J Rehabil Med 2006; 38: 50–56

Taylor & Francis
Taylor & Francis Group

DEVELOPMENT OF A SHORT VERSION OF THE MOTOR FIM™ FOR USE IN LONG-TERM CARE SETTINGS

Shin Yamada, MD, Meigen Liu, MD, PhD, Kimitaka Hase, MD, PhD, Naofumi Tanaka, MD, PhD, Toshiyuki Fujiwara, MD, PhD, Tetsuya Tsuji, MD, PhD and Jun-ichi Ushiba, PhD

From the Department of Rehabilitation Medicine, Keio University School of Medicine, Shinjyuku, Tokyo, Japan

Objective: **To develop a short version of the motor Functional Independence Measure (FIM™) for use in long-term care settings.**
Participants: **For model construction, the participants were 398 community-dwelling persons with disability (mean age 79.3 years (SD 10.3)) who were receiving visiting nurse services. For cross-validation, 169 patients with stroke (mean age 78.0 years (SD 11.2)) in the chronic phase and 187 patients with stroke (mean age 63.4 years (SD 12.7)) in the recovery phase.**
Design: **Model construction and cross-validation study.**
Main outcome measures: **The second power of correlation coefficient (R^2) was used for agreement analysis between the short and the full version. Cross-validation of the models was estimated with the intraclass correlation coefficient (ICC).**
Results: **Five to 7 motor FIM™ items were selected for the models based on Rasch calibration and consideration of internal consistency. Total motor FIM™ was estimated with the 6-item and 7-item models with regression analysis, which yielded high correlations with the original 13-item motor FIM™ score ($R^2 > 0.95$). Regression formulas derived from the models could estimate total motor FIM™ scores accurately in the 2 cross-validation samples (ICC >0.98).**
Conclusion: **The short version of the motor FIM™ developed is a useful measure of functional status, not only in long-term care but in the recovery phase rehabilitation settings.**

Key words: instrument, activities of daily living, stroke, Rasch analysis, community-based rehabilitation.

J Rehabil Med 2006; 38: 50–56

Correspondence address: Shin Yamada, Department of Rehabilitation Medicine, Keio University School of Medicine, 35 Shinanomachi, Shinjyuku, Tokyo, Japan. E-mail: shin@sc.itc.keio.ac.jp

Submitted October 21, 2004; accepted June 1, 2005

INTRODUCTION

With ageing societies, the number of persons who are in need of care is increasing. In Japan, together with health insurance plans that cover acute and recovery phase rehabilitation services, a nation-wide public insurance program called the Public Long Term Care Insurance Program was started in 2000 to cover care and rehabilitation needs after completion of active medical treatment (1). To ensure consistency and continuity of care, it is important for health professionals involved at various phases to have a common language to describe functioning of the patients.

In medical rehabilitation, the Functional Independence Measure (FIM™) is widely used to document patients' functional status and its changes (2). The FIM™ includes 13 motor and 5 cognitive items, and the scoring ranges from 1 (complete dependence) to 7 (complete independence). Originally developed as a unified instrument to evaluate disabilities as a part of a large rehabilitation database called the Uniform Data System for Medical Rehabilitation (UDSMR) (3), it has been shown to be a reliable, valid, practical and responsive instrument to describe functional status at admission, discharge and follow-up for various disabilities (4–7).

Despite its established usefulness in inpatient rehabilitation settings, the FIM™ has the following limitations when used in long-term care settings: (i) it may not be adequate for assessing outpatient rehabilitation outcomes due to the higher levels of functioning and additional areas and domains of importance seen in the outpatient settings (8); (ii) although post-discharge follow-up FIM™ scores are typically obtained by telephone or in-person interview (9), the reliability and validity of the FIM™ in subacute and home health settings have not been well established (6), especially for the cognitive subscale (10); (iii) it is often difficult to carry out full assessment with the FIM™ at home where time is limited for the raters, because it takes approximately 20–30 minutes to complete even for trained assessors (11, 12); (iv) it is often necessary to obtain information from family members to get a complete picture of a person's activities of daily living, and it can be time-consuming to interview carers, especially when they are themselves aged; (v) it is costly and time-consuming to train visiting nurses and carers who are not familiar with the FIM™ reliably to assess the full version of the FIM™.

