

## ORIGINAL ARTICLE

## BERG BALANCE SCALE IS A VALID MEASURE FOR PLAN INTERVENTIONS AND FOR ASSESSING CHANGES IN POSTURAL BALANCE IN PATIENTS WITH STROKE

Kazuhiro MIYATA, PhD, PT<sup>1</sup>, Shuntaro TAMURA, PhD, PT<sup>2</sup>, Sota KOBAYASHI, MSc, PT<sup>3,4</sup>, Ren TAKEDA, PT<sup>4,5</sup> and Hiroki IWAMOTO, MSc, PT<sup>6</sup>

From the <sup>1</sup>Department of Physical Therapy, Ibaraki Prefectural University of Health Science, <sup>2</sup>Department of Rehabilitation, Fujioka General Hospital, <sup>3</sup>Department of Rehabilitation, Public Nanokaichi Hospital, <sup>4</sup>Department of Rehabilitation Sciences, Gunma University Graduate School of Health Sciences, <sup>5</sup>Department of Rehabilitation, Numata Neurosurgery & Heart Disease Hospital and <sup>6</sup>Department of Rehabilitation Center, Hidaka Rehabilitation Hospital, Gunma, Japan.

**Objectives:** After confirming the measurement properties of the Berg Balance Scale (BBS) in patients with stroke by conducting a Rasch analysis, this study sought: (i) to generate a keyform as a tool for goal-setting and intervention-planning; and (ii) to determine the appropriate strata for separating patients' postural balance ability.

**Design:** Methodological analyses of cross-sectional study data.

**Patients:** A pooled sample of 156 patients with stroke: mean (standard deviation) age 74.4 (12.9) years.

**Methods:** This study evaluated the BBS's rating scale structure, unidimensionality, and measurement accuracy (0: unable to perform or requiring help, to 4: normal performance) and then generated a keyform and strata.

**Results:** The BBS rating scale fulfilled the category functioning criteria. Principal component analysis of standardized residuals confirmed the unidimensionality of the test. All items fit the Rasch analysis. Person ability-item difficulty matching was good. Person reliability was 0.96, and the patients were divided into 9 strata. The keyform for the BBS will enable clinicians and investigators to estimate patients' postural balance ability and monitor their progress.

**Conclusion:** The BBS has strong measurement properties. This study generated both a keyform that can contribute to clinicians' decision-making in goal-setting and intervention-planning and strata that can facilitate understanding of patients' abilities.

**Key words:** rehabilitation; postural balance; Rasch analysis; outcome measure; goal-setting; intervention-planning.

Accepted Nov 3, 2022

J Rehabil Med 2022; 54: jrm00359

DOI: 10.2340/jrm.v54.4443

Correspondence address: Kazuhiro Miyata, Department of Physical Therapy, Ibaraki Prefectural University of Health Science, 4669-2, Ami-Machi, Inashiki-gun, Ibaraki, 300-0394 Japan. E-mail: miyatak@ipu.ac.jp

## LAY ABSTRACT

People who have had a stroke often have difficulty maintaining postural balance and controlling their posture. The Berg Balance Scale (BBS) measures a person's ability to maintain postural balance. Several analyses were performed to investigate the measurement properties of the BBS in patients who have had a stroke. Then, a keyform tool was generated and some strata (levels) determined that separate patients according to postural balance ability. A keyform can help therapists to identify items that a patient finds relatively difficult. Use of a keyform can contribute to both rehabilitation goal-setting and planned interventions for patients. The strata can be used to detect and measure changes in a patient's postural balance ability. The findings of this study demonstrate that the BBS has strong measurement properties and provides an appropriate keyform and 9 strata. Use of these tools can facilitate the rehabilitation of patients with stroke through quantification of a patient's postural balance ability.

Measuring and monitoring the postural balance and risk of falling in individuals who have had a stroke are critical to helping maintain their mobility and their activities of daily living (ADL) (1, 2). Suitable interventions to improve an individual's diminished postural balance can be devised, based on an accurate assessment of their postural balance (3). An effective postural balance assessment scale as a core outcome set is essential to determine the effectiveness of rehabilitation interventions. One of the most widely used outcome measures is the Berg Balance Scale (BBS) (4), the measurement properties of which have been analysed in detail in stroke patients by using classical test theory (5). It has been pointed out that floor and ceiling effects are likely to be obtained with the BBS (5); however, the BBS is recommended in neurological disorder guidelines for use as part of a core outcome set to achieve the goals of improving static and dynamic sit-to-stand postural balance (6). BBS results have also been used as clinical trial outcomes (7).

