Scand J Rehab Med 31: 31–41, 1999

CONSTRUCT VALIDITY OF THE FUNCTIONAL INDEPENDENCE MEASURE (FIM): QUESTIONING THE UNIDIMENSIONALITY OF THE SCALE AND THE "VALUE" OF FIM SCORES

Jean-François Ravaud[1], Michel Delcey[2] and Alain Yelnik[3]

From the [1]CERMES – INSERM U 502, Institut Fédératif de Recherche sur le Handicap, Paris, France, [2]Association des Paralysés de France, Paris, France, [3]Service de rééducation fonctionnelle, Hôpital Fernand Widal, Paris, France

ABSTRACT. The Functional Independence Measure (FIM) is one of the most widely used disability and dependence assessment instruments in rehabilitation medicine. As for other similar scales, the expression of results as a unique score raises an important question. Is it legitimate to consider the object being measured (functional independence) as a unidimensional entity? The answer is of major practical importance in justifying the use of the FIM. Having made a critical analysis of the previous validation procedures, the authors then submitted admission FIM items of 127 consecutive patients admitted in a French rehabilitation unit to different multidimensional statistical methods in order to analyse the structure of the FIM. Their findings demonstrate the multidimensional nature of the phenomenon assessed by the scale. This observation raises the question of the relevant use of the FIM total score, currently too widely applied without sufficient precaution, and suggests that preferably subscores should be used.

Key words: assessment scale; construct validity; disability evaluation; Functional Independence Measure; rehabilitation.

Lorsqu'on ne sait pas la vérité d'une chose, il est bon qu'une erreur commune fixe l'esprit des hommes. Car la maladie principale de l'homme réside dans la curiosité inquiète des choses qu'il ne peut savoir, et il ne lui est pas si mauvais d'être dans l'erreur que dans cette curiosité inutile.

[When the truth respective to something is unknown, it is good that men's mind be settled by a common error. Because, as man's chief ailment is his restless curiosity about things he cannot understand, he doesn't feel that bad being wrong than uselessly curious.]

(Pascal (1623–1662) – Les Pensées)

INTRODUCTION

Impact of unidimensionality on the FIM

The Functional Independence Measure (FIM) has emerged among several assessment scales used in rehabilitation medicine. The FIM, which was developed recently (3, 34, 39), is widely used in the U.S.A. and has become the "Barthel of the nineties" (55) for a large number of clinicians in France (8, 12, 21, 22, 42, 56) and Europe. Because of its widespread use, we chose this scale as an example for a more general evaluation of functional assessment since, as with other scales, results are often condensed into a single score. Such a simplification raises the important question of whether it is legitimate to conceive of the object being measured—functional (in)dependence—as a single entity which can be scored on a single unidimensional scale (5, 16, 19, 23, 41). Some of the FIM items concern the "physical" aspects of dependence (e.g. transfers) while other items (e.g. memory) involved a completely different aspect of dependence (Table I). Therefore, we are *a priori* dealing with a multidimensional phenomenon (5, 16).

This empirical hypothesis has a major impact on the consistent use of the scale, especially since the FIM was developed to incorporate the cognitive sphere lacking in the Barthel Index (27, 49). "When items do not fit a common unidimensional continuum, total scores provide uninterpretable information" (62). Unless a unidimensional scale is used, it is impossible to state that one patient is functionally more "independent" than another patient and all calculations using a total score, with the exception of statistical correlations with other quantitative criteria (predictive validity), become uninterpretable (47). Indeed, showing that the score is "statistically" correlated with other data (e.g. length of stay) does not imply anything about the "significance" of the instru-

Table I. *The Functional Independence Measure (FIM)*

7—Complete independence			
L 6—Modified independence		Independence	
E 5—Supervision			
V 4—Minimal contact assistance			
E 3—Moderate assistance			
L 2—Maximal assistance			
1—Total assistance		Dependence	
Self-care	Admission	Discharge	Change
A. Eating	☐	☐	☐
B. Grooming	☐	☐	☐
C. Bathing	☐	☐	☐
D. Dress upper body	☐	☐	☐
E. Dress lower body	☐	☐	☐
F. Toileting	☐	☐	☐
Sphincter control			
G. Bladder management	☐	☐	☐
H. Bowel management	☐	☐	☐
Transfers (mobility)			
I. Bed/chair	☐	☐	☐
J. Toilet	☐	☐	☐
K. Tub/shower	☐	☐	☐
Locomotion			
L. Walk/wheelchair	☐	☐	☐
M. Stairs	☐	☐	☐
Communication			
N. Comprehension	☐	☐	☐
O. Expression	☐	☐	☐
Social cognition			
P. Social interaction	☐	☐	☐
Q. Problem solving	☐	☐	☐
R. Memory	☐	☐	☐

ment but this rather verifies its relationship with other criteria. Similarly, using the score as a purely descriptive instrument for a given individual is unquestionable as the standardized observation scale is used to follow one patient without any reference to a total score (particularly with the star diagram included in the FIM where each item is individualized so its progression can be followed).

The problem arises when the score refers to a precise sense, implying that the cluster of items contributing to a score has a unique meaning, i.e. that "functional independence" is not an agglomeration of distinct notions, but rather one single concept involving, for instance, independence in memory or in eating. Without this prerequisite, "adding" elementary independence items is meaningless and cannot be used for any valid comparison with scores in other individuals. If the scale is not unidimensional, both comparisons between scores and their progression are invalid. Unless the scoring scale is unidimensional, it would be equally unjustified to use it to calculate the efficacy of a rehabilitation programme or to make comparisons between institutions (46). Under these conditions, it is particularly important to know whether earlier studies validating the FIM addressed this question and demonstrated that the total score can be correctly used as the measurement of a single object.

Former studies on the FIM construct

The FIM is a measurement of disability introduced in 1984 by Granger and a joint force representing different organizations in rehabilitation medicine in the U.S.A. It was then further developed in three phases from 1985 to 1987 (25, 28). The FIM includes 18 items which are grouped in 6 sections (Table I), each one including 2 to 6 items. Each item is scored on a detailed scale from 1 to 7. Total scores ranging from 18 to 126 are obtained by simply adding the raw unweighted scores for each item (17). The FIM has also been tested in telephone surveys (10) and adapted for children (4, 7, 24). It was designed for clinicians or hospital managers (3) for assessing the degree of disability, evaluating functional gains, improving training and research methods as well as measuring the cost/benefit ratios in rehabilitation. As summarized by the originators (30) "*the FIM, developed to provide uniform assessment of patient disability and medical rehabilitation outcome, is an 18-item, seven level scale designed to estimate burden of care*".

