Validity and Predictive Value of the FOUR Score Coma Scale in Stroke Patients: A Systematic Review

Alireza Ala ¹, Samad Shams Vahdati ², Nafiseh Vahed ³, Sahar Mohammadi ⁴, Hoorolnesa Ameli ⁴*  
¹Emergency Medicine Research Team, Tabriz University of Medical Sciences, Tabriz, Iran.  
²Research Center for Evidence Based Medicine, Iranian EBM Centre: A Joanna Briggs Institute Affiliated Center, Tabriz University of Medical Sciences, Tabriz, Iran.  
³Department of Emergency Medicine, Zanjan University of Medical Sciences, Zanjan, Iran.  
⁴Clinical Research Development Center, Shahid Beheshti Hospital, Qom University of Medical Sciences, Qom, Iran.

Abstract

Objective: Stroke is known as the second common cause of mortality and the third most common cause of morbidity all over the world. The aim of this study was to evaluate and systematic review the validity and predictive value of the Full Outline of UnResponsiveness (FOUR) Score Coma Scale in stroke patients.

Materials and Methods: In this systematic review, literature from Medline (PubMed, Ovid), The Cochrane Library, Scopus, Embase, ProQuest, published between 1988 and 2020 were included. The search strategy included a combination of free and MeSH terms such as “FOUR Score”, “Full Outline of UnResponsiveness”, Stroke, and “cerebrovascular accident”.

Results: The total number of patients included was 5515, with sample sizes ranging from 53 to 1645 patients. FOUR score was a reliable scale for estimating prognosis in stroke patients. The predictive value of FOUR score in the outcomes of traumatic patients was better than the Glasgow coma score. However in some studies, FOUR Score was not superior to the Glasgow coma score.

Conclusion: In conclusion, FOUR score coma scale with an appropriate validity and predictive value can be proposed as a valid substitute for Glasgow coma score.

Keywords: Stroke; FOUR Score; Validity; Predictive Value.
consciousness deteriorates [6]. The validity of FOUR score has been already evaluated in many populations such as children, patients in intensive care unit and neurosurgical patients [7-9]. So in order to evaluate prognosis in stroke patients, in the present study, it was aimed to estimate validity and predictive value of FOUR score in the emergency department.

Materials and Methods

Search Strategy

In this systematic review, the literature from Medline (PubMed, Ovid), The Cochrane Library, Scopus, Embase, ProQuest, Web of Science, and Google Scholar published between 1988 and 2020 were searched. The search strategy included a combination of free and MeSH terms such as “FOUR Score”, “Full Outline of UnResponsiveness”, Stroke, and “cerebrovascular accident” in all fields. In this systematic review, PRISMA guidelines were followed.

Inclusion and Exclusion Criteria

The publication types of review and articles were included, but meetings abstract, editorials, brief commentaries, letters, errata and subtopics were excluded. Patients with non-relevant topics and studies of participants were assessed by the pediatric FOUR Score were excluded in this systematic review.

Study Selection

The selected papers extracted from the databases were assessed by two investigators using with JBI checklist for studies. Discrepancies between the two raters were referred to the third investigator. First, the titles of all articles were reviewed to screen for eligibility and those found to be inconsistent with the objectives of the study were excluded from the survey. In the later stages, the abstracts and full text articles were, respectively, examined to identify and exclude those that did not satisfy the inclusion, or had a weak correlation with the objectives of the study. The primary search yielded a total of 225 publications, after exclusion of duplicated, 120 articles were not applicable to our aim were excluded by title and abstract screening. The full-texts of remaining articles were reviewed comprehensively. During the review, 15 were excluded because they were commentary, non-randomized, or review rather than an original article. Also, 3 studies assessed FOUR score in children and were excluded. Finally, 28 articles were eligible for the present systematic review (Figure 1).

One review author screened titles and abstracts obtained from database searches. Two of the authors independently reviewed the full texts articles for relevant abstracts against inclusion criteria. Finally, disagreements resolved in inclusion by discussion.

![Figure 1. Search and selection process of systematic review.](image-url)
Data extraction and Statistical Analysis

One reviewer extracted the data from the included studies while a second author checked the results. Any disagreements were resolved by discussion. Data for the primary objective of the review was collected from the full text of each publication and included the trial name, year of publication, type of study, sample size, mean age and results.