Thus it has not been practical to use the FIM™ in long-term care. In the USA, the Minimum Data Set (MDS) (13) has been used widely in nursing homes, and an attempt has been made to bridge the gap between acute rehabilitation and long-term care by developing a pseudo-FIM by selecting and re-scaling 12 items from the MDS that corresponded to the FIM™ items (14). However, the MDS is not widely used in Japan, and no studies are available examining reliability and validity of the pseudo-FIM among the Japanese population. Although 2 studies are reported describing short versions of the FIM™, 1

ISSN 1650-1977
DOI: 10.1080/16501970510044034

for acute trauma care (15) and the other for spinal cord injury (16), no attempt has yet been made to develop one for use in long-term care settings.

Because disability assessment is indispensable for planning care services, predicting outcomes and tracking changes in functional status, there is a strong need for a standardized common scale that can be used practically in long-term care in continuity with acute and recovery phase rehabilitation. The purpose of this study is therefore to develop and cross-validate a minimum set of the FIM™ motor items that can reliably and accurately estimate total motor FIM™ scores in long-term care settings.

METHODS

Although the FIM™ consists of 13 motor and 5 cognitive items, we used only the 13 motor items because: (i) it is known with Rasch analysis (17) that the FIM™ conforms to a unidimensional model if the motor and cognitive subscales are analysed separately (6); (ii) somewhat lower test-retest reliability is reported for the cognitive items than for the motor items in aged persons living in the community (10); (iii) a higher degree of experience is necessary reliably to assess cognitive function with the FIM™ (9, 10). Thus we excluded cognitive items to simplify model construction. We did not consider reducing the number of categories for scaling, because this would make the instrument less responsive to changes. We intended to make the assessment less laborious and more practical to administer in long-term care settings by reducing the number of items.

This study involved the following 3 steps: (i) construction of 3 models consisting of 5–7 motor FIM™ items by analysing the structure of functional status of community-dwelling people with disability and selecting proper items from the 13 motor FIM™ items; (ii) analysis of the performance of the above model subsets to predict the original 13-item motor FIM™; (iii) cross-validation of the models by applying them to different samples.

Participants

For model building, we originally recruited 1710 community-dwelling elderly persons who were receiving visiting nurse services covered by the Public Long-term Care Insurance Program from 32 visiting nurse service stations (11 in Tokyo, 3 in Sendai and the remaining 18 stations in Hokkaido, Kanto, Tokai and Kansai areas) belonging to a same provider group (SECOM Co. Ltd). This was because the services provided were more standardized across the stations and it was easier to assure uniformity of assessment through periodic training sessions. A total of 127 visiting nurses belonging to these stations, who had been well trained in the FIM™ assessment in advance, collected FIM™ data about their clients from December 2003 to January 2004. The Japanese version of the FIM™ has culturally relevant modifications for some of the items. The principal modification is in eating. Use of a spoon instead of chopsticks does not lower the score (18). Before data collection, the purpose and procedures were explained to the clients and their family carers, and written informed consent was obtained. After excluding patients who refused to participate in the study (1174) and patients receiving terminal care (80), 456 patients were enrolled (group M). Among them, 58 patients were excluded because of incomplete data, and the final sample comprised 398 patients (168 men and 230 women) with a mean age of 79.3 (SD 10.3) years and mean length of service period of 514.5 (SD 404.5) days (median 440 days). Among them, 256 patients suffered from stroke, 132 from diseases of internal organs, such as chronic heart failure or diabetes, and 93 from bone and joint diseases (duplicates permitted).

For cross-validation of the model developed, we used data from 2 samples of patients with stroke. One was a group of patients in long-term care settings (group L). Trained rehabilitation professionals assessed the FIM™ cross-sectionally in 169 patients recruited from 6 participating institutions including 1 long-term care hospital ward, 1 general ward, 2 visiting rehabilitation service facilities and 2 health service facilities for the elderly (68 men, mean age 78.0 (SD 11.2) years, mean duration of stroke 1337.2 (SD 1491.9) days (median 843 days)). There were 122 with cerebral infarction, 37 with cerebral haemorrhage and 10 with subarachnoid haemorrhage. Fifty-six patients had right brain damage, 61 had left brain damage and 52 had bilateral or multiple lesions. The second sample consisted of patients with stroke hospitalized for recovery phase rehabilitation. The admission (group A) and discharge (group D) FIM™ data of 187 consecutive patients (98 males) admitted to Tsukigase Rehabilitation Center, one of the affiliated hospitals of Keio University, from May 1998 to August 2001 were available as a part of a structured rehabilitation database, and these data were used for analysis. The mean age of the patients was 63.4 (SD 12.7) years, the mean time from onset to admission was 44.1 ± (SD 23.4) days (median 42 days), and mean length of stay was 99.1 (SD 52.6) days (median 95 days). One hundred suffered from cerebral infarction, 75 from cerebral haemorrhage and 12 from subarachnoid haemorrhage. Eighty-two patients had right brain damage, 88 had left, 5 had brainstem and 12 had bilateral or multiple lesions.

