Critical Care Nursing

Iran J Crit Care Nurs 2012; 5(2):95 -102

Journal home page: www.inhc.ir

Validity and reliability of coma scale (four score) in adult patient hospitalized in Critical Care Units

Seyyed Davood Tadrisi*, Nasrin Bahari[1], Abass Ebadi[1], Seyyed Jalal Madani[1]
*1. Faculty of Nursing, Baqiyatallah University of Medical Sciences,Tehran, Iran

ARTICLE INFO

Article type:
Original article

Article history:
Received: 26 Jul 2011
Revised: 27 Nov 2011
Accepted: 9 Jan 2012

Keywords:
Orientation
Reliability
Validity
Intensive Care Unit
Adult
Four Score

ABSTRACT

Aims: in order to review and determine the level of orientation, a standard scale is needed that shows ability the best type level of orientation. According to that in Iran, there are so far no measures the level of orientation Four Score made native. The aim of this study was to confirm the validity and reliability of Four Score scale in adult patients hospitalized in critical care units of Iran.

Methods: This study is methodological. After translating the English version of Four Score scale to Persian, 155 patients in two trained and untrained evaluators group (120 patients in trained evaluators group and 35 patients in untrained group) hospitalized in critical care units of two hospitals of Tehran city were selected by no probable accessible purposeful sampling method. Six evaluators (main researcher, four nurse and anesthesia resident) independently and at the same time evaluated the patients. Main researcher and other investigators recorded the level of orientation with Four Score and Glasgow Coma scores.

Results: Inter-rater reliability for Four score was "excellent". Interclass correlation coefficient in trained and untrained group was (0.998, 0/993) and the weighted kappa score for inter-rater agreement was (0.981, 0/986). Validity tests showed a high correlation between Four Score with Glasgow Coma scale. (r=0.980, 0/925) (p<0.001)

Conclusion: Validity and Reliability of Coma Scale (Four Score) are confirmed for assessing in Adult Patient Hospitalized in Critical Care Units in Iran.

Please cite this paper as:
Seyyed Davood Tadrisi, Nasrin Bahari, Abass Ebadi, Seyyed Jalal Madani. Validity and Reliability of Coma Scale (Four Score) in Adult Patient Hospitalized in Critical Care Units. Iran J Crit Care Nurs 2012;5(2):95-102.

1. Introduction

Consciousness is concerned to the complex network in the brainstem that is called the reticular activating system. This network establishes connection between thalamus, cortex and feedback system. Consciousness is one of the most complex words in the definition. In terms of psychology, Consciousness is "awareness of one's self and one's surroundings". Awareness of own self include feelings, attitudes, emotions, impulses, decision and active aspect of behavior. In short, consciousness is awareness of mental behavior, especially cognitive processes. This can be understand from patient's words about own self and indirectly from patient's behavior. So, to evaluate a

* Correspondence Author: Seyyed Davood Tadrisi
Faculty of Nursing, Baqiyatallah University of Medical Sciences.Iran.Tehran. Tel:+982122289941(342)
Email: sdt1343@yahoo.com

[Downloaded from mail.jccnursing.com on 2026-11-14]

patient's level of consciousness, doctors and nurses have to consider the patient's behavior instead at his speech. Hence the most common and simplest means of consciousness is awareness of own self and his surroundings. One of the other definitions that has clinical aspect and can use in patient's bedside, is the following definition: Consciousness is the awake and natural state of one self that the responses completely to all stimulations and his speech and behavior define the patient's awareness of his self and his surroundings. Throughout the day, such natural state of alertness fluctuated from full awareness to exact concentration with obvious limitation of attention to mild willful neglect and sleepiness. [1, 2, 3]Today, there are various tools to measure the level of consciousness: A comprehensive scale of consciousness, Ranko scale, Coma scale, Coma-like scale, Rador scale, the scale attached to Glasgow and Glasgow Coma Scale. [4, 5] The most known and valid scale to determine the level of consciousness is Glasgow Coma Scale (GCS), which was invented in 1974 by Tysdal and his colleagues, that has three behavioral components used to check the status of the patient's reactivity and include open eyes, verbal response and motor response to word and painful stimuli, this score is between 3-15 that 3 is related to the lowest-rated alertness and shows server neurological disorders and 15 represents full consciousness and natural reaction in person. [6, 7]