For the data synthesis studies were grouped according to the predictive value of FOUR Score and Glasgow Coma Scale (GCS).

Results

Study characteristics

The main features of the 25 articles were shown in table 1. The studies were published between 2005 and 2019 with the designs of cohort, observational and cross-sectional. The total number of patients included was 5515, with sample sizes ranging from 53 to 1645 patients. Eighteen studies compared the results of FOUR score with Glasgow Coma Scale (GCS). Reliability of FOUR score in Italian, French and Chinese versions was assessed in three studies.

The predictive value of FOUR score was categorized based on the patient’s conditions that were mentioned in the studies. There are six categories of neurological and neurosurgical patients, critically ill patients, traumatic and stroke patients, patients in emergency departments and other conditions. FOUR score is reliable and provides more information in neurological and neurosurgical patients compared to GCS [9-12]. Also it is powerful for the estimation of mortality risk in these patients [9, 12].

In critically ill patients in ICU, FOUR score performed well as a prognostic tool for brain death [7], mortality [13] and level of consciousness [14]. On the other hand, results of an article showed that the FOUR score performed better than the GCS for exact inter-rater agreement, but not for the clinically more relevant agreement within the range of ± 1 score point [15].

The predictive value of FOUR score in the outcomes of traumatic patients was better than the GCS [16-20]. However, in another study, FOUR Score was not superior to the GCS.

Although the combination of the eye and motor components of FOUR Score is a valuable tool that can be used instead of either the FOUR Score or GCS [21]. The predictive value of FOUR score and GCS in outcome of multiple trauma patients admitted to the emergency department was similar [22].

FOUR score was a reliable scale for estimating prognosis in stroke patients [23, 24], and in emergency departments [25, 26]. In the two out of three remaining included studies with other patient conditions, the predictive value of FOUR score was similar to GCS [27, 28], whereas in one study it was better than GCS [6].

Furthermore, the reliability of French, Italian and Chinese versions of FOUR score to assess consciousness of patients has been proved in three studies [29-31].

Discussion

Loss of consciousness is defined as the suspension of patient’s conscious relationships with the outside world. Physical examination is a prominent step in process of approaching patients with loss of consciousness. As these patients are unable to cooperate for a physical examination with physicians, so physical examinations are mostly single-handedly operated by the examiner. Glasgow coma score is the most common method used to assess the level of consciousness designed for traumatized patients, although its usage has extended too many other medical conditions. So, as far as Glasgow coma score is not exclusively designed for other conditions with loss of consciousness, many other methods have been proposed as a substitution. FOUR Score coma scale is a method of examining the level of consciousness composed of eye responses, motor responses, brainstem reflexes, and breathing pattern. Many studies have been conducted in order to validate FOUR score coma scale in comparison to other methods such as Glasgow coma score. In presents study it was shown that FOUR score coma scale presents an acceptable level of validity with similar interobserver and intraobserver variety as Glasgow coma score did, also FOUR score coma score had superior predictive value than Glasgow coma scale.

In a study by Cohen investigating compare the interobserver reliability and predictive validity of the FOUR score and the Glasgow coma score in pediatric patients, it was concluded that the interobserver reliability for the Glasgow coma score was good and that for the FOUR score was excellent, also predictive value for the FOUR score and the Glasgow coma score were both able to predict poor outcome at the end of hospitalization and in-hospital morbidity [8], the results of this study were similar to the present study.

To investigate the interobserver validity of FOUR score in different sections some studies have been conducted. In a study by Van der Berge et al. about Interobserver agreement in assessment of ocular signs in coma, the interobserver agreement the parameter for equality of pupils, reaction of pupils, spontaneous eye movements, and oculocephalic responses was assessed and it was concluded that the agreement in the assessment of the pupils to light and equality of pupils were satisfactory, but in assessing spontaneous eye movements and oculocephalic responses the disagreement was more [32]; as far as the first step in FOUR score coma scale is assessing eye movements in response to command or stimulus and, as our study presents, FOUR score coma scale has an acceptable interobserver agreement rate.