Model building

To reduce the number of motor FIM™ items, we selected 5–7 items based on a statistical approach and clinical judgement. Because it is important for a good instrument to have its item difficulty spread at equal intervals, we performed Rasch analysis using the data obtained from the model-constructing sample to evaluate item difficulty levels. As a result of this calibration, we thinned out items shown to have closer difficulty levels. Rasch analysis is a specific item-response theory technique to investigate the difficulty level of items included in a scale (19). An output parameter of Rasch calibration called "logit", which is allocated to scale items and individual subjects, means the relative difficulty level among them. Ideally, the separation gap between each item is 0.15 logits or more (20). Including both extremes of the difficulty levels, we selected 5 items whose logit values were close to the ideally separated points so as to ensure maximum distribution. For items exhibiting similar difficulty levels, we selected items for the model based on our clinical judgement of their importance in rehabilitation practice.

Next, we added 1 or 2 items to the 5-item subset to reinforce internal consistency based on consideration of the 4 subcategories of the motor FIM™ items (i.e. self-care, sphincter control, mobility, locomotion). Thus, 3 subsets consisting of 5–7 items were constructed.

Analysis of agreement between the 13-item motor FIM™ and the short subsets

To investigate agreement between the 13-item motor FIM™ and the short subsets, we used multivariate regression analysis and secondary Rasch calibration.

Regression analysis. The 13-item motor FIM™ scores (range 13–91 points) were estimated from the 5–7-item subset scores using multivariate regression analysis (21). The dependent variable in the equation was the actually measured 13-item motor FIM™ score and the independent variables were individual 5–7 item subset scores. We calculated coefficient of determination (R^2) as an index of agreement between the original simple summation of the 13 items and the estimated total score.

Secondary Rasch calibration. Although summation of item scores has been widely used for research and clinical practice, this is not theoretically adequate because the FIM™ is essentially an ordinary scale. To avoid this theoretical contradiction inherent in the regression approach, we used an additional method to estimate total motor FIM™ score by converting it into an interval scale. Using 5–7-item scores, we performed secondary Rasch calibration to derive individual logit score. The logit score reflects the relative level of functional independence in the group and can be adjusted to optimal point scale linearly, and handled as an interval scale (20). We reconstructed the 13–91 point interval scale from logit scores for comprehensive and easy comparison with the regression approach. Correlation coefficient (R) between reconstructed score derived from "primary" full 13-item and "secondary" limited-number-item Rasch calibrations in the same subject was calculated. To compare the accuracy of the 2 methods of estimation, the second power of correlation coefficient ($R^{2'}$) was raised.

Cross-validation studies

For cross-validation, the performance of the models developed was evaluated in the long-term care sample (group L) and the admission and discharge data of the recovery phase rehabilitation sample (group A and D). Total FIM™ score was estimated from the 5–7 item scores for each regression formulas derived from the model-constructing sample. Rasch calibration for subset score was performed to estimate total FIM™ score in the same way as model building agreement analysis. Reliability of the subsets was assessed by calculating intraclass correlation coefficients (ICC 3.1) (22).

Statistics

We performed Rasch calibration using a statistical software BIG-STEPS^{TR} (version 2.82 for DOS). ICC was figured out with a macro function written by one of the authors (SY) for Excel^{TR} (version 2002 for Windows^{TR}). Other statistical calculations including regression analysis were performed with Statview^{TR} (version 5.0 for Windows^{TR}).

RESULTS

The characteristics of our samples are listed in Table I. The mean age of the patients for model construction was higher than those of the stroke patients in the recovery phase (ANOVA, $p < 0.001$). The total motor FIM™ score was lower for this sample ($p < 0.001$). Differences in age and total FIM™ score between the model construction sample and the long-term care sample were not significant. Patients with stroke in the recovery phase were younger ($p < 0.001$) and total motor FIM™ score improved by approximately 6 points during inpatient rehabilitation.

Table II shows the results of primary Rasch calibration. The 13 motor FIM™ items were ordered according to their logit scores. Items at the negative end of the scale were considered "easier" and items at the positive end were regarded as more "difficult". "Feeding" was the easiest (−0.58 logits) and "Stairs" was the most difficult item (0.65 logits) for the model construction sample. The difficulty pattern corresponded to that observed in our previous study (18). Each item fitted to the Rasch model acceptably except bladder and bowel management items whose mean squares were more than 1.3.

