Translation and Validation of the FOUR Scale for Children and its Use as Outcome Predictor: A Pilot Study

Sofia Simões FERREIRA, Daniel MEIRELES, Alexandra PINTO, Francisco ABECASIS
Acta Med Port 2017 Sep;30(9):599-607 • https://doi.org/10.20344/amp.8052

ABSTRACT

Introduction: The Full Outline of UnResponsiveness - FOUR scale has been previously validated to assess impaired consciousness in the adult population. The aim of this study is the translation into Portuguese and validation of the FOUR scale in the pediatric population. The study also compares the FOUR scale and Glasgow coma scale score ratings and the clinical outcome of patients hospitalized in Pediatric Intensive Care Units.

Material and Methods: This study prospectively rated patients admitted to the Pediatric Intensive Care Units with impaired consciousness during one year. Both scales were applied daily to patients by three types of examiners: intensivists, residents and nurses, from the moment of admission until clinical discharge. Neurological sequelae was evaluated using the King's Outcome Scale for Childhood Head Injury - KOSCHI.

Results: Twenty seven patients between one and 17 years of age were included. Both scales are reliable and inter-rater reliability was greater for the FOUR score. Glasgow coma scale showed a minimum score in eight evaluations, whereas the FOUR scale obtained the minimum score in only two of these evaluations. In both scales there was a strong association between the admission score and the patient's outcome (area under curve FOUR = 0.939, versus Glasgow coma scale = 0.925).

Discussion: The FOUR scale provides more neurological information than Glasgow coma scale in patients with impaired consciousness and has prognostic interest.

Conclusion: The FOUR scale can be applied in patients admitted with impaired consciousness in Pediatric Intensive Care Units. We think that a multicenter study would be very beneficial for confirming and generalizing these results.

Keywords: Child; Coma/diagnosis; Consciousness Disorders; Glasgow Coma Scale; Intensive Care Units, Pediatrics Translating

INTRODUCTION

Glasgow coma scale (GCS) was first published in The Lancet journal in 1974 by Teasdale and Jennett and has been developed as an assessment tool of the depth and duration of impaired consciousness and coma.1 It was initially designed as a research tool for patients having suffered a severe head injury, aimed at improving the communication among healthcare professionals.1 It has been gradually approved worldwide as an important tool for the assessment of patients with head injury and became a relevant part of the advanced life support courses issued by the American College of Surgeons.2 The use of GCS has been aimed at describing the level of consciousness, at the...
comparison between treatment approaches and outcome prediction. The need for endotracheal intubation for airway protection in a patient with a GCS score of eight or less is a clear example of its use as an indication for the management of neurological pathology. Its widespread popularity is due to the fact that it is easy to apply, showing good inter-rater reliability and has been proven crucial for the assessment of cases with impaired consciousness.

Nevertheless, different limitations and disadvantages are worth mentioning.

As verbal response cannot be assessed in intubated patients, potential inconsistencies regarding the score assigned in this category may exist, i.e. healthcare professionals can either assign the lowest possible score or extrapolate the score in this category depending on the presence of other neurological signs.

No relevant indicators for the adequate assessment of coma severity are included in the GCS, i.e. only the cortical function is assessed and no parameters regarding the assessment of brainstem functions are included, such as breathing patterns, pupil size and pupillary reflexes, as well as eye movements; only the last two are assessed by the GCS and depth of coma may not be accurately detected.

This scale tends to be better applied by experienced professionals and the interpretation of intermediate values is challenging even for intensive care physicians. Considering its limitations, the need for the development of new scales for the assessment of the level of consciousness has emerged, allowing for a better detailed assessment of the neurological status, with similar ease of use and predictive value.

The FOUR (acronym for Full Outline of UnResponsiveness) score coma scale has one of the best performances. Initially developed at the Mayo Clinic College of Medicine, it has been widespread and is currently being used worldwide in different neurotrauma centres and intensive care units (ICU) in the adult population.

The FOUR scale (Table 1) consists of four components: eye response, motor response, brainstem reflexes and breathing patterns. Each component is graded on a scale from 0 (worst) to 4 (best response).

The acronym for this scale – FOUR – relates to the four items and a maximum score of four in each item.

