The Journal of Physical Therapy Science

Original Article

Evaluation of psychometric properties of the activities of daily living scale of motor function used by caregivers using Rasch analysis

Toshikatsu Kaneda, MS, OTR[1, 2]*, Shinichi Takabatake, PhD, OTR[3],
Yasuhiro Higashi, PhD, OTR[1, 4], Yuka Horishima, OTR[1], Yuta Somei, OTR[1],
Kazuyo Nakaoka, MS, OTR[2, 3], Hiroshi Warabino, MS, OTR[4, 5],
Kousuke Kuwada, RPT[1], Tomosaburo Sakamoto, MD[1], Kazuhisa Domen, PhD, MD[6]

[1] Department of Rehabilitation, Kansai Rehabilitation Hospital: 3-11-1 Sakuranocho, Toyonaka, Osaka 560-0054, Japan
[2] Doctoral Program in Functioning, Disability and Health Course, Graduate School of Comprehensive Rehabilitation, Osaka Prefecture University, Japan
[3] Graduate School of Comprehensive Rehabilitation, Osaka Prefecture University, Japan
[4] Faculty of Health Sciences, Morinomiya University of Medical Sciences, Japan
[5] Medicare-Rehabili Home-visit Nursing Station, Japan
[6] Department of Rehabilitation Medicine, Hyogo College of Medicine, Japan

Abstract. [Purpose] We developed a prototype version of the Self-Assessment Burden Scale-Motor as a screening questionnaire for the need of care based on the evaluation of patients' activities of daily living. The questionnaire is comprised of seven items, each of which is scored on a 7-point Likert scale. This study aimed to examine the psychometric properties of the questionnaire using Rasch analysis. [Participants and Methods] A total of 200 individuals completed the questionnaire. Rasch analysis to investigate the rating scale structure and examine the structural validity and reliability of the scale. The unidimensionality of the items was examined using the mean square infit values and principal component analysis of residuals. The separation reliability of the scale was also examined. [Results] The rating scale structure can be improved by collapsing several categories (from seven categories to four categories). Unidimensionality was found for seven items. The separation reliability was acceptable for item calibrations and persons. [Conclusion] Inclusion of seven items with a 4-category rating scale was appropriate for the Self-Assessment Burden Scale-Motor questionnaire. Future studies should examine the intra-rater reliability and the criterion-related validity in more depth and develop a new scale to evaluate cognitive function.
Key words: Caregivers, Activities of daily living, Validation studies

INTRODUCTION

In Japan, the aging population means that an increasing number of individuals require care. In 2000, a rehabilitation hospital system was established for patients who could not return home because of their disabilities. At these rehabilitation hospitals, intensive rehabilitation is provided with the aim of regaining the ability to perform activities of daily living (ADL). After being discharged from hospital, patients live in their community with help from several services that are financed by the long-term care insurance system[1].

It is important to evaluate patients' ADL ability after returning home; if their ADL abilities decline, rehabilitation services

---

*Corresponding author. Toshikatsu Kaneda (E-mail: posikati@gmail.com)

can be provided at the appropriate time. The Functional Independence Measure (FIM) is mainly used to assess ADL in patients in acute and rehabilitation hospitals. However, the FIM is difficult to implement after patients return home because of time constraints and personnel cost involved[1]. The Barthel Index (BI) was developed for caregivers to evaluate ADL at home[2]. However, Rasch analysis has revealed that the BI has no psychometric properties that enable the measurement of patients' ability[3]. In Japan, several evaluation instruments have been developed for caregivers to assess patients' ADL ability at home, such as the questionnaire version of the FIM[4], the short version of the FIM[1, 5], the computer program to assist evaluating the FIM[6], and the University of Occupational and Environmental Health self-assessment version of the BI[7]. The reliability and validity of these instruments have been examined. To simplify the FIM evaluation, the short version of the FIM excludes cognitive items whose reliability is inferior to motor items, and consequently contains only seven motor items. The short version of the FIM involves a complex scoring process and requires expert knowledge to administer, so is not suitable for home use[1, 5]. The FIM questionnaire, a computer program to assist evaluating the FIM, and the University of Occupational and Environmental Health self-assessment version of the BI have not yet been examined using Rasch analysis. In traditional methods, total scores are calculated by summing the ordinal scores of the scale. However, because treating ordinal data as interval data can result in incorrect conclusions, methods in which ordinal scores are added to generate a total score have received criticism in recent years. Thus, summing the raw scores for use as the outcome measure is considered problematic.

