

## ORIGINAL REPORT

# INTER-RATER RELIABILITY OF THE EARLY FUNCTIONAL ABILITIES SCALE

Kristin Alvsåker, MD<sup>1,3</sup>, Sten M. Walther, MD, PhD<sup>1,2</sup>, Ingerid Kleffeltgård, PT<sup>1</sup>, Malin Mongs, OT<sup>1</sup>, Randi Aas Drægebø, NT<sup>1</sup> and Anne Keller, MD, PhD<sup>3</sup>

From the <sup>1</sup>Intensive Care Unit, Oslo University Hospital Ullevål, Oslo, Norway, <sup>2</sup>Linköping University, Faculty of Health Sciences, Department of Medicine and Health, Division of Cardiovascular Medicine and <sup>3</sup>Department of Physical Medicine and Rehabilitation, Oslo University Hospital Ullevål, Oslo, Norway

**Objective:** To investigate the inter-rater reliability of the Early Functional Abilities (EFA) scale.

**Design:** An observational study of inter-rater reliability in an open cohort.

**Patients:** Twenty-four patients with traumatic brain injury in need of medical or surgical intervention in the early rehabilitation section of the intensive care unit.

**Methods:** The EFA was assessed by 4 different professions in the rehabilitation team. Inter-rater reliability was assessed using linear weighted kappa statistics.

**Results:** The overall weighted kappa values of the different EFA items varied from 0.27 to 0.60. The items in the sensorimotor functional area had the highest pairwise agreement, with a mean kappa range of 0.68–0.76. The vegetative stability, position tolerance and wakefulness items had the lowest mean kappa values (0.49, 0.33 and 0.49, respectively). Agreement was good to excellent between the occupational therapist and physiotherapist across the majority of the items, whereas the physician and nurse agreed less with one another.

**Conclusion:** The inter-rater reliability of the EFA scale was good for most items among all the raters. The scale may be used by all members of the interdisciplinary team after training in administration and scoring. A reduction in the number of items in the vegetative functional domain is recommended.

**Key words:** traumatic brain injury; rehabilitation outcome; assessment.

J Rehabil Med 2011; 43: 892–899

*Correspondence address:* Kristin Alvsåker, Department of Physical Medicine and Rehabilitation, Oslo University Hospital Ullevål, NO-0407 Oslo, Norway. E-mail: Kristin.Alvsaker@oslo-universitetssykehus.no

Submitted November 27, 2010; accepted June 17, 2011

## INTRODUCTION

Traumatic brain injury (TBI) is the leading cause of death and long-term disability in young adults (1). The severity of TBI can be classified by the level of consciousness at the injury site using the Glasgow Coma Scale (GCS) (2). GCS scores of 3–8 represent severe TBI, and GCS scores of 9–12 represent moderate TBI. Depending on the severity of the primary brain injury and secondary complications, patients will have widely varying levels of

consciousness during early recovery. Severe disability during the acute phase in the intensive care unit (ICU) is seen both in severe TBI and moderate TBI with secondary neurological deterioration needing medical or surgical interventions (3, 4).

Several studies have shown evidence of improved outcomes and functional benefits during early rehabilitation for moderate and severe brain injury, starting within 35 days post-injury (5–7). However, these studies were performed partly in acute hospitals and partly in rehabilitation centres.

Thus far, to the authors' knowledge, there has been no study of rehabilitation outcome measures during early rehabilitation in the ICU, while the patient is being weaned from organ support and awakening from a coma. In this phase, the patient is still totally dependent on personal assistance for activities of daily living (ADL), and the two most widely used functional independence measures, Barthel Index (BI) (8) and the Functional Independence Measure (FIM) (9), tend to provide low scores (the floor effect) in this phase. These two observer-rated generic scales only measure broad ADL and mobility domains and do not measure more specific functional skills, such as fine motor ability and other fine neurological skills (10). In contrast, assessment with coma scales, such as the GCS and the Coma Recovery Scale (CRS) (11), in the early phase of recovery, typically provides high scores (the ceiling effect).

During rehabilitation in the ICU, patients' functional levels and neurological conditions change rapidly. Patients have short attention spans, and their functional abilities fluctuate during the day. A reliable clinical evaluation scale for the assessment of early neurological recovery is needed to evaluate the effect of the rehabilitation programme during the ICU phase.

The Early Functional Abilities (EFA) scale was developed by Heck et al. (12) to close the gap between coma scales and functional independence scales, and was intended for use during the early phase of neurological rehabilitation in patients with severe brain injury. Like the BI and FIM, the EFA scale is an observer-rated generic measure. In their validation study, Heck et al. (12) found that the EFA scale was more sensitive than both CRS and FIM when describing non-comatose patients with severe brain impairments. The study concluded that the EFA scale is a useful tool for assessing different capabilities at a very early stage of rehabilitation in patients with rising alertness, despite ongoing severe motor-functional impairment.

The assessments in the study by Heck et al. (12) were performed by physiotherapists (PTs) and occupational therapists

(OTs). Heck et al. (12) investigated the inter-rater reliability of the EFA scale using correlations, and concluded that there was high inter-rater reliability ( $r=0.80$ ). However, correlations between measurements are uncertain, because correlations might increase with a greater variability between observed subjects (13). Therefore, reliability should be based on analysis of the agreement between the raters, rather than on correlations (14).