The eye response component allowed for the differentiation between a vegetative state (eyes open but do not track) and a locked-in syndrome (eyes open, blink and vertically track on command). This is not possible with the GCS, as only eye opening is assessed with the scale.

Three different movements of the hand are included into the motor response component (thumbs up, fist or peace...
The two following components were developed to replace GCS verbal response, which is not applicable to intubated patients and has to be modified for patients having not yet acquired reliable language skills.

Three brainstem reflexes (pupillary, cornea and cough reflex) are tested with the third component, allowing for testing of mesencephalon, pons and medulla oblongata. The presence of an acute third nerve injury, a sign of brainstem compression, is tested with the FOUR scale by the presence of unilateral dilated pupil, easily detected through pupillary observation and in need for an immediate action. The cough reflex is usually absent when pupillary and cornea reflexes are also absent.5

Finally, the breathing pattern allows for the assessment of the indication for ventilation. Cheyne-Stokes respiration and irregular respiration are signs of bi-hemispheric or brainstem dysfunction. In already intubated patients, overbreathing the ventilator corresponds to functioning respiratory centres.5

With all categories graded the minimum score (0), brain death evaluation should be considered.5

According with the authors, the FOUR scale is easy to use. It is specific for the detection of certain states of impaired consciousness, allowing differentiation of a locked-in syndrome from a vegetative state and providing for important details regarding the brainstem function, which are crucial in patients with impaired consciousness.5

This prospective study aimed at the translation, validity and clinical application of the FOUR scale to the patients admitted to a paediatric intensive care unit (PICU), as well as the comparison between the results obtained with both scales and its relationship to mortality and morbidity prediction at hospital discharge.

**MATERIAL AND METHODS**

This was a prospective and observational study of children admitted to a PICU from May 2012 to May 2013 with impaired consciousness (GCS <15) at admission; the patients presenting with chronic disorders of consciousness and those under the influence of neuromuscular blocking agents were excluded from the study.

Nurses, residents and intensivists working at a PICU were involved in the study. Before the study started, a presentation session of the FOUR scale, during which a document with the scale, instruction manual and a record sheet for the use of both scales has been handed to study raters. An information session has been subsequently carried out, as well as a second training session aimed at the application of the FOUR scale.

In order to get as closest as possible to daily clinical practice, each patient was daily assessed by the nurse, the intensivist and the resident in charge of the patient, regardless of the experience of each rater. Depending on the clinical status, some patients were more than once a day assessed by each rater. The assessments were obtained with a maximum one-hour interval between raters, in order to insure that these were unaware of the results of other raters and reducing the chance of a change in patient’s clinical status. The minimum score in GCS verbal response has been assigned to intubated patients. The maximum score or the one immediately below in the motor component with both scales was at rater’s discretion to assign to patients under the age of two presenting with normal spontaneous movements.

The neurological status of the patients at discharge from hospital has been documented over the data collection period, for subsequent analysis of the relationship between the assessments obtained with both scales as well as for

### Table 2 - KOSCHI (King's Outcome Scale for Childhood Head Injury) scale

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category 1</td>
<td>Morte</td>
</tr>
<tr>
<td>Category 2</td>
<td>Estado vegetativo</td>
</tr>
<tr>
<td>Category 3: Incapacidade grave</td>
<td>a) A criança é capaz de, pelo menos, movimentar de um modo intermitente parte do corpo/ olhos espontaneamente ou a pedido</td>
</tr>
<tr>
<td>Category 4: Incapacidade moderada</td>
<td>b) Implica grande nível de dependência, mas a criança é capaz de realizar algumas tarefas diárias</td>
</tr>
<tr>
<td>Category 5: Boa recuperação</td>
<td>a) Nova condição patológica que não interfere com o funcionamento global da criança (ex. cefaleias)</td>
</tr>
<tr>
<td></td>
<td>b) Recuperação total, sem sequelas detetáveis</td>
</tr>
</tbody>
</table>
The FOUR scale was translated by a translation service company and was subsequently reviewed by two paediatricians.