Most studies aimed at validating the FIM have dealt with acceptability and reliability (28), especially inter-rater agreement (6, 14, 35, 36) and criteria validity by comparison with the Barthel Index (40, 50, 60, 65) or other scales (13, 32) or again with different indicators (12, 26, 38, 58) while less emphasis was laid on evaluation of the content (Delphi method or empirical evaluation by users) (6, 21, 25). Though these studies can be criticized (1, 5, 14, 15, 31), the most important point is the lack of studies evaluating construct validity and unidimensionality (62). Another point is the question of the scaling properties of the FIM. Some authors have attempted to transform it into an equal-interval scale. Because scales such as the FIM are ordinal scales (9, 14, 20, 52–54), the promoters of the FIM, who admit this drawback (28), justify unlimited use (37) on the basis of Rasch analysis (37, 43, 44, 62, 63, 66, 67). This type of analysis has been used for a long time (11, 48, 51, 57, 59) but it cannot provide an answer to the current problem, especially since it has been shown that scale unidimensionality, a prerequisite for such analysis (16, 53), is lacking in the FIM (44, 45) where at least two dimensions, motor and cognitive, are involved. Thus, the promoters applied the Rasch analysis to the first 13 items and to the last 5 items. Hall et al. also performed the Rasch analysis to compare three scales, including the FIM (32). Thus, while attempting to prove the scalability of the FIM, these authors have demonstrated that the FIM is not a unidimensional scale, but they did not draw the obvious conclusions (33, 37) relative to its founded use. Recently, other authors have contributed to further evidence showing the inadequacy of the FIM to fit the Rasch model and the limits of the transformation of the FIM into an equal-interval scale (15, 16). Even if scalability could be demonstrated, this would not allow use of the FIM for all the proposed indications (41, 52, 53). The total score should no longer be used to express a sole entity and extreme caution should be applied when proposing its use as a mathematical parameter.

One major point is the fact that the empirically recognized heterogeneous nature of dependence has already been confirmed by these analyses. The question of the unidimensionality of the two subscores, addressed in certain studies (44), remains unanswered, as these authors did not complete their analysis after stating that the subdivision of the items into two groups is probably insufficient. In a recent report (16), Dickson & Köhler questioned the unidimensionality of the motor score. The heterogeneous nature of the instrument has already been demonstrated using factor analysis (19), but, rather than focusing on the analysis itself, these authors made the assumption that functional independence involves two concepts they themselves defined. They thus argue against the assertion that the notion of disability assessed by the FIM is that defined by the WHO in the International Classification of Impairments, Disabilities and Handicaps (ICIDH). This point has also been made by other authors (5). Here, the criticism is more theoretical (mismatch between the definition of the items and the chosen reference, ICIDH) than technical (measurement validity). In another work, interpretation of factor analysis was focused less on the examination of FIM items (37) than on pathogenic categories of the sample chosen for Rasch analysis. Thus the multidimensional analysis of the instrument has not been completed; the structure of the FIM remains to be explored (5). Furthermore, in France (7, 8, 10, 12, 21, 22, 42, 56) as well as in other countries (9, 18, 25, 32, 40, 64) the total score of the FIM is still widely applied or recommended. In order further to investigate this question of utmost importance for the meaning of the score, we used multidimensional statistical analysis techniques to investigate the internal consistency of the FIM, avoiding any specific hypothesis concerning dimensions (16) at study onset.

METHODS

Subjects

The study included 127 consecutive patients admitted between May 1991 and February 1993 to a medical rehabilitation unit (Paris, Fernand Widal Hospital). FIM data were collected at admission. The FIM is usually used in this unit. The scores were attributed by well-trained personnel (physicians and other healthcare workers). No disease was considered to be an exclusion criterion. The few patients with a complete independence for all the 18 items (i.e. total score = 126 at admission) were the only ones excluded from the study.

Almost two-thirds of the sample subjects (Table II) had brain damage (either predominant or exclusive) which could produce sequellae scorable on the last 5 items of the FIM. By comparison, the 1991 UDS report, which centralized FIM data in the US (29), indicated the following distribution for the main diagnoses: stroke 33%, orthopaedic condition 29%, brain injury 59%, neurologic condition 5%, spinal cord injury 3%, non-traumatic spinal cord 3%, other brain dysfunction 2%. There were more strokes and fewer orthopaedic conditions in our sample.

Statistics

Data analyses were performed on an IBM 486 computer using the SAS (Statistical Analysis System) software (61). The following analyses were performed successively: descriptive statistical analysis of the different variables, analysis of variable interdependence (Cronbach's alpha coefficient, factor analysis of correspondences) and an analysis of the instrument's structure (principal components analysis using orthogonal transformation, then the rotation method (Varimax)).

Table II. *Diagnoses in the sample population*

	n	%
Stroke	59	46.5
Non-traumatic spinal cord dysfunction	18	14.2
Traumatic brain dysfunction	14	11.0
Arthritis	10	7.9
Orthopaedic conditions	7	5.5
Other brain dysfunction	5	3.9
Peripheral neuropathy	4	3.1
Cerebral tumour	3	2.4
Spinal cord injury	2	1.6
Friedreich ataxia	1	0.8
Myopathy	1	0.8
Miscellaneous	3	2.3
Total	*127*	*100%*

Table III. *Scores at admission to the rehabilitation unit (*n *= 127)*

Items	Mean	Standard deviation
Self care (S 1)		
1. Eating*	6.08	1.49
2. Grooming	4.36	2.39
3. Bathing	3.59	2.16
4. Dress upper body	4.02	2.36
5. Dress lower body	3.61	2.34
6. Toileting	4.09	2.48
Sphincter control (S 2)		
7. Bladder management	5.08	2.79
8. Bowel management	5.52	2.56
Transfers (mobility) (S 3)		
9. Bed/chair	3.64	2.52
10. Toilet	3.61	2.55
11. Tub/shower	2.46	2.35
Locomotion (S 4)		
12. Walk/wheelchair	2.98	2.49
13. Stairs	2.02	2.07
Communication (S 5)		
14. Comprehension	6.14	1.72
15. Expression	5.97	1.98
Social cognition (S 6)		
16. Social interaction	6.24	1.74
17. Problem solving	4.98	2.69
18. Memory	5.79	2.27

* Each item is scored from 1 to 7.