The EFA scale is used in neurorehabilitation for patients with severe brain injury in German-speaking countries (15) and Denmark (16), but is less commonly used than FIM. Because of the ceiling effect, as the patients' functional levels improve, it is common to use FIM when the EFA is over 70 points (17).

There is sparse data on the clinimetric properties of the EFA scale. An evaluation of the EFA scale by Danish PTs concluded that Heck's results were not valid for general use of the EFA scale, since the scale is used differently by different professions on the interdisciplinary teams at different hospitals (17).

The EFA scale has been used in the Early Rehabilitation Section of the ICU (ERSICU) at Oslo University Hospital Ullevål in Norway since 2005 (18, 19). The interdisciplinary rehabilitation team consists of 4 professionals: a PT, OT, nurse and physician. The EFA scale is assessed for each patient by 2 or more professionals in the interdisciplinary team. However, use of the EFA presupposes good reliability of the 4 professionals' scoring.

The aim of the present study was to formally assess the inter-rater reliability of the EFA scale. This investigation was undertaken in an interdisciplinary rehabilitation programme for patients with severe and moderate TBI in the acute phase, and the raters in the interdisciplinary rehabilitation team belonged to 4 different professions.

## METHODS

### Patients

This is an observational study of inter-rater reliability in an open cohort based on the assessment of 24 patients admitted to ERSICU between May 2006 and March 2008.

The inclusion criteria for patients were age greater than 16 years, severe TBI (a GCS score of 3–8 at the injury site) or moderate TBI (a GCS score of 9–12 at the injury site) with neurological deterioration defined as a reduction in GCS; and a need for mechanical ventilation, intracranial pressure (ICP) monitoring or craniotomy. The exclusion criteria were chronic subdural haematoma and non-traumatic head injuries.

All patients had passed the immediate neuro-intensive care phase. They were studied during interdisciplinary early neurorehabilitation and as they were weaned from organ support. Eighteen patients with severe TBI (90%) and all 4 patients with moderate TBI (100%) were under ICP monitoring or craniotomy. On admission to ERSICU, 23 patients were still ventilator-dependent. All patients stayed one week or more in the ERSICU.

### Interdisciplinary rehabilitation programme

The rehabilitation programme consisted of guiding and training in daily activities (20), positioning, transferring, mobilizing the patient into sitting and standing positions (21), and promoting normal swallowing, eating and drinking (22). The therapeutic efforts were carefully and individually adjusted to each patient's needs and medical status at the

time. The efforts most frequently used during early rehabilitation were mobilization to sitting, guiding, and positioning (23). The duration of therapeutic efforts varied from 30 min to 3 h/day depending on other medical interventions and the infection status of the patients.

The physician was responsible for the profile of the programme, and the PT and OT were responsible for planning and performing the therapeutic activities. The nurses were responsible for monitoring the patient's vital parameters (consciousness, respiration, blood pressure, pulse and temperature), but they also assisted during the various therapeutic activities.

### Procedure

The EFA scale was assessed on 31 occasions in 20 patients with severe TBI and 4 patients with moderate TBI. Seventeen patients were assessed once. Seven patients were assessed twice within intervals ranging from 3 to 16 days (mean 8 days). Assessments were carried out within a range of 1–44 days (mean 12 days) after discontinuation of short-acting analgesic infusion (fentanyl). Sedative medication (low-dose propofol, clonidine or neuroleptics) had been given within 24 h before 18 of the assessments (58%). Most patients received zolpidem or zopiclone for sleeping. The patients had a tracheostomy in 10 assessments (32%).

Four professionals on the interdisciplinary rehabilitation team (nurse, PT, OT and physician) were involved in the assessments. Based on clinical observations on a particular day, one rater from each profession assessed the EFA score independently of one another on the same day. In advance, the raters agreed to base their assessments on the highest level of reproducible performance that they observed during the day.

One PT, 2 physicians, 3 OT and approximately 20 nurses were raters in the study.

The physicians were specialists in physical medicine and rehabilitation. The physicians, PT and OT had several years of clinical experience in the rehabilitation of patients with brain injury. Some of the nurses had several years of clinical experience in neuro-intensive care, but none had clinical experience in rehabilitation. Apart from oral instructions on how to perform the assessment using the supplementary text, the raters received no training in advance.

### Instrument

The EFA scale includes 20 items in 4 domains: vegetative, fascio-oral, sensorimotor and cognitive function. It uses a 5-category ordinal scale from 1 (no function) to 5 (practically full function). The minimum total score is 20 points, and the maximum score is 100 points.

The original language of the EFA scale is German. A Danish translation was used (Salling 2000) (24). Each of the 5 categories for the 20 items was described with a few keywords in a supplementary text. All raters used the supplementary text when completing their ratings. For this article, the EFA items were translated into English. The translation is given in Appendix I.

### Statistical analysis

Descriptive statistics were used for the sociodemographic variables (Table I).

The total mean EFA score for the 4 raters was calculated. The range and distribution of scores from all raters were examined for each item. The floor and ceiling effects were determined by calculating the percentage of the sample that achieved the lowest and highest possible scores, respectively. Percentages greater than 20% were considered substantial (25).