Rasch analysis was developed by Georg Rasch in the 1960s[8]. The Rasch model assumes that the items of a scale sit on a unidimensional line in order of difficulty, and confirms the degree of matching between the expected scores under the Rasch model and the observed score. In Rasch analysis, unidimensionality is examined by goodness-of-fit analysis and principal component analysis (PCA) of the residuals. Unidimensionality refers to a single structure from items that are easily performed to items that are difficult to perform. Thus, unidimensionality indicates the structural validity of a scale[8]. In the current study, we used a goodness-of-fit analysis to examine how well the data fit the Rasch model. The Rasch model for the ADL instruments was based on the following two assumptions: (1) The easier the items, the more likely it is for people to obtain higher scores, and (2) more able people are more likely to obtain higher ratings on difficult items compared with less able people. PCA tests unidimensionality and its underlying assumption that all data can be explained by the latent variable measured[9]. If goodness-of-fit analysis and PCA identify unidimensionality of a scale, the total score can be calculated as interval data[10]. Structural validity can also be assessed by examining the logical hierarchical ordering of items, whereby the hierarchical order of item difficulties can be compared with that of other scales.

Goodness-of-fit analysis can also be used to examine the psychometric properties of a rating scale. Thus, it is possible to examine whether a category scale can classify people appropriately[8, 11]. Reliability can also be tested using Rasch analysis[8], which is evaluated by resolution, defined as the ratio between the true standard deviation of a person (or item) and the standard deviation of the error. Item separation is used to validate item hierarchies and reflects the number of hierarchies of measures that can be statistically identified[11, 12].

There is not currently an evaluation instrument that caregivers can conveniently and precisely implement to evaluate ADL ability. Therefore, we developed a prototype version of the Self-Assessment Burden Scale-Motor (SAB-M-P) to assess patients' need of care by assessing their ADL. The SAB-M-P was developed with reference to a range of ADL evaluation instruments, including the Katz Index of Independence in Activities of Daily Living[13], the Physical Self-Maintenance Scale[14], and the Rapid Disability Rating Scale[15] reported by McDowell[16], as well as the University of Occupational and Environmental Health self-assessment version of the BI, and the short version of the FIM[1, 4–7]. The SAB-M-P consists of seven items: Feeding, Bathing, Dressing the lower body, Bed/Chair/Wheelchair transfer, Walk/Wheelchair, Stairs, and Bladder management. For each item, the caregivers assess the patient's degree of need for assistance (not their degree of independence) on a 7-point ordinal rating scale that ranges from 1 (full assistance needed) to 7 (no assistance needed)[17]. More detailed information about the SAB-M-P is described in the Methods section.

When the SAB-M-P is used as an outcome measure to determine the effect of an intervention, the sum of all items' scores is calculated. However, adding ordinal scores to generate a total score has received criticism in recent years, as mentioned earlier[8]. Thus, many ADL instruments used for evaluating the effect of rehabilitation, such as the FIM[18–20], the ADL-focused Occupation-based Neurobehavioral Evaluation[21, 22], and the Japanese version of the ADL-focused Occupation-based Neurobehavioral Evaluation[11], have used Rasch analysis to convert the ordinal score into an interval score. Rasch analysis can therefore be used to examine the psychometric properties of a rating scale and its individual items.

The aim of this study was to examine the psychometric properties of the SAB-M-P using Rasch analysis. After that, we develop the Self-Assessment Burden Scale-Motor (SAB-M). Four research questions concerning the validity (the first three questions) and reliability (the fourth question) were posed, as follows:

1. Does the rating scale of the SAB-M-P demonstrate good psychometric properties, as evidenced by the ability to classify the subjects appropriately?
2. Do the items of the SAB-M-P define a single unidimensional construct, as evidenced by goodness-of-fit and PCA?
3. Do the SAB-M-P item difficulties show a logical hierarchical ordering compared with other ADL scales?
4. Do the SAB-M-P items separate participants into different levels of ability, and do the participants tested separate the items into different levels of difficulty?