Statistical analysis

The values of Cohen’s kappa (κ) coefficient were obtained for both scales, in order to assess the agreement between selected staff pairs: 1) within each pair of raters, i.e. intensivist/resident (I/R), intensivist/nurse (I/N), resident/nurse (R/N); 2) regarding each component of the scales; 3) regarding the total score of each assessment. This coefficient allows for the assessment of interrater agreement and an excellent agreement has been considered for values >0.8.

Total FOUR scale score was graded as Severe (total score ≤ 7), Moderate (8-14) and Mild (15-16) in order to obtain kappa coefficients.

Total score with the GCS (n = 82) was also ranked as Severe (total score ≤ 8), Moderate (9-12) and Mild (13-14).

Cronbach’s alpha coefficient was used for the assessment of the internal consistency of both scales and a reliable scale has been considered for alpha values over 0.8.

Finally, Spearman’s correlation coefficient has been used for the assessment of the correlation between both scales and whether the FOUR scale allowed for the assessment of paediatric patients, as it happened with the GCS. A sample including all the assessments carried out by the staff pairs with both scales has been considered for the validity of the FOUR scale (n = 82).

A sample including the first assessment of each patient (n = 27) has been used for the comparison of the results obtained with both scales and their relationship with predicted mortality and morbidity at discharge from hospital. The correlation between each scale and the KOSCHI scale has been calculated by Spearman’s correlation coefficient. Subsequently, the analysis of the sensitivity and specificity has been carried out, as well as the association between the score assigned to each patient at admission to the PICU and patient’s clinical progression. Each patient’s score at admission to the PICU was graded as Severe / Non-severe. This classification was obtained for both scales: FOUR 0 - 8 (Severe) and 9 - 16 (Non-severe); GCS 3 - 9 (Severe) and 10 - 15 (Non-severe).

A ROC-curve and the area under the curve (AUC) have been used for the scores assigned at admission to the PICU with each patient’s predicted outcome.

IBM SPSS® version 21 software has been used for data analysis.

Survey of raters

Upon completion of the study, each rater was asked to assess the scale by responding to a five-item questionnaire:

1) The FOUR scale is clinically relevant and easy to use;
2) The FOUR scale can be applied within a few minutes;
3) The FOUR scale allows for a more comprehensive assessment of patients when compared to the Glasgow coma scale;
4) The FOUR scale allows for a more comprehensive assessment of the depth of coma and the identification of clinical worsening when compared to the Glasgow coma scale;
5) The FOUR scale is a coma assessment scale that I would apply in case that it was widely acknowledged.

The statistical analysis of these items was aimed at the applicability of the FOUR coma scale to daily clinical practice according to the healthcare professionals who used it.

RESULTS

In total, 27 patients admitted to the PICU during the study period with impaired consciousness and complying with the inclusion criteria were included. An assessment has been considered when both scales were used by the intensivist, the resident and the nurse. As each patient has been assessed daily while staying at the PICU and sometimes more than once, a total of 82 assessments have been obtained. Both scales were used three times within each assessment; as the scales were used by only two raters in some assessments, a total of 221 assessments with each scale have been obtained.

Mean age of the patients was 9.5 years (median of 11 years; male patients – 7.7 and female - 13.2 years) and patients were aged between 1 and 17. Five patients under the age of 2 were included.

As regards the reason for admission, 16 patients had suffered from a traumatic head injury, two presented with hypoxic-ischaemic encephalopathy caused by drowning and the remaining patients were admitted with several conditions, namely with haemolytic-uraemic syndrome, haemorrhagic stroke, polytrauma after being run over, C0-C1 spinal cord injury caused by firearm projectile, brain arteriovenous malformation, hydrocephalus, meningitis, herpetic encephalitis and eclampsia; 14 were intubated.

A maximum score of 16 was assigned with the FOUR scale in 34 of the 221 assessments (15.5%), more frequently assigned by nurses (13 / 77, 16.9%).
A maximum total score in the GCS was assigned in 15 of the 221 assessments (6.8%). The minimum score has been most frequently assigned to the verbal component, which is difficult to obtain in intubated patients, in 132 of 221 assessments (59.7%). In line with the FOUR scale, the maximum score in GCS was more frequently assigned by nurses (seven out of 78 assessments, 9%).

The frequency of each score with each scale is shown in Figs 1 and 2.