In another study by Born et al. studying to what extent agreement exists among different raters assessing brain stem reflexes and to compare the results with those observed from motor responses, it was concluded that good agreement achieved by different examiners in the evaluation of brain stem reflexes. Also, Brain stem reflexes offered a slightly higher agreement than that of the study of motor response [33]; considering motor and brain-stem reflexes examination as a crucial part of FOUR score coma scale, this study is similar to presents study, as it concluded that an appropriate rate of agreement existed in assessing motor and brain-stem functions. Almojuela studied that the FOUR score has been shown to be a useful outcome predictor in many patients with depressed level of consciousness and it displays good inter-rater reliability among physicians and nurses [34].
### Table 1. Studies included in study and their main characteristics

<table>
<thead>
<tr>
<th>Study, year</th>
<th>Type of study</th>
<th>Sample size</th>
<th>Age</th>
<th>Patients population</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Akavipat 2009 [9]</td>
<td>Observational</td>
<td>100</td>
<td>Mean age: 53.7±15.93</td>
<td>Neurosurgical</td>
<td>FOUR score is reliable for consciousness evaluation and powerful for the estimation of mortality risk.</td>
</tr>
<tr>
<td>Baratloo 2016 [22]</td>
<td>Cross-sectional</td>
<td>89</td>
<td>Mean age: 31.9±19.9</td>
<td>Multiple trauma</td>
<td>The predictive value of FOUR score and GCS in predicting the outcome of multiple trauma patients admitted to the emergency department was similar.</td>
</tr>
<tr>
<td>Bruno 2011 [16]</td>
<td>Cohort</td>
<td>176</td>
<td>Mean age: 63±15</td>
<td>Severely brain damaged</td>
<td>The FOUR score may offer the additional advantage to be performable in intubated patients and to identify nonverbal signs of consciousness by assessing visual pursuit, in compared to GCS and GLS.</td>
</tr>
<tr>
<td>Chen 2013 [12]</td>
<td>Cohort</td>
<td>101</td>
<td>Mean age: 64</td>
<td>Neurosurgical patients with severely impaired consciousness</td>
<td>The FOUR score was more robust than the GCS in predicting mortality after 30 days in neurosurgical patients with severely impaired consciousness.</td>
</tr>
<tr>
<td>Eken 2009 [21]</td>
<td>Observational</td>
<td>185</td>
<td>Median age: 59</td>
<td>Head trauma or neurological complaints</td>
<td>FOUR Score, is not superior to Glasgow Coma Scale.</td>
</tr>
<tr>
<td>Fischer 2010 [15]</td>
<td>Observational</td>
<td>267</td>
<td>Mean age: 62±17</td>
<td>Unselected critically ill patients</td>
<td>The small advantage in inter-rater reliability of the FOUR score is most likely insufficient to replace the GCS.</td>
</tr>
<tr>
<td>Gorji 2014 [19]</td>
<td>Cohort</td>
<td>53</td>
<td>Mean age: 33.8±12.60</td>
<td>Traumatic brain injury</td>
<td>The FOUR score is an accurate predictor of discharge outcome in TBI patients.</td>
</tr>
<tr>
<td>Gorji 2015 [20]</td>
<td>Cohort</td>
<td>80</td>
<td>Mean age: 33.8±12.60</td>
<td>Traumatic brain injury (intubated)</td>
<td>FOUR score is more exact and more practical in intubated patients regarding lack of verbal response factor in early mortality prediction in GCS.</td>
</tr>
<tr>
<td>Gujjar 2013 [28]</td>
<td>Observational</td>
<td>100</td>
<td>Mean age: 62±17</td>
<td>Patients with altered mental status due to medical conditions</td>
<td>FOUR score performs comparably with GCS but better than SOFA score in predicting outcomes at 3 months. FS could reliably be used in the medical wards for evaluation of mental status.</td>
</tr>
<tr>
<td>Idrovo 2009 [23]</td>
<td>Cohort</td>
<td>60</td>
<td>Median age: 75</td>
<td>Acute stroke</td>
<td>The FOUR score is a reliable scale for evaluating the level of consciousness in acute stroke patients, showing a good correlation with the GCS and the NIH stroke scale.</td>
</tr>
<tr>
<td>Iyer 2009 [7]</td>
<td>Observational</td>
<td>100</td>
<td>Mean age: 63.0±18.4</td>
<td>Critically ill patients</td>
<td>The FOUR score is a good predictor of the occurrence of brain death.</td>
</tr>
<tr>
<td>Johnson 2013 [17]</td>
<td>Observational</td>
<td>57</td>
<td>&gt;18 years old</td>