PARTICIPANTS AND METHODS

This study was conducted with patients of the rehabilitation hospital to which the first author belongs, from November 2017 to July 2018. The authors and attending physicians selected participants according to the following inclusion criteria: (a) they had been diagnosed with cerebrovascular disease, orthopedic disease, or disuse syndrome after surgery or pneumonia; (b) they were allowed to discharge themselves from the hospital, as decided by their attending physician; and (c) a medical examination demonstrated that participants were medically stable, and their attending physician had granted them permission to participate in the study. The exclusion criteria were as follows: (d) those who had transferred from other hospitals to receive treatment; and (e) those who had no family members. Each participant and their family gave informed written consent before participating in this study, and this study was approved by the Research Ethics Committee of the Graduate School of Comprehensive Rehabilitation, Osaka Prefecture University (Approval number for the research: 2016-208).
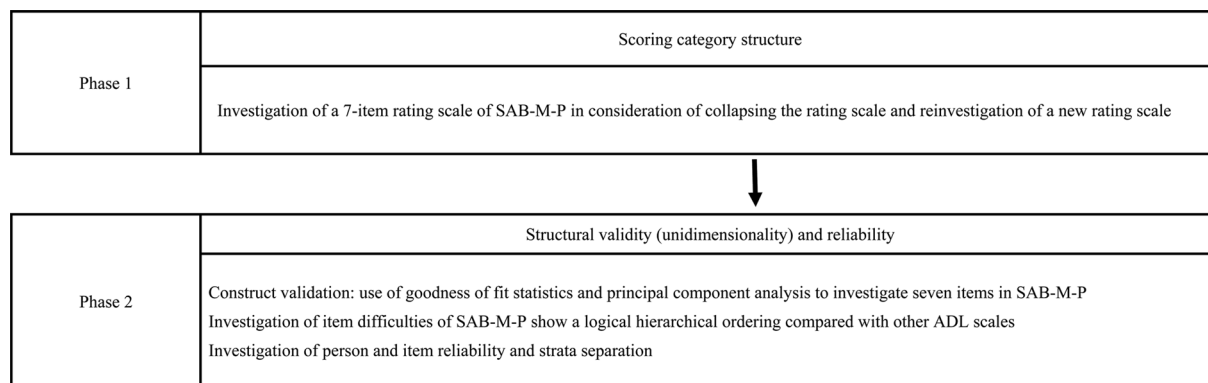
One week before the participants' discharge, the SAB-M-P was given to the participants' families and they responded according to their own observations. In the rehabilitation hospital, patients' families had the chance to care for patients before their discharge, and could therefore grasp the patients' ADL ability. We also extracted the following information from medical records: participants: age, gender, diagnosis, hospitalization period, FIM-Motor score at discharge, FIM-Cognitive score at discharge, and family composition; families: age, gender, and their relationship to the participants.

The SAB-M-P is an ADL evaluation questionnaire that is based on several ADL evaluation instruments. Caregivers complete the SAB-M-P according to their observations about the patient. Four occupational therapists with more than 5 years of clinical experience examined the SAB-M-P items and contents to decide whether the items appropriately evaluated ADL, and a focus group meeting was used to help determine whether the caregiver would understand the questions.

The SAB-M-P consists of seven items (Feeding, Bathing, Dressing the lower body, Bed/Chair/Wheelchair transfer, Walk/Wheelchair, Stairs, Bladder management), and each item is scored on a 7-point Likert scale. Caregivers rate patients' degree of need for assistance rather than their degree of independence. The contents of each item were as follows: "Feeding" refers to carrying the food to the mouth; "Bathing" refers to washing the body under the head; "Dressing the lower body" refers to putting pants on; "Bed/Chair/Wheelchair transfer" refers to moving oneself from the bed to a chair or wheelchair; for the "Walk/Wheelchair" item, "Walk" or "Wheelchair" should be chosen according to the primary way of moving in daily life; "Stairs" refers to going up and down stairs; and "Bladder management" refers to both urination failure and the amount of assistance required to urinate. Each item is scored using an ordinal scale from 1 to 7, including "1: total care needed", "2: considerable care needed", "3: moderate care needed", "4: mild care needed", "5: supervision needed", "6: physical care needed", and "7: no need for care". We examined the criterion-related validity of the SAB-M-P compared with the FIM[17].