RESULTS

Descriptive analysis

We analysed consecutively the scores obtained for the 18 items (Table III); then addition of scores within each of the 6 sections to provide subscores, designated S1 to S6, and the total score, St. All the extreme values (1 and 7 for the items, 6 to 42 for S1 to S6) were present. The total scores ranged from 18 to 123. In order to make comparisons with U.S. reports, means and SD are presented, though median and quartile values would have been better for ordinal scales. The mean $\pm$ SD (80.17 ± 27.06) was at a level usually observed in a rehabilitation facility which does not manage major disabilities. These results were quite similar to those of the U.S. reports (28, 29) (Table IV). Means for items within a given subscore were neighbouring, suggesting that the "difficulty" of these items was similar. However, this similarity reveals the clinical association between deficiencies and disabilities rather than the homogeneous nature of the measured dependence (see below).

Analysis of inter-item relationships

Cronbach's alpha coefficient. This coefficient is interpreted as the part of score variance attributable to a common source which is postulated to have the "true value" of the dimension to be measured (1). According to the authors, alpha should be > 0.7 or 0.8 (18). We observed $\alpha = 0.93$, as did Dodds et al. (18). This figure is close to the 0.94 reported by Fourn et al. (19) or Brosseau et al. (6).

Correlation matrix. The 18 variables contributing to the same total score should be significantly cross-correlated. This was not always the case despite a satisfactory mean correlation, 0.45. Variable 14 (comprehension) was correlated with only 6 of the 17 other variables and variable 15 (expression) was correlated with only 11 out of 17 (Table V). These two variables showed the least cross-correlation. There was an almost complete set of non-correlated variables between items 9 to 13 (mobility and locomotion) and items 14 to 18 (communication and social cognition), as seen in subscores S3-S4 and S5-S6 (Table VI). Subscore S5 (communication) was only significantly correlated with S2 (the weakest significant relationship) and S6.

Table IV. *Overall results by FIM subscores and comparison with UDS reports (28, 29)*

Subscores		French study (*n* = 127)			UDS 90 (28)	UDS 91 (29)
		Total	S.D of total	Item mean	Mean	Mean
S1	Self care (6 items)	25.76	11.02	4.29	4.1	4.0
S2	Sphincter control (2 items)	10.60	4.93	5.30	4.5	4.4
S3	Mobility (3 items)	9.70	6.91	3.23	3.2	3.2
S4	Locomotion (2 items)	4.99	4.19	2.49	2.3	2.2
S5	Communication (2 items)	12.11	3.56	6.05	5.4	5.5
S6	Social cognition (3 items)	17.02	5.97	5.67	4.9	5.0
St	*FIM total*	*80.17*	*27.06*	*4.45*	*73.1*	*72.7*

S.D.: standard deviation.

Table V. *Correlation matrix (items and subscores)*

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18
V1	1 0.0																	
V2	0.49 ***	1 0.0																
V3	0.41 ***	0.78 ***	1 0.0															
V4	0.45 ***	0.71 ***	0.74 ***	1 0.0														
V5	0.35 ***	0.58 ***	0.76 ***	0.83 ***	1 0.0													
V6	0.44 ***	0.48 ***	0.64 ***	0.68 ***	0.79 ***	1 0.0												
V7	0.30 ***	0.28 **	0.44 ***	0.35 ***	0.42 ***	0.55 ***	1 0.0											
V8	0.28 **	0.33 ***	0.49 ***	0.43 ***	0.51 ***	0.59 ***	0.70 ***	1 0.0										
V9	0.37 ***	0.49 ***	0.67 ***	0.68 ***	0.81 ***	0.84 ***	0.49 ***	0.51 ***	1 0.0									
V10	0.36 ***	0.47 ***	0.65 ***	0.67 ***	0.80 ***	0.83 ***	0.47 ***	0.50 ***	0.99 ***	1 0.0								
V11	0.27 **	0.31 ***	0.42 ***	0.41 ***	0.54 ***	0.56 ***	0.29 ***	0.28 **	0.70 ***	0.70 ***	1 0.0							
V12	0.26 **	0.34 ***	0.49 ***	0.56 ***	0.66 ***	0.67 ***	0.40 ***	0.39 ***	0.84 ***	0.84 ***	0.66 ***	1 0.0						
V13	0.19 *	0.22 *	0.30 ***	0.37 ***	0.48 ***	0.48 ***	0.26 **	0.22 *	0.59 ***	0.59 ***	0.57 ***	0.69 ***	1 0.0					
V14	0.27 **	0.24 **	0.19 *	0.20 *	0.12 NS	0.20 *	0.29 ***	0.32 ***	0.12 NS	0.11 NS	−0.04 NS	0.07 NS	−0.01 NS	1 0.0				
V15	0.22 *	0.09 NS	0.03 NS	0.06 NS	−0.02 NS	0.04 NS	0.14 NS	0.18 *	−0.03 NS	−0.04 NS	−0.09 NS	−0.06 NS	−0.09 NS	0.86 ***	1 0.0			
V16	0.36 ***	0.32 ***	0.25 **	0.28 **	0.21 *	0.26 **	0.33 ***	0.28 **	0.14 NS	0.13 NS	0.03 NS	0.11 NS	0.10 NS	0.64 ***	0.51 ***	1 0.0		
V17	0.36 ***	0.38 ***	0.28 **	0.35 ***	0.24 **	0.34 ***	0.39 ***	0.31 ***	0.27 **	0.26 **	0.10 NS	0.15 NS	0.18 NS	0.56 ***	0.49 ***	0.58 ***	1 0.0	
V18	0.40 ***	0.38 ***	0.28 **	0.30 ***	0.22 *	0.29 ***	0.37 ***	0.33 ***	0.18 *	0.17 NS	0.05 NS	0.06 NS	0.04 NS	0.58 ***	0.43 ***	0.79 ***	0.70 ***	1 0.0
S1	0.59 ***	0.82 ***	0.88 ***	0.90 ***	0.89 ***	0.83 ***	0.48 ***	0.54 ***	0.79 ***	0.78 ***	0.51 ***	0.61 ***	0.42 ***	0.24 **	0.07 NS	0.33 ***	0.39 ***	0.37 ***
S2	0.32 ***	0.33 ***	0.51 ***	0.43 ***	0.51 ***	0.62 ***	0.93 ***	0.91 ***	0.54 ***	0.53 ***	0.31 ***	0.43 ***	0.26 **	0.33 ***	0.18 *	0.34 ***	0.38 ***	0.39 ***
S3	0.36 ***	0.46 ***	0.62 ***	0.64 ***	0.78 ***	0.80 ***	0.45 ***	0.47 ***	0.97 ***	0.97 ***	0.85 ***	0.84 ***	0.63 ***	0.07 NS	−0.05 NS	0.11 NS	0.23 **	0.14 NS
S4	0.25 **	0.31 ***	0.44 ***	0.51 ***	0.63 ***	0.64 ***	0.36 ***	0.34 ***	0.79 ***	0.79 ***	0.67 ***	0.93 ***	0.90 ***	0.04 NS	−0.09 NS	0.11 NS	0.17 *	0.05 NS
S5	0.25 **	0.16 NS	0.11 NS	0.13 NS	0.05 NS	0.12 NS	0.22 *	0.26 **	0.04 NS	0.03 NS	−0.07 NS	−0.03 NS	−0.06 NS	0.97 ***	0.97 ***	0.59 ***	0.54 ***	0.52 ***
S6	0.42 ***	0.41 ***	0.31 ***	0.35 ***	0.25 **	0.34 ***	0.42 ***	0.35 ***	0.23 **	0.22 *	0.08 NS	0.12 NS	0.12 NS	0.66 ***	0.53 ***	0.85 ***	0.89 ***	0.93 ***
St	0.55 ***	0.67 ***	0.76 ***	0.78 ***	0.81 ***	0.84 ***	0.66 ***	0.67 ***	0.85 ***	0.83 ***	0.59 ***	0.71 ***	0.54 ***	0.45 ***	0.28 **	0.51 ***	0.58 ***	0.54 ***