Because GCS in Coma toe's patient's, according to absence of speech, is reduced from 15 to 10 and has not required efficiency, so:

1. Since most comatose patients are intubated, verbal components cannot be evaluated in these patients. Some clinics are using the lowest possible score and the verbal response is a neurological basis for other studies.
2. Abnormal brainstem reflexes alter the breathing patterns and need for mechanical ventilation can be a reflection of the severity of coma. But GCS cannot be included in these clinical implications.
3. The GCS might not be able to determine the exact change in neurological tests.

For this purpose, a new tool for determining the level of consciousness in the name of a new coma scale, the four score has been developed that has 4 components and include eye opens, motor response, brainstem reflexes and respiration. Advantages of this tool versus GCS are that in GCS: each part has various score but in 4 score each part have four components that reduce care provider's errors. Furthermore, because this scale is a schematic of not responsiveness, it does a closer examination in comatose, voiceless or intubated patients. Studies, that compared these two scales on different patients, have shown that patients who had lower four score than GCS had more deaths [8]. Therefore in this study, we decided to translate the four score scale, in Persian with correct translation and achieve the validity and reliability of this instrument proportional to our culture. Up to now, numerous of tools were designed in current languages and have been used in other countries. The important point is paying attention to process of translation and culturally adaptation of tools to proper meaning of terms.

Psychometric experts of research tools emphasize on correct translation of tools.

Consistence of tools with culture of target country provides possibility of comparison the results of research with other countries and on the other hand, due to cultural differences and the meaning of words, if the meaning of words were not equal in translates tool and original tool, the validity and reliability of tool would be reduced. So, the proper process of translation and cross-cultural adaptation will lead to preserve the validity and reliability of instrument. [9]Validity is consists of the ability of an instrument to measure the actual size of what should measure and is designed for. Reliability means repeatability and continuity of measurement. Reliability is an index that shown the tools is effective and if the tool is not reliable, it makes error in results [10]. Furthermore, reliability is the ability of a

[Downloaded from mail.jccnursing.com on 2026-11-14]

Table 1: The scale of the original English version and the final version of the scale was translated into Persian consciousness of four scoring methods Quality of Life International.

Characteristic	Score
Eyes open	
Eyelids are open, chases things with eyes or flashes to command	4
Opens eyes but don't chase	3
Eyelids closed but open to laud voice	2
Eyelids closed but open to painful stimulus	1
Eyelids closed ever in painful stimulus	0
Motor Response	
Hand-Shaking in command (Thumbs up, fist, victory sign)	4
Localizes painful stimulus	3
Flexion in response to painful stimulus	2
Extension in response to painful stimulus	1
No response to painful stimulus	0
Brain Stem Response	
Presence pupil response or corneal response	4
One of pupil/corneal response	3
No pupil/corneal response	2
No pupil/corneal response	1
No pupil/corneal/cough response	0
Respiration	
Not intubated, Ordinal respiratory pattern	4
Not intubated, Cheney-Stokes breathing pattern	3
Not intubated, irregular breathing pattern	2
Breathes more than ventilator rate	1
Breathes according to ventilator rate/Apnea	0

test to obtain similar measurement with different observer (inter-rater reliability) or by a person in different times (intra- rater reliability). In clinical studies, correlation between internal viewers is the best indicator for measuring a tool. [11]By using valid and reliable scale, nurses will be able to make the best decision for patient in the least time and can increase their abilities, professional autonomies and provide the main goal of patient care that is the most comfort, in the ideal way. [12]

2. Methods

This is a methodological study that its purpose is translation, examination of reliability and validity of Persian version of four score scale. The research community includes adult male and female patients hospitalized in intensive care units in two selected hospitals in Tehran. For this study, 155 patients were selected by non-probable accessible purposeful sampling method. (Inclusion criteria were:

Patients should not be low vision, impaired hearing, extreme old age (above 80 years), receive neuromuscular blockers and also, without upper and lower limb paralysis. If not, they were not selected) and the research was conducted in four stages.