The BBS aggregate score has limited clinical interpretability, because it does not indicate which of the scale's items were easy, difficult, or optimally challenging for subjects (8–12). One way to obtain more information for interpreting BBS results is to examine both the overall score and the responses to individual test items. Rasch analysis is a model-driven process; it can assess how well a given test item performs in terms of its relevance or usefulness for measuring the underlying construct (13). Rasch analysis can assess aspects of a scale, such as dimensionality, item fit, item difficulty and person ability, and reliability; and, whether the data obtained fit the Rasch model. Rasch analysis provides a framework and method for enhancing the clinical interpretability of evaluations through output termed a “keyform” and “strata” (14).

A keyform is used to assist goal-setting and intervention-planning by drawing attention to missing items that correspond to items that examinees find difficult to complete (15–17). Strata are used to create a conversion table that transforms the raw scores of a scale into Rasch-based interval level measurements and then divides them into statistically detectable groups (18, 19).

The aims of the current study were two-fold. After the study confirmed that the BBS data obtained herein fit the Rasch model, the first goal was to generate a BBS keyform that enables therapists to compare each patient's response with the response predicted by the Rasch model and to highlight specific items or tasks that are relatively difficult for the individual. This can assist in rehabilitation goal-setting and intervention-planning. The second goal was to generate strata with a conversion table to statistically detect changes in patients' postural balance ability.

## PATIENTS AND METHODS

### Patients

This was a multicentre clinical observational study of individuals with subacute stroke who participated in a rehabilitation programme at the convalescent rehabilitation wards of 3 hospitals in Japan during April 2018 to May 2020. Inclusion criteria were: (i) diagnosis with a cerebral haemorrhage or cerebral infarction; (ii) first-ever supratentorial hemispheric lesion; and (iii) stable medical condition. Exclusion criteria were: (i) any neurological or musculoskeletal disorder; (ii) a missing BBS score on admission or at discharge; (iii) unable to follow the indicated behaviour; or (iv) declined to be included in the study.

The study was approved by the Ethics Review Committee of Fujioka General Hospital, Public Nanokaichi Hospital, and Hidaka Rehabilitation Hospital (approval

numbers #194, #20200020, #20200503, respectively) and was conducted in compliance with the Declaration of Helsinki. Because the study was a retrospective observational analysis and used only existing data, the patients' written or verbal consent was not required and was not obtained. To provide the patients with the opportunity to decline being part of the study, an opt-out option for study information was posted on each hospital's website or bulletin board.

### Assessment tools

The BBS is postural balance scale containing 14 items including standing and sitting unsupported, reaching forward, and placing the alternating foot on a stool. Administering the BBS takes approximately 15 min. Each of the 14 items are scored on a 5-level ordinal scale from 0 (“unable to perform or requiring help”) to 4 (“normal performance”), thus providing a potential maximum score of 56 points (4). A review of the measurement properties of the BBS showed them to be valid and reliable (5).

The patients' demographic and clinical characteristics were collected from their medical records and the rehabilitation centres' databases. The BBS was administered to each patient by a well-trained physical therapist as part of the routine clinical assessment at the patient's admission to and discharge from the convalescent rehabilitation ward. Data from the patients' admission evaluations were used for the main analysis, and data from the patients' discharge session were used to assess the differential item functioning (DIF) related to the interventions provided to the patients.

### Rasch analysis

Rasch analysis of the BBS items was conducted according to the framework of the Rasch Reporting Guidelines in Rehabilitation Research (RULER) (20, 21). The Rasch model is a probabilistic, mathematical model that provides an opportunity to transform ordinal scales into interval-like scores to account for differences in item difficulty and differences across categories in an ordinal response scale. The current study sample size ( $n=156$ ), item calibration, and person measures were expected to be stable within 0.5 logits with a 99% confidence interval (22). The patients' data were analysed using WINSTEPS software (ver. 5.2.3; Winsteps Rasch measurement computer program, Beaverton, OR, USA).