For each two by two crossing of variables:
– Correlation coefficient
– p value. * if $p \leq 0.05$. ** if $p \leq 0.01$. *** if $p \leq 0.001$. **NS (non-significant)** if $p > 0.05$.

Factor analysis of correspondences (FAC). This type of analysis is particularly relevant for ordinal scales (2). The total score followed a smooth reversed U-shaped curve, the Gutmann effect, as did the S1 (self-care), the S3 (mobility), and to a lesser extent and shifted, the S4 (locomotion) subscores, i.e. all of the

Table VI. *Correlation matrix (subscores)*

	S1	S2	S3	S4	S5	S6	St
S1	1 0.0						
S2	0.55 ***	1 0.0					
S3	0.75 ***	0.50 ***	1 0.0				
S4	0.57 ***	0.38 ***	0.81 ***	1 0.0			
S5	**0.16** **NS**	0.26 **	**0.002** **NS**	**−0.03** **NS**	1 0.0		
S6	0.41 ***	0.42 ***	0.19 *	**0.13** **NS**	0.62 ***	1 ***	
S tot	0.90 ***	0.72 ***	0.82 ***	0.69 ***	0.37 ***	0.61 ***	1 0.0

For each two by two crossing of variables:
– Correlation coefficient
– p-value: * if $p \leq 0.05$. ** if $p \leq 0.01$. *** if $p \leq 0.001$. **NS (non-significant)** if $p > 0.05$.

Table VII. *Factors in the principal component analysis*

Factors	Eigenvalues	% of explained variance	Cumulative %
1	8.05	44.7	44.7
2	3.42	19.0	63.7
3	1.28	7.1	70.8
4	1.04	5.8	76.5
5	0.79	4.4	80.9
6	0.65	3.6	84.5
7	0.45	2.5	87.0
8	0.44	2.5	89.5
9	0.43	2.4	91.9
10	0.31	1.7	93.6
11	0.26	1.5	95.1
12	0.23	1.3	96.4
13	0.17	0.9	97.3
14	0.16	0.9	98.2
15	0.12	0.7	98.8
16	0.10	0.6	99.4
17	0.09	0.5	99.9
18	0.01	0.1	100.00

motor items. S2 (continence), S5 and S6 (communication and social cognition) followed a much more erratic curve.

Structure analysis

Principal component analysis (PCA). Two criteria are classically used to determine the number of factors to retain: the gap between consecutive eigenvalues and the ratio between the percentage of explained variance over the variance for a given item. Factoring should cease when the new factor's percentage of explained variance is less than the variance of one variable (in this case, item) in the analysis (63). As shown in Table VII, the first factor explained 44.7% of the total variance, and the first four factors explained 76.5% of the variance. There was a gap between the 4th and 5th factors whose eigenvalue was less than 1. Thus, unlike Fourn et al. (19) or Heinemann et al. (38), we chose to work with four factors instead of two (explaining less than two-thirds of the total variance).

Factor analysis with orthogonal transformation. Two principal dimensions appeared: the first factor resulted from the contribution of the first 13 FIM items (all saturation coefficients were > 0.5) and the second factor was explained by the last five items alone (Table VIII).

Factor analysis with the rotation method (Varimax). The first four factors (Table IX) contributed much more homogeneously to the explanation of total variance and four independent dimensions clearly appeared. Independently of method of factoring, the cognitive dimension (last 5 items) which has been postulated in earlier work with the Rasch method retains its integrity. On the other hand, the first 13 items could no longer be considered as a quasi-unique dimension.

The pattern of variable contribution was striking: the first factor corresponded to mobility and locomotion items (subscores 3 and 4); the second factor corresponded to cognitive items (subscores 5 and 6) ; the third factor corresponded to the first subscore (self-care) ; and the fourth factor was explained by the main contribution of sphincter items (subscore 2).