The result of item selection for the subsets is illustrated in Fig. 1. Five items (Feeding, Bathing, Dressing lower-body, Bed/Chair/Wheelchair transfer, Stairs), which were located closely to the ideal distribution represented by 5 lines dividing logit range equally, were selected. "Grooming" and "Dressing upper-body" were excluded because carers had a tendency to help with these activities to save time. "Toilet transfer" was omitted because it depended considerably on circumstances in homecare settings in Japan. We added "Bladder management" to the above 5-item subset to cover the sphincter subcategory. Considering the clinical importance of locomotive function, "Walking/Wheelchair" was adopted for the 7-item subset. For this item, although walking and wheelchair abilities were assessed separately, we adopted either of the more commonly used ones for the patient as the final score following the UDSMR guideline (3).

The results of total score estimation are illustrated in Fig. 2. The subsets fit the regression model excellently (Fig. 2(a)). Multi-colinearity was not observed. Maximum correlation coefficient between variables was 0.90 (between 13 item total and Feeding), and the maximum variance inflation factor derived from this figures was 5.263 (< 10). Analysis of variance with the variables showed acceptable p values (< 0.001). Regression analysis revealed that evaluation with 5 and more motor FIM™ items could predict the total score accurately ($R^2 > 0.95$). Scores derived from primary and secondary Rasch calibrations indicated linear distribution (Fig. 2(b)). $R^{2'}$ was as high as that derived from regression analysis with the 6-and-7 item subsets ($R^{2'} > 0.95$). The 5-item subset had lower correlation with the original estimation compared with other subsets ($R^{2'} = 0.927$).

The results of cross-validation studies are summarized in Table III. The subsets estimated total motor FIM™ scores accurately in the 2 cross-validation samples, particularly those derived from regression adjustment (ICC > 0.98). In general, estimation with Rasch calibration tended to show lower ICC values and the ICCs were higher with greater numbers of items.

DISCUSSION

Previous to this study, 2 other studies were available describing a short version of the FIM™. Mortifee et al. (15) reported a limited version of the FIM™ for patients with acute trauma. The set consisted of 3 items (Feeding, Walking, and 1 of the cognitive items "Expression"), and the scaling was simplified to 4 levels from the original 7 levels (7 = 4, 6 = 3, 5/4/3 = 2, 2/1 = 1). However, the set had poor consistency with the original FIM™ (ICC = 0.11) and was not useful practically. Dijkers & Yavuzer (16) developed another short version of the motor FIM™ for use in patients with spinal cord injury. They used 5 strategies to reduce the number of the motor FIM™ items from 13 to 5–7: random, coefficient alpha maximization, spread across the range of item difficulties, optimization by neurological category and individual optimization. The best performance was achieved by individual optimization 7-item subsets that selected the best-fit 7 items according to the disability level of each patient. The ICC between the estimated and the original data was > 0.98, and they concluded that the short version based on this algorithm approach was reliable and useful. However, there are several drawbacks with this algorithm approach for practical use in long-term care settings. First, the target population is different. Stroke occupies a significant proportion in long-term care instead of spinal cord injury. Secondly, understanding of the rating system of all the 13 motor FIM™ items is required to use the individual optimization 7-item subsets, which renders their model less practical.

Our study is the first to develop a short version of the FIM™ for use in long-term care settings. We demonstrated satisfactory performance of the short subsets; the ICCs in our model were as high as those of a previous study (16). The items included in our model are fixed in contrast to the individual optimization 7-item subsets adopted by Dijker & Yavuzer (16) that require selecting the best-fit 7 items according to the disability level

Table I. *Data for FIM items of the model construction and the cross-validation samples*