Stage 1: After taking permission from Mr.Wijdicks (designer of the scale), scale was translated from original language to target language (Persian). Then, the translation were compared with content index tool and after achieving at least 75 points, the scale was translated from target language to original language again, then we asked Mr.Wijdicks opinion and final version of the scale was prepared. (Table 1)

Stage 2: We asked the opinion of 10 professional (doctors and nurses) for reassessment of the translation of current scale. Can the scale determine the patient's level of consciousness?

All the professionals that participated in this experimental panel were agreed.

Stage 3: The nurses were educated about how to use the scale, after safety training, the

[Downloaded from mail.jccnursing.com on 2026-11-14]

assessment was conducted on 120 patients. (Table 2)

Stage 4: To ensure the easiness, the research was re-evaluated on 35 patients with untrained nurses out of research environment. (Table 3)

All patients were evaluated between 7 AM to 1 PM. Main researcher and other investigators recorded the level of orientation with Glasgow coma and four score, independently and simultaneously. Also, main researcher measured APACHE II score at first 24 hour of ICU admission from records, ICU sheet, laboratory reports and nursing reports. Data analyzed with SPSS and Med Calc 9.2.2 software. Four score data was shown with mean, standard deviation, median and fourth percentile (quarter).Although four score is a rating variable, can also be used as a quantitative continuous variable. Therefore, the interclass correlation test was used to measure consensus among evaluators about four score. The kappa test that is used for qualitative variables was used to assess the reliability between four score evaluators. (Kappa greater than 0.8= excellent or complete agreement, kappa 0.6-0.8= substantial agreement, kappa 0.4-0.6= moderate agreement, kappa less than 0.4= poor agreement).

Goodness of fit test was used for inter-rater agreement of studied variables with normal theorical distribution, but the studied population was not normally distributed. So, the relationship between four score and Glasgow Coma Scale was evaluated with spearman's test.

Table 2: Validity and reliability of the Full Outline of Unresponsiveness of patients by trained evaluators. Determine the relationship between scale alertness, Full Outline of Unresponsiveness with the current level of consciousness scale with Spearman test (rank correlation coefficient) in all cases $p < 0.001$

	Full Outline of Unresponsiveness (FOUR) Scale				Reliability between evaluators		correlation coefficient
Population	Total	Mean ± (SD)	median	IQR	correlation coefficient Between classes	Cohen's kappa coefficient (κ)	GCS4S Vs GCS
Population	120	10.44 ±(5.18)	11	9.71	0.998	0.981	0.980
				Gender			
male	66	10.69 ± (5.27)	11.50	9.46	0.998	0.983	0.980
female	54	10.12 ± (5.10)	9.58	10.21	0.996	0.971	0.976
				Age			
< 40	15	11.15± (5.05)	11	9.33	0.997	0.932	0.962
40-60	24	8.26 ± (6.70)	6.91	15.25	0.999	0.984	0.983
60-80	81	10.95 ± (4.56)	11	8.25	0.997	0.969	0.978
				Acute Physiology and Chronic Health Evaluation II			
10<	21	14.07 ± (2.69)	16	1.5	0.999	0.901	0.996
10-15	25	13.03 ±(2.61)	16	6	0.998	0.964	0.974
>15	74	8.34 ± (4.78)	8.33	8.33	0.999	0.979	0.959
				Surgical Services			
No	67	8.72 ± (4.70)	8.66	7.17	0.998	0.977	0.954
Yes	53	12.61± (3.40)	16	5.07	0.999	0.980	0.996
				Mechanical Ventilation			
No	63	3.21 ± (11.83)	12	6	0.998	0.971	0.977
Yes	57	3.85 ± (5.67)	6	5	0.977	0.974	0.948

[Downloaded from mail.jccnursing.com on 2026-11-14]

3. Results

The quality of translation basis on four detailed concept (clarity, common language, the same concept and overall quality) was sufficient and 100% of expert confirmed the formal validity. Patient have ranged from 18 to 80 years and consisted of 66 male and 54 female. The average of interval consistency coefficient (ICC) between each of investigators and main researcher was 0.998, kappa coefficient was 0.981, correlation coefficient was r=0.980, between untrained