DISCUSSION

The high Cronbach coefficient would suggest that the scale has a good internal consistency. However, it has been widely demonstrated that a high alpha coefficient can be consistent with two absolutely independent dimensions (11). Thus, two items are less correlated due to the expression of the same dimension (dependence) rather than because of their frequent clinical association and their simultaneous degree of severity. Therefore, statistical methods which take into account similar variations in the item scores cannot make any distinction between the contribution of the measurement method and that of the object being measured.

The correlation matrix suggests that motor items

Table VIII. *Initial factor method—orthogonal transformation matrix*

Item		Factor 1	Factor 2	Factor 3	Factor 4	Communality
1.	Eating	0.54 *	0.24	−0.26	0.18	0.45
2.	Grooming	0.68 *	0.10	−0.58	0.11	0.83
3.	Bathing	0.79 *	−0.06	−0.42	−0.10	0.82
4.	Dress upper body	0.81 *	−0.06	−0.36	0.06	0.80
5.	Dress lower body	0.86 *	−0.23	−0.16	−0.03	0.82
6.	Toileting	0.87 *	−0.14	0.06	−0.11	0.79
7.	Bladder management	0.63 *	0.16	0.24	−0.57	0.80
8.	Bowel management	0.65 *	0.13	0.15	−0.61	0.83
9.	Bed/chair	0.89 *	−0.32	0.13	0.04	0.92
10.	Toilet	0.88 *	−0.33	0.14	0.04	0.91
11.	Tub/shower	0.63 *	−0.39	0.24	0.24	0.67
12.	Walk/wheelchair	0.76 *	−0.38	0.30	0.15	0.83
13.	Stairs	0.57 *	−0.34	0.38	0.35	0.71
14.	Comprehension	0.35	0.78 *	0.21	0.08	0.79
15.	Expression	0.17	0.77 *	0.25	0.15	0.70
16.	Social interaction	0.43	0.71 *	0.05	0.14	0.71
17.	Problem solving	0.50	0.61 *	0.08	0.12	0.65
18.	Memory	0.45	0.71 *	−0.03	0.05	0.72
	Eigenvalues	8.05	3.42	1.28	1.04	Total: 13.78
	Percent of explained variance	44.70%	18.98%	7.10%	5.75%	76.54%

* Denotes the most important contribution of the items to the four factors.

involving the lower limbs (mobility, locomotion, dressing lower body) are independent of the cognitive sphere, in particular “communication” items. The matrix suggests a certain intervariable independence that cannot be expressed by the global Cronbach’s alpha coefficient, and which contradicts the idea that the summation of

Table IX. *Rotation method : Varimax—rotated factor pattern*

Item		Factor 1	Factor 2	Factor 3	Factor 4	Communality
1.	Eating	0.18	0.37	0.53 *	0.01	0.45
2.	Grooming	0.14	0.20	0.87 *	0.05	0.83
3.	Bathing	0.29	0.07	0.80 *	0.31	0.82
4.	Dress upper body	0.40	0.13	0.77 *	0.18	0.80
5.	Dress lower body	0.56	0.02	0.63 *	0.31	0.82
6.	Toileting	0.61 *	0.14	0.45	0.44	0.79
7.	Bladder management	0.24	0.25	0.13	0.81 *	0.80
8.	Bowel management	0.21	0.21	0.22	0.84 *	0.83
9.	Bed/chair	0.80 *	0.04	0.42	0.32	0.92
10.	Toilet	0.80 *	0.03	0.40	0.32	0.91
11.	Tub/shower	0.80 *	−0.04	0.19	0.06	0.67
12.	Walk/wheelchair	0.86 *	0.01	0.21	0.20	0.83
13.	Stairs	0.84 *	0.05	0.04	−0.03	0.71
14.	Comprehension	0.01	0.87 *	0.03	0.15	0.79
15.	Expression	−0.06	0.83 *	−0.10	0.03	0.70
16.	Social interaction	0.03	0.81 *	0.21	0.09	0.71
17.	Problem solving	0.12	0.75 *	0.22	0.14	0.65
18.	Memory	−0.03	0.78 *	0.28	0.16	0.72
Eigenvalues		4.48	3.62	3.55	2.11	Total: 13.78
Percent of explained variance		24.92%	20.13%	19.75	11.75	76.55%

* Denotes the most important contribution of the items to the four factors.

these independent variables can contribute to a single dimension result. The subscores also show the seeming independence of the cognitive sphere (S5, S6) from the motor subscores (S3 and S4).

Factor analysis of correspondences (FAC) is clearly favourable toward a heterogeneous nature of functional dependence measured by the scale items as well as the independence of the three subscores, sphincter, communication and social cognition compared with the other three subscores, namely self-care, mobility and locomotion, which constitute the main part of the FIM.

The orthogonal transformation matrix appeared to corroborate the results reported by Linacre et al. (44). The fact that the first factor accounted for 44.7% of the total variance would be satisfactory. The ratio between the first two eigenvalues, which expresses the emergence of a principal dimension favouring unidimensionality, was 2.35, a relatively modest figure (11). The first analysis thus confirmed one major point: there are at least two distinct dimensions within the phenomenon measured by the FIM.

Nevertheless, further analysis is required. If one accepts a threshold level of 0.3 for saturation coefficients (11), then it can be seen that most of the items in subscores S5 and S6 contribute to the first factor, which consequently is not a “pure” motor factor. In addition, several items contribute to different factors, for example in subscores S2 and S4. This may reflect the differential associations among the items depending on pathologies. Moreover, the first two factors can only weakly explain the communality of several items. This indicates that the analysis must involve more than two factors to account for the scale. Last but not least, when the analysis is continued with the rotation method, Varimax, the results are described even more explicitly (63). It depicts a four-dimensional FIM, although it must be recalled that the previous remarks indicate that these four distinct and coherent dimensions are not absolutely “pure” dimensions. The “dressing lower body” contributes highly to the first factor which groups activities involving the capability of using the lower limbs (locomotion, transfers). More surprisingly, the “toilet” item appears to be set apart from the “self-care” group and to come closer to the “transfer/locomotion” group; this may perhaps express a greater proximity between these items but it may also be an erroneous interpretation of this item as a non-exclusive transfer-toilet item (utilization of the toilet implying a transfer).