<td>Patients in neuro/trauma ICU</td>
<td>The FOUR Score scale provides a reliable neurological assessment of intubated patients where the GCS does not differentiate patient status once intubated.</td>
</tr>
<tr>
<td>Kasprowicz 2016 [18]</td>
<td>Cohort</td>
<td>162</td>
<td>Median age: 52</td>
<td>Traumatic brain injury</td>
<td>Incorporating the FOUR score at admission enables further differentiation of in-ICU mortality prediction for patients with low GCS.</td>
</tr>
<tr>
<td>Kevric 2010 [26]</td>
<td>Observational</td>
<td>203</td>
<td>Mean age: 54.9</td>
<td>Emergency department patients</td>
<td>The FOUR Score showed greater reliability than the GCS in ED patients using ED clinicians as raters.</td>
</tr>
<tr>
<td>Khanal 2016 [10]</td>
<td>Cohort</td>
<td>97</td>
<td>&gt;16</td>
<td>Neurological and neurosurgical patients</td>
<td>FOUR score was better than GCS.</td>
</tr>
<tr>
<td>Kocak 2012 [24]</td>
<td>Cohort</td>
<td>100</td>
<td>Mean age: 70.49±12.42</td>
<td>Stroke patients</td>
<td>FOUR score is a beneficial and reliable scale for estimating prognosis in stroke patients in intensive care units.</td>
</tr>
<tr>
<td>Kramer 2012 [14]</td>
<td>Observational</td>
<td>907</td>
<td>Median age: 60</td>
<td>Critically ill patients</td>
<td>The FOUR score can measure the level of consciousness.</td>
</tr>
<tr>
<td>Marcati 2011 [30]</td>
<td>-</td>
<td>87</td>
<td>Mean age: 70.2±13.9</td>
<td>-</td>
<td>The Italian version of the FOUR score can be used to reliably assess patients with impaired consciousness.</td>
</tr>
<tr>
<td>Peng 2015 [31]</td>
<td>-</td>
<td>120</td>
<td>Mean age: 47.9±14.8</td>
<td>-</td>
<td>The Chinese version of the FOUR score can be used to reliably assess patients with impaired consciousness in acute brain injury.</td>
</tr>
<tr>
<td>Stead 2009 [25]</td>
<td>Cohort</td>
<td>69</td>
<td>Emergency department patients</td>
<td>-</td>
<td>Both FOUR score and GCS performed equally well, but the neurologic detail incorporated in the FOUR score makes it more useful in management and triage of patients.</td>
</tr>
<tr>
<td>Weiss 2009 [29]</td>
<td>-</td>
<td>178</td>
<td>Mean age: 62</td>
<td>-</td>
<td>The French version of the FOUR score has an excellent interobserver reliability.</td>
</tr>
<tr>
<td>Wijdicks 2005 [6]</td>
<td>-</td>
<td>120</td>
<td>Mean age: 58.9</td>
<td>-</td>
<td>The FOUR score provides greater neurological detail than the GCS.</td>
</tr>
</tbody>
</table>
The FOUR score is useful than the Glasgow coma score for patients who have undergone a serious neurologic event as a complication of surgery or medical illness. Also, the Glasgow coma score performs poorly in assessing patients with less severe degrees of loss of consciousness, such as those seen in the medical ICU. Use of mild sedation in the medical and surgical ICU affects eye opening and motor response but not brainstem reflexes and respiration. In the other hand in patients with aphasia, calculation of GCS will be incorrect in evaluation of level of consciousness; so, all 3 components of the Glasgow coma score are affected by sedation but not in FOUR score coma scale.

**Conclusion**

In conclusion, the present study in a systematic review concludes that FOUR score coma scale with an appropriate validity and predictive value can be proposed as a valid substitute for Glasgow coma score in the stroke patients and patients with loss of consciousness. So, further national and international guidelines regarding approaching patients with stroke and also, loss of neurological outcome in acute stroke patients can be modified using results of the present study. Further studies are encouraged to be performed to estimate validity and predictive value of FOUR score coma scale among different population of patients.

**References**

10. Khanal K, Bhandari S, Shrestha N, Acharya S, Marhatta M. 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 Journal of...


45. Gorji MAH, Gorji AMH, Hosseini SH. Which score should be used in intubated patients’ Glasgow Coma Scale or full outline of unresponsiveness? International Journal of Applied and Basic Medical Research. 2015; 5(2):92.


49. Kocak Y, Ozturk S, Ege F, Ekmecki A. A useful new coma scale in