The raw SAB-M-P scores were analyzed using the WINSTEPS Rasch computer software program (Version 4.0.0)[23]. The analysis was divided into two phases in accordance with a previous Rasch analytic study[11]. First, we examined the psychometric properties of the 7-category rating scale in the SAB-M-P, as suggested by Bond and Fox[8], then investigated the validity and reliability of all items. Then, with the psychometrically suggested rating scale, we examined the unidimensionality of the SAB-M-P using a goodness-of-fit analysis and PCA, and examined the reliability using a separation index and reliability coefficient (Fig. 1).

For the rating scale analysis, a minimum of 10 observations were required for accurate estimation of category measures, and we used the criteria of the outfit mean square (MnSq) value within 2.0 and advancement by at least 1.4 logits of the calibration thresholds between the rating scale categories[8, 11, 24]. If these criteria were not met, we planned to collapse non-

Phase 1	Scoring category structure
	Investigation of a 7-item rating scale of SAB-M-P in consideration of collapsing the rating scale and reinvestigation of a new rating scale

Phase 2	Structural validity (unidimensionality) and reliability
	Construct validation: use of goodness of fit statistics and principal component analysis to investigate seven items in SAB-M-P
	Investigation of item difficulties of SAB-M-P show a logical hierarchical ordering compared with other ADL scales
	Investigation of person and item reliability and strata separation

SAB-M-P: Prototype version of Self-Assessment Burden Scale-Motor

ADL: Activities of daily living

SAB-M: Self-Assessment Burden Scale-Motor

Fig. 1. Rasch analysis procedure for SAB-M development.

J. Phys. Ther. Sci. Vol. 32, No. 2, 2020

150

advancing categories and subsequently reanalyze the data.

For the goodness-of-fit analysis of items, the range of fit statistics differs according to test characteristics. In survey assessments, MnSq >1.4 associated with standardized Z (Zstd) >2.0 indicates a misfit, that is, a problem with the internal consistency of test items[8]. We removed items that had infit MnSq >1.4 associated with infit Zstd >2.0, whereby "infit" is an information-weighted indicator of misfit. Aberrant infit statistics usually cause more concern than do large outfit statistics[8].

For PCA, we used the Fisher's[25] five-level quality criteria for rating scale instruments and aimed for a "good" level of quality. Thus, the required proportion of variance explained by the measures (Rasch dimension) needed to be >50% and the proportion of unexplained variance accounted for by the first contrast (the largest secondary dimension) needed to be <15% for the results to support unidimensionality.

Brentani et al. reported that goodness-of-fit analysis and PCA need to be combined to validate unidimensionality[26]; we therefore adopted this approach. Linacre[27] reported that data including items that did not meet the above-described infit MnSq and Zstd criteria in the goodness-of-fit analysis should be re-analyzed without those items. We decided that, if no significant change in PCA was found before and after the item removal, we would retain those items.

A reliability coefficient of >0.8 and a separation index of >2.0 are good criteria according to Fisher's five-level quality criteria[25], as they allow at least three strata to be distinguished, and indicate that the persons or items comprise at least three levels of abilities or difficulties[12, 28].

RESULTS

We selected 205 participants to participate in the study. Missing data were found in five participants, resulting in data from a final total of 200 participants. The valid response rate was 97.6%. Detailed demographic information about participants and their families is presented in Table 1.