This study examined the following measurement properties based on the RULER guidelines to investigate whether the pattern of the patients' responses met the assumptions of the Rasch model:

- *Rating scale structure*: items of the BBS were analysed using a partial credit model that allows

the threshold calibrations for rating scale steps to be determined separately for each item. Each BBS response category performed as intended based on the criteria suggested by Linacre (23). The item response categories of each item are described herein with the use of the item characteristic curve (ICC).

- *Unidimensionality*: a principal component analysis (PCA) of the standardized Rasch residuals was performed to evaluate the scale's dimensionality. If the BBS is unidimensional, all items in the BBS could be interpreted as fitting to the scale construct of postural balance. For a confirmation of the unidimensionality of a scale, the following criteria must be met: (i) the variance explained by the measured construct (i.e. Rasch factor) is >50%, and (ii) the eigenvalue and the percent variance explained by the first contrast is <2.0 and <10%, respectively (21).
- *Local independence of the items*: this study calculated the correlations of standardized residuals; once the Rasch factor has been conditioned out, no residual association among item responses should be observed. A high correlation of residuals (>0.30) for a pair of items indicates their potential local dependence (21).
- *Item fit*: the study verified how well the observed responses to the items were consistent with the responses predicted by the Rasch model. How accurately the data fit the Rasch model is given by the infit and outfit mean square statistics (MnSq) and by infit and outfit standardized z-values (ZSTD). The MnSq is the size of the departure of data from the model's prediction, and the ZSTD gives the statistical significance of this departure. We considered acceptable fit as MnSq values in the range 0.5–1.5, associated with ZSTD values <2.0 (14). Items beyond this range were considered underfitting (MnSq >1.5) or overfitting (MnSq <0.5). More emphasis was placed on infit values than on outfit values for identifying misfitting items.
- *Reliability*: the study estimated 2 types of reliability statistics, i.e. the person separation reliability (PSR) and the internal consistency represented by the traditional Cronbach's alpha. The concept of "person reliability" provides the degree of confidence that can be placed in the consistency of the estimates. The reliability ranges from 0 to 1, and coefficients >0.70 and >0.90 are considered acceptable for group-level and individual-level decision-making, respectively (21). A person separation index (PSI) indicates how well one can differentiate between different individuals' or items' performances along the measurement construct. The PSI value of 2.0 is judged as good because it enables the distinction of 3 [(4G+1)/3] statistically detectable groups (or "strata") of measures (14).

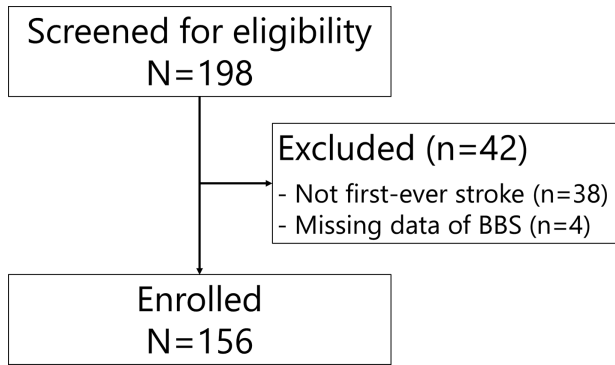
- *Targeting and keyform*: Rasch analysis calculated the item difficulty and person ability measures to a single metric (logits), and the study then examined a person-item map. The figures of the person-item map demonstrated the difficulty (item locations) and relative ability (person location) on the same ruler of logits. The better the match (targeting), the greater the potential for precise person measurement (21). Poor targeting often results in floor or ceiling effects. A keyform was created for the BBS to be used for patients with subacute stroke by using the WINSTEPS software program's table output option "general keyform". The keyform may help in goal-setting and intervention-planning. To illustrate the process of using the keyform to interpret the BBS score, an individual with stroke who has a postural balance deficit (BBS score 40) was randomly selected and the patient's raw data on the keyform displayed, a transition zone identified, appropriately challenging therapy goals described, and examples of how the goals could be linked to intervention(s) provided.
- *Differential item functioning*: the study examined the stability of the item hierarchy across the following subgroups. Intervention (before vs after rehabilitation; admission vs discharge); age (younger vs older, split according to the median age of the sample (77 years)); and sex (female vs male). DIF analysis was conducted separately to investigate the stability of the scale for each subgroup, and the t-statistic was used to identify significant differences in the item difficulty measures between subgroups (2-side t-test, the null hypothesis being that the 2 estimates would be the same).