Group	M: model constructing			L: long-term care		A: recovery phase (admission)			D: recovery phase (discharge)	
n	398 (male 168, female 230)			169 (male 68, female 101)		187 (male 98, female 89)			187 (male 98, female 89)	
Age* (years)	79.3 ± 10.3			78.0 ± 11.2		63.4 ± 12.7			63.4 ± 12.7	
Motor FIM items	Mean ± SD	Median	IQR	Mean ± SD	Median	IQR	Mean ± SD	Median	IQR	Mean ± SD
Self-care										
Feeding	4.18 (2.32)	5.00	2.50	4.94 (2.23)	5.00	1.50	5.34 (1.63)	5.00	1.00	6.02 (1.48)
Grooming	3.48 (2.50)	2.00	2.50	3.88 (2.37)	4.00	2.50	5.04 (1.98)	5.00	1.50	5.86 (1.81)
Bathing	3.29 (1.85)	3.00	1.88	2.81 (2.14)	2.00	1.50	3.11 (1.82)	3.00	1.50	4.46 (2.08)
Dressing upper body	3.29 (2.26)	3.00	2.00	3.36 (2.23)	3.00	2.00	4.14 (2.34)	4.00	2.00	5.55 (2.04)
Dressing lower body	2.78 (2.29)	1.00	2.00	3.20 (2.28)	3.00	2.00	3.81 (2.42)	4.00	2.50	5.27 (2.23)
Toileting	3.36 (2.51)	2.00	2.50	3.51 (2.35)	3.00	2.50	3.86 (2.38)	4.00	2.50	5.24 (2.15)
Sphincter control										
Bladder management	3.41 (2.40)	3.00	2.00	3.92 (2.59)	4.00	3.00	4.52 (2.50)	5.00	2.50	5.39 (2.23)
Bowel management	3.51 (2.37)	3.00	2.50	4.08 (2.51)	5.00	2.50	5.05 (2.21)	6.00	1.50	5.64 (1.84)
Mobility										
Bed/Chair/Wheelchair transfer	3.78 (2.43)	4.00	2.50	4.00 (2.38)	5.00	2.50	4.37 (1.92)	5.00	1.50	5.57 (1.70)
Toilet transfer	3.68 (2.44)	3.00	2.50	3.85 (2.34)	4.00	2.50	4.34 (1.94)	5.00	1.50	5.49 (1.75)
Tub transfer	3.11 (2.29)	2.00	2.50	2.86 (2.02)	3.00	2.00	3.59 (1.79)	4.00	1.50	4.63 (1.84)
Walk/Wheelchair	2.95 (2.22)	2.00	2.00	3.90 (2.40)	5.00	2.50	2.69 (2.26)	1.00	2.00	5.27 (1.83)
Locomotion										
Stairs	2.41 (2.00)	1.00	1.50	2.22 (1.82)	1.00	1.00	1.69 (1.75)	1.00	0.00	3.45 (2.45)
Total score*	43.2 (25.1)	39.00	23.00	46.5 (26.0)	49.00	26.00	51.5 (22.7)	5.00	1.00	67.9 (22.4)

IQR = inter quartile range; SD = standard deviation.
*: $p < 0.001$ ANOVA.

Table II. *Results of the 13-item Motor FIM Rasch Calibration*

Item	Logit value	SE	MNSQ
Stairs	0.65	0.05	1.24
Dressing lower body	0.37	0.04	0.69
Walking/Wheelchair	0.25	0.05	1.22
Tub transfer	0.14	0.04	0.80
Dressing upper body	0.02	0.04	0.92
Bathing	0.02	0.04	1.11
Toileting	−0.03	0.04	0.86
Bladder management	−0.06	0.05	1.36*
Grooming	−0.10	0.04	1.13
Bowel management	−0.13	0.05	1.48*
Toilet transfer	−0.24	0.04	0.59
Bed/chair/WC transfer	−0.31	0.04	0.63
Feeding	−0.58	0.04	0.99
Mean (SD)	0.00 (0.30)	0.04 (0.00)	1.00 (0.27)

SE = standard error; SD = standard deviation; MNSQ = mean square variance ratio statistic (infit).
*: Misfit >1.3.

of each patient. This renders it easier for the rater to master and use the instrument, which is particularly beneficial in home care settings where time for evaluation is limited. Furthermore, the evaluation results can be easily converted to 13-item motor FIM™ scores using the Rasch model or the regression formulae.

With the regression approach, total FIM™ score was predicted accurately using only the 5-item score. Accuracy of the estimation with Rasch calibration was lower than that of the regression approach. Some concern remains regarding the application of the Rasch model. In cases of extremely dependent or independent persons, estimated FIM™ scores using Rasch calibration become inaccurate because of its theoretical feature (21). As shown in Fig. 2(b), both ends of the plots tend to be out of the ideal linear distribution compared with those of regression analysis. $R^{2'}$ for the 5-item estimation was lower than 0.95. Judging from these observations, at least 6 or more items seemed necessary to be able to satisfactorily describe heterogeneity of patients. Although we built a 7-item subset to secure higher consistency, no remarkable difference was observed in the evaluation accuracy between the 6-item and the 7-item subsets. Therefore, we suggest the 6-item subset as a practical solution (Feeding, Bathing, Dressing lower-body, Bladder management, Bed/Chair/Wheelchair transfer, Stairs).