Thus the heterogeneous nature of the motor dimension, previously emphasized (16), becomes very clear, split into three components, one being the sphincter control subscore whose position within the motor items has already been questioned (62). The analysis of the correlation matrix (strong correlations between S3 and S4 and between S5 and S6) and the FAC (independence of S2 and S5-S6 on the main part of the scale) are in agreement with this analysis.

CONCLUSION

The FIM is widely used in the field of rehabilitation throughout Europe and North America with a total score resulting from the summation of the 18 items. This study confirms that the single overall score is not sustainable. Following some works carried out using the Rasch analysis, it would be tempting to accept a two-dimensional interpretation of the instrument. But, when further advancing the analysis, it becomes evident that more heterogeneity is involved. If the correlation matrix and the orthogonal factor solution are supporting the two-dimensional solution, the factor analysis with Varimax rotation suggests four distinct and coherent dimensions which approximately correspond to the FIM subscores.

These findings strongly suggest that a number of utilizations of the FIM, but also other similar scales, should be questioned. The main point is that neither the FIM nor the motor subscore are unidimensional and that practical conclusions must be drawn. It must be clearly understood that the limitation on a “valid” use of the FIM is not related to unavoidable approximations due to this type of instrument construct, but that it concerns the meaning of the measurement itself and the interpretations of the FIM results. After rigorous analysis, it must be stated that the total score does not describe a sole entity and does not measure any defined phenomenon.

In some situations, the instrument can be used as a descriptive tool, without using the total score, i.e. for the individual follow-up of a patient. The FIM can likewise be used as a statistical indicator if it has been proven, in a certain context and in a well-defined population, that the score is correlated with a well-defined criterion. Caution is required, however, when interpreting score differences (which would imply comparisons between scores and thus their meaning) and for comparisons between institutions since the scale cannot account for the differences between these structures.

Likewise, comparing the total scores (for a patient, between patients, between institutions) or subsequent ratios is ungrounded because they refer to a heterogeneous content which, to be interpreted, must call upon the observed complexity of the individual clinical case and the composite nature of dependence.

In making recommendations for clinical use, subscores can be considered, at least in a first approximation, as an expression of one aspect of the functional dependence. Using the subscores assumes, however, that a linear measure can be established by Rasch analysis, as was suggested by some authors (44). This conversion only partly lessens interpretation difficulties in terms of improved independence (52). Under these conditions, and based on our findings, four indicators issued from the FIM, or reorganized within the FIM, could be proposed:

–"self-care" indicator (the first 4 items)

–"overall body mobility" indicator (subscores 3 and 4: transfer and locomotion)

–"sphincter control" indicator (subscore 2): and

–"communication and social cognition" indicator (subscores 5 and 6)

Recognizing that the cognitive sphere of the FIM has provided unsatisfactory results when used alone (13, 38), and that additional studies are required to further define the roles of items 5 "dressing lower body" and 6 "toilets" (which, according to our data, belong to the second group : "overall body mobility").

ACKNOWLEDGMENTS

The authors extend their thanks to Béatrice Beaufils (University of Paris VIII) and Papa Abdou (Association des Paralysés de France) for their help with programming and data processing, and to Françoise Marchand for her assistance with translation. They also thank the anonymous referees for their useful comments.

REFERENCES

1. Beaufils, B. & Paicheler, H.: La dépendance chez les personnes âgées: que mesure-t-on exactement? Santé publique *6:* 23–30, 1992.
2. Benzecri, J.-P. & Benzecri, F. Pratique de l'analyse des données—Analyse des correspondances, exposé élémentaire. Dunod, Paris, 1980.
3. Boulanger, Y., Minaire, P. & Chantraine, A.: Mesure d'Indépendance Fonctionnelle (MIF), système uniforme de données pour la médecine de rééducation et réadaptation (SUDMERR). Guide d'utilisation pour la collecte des données de la MIF à l'aide du système "Uniform data set for medical rehabilitation", French translation, 1991 (copyright Research Foundation, State University of New York, 1987)..
4. Braun, S. L. & Granger, C. V.: A practical approach to functional assessment in paediatrics. Occup Ther Pract *2:* 46–51, 1991.
5. Brosseau, L., Philippe, P., Dutil, E. & Boulanger, Y. -L.: Mesure d'Indépendance fonctionnelle (M.I.F.), Recension des écrits. J Réadapt Méd *16:* 9–21, 1996.
6. Brosseau, L., Wolson, C. & Daoust, J.: The interrater reliability and construct validity of the Functional Independence Measure (FIM) for multiple sclerosis subjects. Clin Rehabil *8:* 107–15, 1994.
7. Brugel, D. G., Ribet-Reinhart, N., Saufier, P. & Laurent-Vannier, A.: Résultats préliminaires d'une application de la MIF-Mômes dans un service de rééducation pédiatrique. Ann Réadaptation Méd Phys *37:* 419–421, 1994.
8. Buia, J. M., Carlier, P. & Belkacemi, A.: Intérêt de la mesure d'indépendance fonctionnelle (MIF) dans un service de rééducation et de réadaptation gériatrique, Ann Réadaptation Méd Phys *38:* 435–441, 1995.
9. Bunch, W. H. & Dvonch, V. M.: The "value" of functional independence measure scores. Am J Phys Med Rehabil *73:* 40–43, 1994.
10. Calmels, P., Vedel, E., Bethoux, F., Charmet, E. & Minaire, P.: Mesure de l'indépendance fonctionnelle (MIF) : intérêt d'une utilisation téléphonique. Ann Réadaptation Méd Phys *37:* 469–476, 1994.
11. Ciadella, P., Guillaud-Bataille, J. M., Gausset, M. F., Terra, L., Gerin, P., Palliard, E. & Jouishomme, J. C.: Étude sur l'unidimensionnalité de l'échelle de dépression gériatrique de Yeasavage et Brink. L'Encéphale, *18:* 537–44, 1992.
12. Daverat, P., Petit, H., Kemoun, G., Dartigues, J. F. & Barat, M.: Suivi à long terme d'une cohorte de 188 blessés médullaires. Ann Réadaptation Méd Phys, *37, Suppl 1:* 14s, 1994.
13. Davidoff, G., Roth, E., Haughton, J. & Ardner, M.: Cognitive dysfunction spinal injury patient, sensitivity of the F.I.M subscales versus neuropsychological assessment. Arch Phys Med Rehabil, *71:* 326–329, 1990.
14. Dickson, H. G. & Köhler, F.: Interrater reliability of the 7-level functional independence measure (FIM). Letter to the editor. Scand J Rehab Med *27:* 253–256, 1995.
15. Dickson, H. G. & Köhler, F.: Correspondence. Disability and Rehabilitation *17:* 456, 1995.
16. Dickson, H. G. & Köhler, F.: The multi-dimensionality of the FIM motor items precludes an interval scaling using Rasch analysis. Scand J Rehabil Med *26:* 159–162, 1996.
17. Dijkers, M. P.: Developments of item weights for the Functional Independence Measure: pilot study. Abstract. Arch Phys Med Rehabil *69:* 755, 1988.
18. Dodds, T. A., Martin, D. P., Stolov, W. C. & Deyo, R. A.: A validation of the functional independence measure and its performance among rehabilitation inpatients. Am J Phys Med Rehabil *74:* 531–536, 1993.
19. Fourn, L., Brosseau, L., Dassa, C. & Dutil, E.: Validation factorielle de la Mesure d'Indépendance Fonctionnelle (M.I.F.) auprès de personnes atteintes de la Sclérose en Plaques (S.E.P.). J Réadapt Méd *14:* 7–16, 1994.
20. Gaito, J.: Measurement scales and statistics: resurgence of an old misconception. Psychol Bull *87:* 564–567, 1980.
21. Galin, L. C., Rode, G., Soler-Michel, P., Eyssette, M. & Boisson, D.: Intérêt de l'étude de la MIF dâns l'évaluation de la rééducation des sujets hémiplégiques âgés de plus de 60 ans. Ann Réadaptation Méd Phys *37:* 281–290, 1994.
22. Gallien, P., Sebert, E., Galanth, E., Brissot, R. & Perrigot, M.: Intérêt de la mesure d'indépendance fonctionnelle (MIF) dans la prise en charge de l'incontinence chez l'hémiplégique. Ann Réadaptation Méd Phys *38:* 21–24, 1995.
23. Gardent, H., Spinga, J. & Bounot, A. Dépendance des personnes âgées et charge de soins. Expérimentation