Inter-rater agreement between the 4 professionals was analysed for each item using Fleiss' weighted kappa statistics for multiple raters (26) and Cohen's linear weighted pair-wise kappa statistics among the 4 raters, with 6 combinations altogether (27). For ordinal scales, a weighted kappa accounts for the fact that, with scores between 1 and 5, there is greater disagreement across several categories than across only one category. Kappa values less than 0.40, between 0.41

Table I. Patient characteristics

Characteristics	
Male/female, <i>n</i>	19/5
Age, years, median (Q1–Q3)	25 (20–37)
Reason for injury, <i>n</i> (%)	
Traffic accident	15 (63)
Fall	5 (21)
Assault	2 (8)
Other/unknown	2 (8)
Initial GCS, median (Q1–Q3)	5 (3–8)
Intracranial pressure monitoring, <i>n</i> (%)	22 (92)
Intracranial surgery, <i>n</i> (%)	11 (46)
Other injuries, <i>n</i> (%)	19 (79)
Face injuries	13 (54)
Orthopaedic injuries	11 (46)
Thoracic and abdominal injuries	12 (50)
Length of stay in ICU, median (Q1–Q3)	15 (9–19)
Length of stay in ERSICU, median (Q1–Q3)	19 (12–24)
Days on ventilator, median (Q1–Q3)	23 (17–31)
Tracheostomy at discharge, <i>n</i> (%)	1 (4)
GCS at discharge, median (Q1–Q3)	13 (12–14)
Discharged to rehabilitation hospital, <i>n</i> (%)	17 (71)
Assessed days after injury, median (Q1–Q3)	24 (18–34)

ICU: intensive care unit; GCS: Glasgow Coma Scale; ERSICU: Early Rehabilitation Section of the ICU: Q1–Q3: 25<sup>th</sup>–75<sup>th</sup> percentile.

and 0.74, and above 0.75 were considered to be in poor, good and excellent agreement, respectively (28). The mean kappa value of the 4 raters was calculated for each item.

The Statistical Package for Social Sciences version 15.0 (SPSS Inc., Chicago, IL, USA), STATA/SE 11.1 (College Station, TX, USA) and MedCalc Version 11.5.1 programs were used to analyse the data.

Ethics

The EFA assessment was done as a part of a clinical routine and without any additional examination of the patients. Therefore, consent from the patient or next of kin was not required. The Norwegian Data Inspectorate Committee and the ethics committee at the hospital approved the study.

RESULTS

The patient characteristics are shown in Table I.

Distribution of scores

The mean total EFA score among the 4 raters was 59. Eight tests in 7 patients with severe TBI, and 2 tests in 2 patients with moderate TBI had a total score above 70. The distribution of the total EFA scores for all raters was bimodal, with a cluster of items with low scores and another cluster with high scores.

The distributions of scores per item for the 4 raters are shown in Fig. 1. Most items had good distribution, although there were some notable exceptions (Table II).

Items with a high proportion of low scores (floor effects) were excretion functions (continence) (70%), movement changes/transfer (42%), wakefulness (39%), standing (35%) and comprehension (30%). The items with a high proportion of high scores (ceiling effects) were tonus (43%), arbitrary movements (38%), head control (30%) and tactile information (30%). The highest level of function was rarely recorded for the movement changes/transfer and comprehension items.

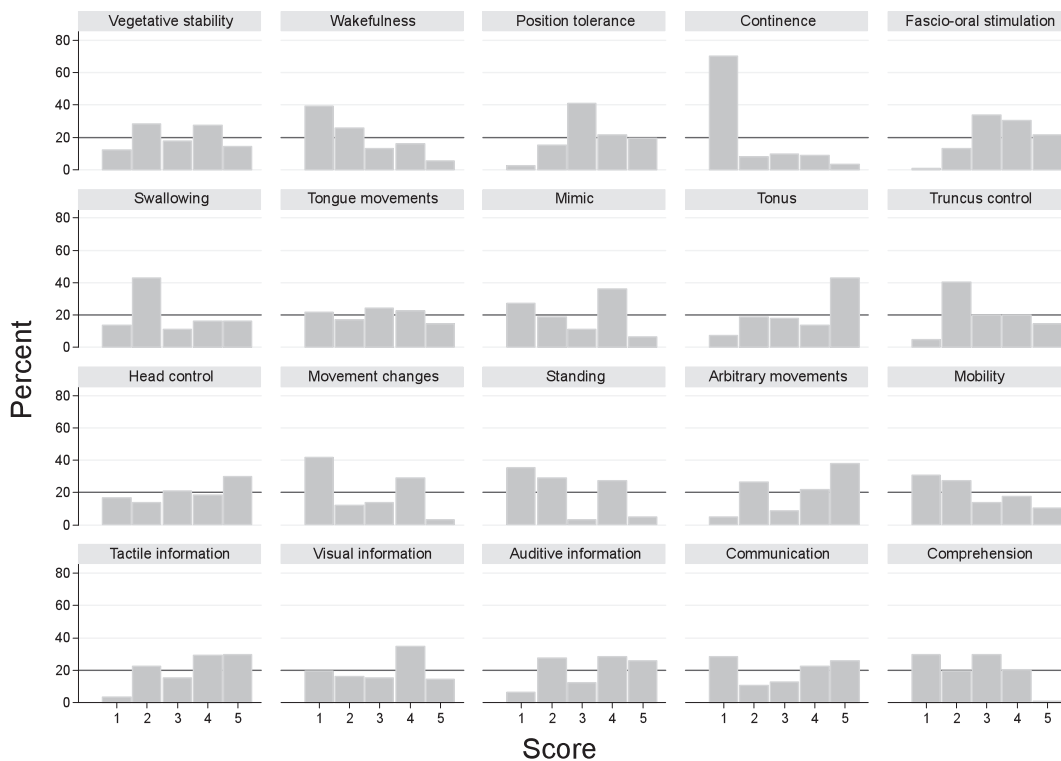


Fig. 1. Distribution and scores from the 4 raters for each of the 20 items in the Early Functional Abilities scale (*n*=31).