## RESULTS

The patient selection flow chart is shown in Fig. 1. The patients' ( $n = 156$ ) main clinical and demographic characteristics are summarized in Table I.

For the Rasch analysis:

- *Rating scale structure*: the 5-level rating scale of the BBS fulfilled the 4 category-function criteria. Each rating category included >10 observations. The mean measures increased monotonically; the category thresholds increased with each category, and the outfit MnSq values were <2.0. Fig. S1 shows the ICC for each item of the BBS.
- *Unidimensionality*: PCA of the standardized residuals revealed that the variance attributable to the Rasch factor was high (88.7%, of which 52.2% was explained by persons and 36.4% was explained by items), whereas the eigenvalue of the first residual factor was 2.03 (1.7%), just marginally above the



**Fig. 1.** Enrolment process flow chart. BBS: Berg Balance Scale.

**Table I.** Main clinical and demographic characteristics of the sample at the admission (n = 156)

Characteristic	Mean (SD)
Age, years	74.4 (12.9)
Male/female, n	101/55
Time since stroke, days	25.6 (15.5)
Stroke type, ischaemic/haemorrhagic, n	119/37
Hemiplegic side, left/right, n	73/83
BRS of lower extremity, II/III/IV/V/VI, n	15/7/20/62/52
BBS, points	30.3 (18.5)

BBS: Berg Balance Scale; BRS: Brunnstrom recovery stage; SD: standard deviation.

predicted noise level of 2.0. These results confirmed the unidimensionality of the BBS.

- *Local independence of the items:* for a single item-pair on the BBS, i.e. item numbers 1 and 4, the standardized residual correlation was slightly high (0.34). This response dependence was considered negligible and thus no action was taken.
- *Item fit:* MnSq and ZSTD values for infit and outfit are shown in Table II. The infit MnSq was 0.64–1.39 and the ZSTD was –2.21 to 2.41, and all items exhibited an adequate infit MnSq, indicating that the underlying construct represents postural balance.
- *Reliability:* the PSR and Cronbach’s alpha were acceptable, at 0.96 and 0.97, respectively. The PSI

was 5.22, and the patients were divisible into 9 statistically distinct levels (Table III).

- *Targeting and keyform:* item difficulty showed a fairly even spread from the easiest item (number 3) to the most difficult item (number 14). Some items had similar item-difficulty measures. For example, item numbers 8, 9, and 10 had similar difficulty levels (0.17, –0.01, and 0.21 logits, respectively). The person-item map (Fig. 2) illustrates that the items were closely matched to the range of the patients’ ability measure, with the exception of the non-extreme person ability. Fig. 3 shows a BBS keyform for an individual with stroke who has a postural balance deficit (BBS score 40) and shows a near-normal performance on the easiest items (at the bottom of the figure), with progressively poorer performance as the items become more difficult. The dotted box is called the transition zone, and items that are difficult to obtain perfect scores appear. The transition zone of the keyform is not perfect, but it shows some competent behaviour that can lead to therapeutic intervention. In this case, item numbers 5 and 9 were considered to reflect the boundary (the point at which the transition occurs from one rating scale category to the next), and item numbers 9 to 11 were designated as the transition zone.
- *Differential item functioning:* no DIF was found across intervention, age, and sex groups.