The performance of the short subset was cross-validated in 2 independent samples, 1 with patients with mixed disabling conditions in the chronic phase and the other with patients with stroke in the recovery phase rehabilitation. In particular, scores adjusted with Rasch calibration showed superior correlation with those in the model building procedure. This was presumably because the influence of the floor effect was minimal in patients with stroke. Our model could therefore be used to document functional status consistently from the recovery phase to the community phase, and the regression formula we described would be a great help to compare short subset scores with the fully assessed 13-item scores. Rasch calibration would be of use when a strict interpretation of scaling is required.

There are several limitations in our study. Because we could only obtain data from a small cluster of the people (398 of 1710 persons), the first limitation concerns with a possible selection bias for the model-building sample. The relatively high refusal rate could be explained partly by the fact that in Japan there is still a tendency to hide persons with disability from society, particularly among the aged population. The representativeness of the sample should therefore be interpreted with caution.

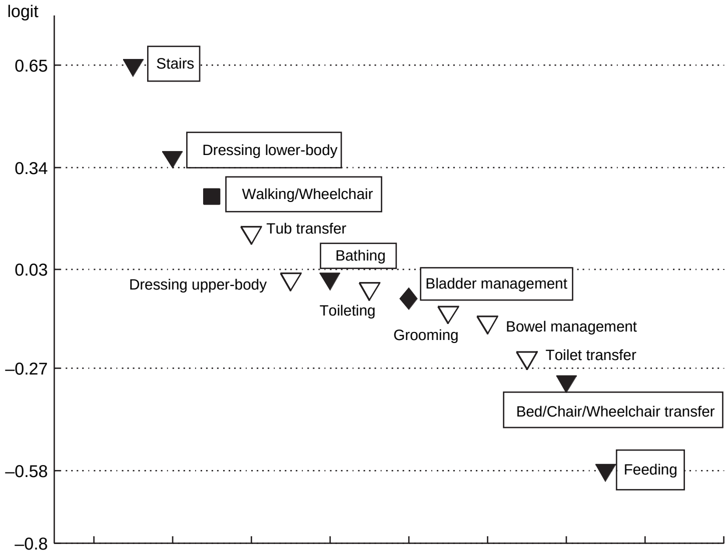

Fig. 1. Spread of item weights and item selection. ▼: 5-item subset; ◆: additional item for 6-item subset; ■: item adopted to 7-item subset; ▽: others. Horizontal dotted lines divides range of logit score into 5 with equal intervals.