Table II. Floor and ceiling effects for the 4 raters for each of the 20 items in the Early Functional Abilities (EFA) scale (n = 31)

Item	Floor %	Ceiling %
1. Vegetative stability	12	15
2. Wakefulness	39	6
3. Position tolerance	3	19
4. Excretion functions (continence)	70	4
5. Fascio-oral stimulation/oral hygiene	1	22
6. Swallowing	14	15
7. Tongue movements/chewing	22	15
8. Mimic	27	7
9. Tonus	7	43
10. Truncus control	5	15
11. Head control	18	30
12. Movement changes/transfer	42	3
13. Standing	35	5
14. Arbitrary movements	5	38
15. Mobility in wheelchair	31	10
16. Tactile information	3	30
17. Visual information	19	15
18. Auditive information	7	26
19. Communication	28	26
20. Comprehension	30	1

Inter-rater reliability

Kappa values for multiple raters are shown in the last column of Table III.

Fig. 2 and Table III present pairwise kappa values for each item (6 combinations and the mean value). Table IV presents actual cross-ratings from assessments performed by a PT and an OT in items with the lowest, median and highest kappa values.

Table III. Inter-rater reliability among the 4 professionals (weighted pairwise kappa values, mean kappa values of pairwise combinations and overall kappa values from multiple raters from 31 assessments)

	PT-OT	PT-MD	OT-MD	PT-Nurse	OT-Nurse	MD-Nurse	Pairwise mean	Multiple raters
1. Vegetative stability	0.65	0.68	0.60	0.39	0.36	0.25	0.49	0.31
2. Wakefulness	0.72	0.39	0.48	0.49	0.49	0.37	0.49	0.28
3. Position tolerance	0.48	0.25	0.20	0.37	0.44	0.22	0.33	0.27
4. Excretion functions (continence)	0.88	0.60	0.58	0.44	0.44	0.62	0.59	0.44
5. Fascio-oral stimulation/oral hygiene	0.73	0.60	0.62	0.71	0.73	0.43	0.64	0.45
6. Swallowing	0.68	0.72	0.83	0.77	0.75	0.77	0.75	0.60
7. Tongue movements/chewing	0.71	0.60	0.58	0.66	0.6	0.67	0.64	0.34
8. Mimic	0.80	0.76	0.69	0.76	0.61	0.66	0.71	0.49
9. Tonus	0.89	0.69	0.76	0.69	0.67	0.61	0.72	0.47
10. Truncus control	0.79	0.74	0.74	0.69	0.64	0.59	0.70	0.48
11. Head control	0.85	0.74	0.75	0.73	0.70	0.78	0.76	0.54
12. Movement changes/transfer	0.80	0.69	0.64	0.70	0.65	0.74	0.70	0.47
13. Standing	0.79	0.73	0.79	0.70	0.59	0.66	0.71	0.54
14. Arbitrary movements	0.79	0.73	0.69	0.62	0.62	0.65	0.68	0.45
15. Mobility in wheelchair	0.78	0.75	0.79	0.69	0.74	0.67	0.74	0.53
16. Tactile information	0.83	0.61	0.56	0.68	0.62	0.6	0.65	0.48
17. Visual information	0.74	0.77	0.68	0.74	0.55	0.69	0.70	0.42
18. Auditive information	0.78	0.75	0.75	0.66	0.62	0.80	0.73	0.51
19. Communication	0.79	0.75	0.76	0.75	0.73	0.79	0.76	0.49
20. Comprehension	0.80	0.66	0.70	0.67	0.76	0.51	0.68	0.50

Overall mean denotes the mean value of all cross-ratings.  
PT: physiotherapist; OT: occupational therapist; MD: physician.

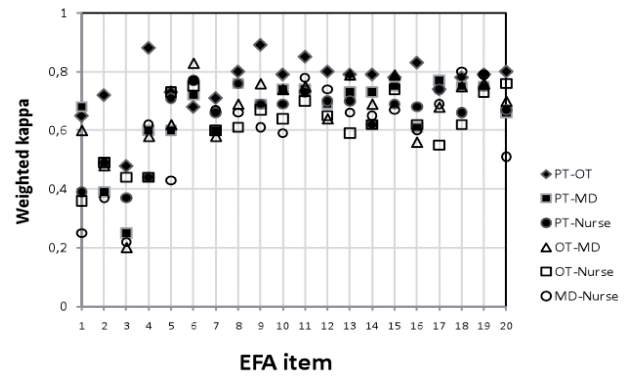


Fig. 2. Inter-rater reliability among the 4 professionals (mean kappa values from 31 assessments). PT: physiotherapist; OT: occupational therapist; MD: physician.

Swallowing had the highest kappa value for multiple raters ( $\kappa=0.60$ ), whereas vegetative stability, wakefulness, position tolerance and tongue movements/chewing had the lowest (0.31, 0.28, 0.27 and 0.34, respectively). The other items had overall kappa values for multiple raters ranging between 0.42 and 0.54.