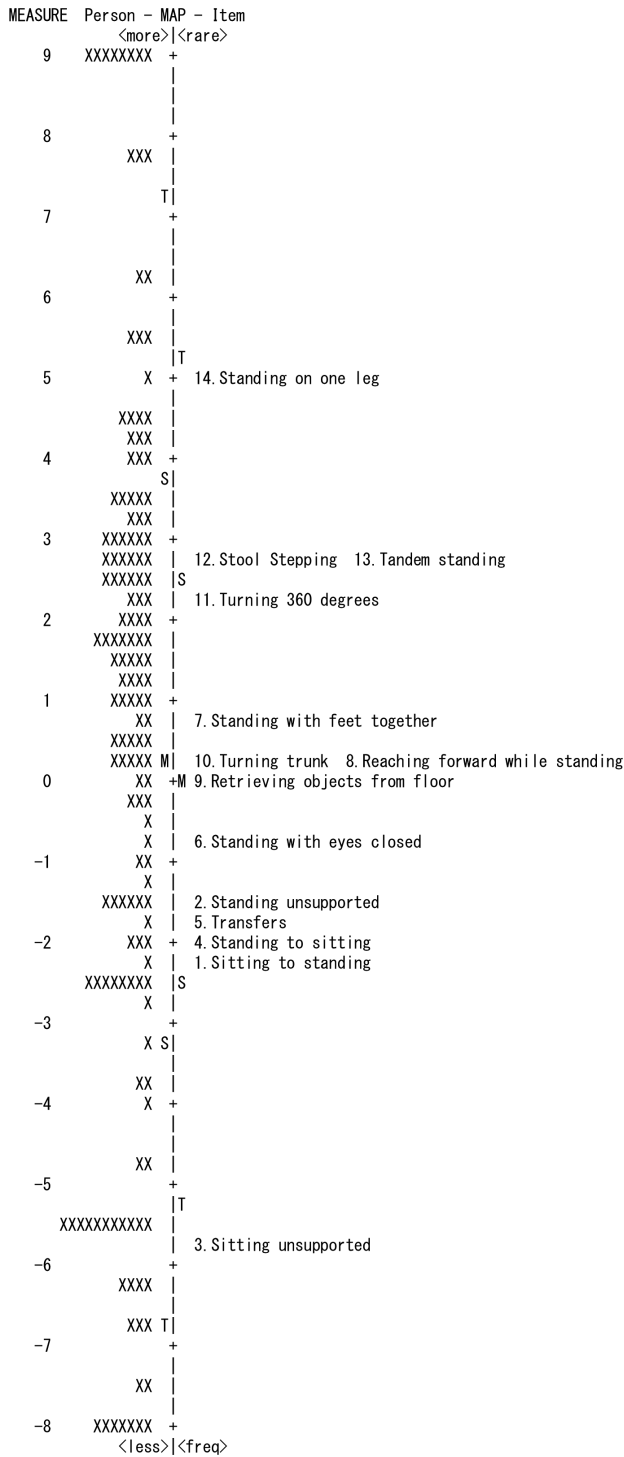
## DISCUSSION

This study evaluated the measurement properties of the BBS by performing Rasch analysis to determine the suitability of each item of the BBS as a general tool in the assessment of patients’ postural balance. The results of the measurement evaluation show that the BBS provides valid and reliable measures of stroke postural balance, and a keyform and strata were developed as tools to support decision-making.

**Table II.** Summary of Rasch analysis results for structural validity, containing item measure and fit information

Item	Measure	SE	Infit		Outfit	
			MnSq	ZSTD	MnSq	ZSTD
1. Change of position: sitting to standing	-2.21	0.17	0.76	-1.47	0.72	-0.77
2. Standing unsupported	-1.41	0.15	0.90	-0.59	3.38	2.91
3. Sitting unsupported	-5.73	0.21	0.99	0.03	2.52	1.77
4. Change of position: standing to sitting	-1.98	0.16	1.23	1.44	0.85	-0.42
5. Transfers	-1.71	0.15	1.07	0.61	1.08	0.39
6. Standing with eyes closed	-0.75	0.15	0.64	-2.21	0.60	-0.60
7. Standing with feet together	0.69	0.14	1.20	1.11	0.85	-0.14
8. Reaching forward while standing	0.17	0.15	1.22	1.38	1.13	0.47
9. Retrieving objects from floor	-0.01	0.15	1.08	0.41	0.78	-0.32
10. Turning trunk (feet fixed)	0.21	0.14	0.80	-1.29	0.59	-0.87
11. Turing 360 degrees	2.35	0.13	0.83	-1.18	0.58	-0.87
12. Stool stepping	2.75	0.13	0.76	-1.69	0.47	-1.39
13. Tandem standing	2.75	0.14	1.39	2.41	0.97	0.08
14. Standing on one leg	4.90	0.18	0.95	-0.27	0.67	-0.76

MnSq: mean square statistics; ZSTD: standardized z values; SE: standard error.



**Fig. 2.** Person-item map of the Berg Balance Scale (BBS). The numbers (-8 to 9) on the left side of the figure represent the logits. Along the vertical line, M=mean, S=1 standard deviation, and T=2 standard deviations. Left side of the vertical line represents the person ability (x); right side represents the item difficulty for each item of the BBS, both in logits. Each "x" represents a single person.