Table 3: Validity and reliability of the Full Outline of Unresponsiveness of patients by untrained evaluators. Determine the relationship between scale alertness, Full Outline of Unresponsiveness with the current level of consciousness scale with Spearman test (rank correlation coefficient) in all cases $p < 0.001$

Full Outline of Unresponsiveness (FOUR) Scale					Reliability between evaluators		correlation coefficient
Population	Total	Min ±(SD)	median	IQR	correlation coefficient Between classes	Cohen's kappa coefficient (κ)	GCS4S Vs GCS
Population	35	7.77 ± (4.67)	8	7	0.993	0.986	0.925
			Gender				
male	23	8.63 ± (5.08)	8.83	7	0.997	0.964	0.941
female	12	5.97 ± (3.24)	6.83	6	0.954	0.893	0.923
			Age				
< 40	10	5.02 ± (4.54)	4.75	9	0.994	0.890	0.831
40-60	11	6.92 ± (4.53)	6	5	0.997	0.913	0.806
60-80	14	10.37 ± (3.63)	9.58	5	0.997	0.926	0.935
			Acute Physiology and Chronic Health Evaluation II				
>15	35	7.51 ± (4.51)	8.08	8	1	0.967	0.918
			Surgical Services				
No	30	7.56 ± (4.60)	8.08	7	0.997	0.963	0.944
Yes	5	8.97± (5.44)	11	9	0.997	0.884	0.895
			Mechanical Ventilation				
No	12	3.07 ± (11.88)	12	6	0.992	0.907	0.972
Yes	23	3.86 ±(5.61)	6	6	0.995	0.958	0.878

Table 4: Llinear correlation between APACHEII score and age and Full Outline of Unresponsiveness scale determine a level of consciousness. The above table using the Spearman correlation coefficient with $P < 0.001$ indicates that the Apache II scale and Full Outline of Unresponsiveness scale a investigator there is a negative relationship such that the scale scores are more Apache II score determines the Full Outline of Unresponsiveness scale decreases. And scale scores between age and the age so that there is direct contact Apache II is Apache II Scale score also increases

linear correlation→ Scale↓		Age	A PA C H E II	(FOUR) Scale investigator
Age	correlation coefficient		0.191	-0.015
	P Significant level		0.036	0.872
	TOTAL		120	120
A PA C H E II	correlation coefficient			-0.620
	P Significant level			0.001
	TOTAL			120
(FOUR) Scale investigator				

[Downloaded from mail.jccnursing.com on 2026-11-14]

Table 5: The agreement between trained evaluators and principal investigator on a scale to determine Full Outline of Unresponsiveness scale Kappa scores and internal consistency coefficient (ICC). The table above shows that the correlation coefficient between the classes in between all evaluators (n =6) based on the scoring scale to determine the Full Outline of Unresponsiveness scale in the group trained over 0.9 and kappa coefficient much higher (over 0.9) with $p < 0.001$. Therefore of high reproducible

Evaluators	Cohen's kappa coefficient	Agreement level	Coefficient of internal consistency (ICC)
Anesthesiologist and investigator	0.987	Excellent agreement	0.999
Nurse1 and investigator	0.991	Excellent agreement	0.999
Nurse2 and investigator	0.979	Excellent agreement	0.998
Nurse3 and investigator	.953	Excellent agreement	0.995
Nurse4 and investigator	0.967	Excellent agreement	0.997

Table 6: The agreement between untrained evaluators and principal investigator on a scale to determine Full Outline of Unresponsiveness scale Kappa scores and internal consistency coefficient (ICC). The table above shows that the correlation coefficient between the classes in between all evaluators (n =6) based on the scoring scale to determine the Full Outline of Unresponsiveness scale in the group Untrained over 0.9 and kappa coefficient much higher (over 0.9) with $p < 0.001$. Therefore of high reproducible

Evaluators	Cohen's kappa coefficient	Agreement level	Coefficient of internal consistency (ICC)
Anesthesiologist and investigator	0.964	Excellent agreement	0.996
Nurse1 and investigator	0.974	Excellent agreement	0.997
Nurse2 and investigator	0.970	Excellent agreement	0.995
Nurse3 and investigator	0.915	Excellent agreement	0.989
Nurse4 and investigator	0.939	Excellent agreement	0.993

evaluator was 0.983 with kappa coefficient and correlation coefficient r=0.986. Also, there was a significant and direct linear relationship between four score scale and patient's age and a significant inverse correlation between four score scale and APACHE II score (Table 4). Inter-rater agreement of four score scale between trained and untrained evaluators, using kappa coefficient and interval consistency coefficient (ICC), was excellent and showed that this measure is repeatable (Table 5, 6).