The items in the sensorimotor functional domain (items 9–15) had the highest pairwise agreement, with a weighted mean kappa range of 0.68–0.76. The vegetative stability, wakefulness and position tolerance items in the vegetative functional domain had the lowest mean kappa values (0.49, 0.49 and 0.33, respectively) (Table III).

Agreement between the OT and PT was good to excellent across all 20 items (mean  $\kappa=0.76$ , range 0.48–0.89), whereas the physician and nurse agreed less with one another (mean  $\kappa=0.60$ , range 0.22–0.80) (Fig. 2).

Table IV. Actual cross-ratings from a physiotherapist (PT) and an occupational therapist (OT) in items with lowest, median and highest kappa values

Patient ID	Item 3 (position tolerance) with the lowest kappa value (0.48)		Item 19 (communication) with the median kappa value (0.79)		Item 9 (tonus) with the highest kappa value (0.89)	
	OT	PT	OT	PT	OT	PT
1	4	3	3	3	5	5
1	5	5	5	5	5	5
2	2	2	2	2	4	4
2	4	5	4	4	5	5
3	2	3	3	3	5	5
4	5	5	4	5	5	5
5	2	3	1	2	2	2
6	5	2	5	5	4	5
6	5	1	5	5	5	5
7	3	3	1	1	2	3
7	4	5	3	3	4	5
11	3	2	4	5	5	5
12	4	3	1	1	3	3
13	4	3	3	3	3	4
14	3	3	3	3	3	4
14	4	4	5	4	5	5
15	5	5	4	4	5	5
17	3	3	1	1	1	1
18	2	5	5	5	5	5
21	2	3	5	4	5	5
22	3	3	1	1	2	2
23	3	3	1	1	2	2
27	3	3	2	1	3	3
29	3	3	1	1	2	2
29	3	3	2	1	2	2
31	2	2	4	5	5	5
32	5	5	4	4	5	5
33	3	3	2	2	3	3
33	3	3	3	4	2	2
34	5	5	4	5	5	5
35	4	4	2	3	3	3

## DISCUSSION

In this study, we found good inter-rater reliability for most items in the EFA scale among the raters as a group, and good to excellent inter-rater reliability in pairwise combinations.

The PT and the OT had the highest agreement with one another, whereas the physician and nurse agreed less. For some items, we found floor and ceiling effects. The results indicate that the scale yields reproducible findings across examiners for these items.

Swallowing was the most reliably scored item among the professionals. The clinical importance of the swallowing function in ERSICU might explain this finding. The interdisciplinary team typically evaluates the swallowing function and the need for percutaneous endoscopic gastrostomy (PEG) before discharging patients from the ERSICU, and the EFA scale is helpful in this process.

The items in the sensorimotor functional domain had the highest pairwise agreement between raters, whereas the items in the vegetative functional domain had the lowest agreement. The finding that the sensorimotor function was the most reliable agrees with another study of the GCS, which found that motor response is the most reliably scored item (29).

The poor agreement of scores of vegetative stability can be explained by the fact that the categories overlap (30). In the supplementary text, both the need for monitoring and the need for sedative or stabilizing medications (e.g. beta-blockers) were used in several categories. All patients in the study were being weaned from organ support in the ICU, and some patients, for example, needed monitoring of peripheral oxygen saturation but did not require sedative medications (e.g. low-dose propofol, clonidine or neuroleptics) or stabilizing medication. Therefore, the raters had difficulties in knowing which parameter to take into account during the assessment. This item intended to describe sympathetic overactivity (dysautonomia) (31) since it included the need for beta-blockers. Dysautonomia might be better diagnosed using another scale. For example, it may be preferable to measure heart rate changes in response to afferent stimuli (31). The position tolerance item had low agreement both for multiple raters and in pairwise combinations. We recommend removing the vegetative stability and position tolerance items from the EFA scale or, alternatively, improving them by adding explanations.

### Floor and ceiling effects

The items with substantial floor effects were excretion functions (continence), movement changes/transfer, wakefulness, standing, mobility in wheelchair and comprehension. One reason for this finding may be that patients were so early in the awakening process that the items reflected their weaning from organ support, rather than their neurological statuses. In particular, the wakefulness item was affected by the use of sedative medications and the excretion functions (continence) item was affected by the use of a Foley catheter to monitor fluid balance. For the EFA scale to be valid in patients being weaned from organ support, the removal of the wakefulness and excretion function items from the scale should be considered.

Transfers, standing and mobility in wheelchair items were not applicable to most patients in the early stages of assessment, and exhibited floor effects. However, these items were found to be valuable later in early rehabilitation.

We found substantial ceiling effects for the tonus, arbitrary movements, head control and tactile information items. During rehabilitation, the majority of patients will eventually exhibit the ceiling effect due to progress in their recovery, and other ADL scales (like FIM) should be used instead. We did not compare EFA scores with FIM scores, but the clinical recommendation to switch to FIM when EFA exceeds 70 point appears reasonable. One-third of the assessments in our study had an EFA score of over 70 points.

The results show that agreement was higher in some items with a floor or ceiling effect. The highest weighted kappa values (over 0.80) were found in one item with a floor effect (excretion functions) and 3 items with a ceiling effect (tonus, head control and tactile information). For the other items with a floor or ceiling effect, the weighted kappa values did not differ from the other items (Tables II and III).

Our findings regarding clusters of items with floor or ceiling effects support the hypothesis that the EFA scale to some



extent overlaps the coma scales and functional independence scales for a few items (12). However, our results confirm the need for a method to assess recovery during the gap between coma and functional independence.