A very high Spearman correlation coefficient has been obtained between the FOUR and the GCS scale (Sp = 0.954 for n = 82).
Internal consistency / Inter-rater reliability

A value of 0.834 has been obtained with the FOUR scale and 0.784 with the GCS in the analysis of internal consistency.

Cohen’s kappa of 0.649, 0.757 and 0.806 have been obtained with the FOUR scale by I/R, I/N and R/N staff pairs, respectively, for the total number of assessments.

A \( \kappa \) of 0.521, 0.536 and 0.767 have been obtained for the total number of assessments with the GCS by I/R, I/N and R/N staff pairs, respectively.

Inter-rater reliability regarding each component of both scales is shown in Table 3.

Figura 2 - Frequency of each score with the GCS

Prognostic value

Five patients in our group of patients died (18.5%; KOSCHI = 1). Only one patient has been assigned at discharge from hospital to 3A, 3B, 4A and 4B categories of the KOSCHI scale, respectively. Three patients were assigned to KOSCHI - 5A category (good recovery) and 15 to KOSCHI - 5B (full recovery). A similar Spearman’s correlation coefficient was obtained between the FOUR and GCS scales and the KOSCHI (Sp GCS = 0.776 vs. Sp
FOUR = 0.765) and both were strong correlations.

Sensitivity and specificity values of both scales are shown in Table 4, as well as the association between patient’s clinical status and the score assigned with each scale and by each type of rater. Specificity values obtained with the FOUR scale were higher when compared to the GCS, while those regarding sensitivity were higher for the GCS. The association between the score with the scale and patient’s predicted outcome was weighed by the contingency coefficient (CC), which was significantly more relevant regarding the FOUR scale.

These results were obtained based on the grading as Severity / Non-severity (obtained from the recodification of the mean score assigned by the three raters) and as Good / Poor Clinical Progression (KOSCHI category > or <4).

The ROC-curve has been obtained for both scales, as well as for the different types of rater, in order to allow for a better knowledge of the performance of the scales as assessment systems. The values of AUC for each rater with each scale are shown in Table 4, allowing for the comparison between the performances of each scale, by using sensitivity and specificity values.

In total, 17 questionnaires were responded by raters and the FOUR scale has been considered by 16 respondents (94.1%) as clinically relevant and easy to use, it can be applied in just a few minutes and it is a coma assessment scale that they would apply in case that it was widely acknowledged. Fourteen respondents (82.4%) considered that the FOUR scale allows for a more comprehensive clinical assessment when compared to Glasgow scale. Finally, 15 respondents (88.2%) considered that the FOUR scale allowed for a more comprehensive assessment of the depth of coma and the identification of clinical worsening when compared to the Glasgow coma scale.

**DISCUSSION**

The FOUR scale has been aimed at rectifying some limitations of the GCS.\(^5\)

Good internal consistency has been shown with the FOUR scale and acceptable with the GCS.

Healthcare professionals working at a PICU were the study raters regardless of their time of experience, which is closer to the reality of clinical practice, unlike previous studies.

A good agreement has been obtained by the pair intensivist / resident and intensivist / nurse with the FOUR scale and an excellent agreement has been obtained by the pair intensivist / nurse, while a reasonable to moderate agreement has been obtained by the pair intensivist / resident and intensivist / nurse and good agreement by the pair resident / nurse. These were impressive results considering that the raters lacked experience with the new

| Table 4 - Sensitivity, specificity, adjusted residuals, contingency coefficient for each scale and values of the area under the ROC curve (AUC) |
|-----------------|-----------------|-----------------|-----------------|-----------------|
|                  | FOUR             | GCS             |                  |                  |
|                  | Total I R N      | Total I R N     |                  |                  |
| S                | 71.4 71.4 66.7   | 71.4 100.0 85.7 | 100.0 100.0 100.0|
| Spe              | 95.0 100.0 88.9  | 100.0 60.0 55.6 | 50.0 55.0        |
| AR               | 3.6 4.0 2.7      | 4.2 2.7 1.9     | 2.2 2.5          |
| CC               | 0.574 0.626 0.486| 0.627 0.468 0.350| 0.408 0.440     |

S: sensitivity; Spe: specificity; AR: adjusted residuals; CC: contingency coefficient; I: intensivist; R: resident; N: nurse; GCS: Glasgow coma scale

<table>
<thead>
<tr>
<th>ROC curve (AUC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOUR</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>I</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>GCS</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>I</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>N</td>
</tr>
</tbody>
</table>

Revista Científica da Ordem dos Médicos 605 www.actamedicaportuguesa.com
scale and that only attended to two training sessions, unlike what happened with the GCS, which is daily applied since when the raters started their professional activity.