Validity of four score scale between trained and untrained evaluators, using the Spearman test, was obtained 0.90 which indicates that this instrument has high reliability between the evaluators.

4. Discussion

Eliminating the problems that hinder the transfer of appropriate informational, emotional content and cognitive style of main manage, is the purpose of translation. Most of questionnaires have been developed in English-spoken countries, but even in these countries, researcher should consider the migrant population in health related studies. Particularly when their removal may lead to systematic deviations in the use of therapeutic services or quality of life [13].

This scale is clear and easy to use, and in comparison with Glasgow Coma Scale, can be used easily to determine the patient's level of consciousness. The scale also has formal validity after translated to Persian. Categorical correlation coefficient showed that the reliability between the evaluators is excellent. The final score result is between our studies evaluators and Wijdicks study that has shown the excellent reliability between four score scale's observer with kappa coefficient kW=0.82 [8] as well as Vyoiek's

[Downloaded from mail.jccnursing.com on 2026-11-14]

Table 7: linear correlation between the evaluations based on scale scores to determine the level of alertness, Full Outline of Unresponsiveness scale (trained). Table above using the Spearman correlation coefficient with p<0.001indicates that the scores of evaluators (n =6) level using the Full Outline of Unresponsiveness scale has high correlation (over 0.9), there are trained evaluators in the scale of reliability is high.

linear correlation→ Scale↓		Nurse investigator	Nurse1	Nurse2	Nurse3	Nurse4	Anesthesiologist
Nurse investigator	correlation coefficient		0.999	0.998	0.996	0.996	0.999
	p Significant level		0.001	0.001	0.001	0.001	0.001
	TOTAL		120	120	120	120	120
Nurse1	correlation coefficient			0.998	0.997	0.997	0.998
	p Significant level			0.001	0.001	0.001	0.001
	TOTAL			120	120	120	120
Nurse2	correlation coefficient				0.997	0.996	0.998
	p Significant level				0.001	0.001	0.001
	TOTAL				120	120	120
Nurse3	correlation coefficient					0.997	0.996
	p Significant level					0.001	0.001
	TOTAL					120	120
Nurse4	correlation coefficient						0.996
	p Significant level						0.001
	TOTAL						120
Anesthesiologist	correlation coefficient						
	p Significant level						
	TOTAL						

study that has shown excellent kappa coefficient for four score scale (eye open= 0.96, motor response =0.97, brainstem response= 0.98 and respiration=1) [14]

The results of this study is matched with the study of comparison of GCS and four score scale on 176 patients admitted within a month and the reliability of this scale in comparison to GCS , with and without regard to age and etiology, have been reported "good". [15]

The reliability of this scale in other study that is done for comparison of GCS and four score on 267 ICU admitted patients has shown that four score indicate more benefit than GCS in the neurological patients. [16]

In other study, the validity and reliability of four score scale in comparison with GCS in emergency department are assessed and reliability of four score scale in all trained evaluators of all ages and both male and female, have reported "excellent". (Kw=0.86, 0.88) [17]

Another study that has been carried out for validity and reliability of four score scale in Indian emergency department represented a better efficiency of this scale than GCS. [18]

Finally, this scale validated by intensive care nurses and the result is reported "good" to "excellent". This scale provides more neurological information than GCS and all nurses, even those who are not experienced enough, can use it easily. [19].

5. Conclusion

Validity and reliability of four score scale is confirmed. Now, this scale can be a valid and reliable scale for evaluation the level of consciousness. This scale is simple to use and requires no training.