In the stroke patients in this study, the BBS was uni-dimensional as a postural balance assessment scale; it fulfilled all of the criteria of the rating scale and was

**Table III.** Raw score-to-measure conversion table for Berg Balance Scale

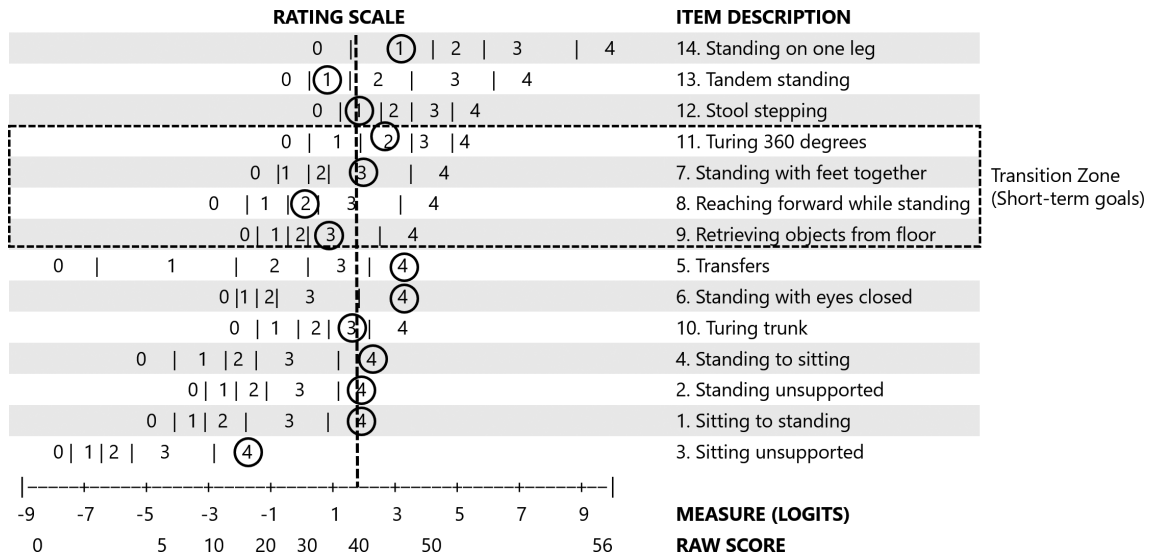
Score	Measure	SE	Strata	Score	Measure	SE	Strata
1	-8.52	1.79	1	29	0.04	0.36	
2	-7.42	0.94		30	0.18	0.37	
3	-6.76	0.74		31	0.31	0.37	
4	-6.2	0.78		32	0.45	0.38	
5	-5.47	0.93		33	0.6	0.38	
6	-4.64	0.85	2	34	0.75	0.39	
7	-4.05	0.70		35	0.9	0.40	
8	-3.63	0.60		36	1.06	0.40	6
9	-3.31	0.54		37	1.23	0.41	
10	-3.04	0.50		38	1.41	0.42	
11	-2.8	0.47		39	1.59	0.43	
12	-2.59	0.45	3	40	1.77	0.44	
13	-2.4	0.43		41	1.97	0.45	
14	-2.22	0.42		42	2.18	0.46	
15	-2.05	0.41		43	2.39	0.47	7
16	-1.88	0.40		44	2.62	0.48	
17	-1.72	0.39		45	2.86	0.49	
18	-1.57	0.39		46	3.1	0.50	
19	-1.42	0.38		47	3.35	0.51	
20	-1.27	0.38	4	48	3.61	0.52	
21	-1.13	0.37		49	3.89	0.54	8
22	-0.99	0.37		50	4.19	0.56	
23	-0.85	0.36		51	4.53	0.60	
24	-0.72	0.36		52	4.92	0.66	
25	-0.59	0.35		53	5.42	0.77	
26	-0.46	0.35		54	6.17	0.99	9
27	-0.34	0.36		55	7.66	1.47	
28	-0.21	0.36	5	56	9.67	2.08	

SE: standard error.

highly reliable, and all items were fitted to the Rasch model. These findings are generally consistent with studies of the use of the BBS for patients with subacute and chronic stroke (8) or Parkinson’s disease (9) and community-dwelling older adults (12). Although it has been pointed out that the ceiling effect and rating redundancy can be obtained easily with the BBS (3), the findings of the current study confirm that the BBS with strong psychometric properties can be used as part of a core outcome set to measure postural balance in specific stroke patients (e.g. excluding very mild and severe strokes).