In the results of the rating scale analysis, the frequencies for all seven categories of the scale exceeded 10 (minimum=25, maximum=801). The rating structure analysis of the 7-category rating scale revealed an interval of less than 1.4 logits for the category measures from 1 to 2, 3 to 4, 4 to 5, and 6 to 7 in Table 2. Therefore, we collapsed the non-advancing categories by combining categories 3 and 4, which had the smallest interval between category measures, resulting in a 6-category rating scale. We re-analyzed data according to this 6-category rating scale and continued to collapse the non-advancing categories in the same way until the values satisfied the criteria. As a result of collapsing the categories, the scale became a 4-category rating scale in Table 2. The 7-category rating scale included "1: complete care needed", "2: considerable care needed", "3: moderate care needed", "4: mild care needed", "5: supervision needed", "6: physical care needed", and "7: no need for care", and the collapsed 4-category scale included "1: complete care needed", "2: partial care needed", "3: physical care needed", and "4: no need for care".

We retained the collapsed 4-category rating scale and proceeded to examine the structural validity and reliability of items. The goodness-of-fit analysis of items revealed that the item "bladder control" failed to demonstrate acceptable goodness-of-

Table 1. Demographic information of participants and caregivers

		Patients (n=200)			Caregivers (n=200)	
Age (years)	Mean	72.6	Age	Mean	60.0
	SD	14.7		SD	12.6
	Range	42–97		Range	36–89
Gender	Male	91	Gender	Male	59
	Female	109		Female	141
Diagnosis	Cerebrovascular disease	104	Relationship	Wife	61
	Bone disease	74		Daughter	54
	Disuse syndrome others	22		Husband	27
Length of hospital stay (days)	Mean	71.2		Son	26
	SD	37.9		Bride	8
FIM-M	Mean	75.4		Others	24
	SD	20.9	Family structure	Couple households	74
FIM-C	Mean	29.1		Two households	64
	SD	7.6		Living alone	47
SAB-M-P	Mean	40.4		Three households	10
	SD	12.1		Others	5

FIM-M: Functional Independence Measure-Motor; FIM-C: Functional Independence Measure-Cognitive; SAB-M-P: Prototype version of Self-Assessment Burden Scale-Motor.

Table 2. Rating scale category statistics (7 items, 7 categories, 6 categories, 5 categories, 4 categories)

Score	Frequency (%)	Outfit MnSq	Calibration threshold	Category measure
7−category scale				
1	152 (11)	1.41	NONE	−1.98
2	25 (2)	1.00	−0.02	−1.14
3	32 (2)	0.53	−1.05	−0.68
4	39 (3)	1.06	−0.46	−0.30
5	110 (8)	1.13	−0.70	0.22
6	241 (17)	0.69	0.41	1.27
7	801 (57)	1.02	1.83	3.08
6−category scale				
1	152 (11)	1.15	NONE	−2.24
2	25 (2)	0.77	−0.11	−1.28
3	71 (5)	0.81	−1.74	−0.61
4	110 (8)	1.10	−0.32	0.12
5	241 (17)	0.69	0.34	1.27
6	801 (57)	1.03	1.83	3.08
5−category scale				
1	152 (11)	0.93	NONE	−2.86
2	96 (7)	0.97	−1.56	−1.20
3	110 (8)	1.15	−0.42	−0.07
4	241 (17)	0.73	0.18	1.17
5	801 (57)	1.06	1.80	3.03
4−category scale				
1	152 (11)	0.93	NONE	−3.64
2	206 (15)	1.11	−2.50	−1.06
3	241 (17)	0.83	0.46	1.27
4	801 (57)	1.10	2.04	3.27

fit to the model (infit MnSq=1.76, z=5.0) in Table 3. PCA results from analyzing the unidimensionality of the seven items revealed that 73.4% of the total variance was explained by the measures and that 8.0% of the unexplained variance was accounted for by the first contrast. Thus, these results indicated that the SAB-M-P satisfied the criteria for unidimensionality.

Since "bladder control" did not fit to the Rasch model with seven items, we performed the goodness-of-fit analysis again with the other six items. All items then fitted to the Rasch model. PCA revealed that 77.3% of the total variance was explained by the measures and that 7.8% of the unexplained variance was accounted for by the first contrast. Therefore, the results revealed that the SAB-M-P satisfied the criteria of unidimensionality. However, there were no clear differences in the PCA results before and after the removal of the "bladder control" item. We therefore retained all seven items in the SAB-M-P, and proceeded to perform the other analyses.

The seven items are presented in order of difficulty in Table 3. The easiest item was "Eating" and the most difficult item was "Stairs".