simultanée de six grilles de dépendance. CTNERHI, Vanves, 1988.

24. Gautheron, V. & Minaire, P. Système uniforme de données pour la médecine de rééducation et réadaptation (SUDMERR) incluant la Mesure de l'Indépendance Fonctionnelle pour enfants (MIF–MOMES), French translation, 1992, (copyright Research Foundation, State University of New York, 1991)..
25. Granger, C. V.: Mesure de l'Indépendance Fonctionnelle. *In* Actes des 6° entretiens de l'Institut Garches, pp 13–20. Arnette, Paris, 1993.
26. Granger, C. V., Cotter, A. C., Hamilton, B. B. & Riddler, R. C.: Functional assessment scales: a study of persons with multiple sclerosis. Arch Phys Med Rehabil *71:* 870–874, 1990.
27. Granger, C. V. & Hamilton, B. B.: Measurement of stroke rehabilitation outcome in the 1980s. Stroke *21, Suppl II:* 1146–1147, 1990.
28. Granger, C. V. & Hamilton, B. B.: The Uniform Data System for medical rehabilitation report of first admissions for 1990. Am J Phys Med Rehabil *71:* 33–38, 1992.
29. Granger, C. V. & Hamilton, B. B.: The Uniform Data System for medical rehabilitation report of first admissions for 1991. Am J Phys Med Rehabil *72:* 108–113, 1993.
30. Granger, C. V., Hamilton, B. B., Gresham, G. E. & Kramer, A. A.: The stroke rehabilitation outcome study: part II. Relative merits of the total Barthel Index score and a four-item subscore in predicting patient outcomes. Arch Phys Med Rehabil *70:* 100–103, 1989.
31. Haley, S. M., Coster, W. J. & Binda-Sunberg, K.: Measuring physical disablement: the contextual challenge. Physical Therapy *74:* 443–451, 1994.
32. Hall, K. M., Hamilton, B. B., Gordon, W. A. & Zasler, N. D.: Characteristics and comparison of functional assessment indices: Disability Rating Scale, Functional Independence Measure and Functional Assessment Measure. J Head Trauma Rehabil *8:* 60–74, 1993.
33. Hamilton, B. B. & Granger, C. V.: Totaled functional score can be valid. Letter to the editor. Arch Phys Med Rehabil *70:* 861–862, 1989.
34. Hamilton, B. B., Granger, C. V., Sherwin, R. S., Zielezny, M. & Lashman, J. S.: A uniform national data system for medical rehabilitation. *In* Rehabilitation Outcomes, Analysis and Measurement (ed. M. J. Further), pp. 137–147. P. H. Brookes, Baltimore, 1987.
35. Hamilton, B. B., Laughlin, J. A., Fielder, R. & Granger, C. V.: Interrater reliability of the 7–level functional independence measure (FIM). Scand J Rehab Med *26:* 115–119, 1994.
36. Hamilton, B. B., Laughlin, J. A., Granger, C. V. & Kayton, R. M.: Interrater agreement of the seven level Functional Independence Measure (FIM). Abstract. Arch Phys Med Rehabil *72:* 790, 1991.
37. Heineman, A. W., Linacre, J. M. & Wright, B. D.: Prediction of rehabilitation outcomes with disability measures. Arch Phys Med Rehabil *75:* 133–153, 1994.
38. Heineman, A. W., Linacre, J. M., Wright, B. D., Hamilton, B. B. & Granger, C. V.: Relationships between impairment and physical disability as measured by the Functional Independence Measure. Arch Phys Med Rehabil *74:* 566–573, 1993.
39. Keith, R. A., Granger, C. V., Hamilton, B. B. & Fielder, R. C.: The functional independence measure: a new tool for rehabilitation. *In* Advances in Clinical Rehabilitation (ed. M. G. Eisenberg & R. C. Grzesiak), pp. 6–18. Springer-Verlag, New York, 1987.
40. Kidd, D., Stewart, G., Baldry, J., Johnston, J., Rossiter, D., Petruckevetch, A. & Thompson, A. J.: The Functional Independence Measure: a comparative validity and reliability study. Disability and Rehabilitation *17:* 10–14, 1995.
41. Krebs, D. E.: Measurement theory. Phys Ther *67:* 1834–1839, 1987.
42. Letessier, J.: Un autre regard sur l'évaluation en rééducation. J Réadapt Méd *14:* 17–22, 1994.
43. Linacre, J. M. FACETS Computer program for many-faceted Rasch analysis. MESA Press, Chicago, 1989.
44. Linacre, J. M., Heinemann, A. W. & Wright, B. D.: The structure and stability of the Functional Independence Measure. Arch Phys Med Rehabil *75:* 127–132, 1994.
45. Linacre, J. M., Heinemann, A. W., Wright, B. D., Granger, C. V. & Hamilton, B. B. Research Report 91–01. Rehabilitation Service Evaluation Unit, Rehabilitation Institute of Chicago, Chicago, 1991.