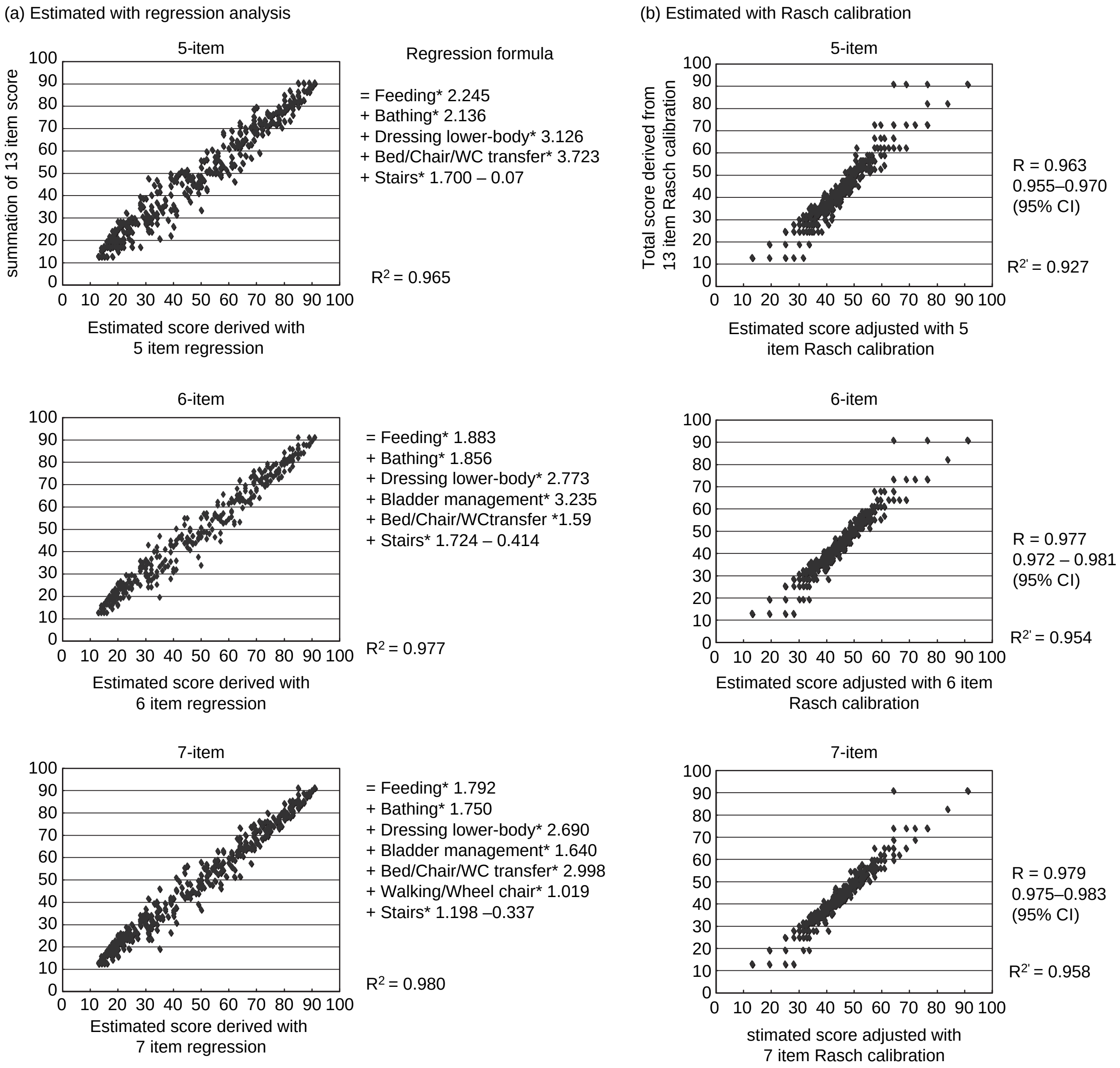

Fig. 2. Results of agreement analysis between the original 13-item motor FIM score and the estimated motor FIM score with the subsets. Scatter plots depicting the relationship between the original 13-item motor FIM score and estimated total FIM score using the subset score. A = simple summation of 13 items and estimated total FIM score derived from regression formulas with 5 to 7 items; B = total FIM scores adjusted with 13 item and 5–7 item Rasch calibration; R^2 = coefficient of determination; R = correlation coefficient; CI = confidence interval; $R^{2\prime}$: 2nd power of correlation to coefficient.

Secondly, we did not examine the influence on the remaining items of removing items when developing short subsets. When some of the items in the original scale are no longer administered, it can affect how the rater scores the remaining items. In future studies, we need to examine validity and reliability of the short versions *per se*.

Thirdly, the number of items (6) we recommended is not necessarily minimum. Focusing on specific population may make it possible further to reduce the number of items needed for reliable estimation.

Finally, although we demonstrated that our model could be used not only in the community phase but also in the recovery phase of rehabilitation, our study was limited because of its cross-sectional nature. In general, a shorter version is not suitable for catching trivial changes, and we need to investigate its responsiveness to changes over time in future longitudinal studies.

Despite these limitations, we consider that the short version of the motor FIM™ we developed is a simple and useful measure of functional ability, not only in long-term care but also in recovery phase rehabilitation settings. It is easier to master and less time-consuming to administer than the full version of the motor FIM™. Based on an evaluation with a common scale, more integrated rehabilitation interventions from the acute to the community phase would become possible.

Table III. *Correlation between the estimated and original total FIM score*

	Subset					
	5-item		6-item		7-item	
Adjustment	ICC	95% CI	ICC	95% CI	ICC	95% CI
Rasch						
group M	0.958	0.969–0.979	0.975	0.973–0.982	0.979	0.980–0.989
group L	0.977	0.969–0.983	0.986	0.981–0.989	0.989	0.985–0.992
group A	0.974	0.965–0.980	0.978	0.970–0.983	0.987	0.982–0.990
group D	0.964	0.952–0.973	0.971	0.961–0.978	0.981	0.974–0.985
Regression formulae						
group M	0.982	0.978–0.985	0.988	0.985–0.990	0.990	0.987–0.991
group L	0.985	0.980–0.989	0.994	0.992–0.995	0.995	0.993–0.996
group A	0.980	0.973–0.984	0.987	0.982–0.990	0.991	0.987–0.993
group D	0.987	0.982–0.990	0.990	0.987–0.992	0.994	0.991–0.995

Group M = model constructing; L = long-term care; A = recovery phase (admission); D = recovery phase (discharge); ICC = intraclass correlation coefficient; CI = confidence interval.