At study onset, no other studies involving the paediatric population had yet been published. Other studies including patients aged 2 to 5 with impaired consciousness caused by a traumatic or non-traumatic condition have emerged meanwhile and found that the FOUR scale has a higher or similar prognostic value when compared to the GCS. However, further studies are still necessary in order to determine the prognostic value of the FOUR scale and its applicability to Paediatrics.9-11

Patients admitted to ICU are often under the effect of sedative drugs which potentially affect motor response and eye opening, even though impaired brainstem or respiration reflexes are less likely. The three components of the GCS are all affected by sedation.6,12 In our study, a more comprehensive assessment of coma in patients with the minimum score of three with the GCS has been obtained with the FOUR scale. In line with what has been described in previous studies, brainstem and breathing pattern components provided additional information on the neurological status regarding those cases.

The GCS, including patient’s verbal response, was not so beneficial in around half of the patients due to the fact that they were intubated.

Similar or even excellent inter-rater reliability was found regarding brainstem and breathing pattern components in each of the three pairs of raters. A poor agreement was found regarding the GCS verbal component in all the pairs of raters, probably due to the presence of intubated patients, to whom the minimum score has been assigned.

A strong and positive correlation has been found between each scale (FOUR and GCS) and the KOSCHI scale, showing that both followed a similar trend, allowing for a similar outcome prediction. Therefore, this study was also in line with different studies in the adult population that have shown that the FOUR scale had the highest predictive value regarding in-hospital outcome as well as at discharge from hospital.4-6 The ROC curves showed good results with both scales, with a slight superiority regarding the FOUR scale. The use of this scale seems quite relevant for the classification of patients and for the definition of their outcome.

Globally, a better association has been obtained between scoring and outcome prediction with the FOUR scale, which showed higher specificity when compared to the GCS, i.e. the lower the scores with the FOUR scale the higher the odds of in-hospital non-survival, when compared to the GCS. The FOUR scale allowed for the detection of signs of neurological worsening as brainstem reflexes (direct and consensual pupilary reflex, cornea reflex and cough reflex) and the breathing pattern were assessed, the latter allowing for the identification of patients in need for assisted ventilation or the identification of ventilated patients with well-functioning respiratory centres.

This study involved some limitations, including the small number of patients (n = 27) despite a reasonable number of assessments (n = 82). Patients under the age of 2 represented a limitation regarding the motor component of the FOUR scale, due to their inability to follow directions. However, this is also a limitation with the GCS.

CONCLUSION

The FOUR scale can be used at the paediatric intensive care unit as it is easy to use and provides information on patient’s neurological status allowing for a better assessment of the state of consciousness and therefore a better management of the patient from the moment of the admission to the PICU. It should be applied in patients over the age of 2 or alternatively an adaptation of the scale to this age group could be developed, such as what happened with the GCS.

This was a pilot study corresponding to the first approach to the validity of the FOUR scale and positive results were found. A multicentric study would be relevant for the confirmation and generalisation of the results as a good alternative to the GCS in paediatric population.

HUMAN AND ANIMAL PROTECTION

The authors declare that the followed procedures were according to regulations established by the Ethics and Clinical Research Committee and according to the Helsinki Declaration of the World Medical Association. The study was approved by the Ethics Committee of the Centro Hospitalar de Lisboa Norte/Faculdade de Medicina de Lisboa and by the Clinical Direction of the Centro Hospitalar de Lisboa Norte.

DATA CONFIDENTIALITY

The authors declare that they have followed the protocols of their work centre on the publication of patient data.

CONFLICTS OF INTEREST

The authors declare that there were no conflicts of interest in writing this manuscript.

FINANCIAL SUPPORT

The authors declare that there was no financial support in writing this manuscript.
REFERENCES