46. Long, W. B., Sacco, W. J. & Coombes, S. S.: Determining normative standards for functional independence measure transitions in rehabilitation. Arch Phys Med Rehabil *75:* 144–148, 1994.
47. Luquet, C., Chau, N., Nadif, M., Guillemin, F., Gavillot, C., Petry, D., Moreau, T., Bourgkard, E., Hecquet, B., André, J. M. & Mur, J. M.: Unidimensionnalité d'une mesure fonctionnelle pour des malades ayant une lésion du membre supérieur. Rev Epidém et Santé Publ *44:* 248–261, 1996.
48. MacArthur, D. L., Cohen, M. J. & Schandler, S. L.: Rasch analysis of functional assessment scales: an example using pain. Arch Phys Med Rehabil *72:* 296–304, 1991.
49. Mahoney, F. I. & Barthel, D. W.: Functional evaluation: the Barthel Index. Maryland State Medical Journal *14:* 61–65, 1965.
50. Marolf, M. V., Vaney, C., König, N., Schenk, T. & Prosiegel, M.: Evaluation of disability in multiple sclerosis patients: a comparative study of the Functional Independence Measure, the extended Barthel Index and the expanded Disability Status Scale. Clin Rehabil *10:* 309–313, 1996.
51. Mellenbergh, G. J. & Vijn, P.: The Rasch model as a loglinear model. Applied Psychological Measurement *5:* 269–376, 1981.
52. Merbitz, C. H., Morris, J. & Grip, J. C.: Ordinal scales and foundations of misinference. Arch Phys Med Rehabil *70:* 308–312, 1989.
53. Merbitz, C. H., Morris, J. & Grip, J. C.: The authors reply to Dr. Johnston/to Drs. Hamilton and Granger. Arch Phys Med Rehabil *70:* 861–862, 1989.
54. Michels, E.: Measurement in physical therapy. On the rules for assigning numerals to observations. Physical Therapy *63:* 209–215, 1983.
55. Minaire, P.: La mesure d'indépendance fonctionnelle (MIF): historique, présentation, perspectives. J Réadapt Méd *11:* 168–174, 1991.
56. Minaire, P.: L'évaluation globale mesurable en médecine de rééducation et réadaptation. *In* Actes des 6° entretiens de l'Institut Garches, pp. 3–12. Arnette, Paris, 1993.
57. Mortifee, R. S., Busser, J. & Anton, H. A.: The performance of a limited set of items from the Functional Independence Measure for use in acute trauma care and rehabilitation. Arch Phys Med Rehabil *77:* 436–439, 1996.
58. Muecke, L., Sheka, R. S., Dwyer, D., Israel, E. & Flynn, J. P. G.: Functional screening of lower-limb amputees: a role in predicting rehabilitation outcome? Arch Phys Med Rehabil *73:* 851–858, 1992.
59. Rasch, G.: Probabilistic models for some intelligence and attainment tests. Danish Institute for Educational Research,

Copenhagen, 1960, and University of Chicago Press, Chicago, 1980.

60. Roth, E., Davidoff, G., Haughton, J. & Ardner, M.: Functional assessment in spinal injury: a comparison of the modified Barthel Index and adapted Functional Independence Measure. Clin Rehabil *4:* 277–285, 1990.
61. SAS Institute : SAS/STAT User's guide. Version 6, 4th ed. SAS institute Inc, Cary NC, 1989.
62. Silverstein, B. S., Fisher, W. P. & Kilgore, K. M. et al: Applying psychometric criteria to functional assessment in medical rehabilitation: II. Defining interval measures. Arch Phys Med Rehabil *73:* 507–518, 1992.
63. Silverstein, B. S., Kilgore, K. M., Fisher, W. P., Harley, J. P. & Harvey, R. F.: Applying psychometric criteria to functional assessment in medical rehabilitation: I. Exploring unidimensionality. Arch Phys Med Rehabil *72:* 631–637, 1991.
64. Stineman, M. G., Shea, J. A., Jette, A., Tassoni, C. J., Ottenbacher, K. J., Fiedler, R. & Granger, C. V.: The Functional Independence Measure: tests of scaling assumptions, structure, and reliability across 20 diverse impairment categories. Arch Phys Med Rehabil *77:* 1101–1108, 1996.
65. Wagner, M. T. & Zucchicagna, L. J.: Longitudinal comparison of the Barthel and FIM during the first six months of recovery from stroke. Abstract. Arch Phys Med Rehabil *69:* 755, 1988.
66. Wright, B. D. & Linacre, J. M.: Observations are always ordinal, measurements, however, must be interval. Arch Phys Med Rehabil *70:* 857–860, 1989.
67. Wright, B. D. & Linacre, J. M. BIGSTEPS, a Rasch-model computer program. MESA Press, Chicago, 1991.

Accepted March 4, 1998

Address for offprints:

Dr. Jean-François Ravaud
CERMES - INSERM U 502
Institut Fédératif de Recherche sur le Handicap
182, Boulevard de la Villette
FR-75019 Paris
France