ACKNOWLEDGEMENTS

This work was supported by a grant from the Ministry of Health, Labor and Welfare, Japan (grant no. 05-046-0310).

REFERENCES

1. Ikegami N, Yamauchi K, Yamada Y. The long term care insurance law in Japan: impact on institutional care facilities. Int J Geriatr Psychiatry 2003; 18: 217–221.
2. Heinemann AW, Linacre JM, Wright BD, Hamilton BB, Granger CV. Prediction of rehabilitation outcomes with disability measures. Arch Phys Med Rehabil 1994; 75: 133–143.
3. Guide for the Uniform Data Set for Medical Rehabilitation (Adult FIM™). Version 5.0. Buffalo (NY): State University of New York at Buffalo; 1996.
4. Hamilton BB, Granger CV, Sherwin FS, Zielezny M, Tashman JS. A uniform national data system for medical rehabilitation. In: Fuhrer MJ, ed. Rehabilitation outcomes: analysis and measurement. Baltimore: Brookes; 1987, pp. 137–147.
5. Dodds AT, 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.
6. Heinemann AW, Linacre JM, Wright BD, Hamilton BB, Granger CV. Relationships between impairment and physical disability as measured by the functional indendence measure. Arch Phys Med Rehabil 1993; 74: 566–573.
7. Granger CV, Hamilton BB, Linacre JM, Heinemann AW, Wright BD. Performance profiles of the functional independence measure. Am J Phys Med Rehabil 1993; 72: 84–89.
8. Granger CV, Ottenbacher KJ, Baker JG, Sehgal A. Reliability of a brief outpatient functional outcome assessment measure. Am J Phys Med Rehabil 1995; 74: 469–475.
9. Cohen ME, Marino RJ. The tools of disability outcomes research functional status measures. Arch Phys Med Rehabil 2000; 81 (suppl 2): s21–s29.
10. Pollak N, Rhault W, Stoecker JL. Reliability and validity of the FIM for persons aged 80 years and above from a multilevel continuing care retirement community. Arch Phys Med Rehabil 1996; 77: 1056–1061.
11. Hamilton BB, Granger CV, Sherwin FF, Zielezny M, Tashman JS. A uniform national data system for medical rehabilitation. In: Fuhrer M, ed. Rehabilitation outcomes: analysis and measurement. Brookes: Baltimore; 1987, pp. 137–147.
12. Gonnella C. Program evaluation. In: Fletcher GF, Banja JD, Jann BB, Wolf SL, eds. Rehabilitation medicine: clinical perspectives. Philadelphia: Lea and Febiger; 1992, pp. 243–268.
13. Morris JN, Fries BE, Steel K, Ikegami N, Bernabei R, Carpenter GI, et al. Comprehensive clinical assessment in community setting: applicability of the MDS-HC. J Am Geriatr Soc 1997; 45: 1017–1024.
14. Williams BC, Li Y, Fries BE, Warren RL. Predicting patient scores between the functional independence measure and the minimum data set: development and performance of a FIM-MDS "crosswalk". Arch Phys Med Rehabil 1997; 78: 48–54.
15. Mortifee PR, Busser JR, Anton HA. 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 1996; 77: 436–439.
16. Dijkers MP, Yavuzer G. Short version of the telephone motor functional independence measure for use with persons with spinal cord injury. Arch Phys Med Rehabil 1999; 80: 1477–1484.
17. Rasch G. Probabilistic models for some intelligence and attainment tests. Chicago: University of Chicago; 1980.
18. Tsuji T, Sonoda S, Domen K, Saitoh E, Liu M, Chino N. ADL structure for stroke patients in Japan based on the functional independence measure. Am J Phys Med Rehabil 1995; 74: 432–438.
19. Merbitz C, Morris J, Grip JC. Ordinal scales and foundations of misinference. Arch Phys Med Rehabil 1989; 70: 308–312.
20. Silverstein F, Kilgore DJ, Harman JG, Harvey T. Applying psychometric criteria to functional assessment in medical rehabilitation, I: exploring unidimensionality. Arch Phys Med Rehabil 1991; 72: 631–637.
21. Armitage P, Berry G. Statistical methods in medical research, 3rd edn. Oxford: Blackwell Scientific Publications; 1994.
22. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979; 86: 420–428.