Adequate interpretations of test results and high clinical utility are critical in rehabilitation. This study approached the issue of clinical interpretability by linking the qualitative content of the BBS (the postural balance tested by each item) to its quantitative aspects (item ratings) via a new scoring method, the keyform. Earlier studies of keyforms focused on the evaluation of upper extremity function (16, 24) in patients with stroke and demonstrated the application of keyforms in clinical practice.

A keyform is a pencil-and-paper scoring template on which a therapist records item ratings and then examines the pattern of item responses (15, 16). A keyform as used herein can be used in 2 ways to support clinical decision-making. The first is to ensure that each examinee presents a pattern of item responses that matches the expectations of the model, thus facilitating the detection of unexpected responses. The second is



**Fig. 3.** Example of clinical use of the keyform in a representative patient with a Berg Balance Scale (BBS) score of 40 points. Latent trait (balance ability) increases toward the right of the graph. The threshold map for each of the items (listed in descending order of difficulty) is in the middle of the figure. Marking a vertical line that starts from the patient’s global score (here, 40 points), the point where this line intersects the rating zones for each item (0–4) indicates the score most probable for that item. The threshold between adjacent categories is marked with “|”. The bottom 2 lines contain the Rasch nomogram, which allows the conversion of the total raw score into a logit measure (centred at the mean item difficulty).

to assist in goal-setting and treatment planning to call attention to missing items that correspond to items that examinees find difficult to complete. To the best of our knowledge, there is a BBS keyform only for community-dwelling older adults (17), and there is no BBS keyform to measure a patient’s postural balance after stroke.

This study has shown how clinicians can design interventions using the BBS keyform. As shown in the example, the therapist can use the keyform to plan an intervention, by first administering the BBS and then circling the patient’s item ratings on the keyform. Next, the therapist finds the transition zone on the BBS keyform. It occurs above easier items on which patients show high performance (rating=3 or 4) and below more difficult items on which the patients show limited performance (rating=1 or 2) or the inability to perform (rating=0). Items in this transition zone represent the expected next steps as patients transition from their current skill level to a higher skill level.

For example, the patient with a moderately severe postural balance deficit represented in Fig. 3 achieved high ratings on easy items and lower ratings on more difficult items, and had fluctuating ability on moderately difficult items, represented by a region of back-and-forth ratings in the middle of the keyform. The BBS items within this zone are neither too easy nor too difficult and indicate postural balance ability at the “just right” challenge level. Postural balance tasks (items) within the transition zone suggest appropriate short-term goals, and exercise actions (items) beyond

the transition zone may form the basis for appropriate long-term goals.

The Rasch analysis in this study demonstrated that the BBS was highly reliable and that the examinees could be divided into 9 strata; the range of scores for each stratum was 4–8 points. This score range is also larger than the BBS’s minimal detectable change (MDC) 3–5 points for patients with stroke (25, 26), except for a single stratum. This suggests that the difference in strata exceeds the measurement error. The cut-off points of the BBS for stroke patients were 42 points for the ability to walk around the household (27) and 47–51 points for fall prediction (28, 29). Thus, if an examinee’s postural balance ability belongs to the either of the upper 2 strata, he or she may be at low risk of falling.

The results of the current study show that the stability of the BBS item hierarchy was robust in DIF analysis. In particular, the finding of no DIF for the intervention means that the BBS can be used safely to compare an individual’s postural balance before and after interventions, thus increasing the chance of drawing the correct conclusions about the effectiveness of interventions. A similar recent analysis of the Falls Efficacy Scale – International suggested that there were no items of DIF before and after rehabilitation (30). There is growing evidence that scales validated by Rasch analysis are more valid than scores validated by classical test theory (31, 32). Rasch analysis can also be used to test the robustness of items to changes in the ordinal score (e.g. before and after intervention) and may provide information about the choice of scales.

*Study limitations*

This study has some limitations. The study did not evaluate the patients’ data collected by the 3 hospitals or the intra-hospital reliability. However, the BBS was regularly conducted at each hospital by well-trained physical therapists who had participated in in-hospital training. The keyform is less useful for examinees with very high/low ability, because there is greater measurement error at these extremes due to floor/ceiling effects. In addition, as with all statistical values, the locations of transition-zone boundaries and the determination of short- and long-term goals are imprecise, and the resulting intervention suggestions are hypothetical. It is now necessary to determine whether interventions generated through the new keyform are more effective than traditionally induced clinical interventions.