The person separation index was 2.37, and the person reliability coefficient was 0.85. The item separation index was 6.96, and the item reliability coefficient was 0.98.

DISCUSSION

Linacre stated that 150 participants are required to perform Rasch analysis for most purposes (99% confidence interval for estimated item difficulty calibrations remaining stable within the absolute value of 0.5 logit)[29]. A total of 200 people participated in the present study, which satisfied the sample size requirements. Because Rasch analysis is a sample-independent analysis, random sampling was not required[30]. Thus, although this study was conducted at a single center, this would not be expected to cause a problem in the selection of the participants.

In this study, we collapsed the several non-advancing categories such that the rating scale was changed from a 7-category scale to a 4-category scale. According to Bond et al.[8] and Linacre[24], there are multiple reasons why categories are collapsed. For example, participants may have difficulty discriminating between certain categories, or there may be a rarely used category. In the current study, the frequencies of the following categories were relatively low: "2: considerable care needed

J. Phys. Ther. Sci. Vol. 32, No. 2, 2020

152

Table 3. Item measurement report (7 items, 6 items)

Item	Item difficulty (logits)	SE	infit	
			MnSq	Zst
7 items				
Stairs	2.05	0.14	0.86	−1.20
Bed/chair/wheelchair transfer	0.92	0.15	0.84	−1.40
Bladder management	0.32	0.15	**1.76**	**5.00**
Bathing	−0.16	0.16	0.94	−0.40
Walk/wheelchair	−0.29	0.16	1.12	0.90
Dressing lower body	−0.74	0.17	0.72	**−2.20**
Feeding	−2.10	0.20	1.07	0.50
6 items				
Stairs	2.57	0.16	0.91	−0.80
Bed/chair/wheelchair transfer	1.18	0.16	0.89	−1.00
Bathing	−0.14	0.17	1.18	1.30
Walk/wheelchair	−0.29	0.18	1.17	1.30
Dressing lower body	−0.85	0.19	0.81	−1.40
Feeding	−2.47	0.22	1.33	1.90

Misfit items appear in bold.

(25)", "3: moderate care needed (32)", and "4: mild care needed (39)".

In addition, "5: supervision" was collapsed into "partial care". Supervision means that the caregiver always accompanied the patient. It could have been difficult for caregivers to distinguish between being beside patients (5: supervision) and providing mild care (4: mild care), in which the caregiver touched and cared for the patient. This may explain why "5: supervision" was collapsed into the "mild care" category.

In a previous study of the FIM, the 7-category rating scale was also collapsed into a 4-category rating scale[18]. This finding suggests that it is reasonable to measure ADL abilities using a 4-category rating scale in quantitative evaluation.

The item "bladder control" did not fit to the Rasch model. In a previous study of the short version of the FIM, the item "bladder control" also did not fit to the Rasch model; however, the researchers retained this item because it was clinically meaningful[1]. In the current study, we found no clear changes before and after the "bladder control" item was removed. As reported in the previous FIM study, the item "bladder control" is one of the most important items in a clinical setting. We therefore retained this item in the SAB-M-P. Thus, it can be concluded that an appropriate scale can be constructed using the total score of the seven items on the 4-category rating scale, as in the SAB-M.

In the ordering of item difficulty, the difficulty level of "dressing the lower body" was lower than that reported in the FIM study[1]. This may be affected by the difference in evaluation content. Specifically, in the SAB-M-P, "dressing the lower body" only referred to putting on pants, whereas in the FIM, socks and leg braces were also included.

In the examination of reliability, the person separation index was 2.37, and the item separation index was 6.96. The results revealed that the SAB-M-P was able to distinguish the participants into at least three strata, and the scale could be divided into at least seven difficulty levels. Both values satisfied the criteria mentioned above (a reliability coefficient of >0.8 and a separation index value of >2.0), indicating that the SAB-M can be considered a reliable evaluation instrument.

The creation of a conversion table is important for using the evaluation as an interval measure, by changing raw scores to logits[21]. By creating a conversion table, it is possible to convert the raw sum of the SAB-M scale to an interval scale, and the SAB-M can then be used more practically. The conversion table is presented in Table 4.