Our general clinical experience was that the EFA scale described our patients' functional abilities quite well, with the exception of the vegetative functional domain. Most patients still needed help in ADL and were unable to take care of themselves when discharged from the ERSICU.

#### *Moderate traumatic brain injury with secondary complications*

The EFA scale was originally developed for severe TBI. Our study also included patients with moderate TBI who developed secondary complications, since their state upon admission to ERSICU was quite similar to that of patients with severe TBI. Due to the small number of patients in the sample, we cannot give any firm recommendation regarding the use of EFA for this subgroup of moderate TBI.

#### *Agreement between professionals*

The PT and OT had the highest level of agreement, whereas the physician and the nurse agreed less. The interpretation of the kappa values between the PT and OT as ranging from good to excellent was based on general rules of thumb. The actual cross-rating of the items (Table IV) shows that the two raters agreed in most tests for all 3 items, except for a few tests in the item with the lowest kappa value, position tolerance, where they had notable disagreements in several categories. We conclude that it is reasonable to label the agreement between the PT and the OT as good to excellent.

Two different characteristics of the EFA scale explain the results. First, the assessments in the EFA scale are based on observations of the patients during the day and not on stimulus-response tests, in which the rater only scores the response observed after performing a stimulus. Our experience was that the patients' functional levels varied during the day due to their short attention spans. In addition, it was difficult to interpret subtle changes in the patients' recovery and evaluate whether these changes were sufficiently significant to advance from one score level to the next. The PT and OT were responsible for the therapeutic actions and worked together much of the time, and they were familiar with the patients' functional levels. Although the nurses spent the whole day at the bedside, they were responsible for observing vital functions and thus had a different focus than the therapists. Some nurses only performed the ratings once and did not familiarize themselves with the scale. The physician was responsible for the rehabilitation efforts and neurological progress, but spent less time at the bedside than the other raters, observing the patients during therapeutic activities during the day.

Secondly, the scale uses words, such as "placing" and "facilitation," that are common in the literature of physiotherapy and occupational therapy, but are less common in the literature for physicians and nurses. Although the nurse and physician agreed less than the PT and the OT, the agreement of the former pair was good in most items. During the group discussion after

the assessment, we received the impression that nurses who had some experiences in the EFA scale had a higher level of inter-rater agreement than the nurses without experience. We think that the scale can be used effectively by all members of the rehabilitation team, after careful education in the text manual and the terminology used in the EFA scale, and training in administration and scoring.

#### *Interdisciplinary teamwork*

Four professions work together during early rehabilitation in the ERSICU. An aim of interdisciplinary teamwork is to allow cross-professional dialogue and create an interdisciplinary plan (32). During the assessment, the raters focused their attention on the patient's functional level. They became more aware of the patient's need for specific stimulation in certain areas when the patient showed a small degree of improvement. Group discussion after the individual assessments of EFA scores was helpful in identifying and adjusting short-term goals, such as evaluating the swallowing function and the need for PEG.

The EFA scale helped the team to tailor therapeutic efforts and create a common understanding among the different professions regarding the patients' neurological and functional progress and further need for brain injury rehabilitation after discharge. However, the assessments were done very early in the rehabilitation process, making it premature to set long-term goals based on EFA item scores or sum scores in the ERSICU.

#### *Prediction of outcome*

Prognosticating the outcome after TBI is an integral part of TBI rehabilitation. It can be used to aid functional goal-setting and estimate the resources needed for the rehabilitation effort and length of stay. No single set of indicators has been demonstrated to accurately predict outcomes (33). A reliable clinical scale during early recovery after severe and moderate TBI would be a valuable tool to evaluate the rehabilitative interventions and the patient's prognosis. For example, an improvement in the total EFA score during the stay at the ERSICU might have prognostic value, and further studies should investigate this possibility. Thus far, the EFA scale needs further improvement to be used as a reliable prognostic indicator.

#### *Study limitations*

This study has several limitations. The generalizability of the findings is uncertain because the number of raters was small for 3 of the 4 professions. The nurses who participated in the study did not have experience in rehabilitation, and only a few nurses had training in EFA scoring. Our results indicate that the raters need training in administration and scoring with the EFA scale for their assessments to be reliable.

This study was performed in an ICU setting, and the results are not directly generalizable to other rehabilitation hospital settings.

We used a Danish version of the EFA scale. In written texts, there are only minor dialectic differences between the Danish and Norwegian languages, and Danish texts are commonly used in Norway. Therefore, most raters did not consider using the

Danish version of the EFA scale to be a problem. However, the Danish version of the EFA scale was not validated. During the study, we found that the Danish versions of 3 items had slightly different meanings from the original German text.

Another limitation of the EFA scale is that there has been only one study examining its validity (12). In the past few years, the EFA scale has become integrated in TBI research in Denmark (34). In 2007, a validated Danish translation of the EFA manual was published (35). There is now an ongoing study of the validity of the EFA scale in Hvidovre Hospital in Denmark. There are several potential reasons for the lack of research on the EFA scale so far: one being the language barrier in English-speaking countries, and another might be different national strategies in the rehabilitation of severe brain injuries.

