STROBE Statement

|  |  |  |
| --- | --- | --- |
|  | Item No | Recommendation |
| **Title and abstract** | 1 | FOUR score versus GCS in patients with traumatic brain injury in the prehospital setting |
| 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 outcomes. In our study, the results of our research confirm that there are no practical or clinical differences between the GCS and FOUR scores in terms of predicting morality outcomes 24 hours, one month, and three months after injury. No statistically significant differences were found in the Youden index or the area under the ROC curve 24 hours, one month or three months after the injury. |
| Introduction |
| Background/rationale | 2 | The intent of this study was to compare GCS and FOUR scores and to verify their predicting power outcome in TBI coma patients out of hospital setting. |
| Objectives | 3 | GCS predict some chances of neurological recovery, but presents several weaknesses, such as limited utility in intubated patients as well as inability to estimate brainstem reflexes. Considering these limitations, FOUR score has been developed to overcome the particular shortcomings and to provide further neurological details that might lead to a better prediction of outcomes in coma patients, also in pre-hospital setting. |
| Methods |
| Study design | 4 | We performed a prospective observational cohort study. |
| Setting | 5 | TBI patients were treated by emergency prehospital medical unit personnel (team formed by a physician specialized in emergency medicine, 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. |
| Participants | 6 | 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. 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 and one and three months after the injury in regard to outcomes. |
| We included 200 patients with TBI in our study (133 men and 67 women). Within 24 hours after the injury, 6 patients (3%) died. One month after the injury, we registered 23 deceased patients (11,5%), and three months after the injury, we recorded a cumulative total of 25 deceased patients (12.5% of the total number of studied patients). Among all patients, 33 out of 200 (16,5%) were ET intubated (with appropriate sedation, analgesia and muscular relaxation), mechanically ventilated and monitored with end-tidal CO2. No missing data were present in the final dataset. |
| Variables | 7 | The sensitivity, specificity, correct prediction and Youden’s J-statistic was used to assess the performance of a dichotomous diagnostic test. Youden index were obtained with two-by-two table. For each score, ROC curves were obtained. The outcome prediction data were analysed with McNemar’s test. P < 0.05 was considered statistically significant. |
| Data sources/ measurement | 8 | 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.  |
| Bias | 9 | All physicians that performed FOUR score in our study, were properly instructed.  |
| Study size | 10 | Study size was obtained with power analysis. |
| Quantitative variables | 11 | 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 accessed 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 outcome were obtained according to these cut-off points. The greater the area under the ROC curve, the better the scoring system. |
| Statistical methods | 12 | We used the χ² 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.The sensitivity, specificity, correct prediction and Youden index were obtained with two-by-two table. We calculated sensitivity (true positive) by choosing survivors with equal or more points according to the selected cut-off point. We also calculated specificity (true negative), where we choose non-survivors with fewer points according to the selected cut-off point. The outcome with four severity scores accessed 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 outcome were obtained according to these cut-off points. The greater the area under the ROC curve, the better the scoring system.The outcome prediction data were analysed with McNemar’s test. P < 0.05 was considered statistically significant. Data were analysed with IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp. |
| Results |
| Participants | 13 | We included 200 patients with TBI in our study (133 men and 67 women). Within 24 hours after the injury, 6 patients (3%) died. One month after the injury, we registered 23 deceased patients (11,5%), and three months after the injury, we recorded a cumulative total of 25 deceased patients (12.5% of the total number of studied patients). No missing data were present in the final dataset. |
| Descriptive data | 14 | 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.We used the χ² method with the best cut-off point for each model in three different moments at the scene, than we reassessed the scores scales in survived patient 24 hours, one and three months after the injury, and later we compared the outcomes.  |
| Outcome data | 15 | In terms of the predictions of positive outcomes, our study showed the following:1. Twenty-four hours after the injury, the best cut-off points obtained were from the FOUR 1 and GCS 1 models.2. One month after the injury, the best cut-off points obtained were from the following models: FOUR 2 and GCS 3 as well as FOUR 3 and GCS 3.3. Three months after the injury, the best cut-off points obtained were from the following models: FOUR 2 and GCS 1, FOUR 3 and GCS 1, and GCS 1 and GCS 2. |
| Main results | 16 | The best cut-off points for predicting outcomes within 24 hours after the injury were 12 for the FOUR 1 model and 8 for the GCS 1 model, evaluated immediately upon first contact at the scene (1: first assessment). We observed no statistically significant differences within 24 hours after the injury between the FOUR score and the GCS for assessments 2 and 3 (2: after initial management and intervention of the patient, 3: during patient handover by the ambulance staff at the hospital).These results, according to McNemar's test, confirmed that the FOUR 1 model and the GCS 1 model showed slightly better predictive power in terms of patient outcome. The best cut-off values for the Youden index after 24 hours were 0.85 for the FOUR 1 model and 0.88 for the GCS 1 model. The area under the ROC curve (area ± standard error) obtained after 24 hours was 0.94 ± 0.02 for FOUR 1 and was almost the same at 0.93 ± 0.02 for GCS 1. We conclude that no differences in the Youden index or area under the ROC curve 24 hours after the injury were found.Analysis performed one month after the injury showed that the best cut-off point was 8 for the FOUR 2 and FOUR 3 models and 11 for the GCS 3 model. According to McNemar's test, these three models had the highest outcome prediction value among all combined models. The cut-off value for the Youden index was the same at 0.81 for the FOUR 2, FOUR 3 and GCS 3 models. The ROC curve was the same at 0.91 ± 0.04 for both the FOUR 2 and FOUR 3 models and was 0.95 ± 0.05 for the GCS 3 model. No differences in the 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 models and according to McNemar's test, the best models were the FOUR 2 and FOUR 3 (both cut-off values of 12) as well as GCS 1 (cut-off value of 12) and GCS 2 (cut-off value of 9) models. The Youden index values were 0.77 for the FOUR 2, FOUR 3 and CGS 2 models and 0.78 for the GCS 1 model. The ROC was 0.89 ± 0.04 for the FOUR 2 and FOUR 3 models, 0.93 ± 0.02 for the GCS 1 model and 0.95 ± 0.02 for the GCS 2 model. No significant differences in the Youden index or area under the ROC curve were found. |
| Other analyses | 17 | In addition to the 24-hour follow-up outcomes, we also evaluated the GCS and FOUR scores in surviving patients one and three months after the injury.The two scores were also compared by drawing ROC curves to avoid bias of arbitrary cut-off points. The comparison of the GCS score and the FOUR score in patients who survived for 24 hours and for one and three months after the injury in regard to outcomes showed only a minimal significant difference in the correct prediction of outcomes in surviving patients 24 hours after the injury.  |
| Discussion |
| Key results | 18 | The results of our research confirm that there are no significant statistical differences between the GCS and FOUR scores in terms of predicting morality outcomes 24 hours, one month, and three months after injury. |
| Limitations | 19 | Our power analyses showed that our main limitations to the study is related to a limited time necessary to achieve statistically sufficient number of cases. Data collected on a larger metropolitan area, or study performed simultaneously in different urban area, should be advisable.  |
| Interpretation | 20 | 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. |
| Generalisability | 21 | We believe that the FOUR score has promising predictive outcome potential and could be regularly performed in the prehospital setting, especially in intubated patients with brain injuries. |
| Other information |
| Funding | 22 | The intent of this study was to compare the GCS and FOUR score and to verify their ability to predict outcome in TBI coma patients out of the hospital setting, specifically in the prehospital scenario. Until now, these two scoring scales have not been compared in TBI patients in the prehospital setting. |