References

1. Ahmadi A, Shadvar K, Shahrami R, Shahrami A, Madani Sj, Abbasi S. Crit Care Med. 1st ed. Tehran. Omidan. 2007.
2. Ahmadi K. Neurology 3rd ed. Tehran. Farhang Farda. 2007.
3. Lotfi j. Neurology Aminoff. 1st ed. TEHRAN. Mir. 2007.
4. Chowdhury AH, Rahman Z, Sharma JD. Assessment of coma and impaired consciousness: A critical review ORION. 2001;10:1-12.

[Downloaded from mail.jccnursing.com on 2026-11-14]

Table 8: Linear correlation between the evaluations based on scale scores to determine the level of alertness, Full Outline of Unresponsiveness scale (untrained). Table above using the Spearman correlation coefficient with p<0.001indicates that the scores of evaluators (n =6) level using the Full Outline of Unresponsiveness scale has high correlation (over 0.9), there are trained evaluators in the scale of reliability is high.

linear correlation→ Scale↓		Nurse investigator	Nurse1	Nurse2	Nurse3	Nurse4	Anesthesiologist
Nurse investigator	correlation coefficient p Significant level TOTAL		0.997 0.001 35	0.995 0.001 35	0.983 0.001 35	0.980 0.001 35	0.994 0.001 35
Nurse1	correlation coefficient p Significant level TOTAL			0.997 0.001 35	0.985 0.001 35	0.993 0.001 35	0.994 0.001 35
Nurse2	correlation coefficient p Significant level TOTAL				0.985 0.001 35	0.994 0.001 35	0.994 0.001 35
Nurse3	correlation coefficient p Significant level TOTAL					0.988 0.001 35	0.982 0.001 35
Nurse4	correlation coefficient p Significant level TOTAL						0.990 0.001 35
Anesthesiologist	correlation coefficient p Significant level TOTAL						

5. Hanks RA, MillisSR, RickrJH, Giacino JT, Nakese-Richardson R, Frol AB, et al. The Predictive Validity of a Brief Inpatient Neuropsychologic Battery for Persons With Traumatic Brain Injury. Arch Phys Med Rehabil. 2008;89(5):950-7.
6. Marino PL. The ICU Book. 3 rd ed ed. New York: Lippincott Williams & Wilkins. 2007.
7. Schnakers C, Vanhaudenhuyse A, Giacino J, Ventura M, Boly M, Majerus S, et al. Diagnostic accuracy of the vegetative and minimally conscious state: Clinical consensus versus standardized neurobehavioral assessment. BMC Neurol. 2009;9:35.
8. 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.
9. Afrasiabai Far A, Yaghmaie F, Abdoli S, Abed Saiedi Zh. Research tool translation and cross-cultural adaptation. Medical Education. 2007;16(54):58-67.
10. Abedsaeidi Zh, Amir Ali Akbari S. Research in Medical Science. 2nd ed. Tehran: Salemi. 2005.
11. Carrasco G. Instruments For Monitoring Intensive Care Unit Sedation. Crit Care. 2000;4(4):217-25.
12. Reschreiter H, Maiden M, Kapila A. Sedation practice in the intensive care unit :a UK national survey. Crit Care. 2008;12(6):R152.
13. Karma D, Rawat AK. Effect of Stimulation in Coma. Indian Pediatrics. 2006;43(10):856-60.
14. 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.
15. Bruno MA, Ledoux D, Lambermont B, Damas F, Schnakers C, Vanhaudenhuyse A, et al. Comparison of the Full Outline of UnResponsiveness and Glasgow Liege Scale/Glasgow Coma Scale in an Intensive Care Unit Population. Neurocrit Care. 2011.
16. Fischer M, Rüegg 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.
17. Stead LG, Wijdicks EF, Bhagra A, Kashyap R, Bellolio MF, Nash DL, et al. Validation of a new coma scale, the FOUR score, in the emergency department. Neurocrit Care. 2009;10(1):50-4.
18. Murthy T. A new score to validate coma inemergency department - FOUR score. Indian J Neurotrauma (IJNT). 2009;6(1):59-62.
19. Wolf CA, Wijdicks EF, Bamlet WR, McClelland RL. Further Validation of the FOUR Score Coma Scale by Intensive Care Nurses. Mayo Clin Proc. 2007;82(4):435-8.

[Downloaded from mail.jccnursing.com on 2026-11-14]