In conclusion, the inter-rater reliability of the EFA scale was good for most items among all the raters. Some items were affected by floor and ceiling effects. The inter-rater reliability was good to excellent between the PT and OT, whereas the physician and nurse agreed less. All members of the interdisciplinary team can use the scale, but training in administration and scoring is necessary to increase the scale's reliability. A reduction in number of items in the vegetative functional domain is recommended. The interdisciplinary rehabilitation team in this study found the EFA scale to be clinically helpful in the rehabilitation process when the patients were being weaned from organ support.

#### ACKNOWLEDGEMENTS

The authors thank the nursing staff in the ICU and OT Beate Klungland for their assistance in data collection. The authors also thank physician Ute Heller for her translation of the items in the EFA scale from German to English.

The authors report no conflicts of interest.

#### REFERENCES

1. US Dept of Health and Human Services, National Institutes of Health, Office of the Director. Rehabilitation of persons with Traumatic Brain Injury: NIH Consensus Statement. October 26–28 1998. Bethesda, MD: National Institutes of Health; 1998; 16, p. 1–41.
2. Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. *Lancet* 1974; 2: 81–84.
3. Lobato RD, Rivas JJ, Gomez PA, Castaneda M, Canizal JM, Sarabia R, et al. Head-injured patients who talk and deteriorate into coma. *J Neurosurg* 1991; 75: 256–261.
4. Brown CV, Zada G, Salim A, Inaba K, Kasotakis G, Hadjizacharia P, et al. Indications for routine repeat head computed tomography (CT) stratified. *J Trauma* 2007; 62: 1339–1344.
5. Cope DN, Hall K. Head injury rehabilitation: benefit of early intervention. *Arch Phys Med Rehabil* 1982; 63: 433–437.
6. Mackay LE, Bernstein BA, Chapman PE, Morgan AS, Milazzo LS. Early intervention in severe head injury: long-term benefits of a formalized program. *Arch Phys Med Rehabil* 1992; 73: 635–641.
7. Zhu XL, Poon WS, Chan CC, Chan SS. Does intensive rehabilitation improve the functional outcome of patients with traumatic brain injury (TBI)? A randomized controlled trial. *Brain Inj* 2007; 21: 681–690.
8. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. *Md State Med J* 1965; 14: 61–65.
9. Hamilton BB, Laughlin JA, Fiedler RC, Granger CV. Interrater reliability of the 7-level functional independence measure (FIM). *Scand J Rehabil Med* 1994; 26: 115–119.
10. Dodds TA, Martin DP, Stolov WC, Deyo RA. A validation of the functional independence measurement and its performance among rehabilitation inpatients. *Arch Phys Med Rehabil* 1993; 74: 531–536.
11. Giacino JT, Kalmar K, Whyte J. The JFK Coma Recovery Scale-Revised: measurement characteristics and diagnostic utility. *Arch Phys Med Rehabil* 2004; 85: 2020–2029.
12. Heck G, Steiger-Bächler G, Schmidt T. Early Functional Abilities (EFA) – eine Skala zur Evaluation von Behandlungsverläufen in der neurologischen Frührehabilitation. *Neurol Rehabil* 2000; 6: 125–133.
13. Bland JM, Altman DG. Measurement error and correlation coefficients. *BMJ* 1996; 313: 41–42.
14. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; 86: 420–428.
15. Frank U. Treating tracheotomized patients with severe dysphagia: an explorative evaluation of a multidisciplinary treatment protocol. Potsdam: Universitätsverlag Potsdam; 2008.
16. Engberg AW. A Danish national strategy for treatment and rehabilitation after. *J Head Trauma Rehabil* 2007; 22: 221–228.
17. Lorentzen J, Poulsen I, Maribo T. Evaluation of the Early Functional Abilities (EFA). *Danske Fysioterapeuter, Projekt Måleredskaber*; 2005 September 23. Available from: <http://fysio.dk/fafo/Maleredskaber/Maleredskaber-alfabetisk/Early-Functional-Abilities-EFA-/Videnskabelig-vurdering-af-EFA/> (in Danish).
18. Hanoa R, Alvsaker K. [Early phase rehabilitation of brain injuries]. *Tidsskr Nor Laegeforen* 2008; 128: 66–67 (in Norwegian).
19. Rehabilitation for persons with severe brain injury. Report from the Norwegian Directorate of Health. Sosial- og helsedirektoratet; 2005 Jul. Report No.: IS-1279.
20. Davis JZ. The Affolter methode: a model for treating perceptual disturbances in the hemiplegic and brain-injured patient. *Occup Ther Pract* 1992; 3: 30–38.
21. Gjelsvik BEB. The Bobath concept in adult neurology. Stuttgart: Georg Thieme Verlag; 2008.
22. Coombes K. Facial Oral Tract Therapy (F.O.O.T.). Jubiläumsschrift – 10 Jahre Sculungszentrum. Burgau; 2001, p. 47–57. The International Brain Injury Association.
23. Kleffellgård I, Mongs M, Alvsåker K. Early rehabilitation of severe traumatic brain injury patients in the intensive care unit. Abstract 0296. 7th World Congress on Brain Injury; 2008 April 9–12 Lisabon, Portugal.
24. Salling B. Danish translation of the Early Functional Abilities (EFA). Hammel Neurocenter, Voldby vej 15, 8450 Hammel, Denmark. 2000 October. Report. Available from: [www.vfhj.dk/default.asp?PageID=864](http://www.vfhj.dk/default.asp?PageID=864).
25. McHorney CA, Ware JE Jr, Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care* 1994; 32: 40–66.
26. Fleiss JL, Cohen J. Statistical methods for rates and proportions. 2nd edn. New York: John Wiley; 1981.
27. Fleiss J.L., Cohen J. The equivalence of weighted kappa and intraclass correlation coefficient as measures of reliability. *Educ Psychol Meas* 1973; 33: 613–619.
28. McCluggage WG, Bharucha H, Caughley LM, Date A, Hamilton PW, Thornton CM, et al. Interobserver variation in the reporting of cervical colposcopic biopsy specimens: comparison of grading systems. *J Clin Pathol* 1996; 49: 833–835.
29. Prasad K. The Glasgow Coma Scale: a critical appraisal of its clinimetric properties. *J Clin Epidemiol* 1996; 49: 755–763.
30. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960; 20: 37–46.
31. Baguley IJ, Nott MT, Slewa-Younan S, Heriseanu RE, Perkes IE. Diagnosing dysautonomia after acute traumatic brain injury: evidence for overresponsiveness to afferent stimuli. *Arch Phys*