The results indicated that a scale with seven items using a 4-category rating scale was appropriate to use in the SAB-M. Future studies should examine the reliability and validity of the SAB-M in more depth, including its intra-rater reliability and criterion-related validity. In the current study, we did not conduct an analysis according to different disease types. Therefore, it will also be necessary for future studies to increase the number of participants and re-examine the results with consideration of disease type. Finally, while the SAB-M can evaluate the motor aspects of ADL ability, it cannot evaluate participants' cognitive function. In future, it will be necessary to develop a scale to evaluate cognitive function.

This study provides an important first step to explore the possibility of revising the SAB-M-P to the SAB-M. The results revealed that the SAB-M can function appropriately with a 4-category rating scale. It will be necessary for future studies to examine measures of the reliability and validity of the SAB-M, such as intra-rater reliability and the criterion-related validity, and to develop a new scale to test cognitive function.

Table 4. Conversion table (7 items)

Score	Logit	SE
7	−6.24	1.92
8	−4.80	1.15
9	−3.80	0.89
10	−3.11	0.78
11	−2.54	0.72
12	−2.05	0.69
13	−1.59	0.66
14	−1.17	0.64
15	−0.77	0.62
16	−0.40	0.61
17	−0.04	0.59
18	0.31	0.59
19	0.65	0.58
20	0.99	0.58
21	1.33	0.59
22	1.69	0.60
23	2.06	0.63
24	2.48	0.66
25	2.96	0.73
26	3.56	0.84
27	4.46	1.10
28	5.81	1.88

Presentation at a Conference

The 47th Japanese Occupational Therapy Congress and Expo volume 47, Page O252.

The Japanese Journal of Rehabilitation Medicine JARM 2016, Page I179.

The 51th Japanese Occupational Therapy Congress and Expo, PA-1A02. http://jotc51.mas-sys.com/site/searchtop.html

The 52th Japanese Occupational Therapy Congress and Expo, ON5-2. https://www.mas-sys.com/JOTC52_Abstract/site/searchtop.html

The 53th Japanese Occupational Therapy Congress and Expo, ON-3-3. https://www.mas-sys.com/JOTC53_Abstract/site/searchtop.html

The 53th Japanese Occupational Therapy Congress and Expo, ON-3-4. https://www.mas-sys.com/JOTC53_Abstract/site/searchtop.html

The 53th Japanese Occupational Therapy Congress and Expo, ON-3-5. https://www.mas-sys.com/JOTC53_Abstract/site/searchtop.html

Conflict of interest

The authors report no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

REFERENCES

1) Yamada S, Liu M, Hase K, et al.: Development of a short version of the motor FIM for use in long-term care settings. J Rehabil Med, 2006, 38: 50–56. [Medline] [CrossRef]

2) Schlote A, Krüger J, Topp H, et al.: [Inter-rater reliability of the Barthel Index, the Activity Index, and the Nottingham Extended Activities of Daily Living: the use of ADL instruments in stroke rehabilitation by medical and non medical personnel]. Rehabilitation (Stuttg), 2004, 43: 75–82 (in German). [Medline] [CrossRef]

3) de Morton NA, Keating JL, Davidson M: Rasch analysis of the barthel index in the assessment of hospitalized older patients after admission for an acute medical condition. Arch Phys Med Rehabil, 2008, 89: 641–647. [Medline] [CrossRef]

4) Ota T, Domen K, Liu M, et al.: Trial of scoring the functional independence measure using a questionnaire. Gen Rehabil, 1997, 25: 449–454.

5) Yamada S, Liu M, Fujimoto M, et al.: Identification of quasi-in-need-of-care state (QUINOCS) among community dwelling elderly people using a seven-item subset of the Functional Independence Measure (FIM). Disabil Rehabil, 2009, 31: 381–386. [Medline] [CrossRef]

J. Phys. Ther. Sci. Vol. 32, No. 2, 2020

154

6) Yamada S, Ota T, Liu M, et al.: Development of a computer program to assist evaluating the FIMTM (iFIM). Gen Rehabil, 2006, 34: 69–76.

