss M. Comparison of the Full Outline of UnResponsiveness score and the Glasgow Coma Scale in predicting mortality in critically ill patients. *Crit Care Med*. 2015;43(2):439-44. doi: 10.1097/CCM.0000000000000707. [PubMed](#)
11. Saika A, Bansal S, Mariamma P, Bhagavatula ID, Dhaval P. Shukla. Prognostic value of FOUR and GCS scores in determining mortality in patients with traumatic brain injury. *Acta Neurochirurgica, The European Journal of Neurosurgery* 2015;157(8):1323–28. [Springer](#) [PubMed](#) [ResearchGate](#)
12. Grmec Š, Gašparovic V. Comparison of APACHE II, MEES and Glasgow Coma Scale in patients with nontraumatic coma for prediction of mortality. *Acute Physiology and Chronic Health Evaluation. Mainz Emergency Evaluation System. Crit Care*. 2001;5(1):19-23. Epub 2000 Dec 14. [PubMed](#) [SemanticScholar](#)
13. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3(1): 32-5. [WileyOnlineLibrary](#)
14. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143(1): 29-36. [PubMed](#)

15. McNeil BJ, Hanley JA. Statistical approaches to the analysis of receiver operating characteristic (ROC) curves. *Med Decis Making* 1984;4(2):137-50. [ResearchGate](#)
16. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983;148(3):839-43. [PubMed](#)
17. Fawcett T. An introduction to ROC analysis. *Pattern Recog Lett* 2006;27:861-74. [AcmDigitalLibrary](#)
18. Murthy T. A new score to validate coma in emergency department-FOUR score. *Ind J Neurotrauma* 2009;6(1):59-62. [ScienceDirect](#)
19. Sharath S. Nair¹, Anilkumar Surendran, Rajmohan B. Prabhakar, Meer M. Chisthi. Comparison between FOUR score and GCS in assessing patients with traumatic head injury: a tertiary centre study. Nair SS et al. *Int Surg J.* 2017 Feb;4(2):656-662. [IJS](#)
20. Yesim Serife BAYRAKTAR, Mert SAHINOGLU, Faruk CICEKCI, Inci KARA, Hakan KARABAGLI, Ates DUMAN, Jale Bengi CELIK. Comparison of Glasgow Coma Scale and Full Outline of Unresponsiveness (Four) Score: A Prospective Study. *Turkish Neurosurgery* 2019, 29 (2): 285-288 [Read](#)
21. Kishor Khanal, Sanjeeb Sudarshan Bhandari, Ninadini Shrestha, Subhash Prasad Acharya, and Moda Nath Marhatta. Comparison of outcome predictions by the Glasgow coma scale and the Full Outline of UnResponsiveness score in the neurological and neurosurgical patients in the Intensive Care Unit. *Indian J Crit Care Med.* 2016 Aug; 20(8): 473–476. [Researchgate](#)
22. Ching C. Foo, James J.M. Loan, and Paul M. Brennan. The Relationship of the FOUR Score to Patient Outcome: A Systematic Review. *JOURNAL OF NEUROTRAUMA* 36:2469–2483 (September 1, 2019) Mary Ann Liebert, Inc. DOI: 10.1089/neu.2018.6243 [Mary Ann Liebert, Inc.](#)

Figures

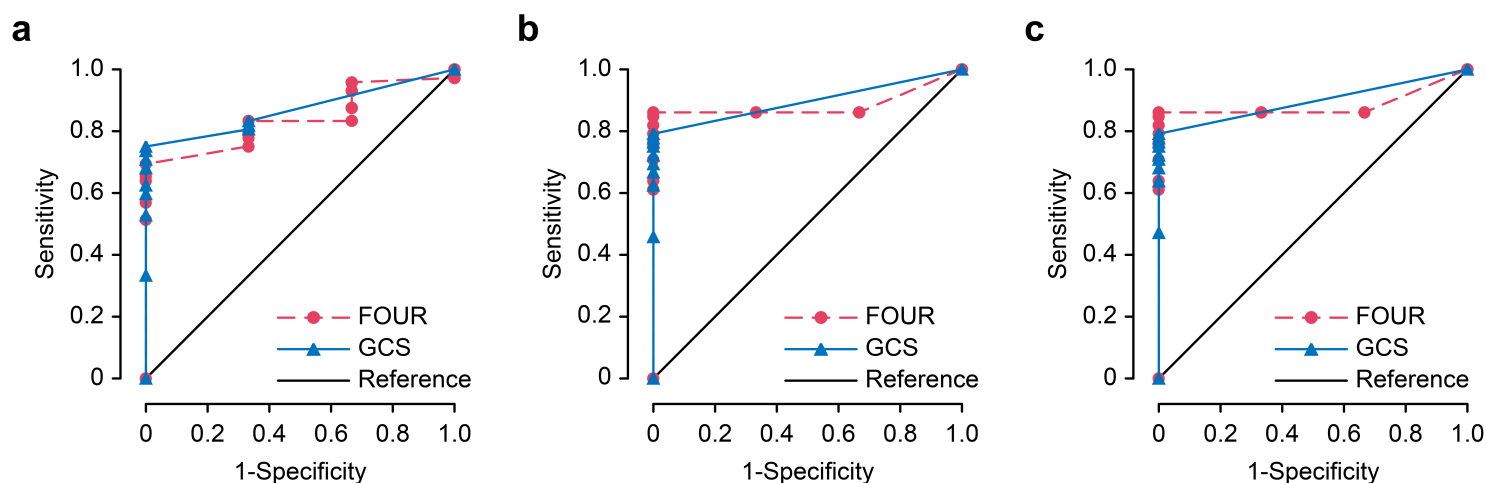


Figure 1

Comparative GCS and FOUR score mortality outcome assessment in patients who survived 24 hours (A) and one month (B) and three months (C) after the injury.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [MedicalethicsconsentoriginalSlovenian.pdf](#)
- [STROBESstatementnew.docx](#)
- [MedicalethicsconsenttranslatedinEnglish.pdf](#)