Med Rehabil 2009; 90: 580–586.  
 32. Huber L. Interdisciplinary teamwork helps quality efforts reach new heights. AORN J 2010; 92: 345–346.  
 33. Chua KS, Ng YS, Yap SG, Bok CW. A brief review of traumatic brain injury rehabilitation. Ann Acad Med Singapore 2007; 36: 31–42.  
 34. Norup A, Kristensen KS, Siert L, Poulsen I, Mortensen EL. Neuropsychological support to relatives of patients with severe.

Neuropsychol Rehabil 2011; 1–16.  
 35. Kristensen K.S., K.Nielsen, Jakobsen D. Translation of the Early Functional Abilities (EFA). Afdeling for Neurorehabilitering, Hvidovre Hospital, Kettegård alle 30, 2650 Hvidovre, Denmark. [cited 2007 Aug 7]. Available from: [http://fysio.dk/Upload/Graphics/PDF-filer/Maaleredskaber/Early\\_Functional\\_Abilities\\_\(EFA\)\\_Testmanual.pdf](http://fysio.dk/Upload/Graphics/PDF-filer/Maaleredskaber/Early_Functional_Abilities_(EFA)_Testmanual.pdf).

APPENDIX I. Early Functional Abilities (EFA) scale

German text	German text translated to English					
<b>A. Vegetativum</b>	<b>A. Vegetative function</b>					
1. Vegetative Stabilität	1. Vegetative stability					
2. Wachheit	2. Wakefulness					
3. Lagerungstoleranz	3. Position tolerance					
4. Ausscheidungsverhalten	4. Excretion functions (continence)					
<b>Sum Vegetativum</b>	<b>Total Vegetative function score</b>					
<b>B. Fazio-oraler Bereich</b>	<b>B. Fascio-oral function</b>					
5. FO-Stimulation/Mundhygiene	5. Fascio-oral stimulation/oral hygiene					
6. Schlucken	6. Swallowing					
7. Zungenbeweglichkeit/Kauen	7. Tongue movements/chewing					
8. Mimik	8. Mimic					
<b>Sum Fazio-oraler Bereich</b>	<b>Total Fascio-oral function score</b>					
<b>C. Sensomotorik</b>	<b>C. Sensorimotor function</b>					
9. Tonusanpassung	9. Tonus					
10. Kopfkontrolle	10. Head control					
11. Rumpfkontrolle/Sitzen	11. Truncus control/sitting					
12. Bewegungsübergänge/Transfer	12. Movement changes/transfer					
13. Stehen	13. Standing					
14. Willkürmotorik	14. Arbitrary movements					
15. Fortbewegung/Mobilität im Rollstuhl	15. Mobility in wheelchair					
<b>Sum Sensomotorik</b>	<b>Total Sensorimotor function score</b>					
<b>D. Sensorisch-kognitive Fähigkeiten</b>	<b>D. Sensoricognitive function</b>					
16. Taktile Information	16. Tactile information					
17. Visuelle Information	17. Visual information					
18. Akustische Information	18. Acoustic information					
19. Kommunikation	19. Communication					
20. Situationsverständnis	20. Comprehension					
<b>Sum Sensorisch-kognitive Fähigkeiten</b>	<b>Total Sensoricognitive function score</b>					
<b>Total sum (min. 20, max. 100)</b>	<b>Total score (min. 20, max. 100)</b>					
The 5 EFA levels	<table border="0"> <tr> <td>1 point: Lack of function, no function observed</td> <td>2 points: A trace of function is present but unstable, not reproducible and not specific or purposeful. Very reduced</td> <td>3 points: Obviously present. Stable, specific, reproducible and purposeful but not subtle. Moderately reduced</td> <td>4 points: Stable, specific, purposeful and nuanced. Lightly reduced</td> <td>5 points: Practically full functioning</td> </tr> </table>	1 point: Lack of function, no function observed	2 points: A trace of function is present but unstable, not reproducible and not specific or purposeful. Very reduced	3 points: Obviously present. Stable, specific, reproducible and purposeful but not subtle. Moderately reduced	4 points: Stable, specific, purposeful and nuanced. Lightly reduced	5 points: Practically full functioning
1 point: Lack of function, no function observed	2 points: A trace of function is present but unstable, not reproducible and not specific or purposeful. Very reduced	3 points: Obviously present. Stable, specific, reproducible and purposeful but not subtle. Moderately reduced	4 points: Stable, specific, purposeful and nuanced. Lightly reduced	5 points: Practically full functioning		