7) Fukuda M, Takemoto A, Shirayama Y, et al.: Inter-rater reliability of the activities of daily living: UOEH self-assessment version of the Barthel index. Jpn J Traumatol Occup Med, 1995, 43: 842–846.

8) Bond TG, Fox CM: Applying the Rasch model: fundamental measurement in the human sciences, 3rd ed. NJ: Erlbaum, Mahwah, 2015.

9) Dickens GL, Rudd B, Hallett N, et al.: Factor validation and Rasch analysis of the individual recovery outcomes counter. Disabil Rehabil, 2019, 41: 74–85. [Medline] [CrossRef]

10) Tennant A, Conaghan PG: The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Rheum, 2007, 57: 1358–1362. [Medline] [CrossRef]

11) Higashi Y, Takabatake S, Matsubara A, et al.: Reliability and validity of the Japanese version of the ADL-focused Occupation-based Neurobehavioural Evaluation (A-ONE J): applying Rasch analysis methods. Hong Kong J Occup Ther, 2019, 32: 32–40. [Medline] [CrossRef]

12) Fisher WP: Reliability, separation, strata statistics. Rasch Meas Trans, 1992, 6: 238.

13) Katz S, Hedrick SC, Henderson NS: The measurement of long-term care needs and impact. Health Med Care Serv Rev, 1979, 2: 1–21, 3–21. [Medline]

14) Lawton MP, Brody EM: Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist, 1969, 9: 179–186. [Medline] [CrossRef]

15) Linn MW, Linn BS: The rapid disability rating scale-2. J Am Geriatr Soc, 1982, 30: 378–382. [Medline] [CrossRef]

16) McDowell I: Measuring health: a guide to rating scales and questionnaires, 3rd ed. New York: Oxford University Press, 2006, pp 55–149.

17) Kaneda T, Takabatake S, Nishikawa T, et al.: Examination of validity activities of daily living scale by primary care giver. J Jpn Occup Ther Assoc, 2013, 32: 194–198.

18) Nilsson ÅL, Sunnerhagen KS, Grimby G: Scoring alternatives for FIM in neurological disorders applying Rasch analysis. Acta Neurol Scand, 2005, 111: 264–273. [Medline] [CrossRef]

19) Lundgren Nilsson Å, Tennant A: Past and present issues in Rasch analysis: the functional independence measure (FIM™) revisited. J Rehabil Med, 2011, 43: 884–891. [Medline] [CrossRef]

20) Linacre JM, Heinemann AW, Wright BD, et al.: The structure and stability of the functional independence measure. Arch Phys Med Rehabil, 1994, 75: 127–132. [Medline]

21) Arnadóttir G, Fisher AG: Rasch analysis of the ADL scale of the A-ONE. Am J Occup Ther, 2008, 62: 51–60. [Medline] [CrossRef]

22) Árnadóttir G, Fisher AG, Löfgren B: Dimensionality of nonmotor neurobehavioral impairments when observed in the natural contexts of ADL task performance. Neurorehabil Neural Repair, 2009, 23: 579–586. [Medline] [CrossRef]

23) Linacre JM: Winsteps (Version 4.0.0) [Computer software]. Chicago: Winsteps.com, 2017.

24) Linacre JM: Optimizing rating scale category effectiveness. J Appl Meas, 2002, 3: 85–106. [Medline]

25) Fisher WP: Rating scale instrument quality criteria. Rasch Meas Trans, 2007, 21: 1095.

26) Brentani E, Golia S: Unidimensionality in the Rasch model: how to detect and interpret. Statistica, 2007, 67: 253–261.

27) Linacre JM: When to stop removing items and persons in Rasch misfit analysis? Rasch Meas Trans, 2010, 23: 1241.

28) Caronni A, Sciumè L, Donzelli S, et al.: ISYQOL: a Rasch-consistent questionnaire for measuring health-related quality of life in adolescents with spinal deformities. Spine J, 2017, 17: 1364–1372. [Medline] [CrossRef]

29) Linacre JM: Sample size and item calibration stability. Rasch Meas Trans, 1994, 7: 328.

30) Hammond FM, Malec JF, Nick TG, et al.: Handbook for clinical research: design, statistics, and implementation. New York: Demos Medical Publishing, 2015, pp 195–197.