**CONCLUSION**

The findings of this study indicate that the BBS has strong measurement properties. A keyform and strata that can be used with this scale were developed (Fig. S2). A keyform can be used to quickly compare response patterns from a clinical examination with those from Rasch analysis, which may help therapists use the BBS as a postural balance assessment scale in clinical practice to aid decision-making in goal-setting and intervention-planning. Moreover, as the conversion table and strata become clear, the postural balance ability of the tested patients could be better understood.

*The authors have no conflicts of interest to declare.*

**REFERENCES**

1. Jørgensen HS, Nakayama H, Raaschou HO, Olsen TS. Recovery of walking function in stroke patients: The Copenhagen Stroke Study. *Arch Phys Med Rehabil* 1995; 76: 27–32. DOI: 10.1016/s0003-9993(95)80038-7
2. Xu T, Clemson L, O’Loughlin K, Lannin NA, Dean C, Koh G. Risk factor for falls in community stroke survivors: a systematic review and meta-analysis. *Arch Phys Med Rehabil* 2018; 99: 563–573. DOI: 10.1016/j.apmr.2017.06.032
3. Mancini M, Horak FB. The relevance of clinical balance assessment tools to differentiate balance deficits. *Eur J Phys Rehabil Med* 2010; 46: 239–248.
4. Berg K. Measuring balance in the elderly: preliminary development of an instrument. *Physiother Can* 1989; 41: 304–311. DOI: 10.3138/ptc.41.6.304
5. Blum L, Korner-Bitensky N. Usefulness of the Berg Balance Scale in stroke rehabilitation: a systematic review. *Phys Ther* 2008; 88: 559–566. DOI: 10.2522/ptj.20070205
6. Moore JL, Potter K, Blankshain K, Kaplan SL, O’Dwyer LC, Sullivan JE. A core set of outcome measures for adults with neurologic conditions undergoing rehabilitation. *J Neurol Phys Ther* 2018; 42: 174–220. DOI: 10.1097/NPT.0000000000000229
7. Yeung LF, Ockenfeld C, Pang MK, Wai HW, Soo OY, Li SW, et al. Randomized controlled trial of robot-assisted gait

- training with dorsiflexion assistance on chronic stroke patients wearing ankle-foot-orthosis. *J Neuroeng Rehabil* 2018; 15: 51. DOI: 10.1186/s12984-018-0394-7
8. Straube D, Moore J, Leech K, Hornby TG. Item analysis of the Berg Balance Scale in individuals with subacute and chronic stroke. *Top Stroke Rehabil* 2013; 20: 241–249. DOI: 10.1310/tsr2003-241
9. La Porta F, Giordano A, Caselli S, Franchignoni F. Is the Berg Balance Scale an effective tool for the measurement of early postural control impairments in patients with Parkinson’s disease? Evidence from Rasch analysis. *Eur J Phys Rehabil Med* 2015; 51: 705–716.
10. La Porta F, Caselli S, Susassi S, Cavallini P, Tennant A, Franceschini M. Is the Berg Balance Scale an internally valid and reliable measure of balance across different etiologies in neurorehabilitation? A revisited Rasch analysis study. *Arch Phys Med Rehabil* 2012; 93: 1209–1216. DOI: 10.1016/j.apmr.2012.02.020
11. Kornetti DL, Fritz SL, Chiu YP, Light KE, Velozo CA. Rating scale analysis of the Berg Balance Scale. *Arch Phys Med Rehabil* 2004; 85: 1128–1135. DOI: 10.1016/j.apmr.2003.11.019
12. Chen H, Smith SS. Item distribution in the Berg Balance Scale: a problem for use with community-living older adults. *J Geriatr Phys Ther* 2019; 42: 275–280. DOI: 10.1519/JPT.0000000000000208
13. Bond TG, Fox CM. *Applying the Rasch Model: fundamental measurement in the human sciences*. 3rd edn. New York: Routledge; 2015.
14. Linacre JM. *A user’s guide to Winsteps Ministep 4.4.5: Rasch model computer programs*. Chicago: Winsteps; 2019.
15. Linacre JM. Instantaneous measurement and diagnosis. *Phys Med Rehabil* 1987: 1605–1610.
16. Velozo CA, Woodbury ML. Translating measurement findings into rehabilitation practice: an example using Fugl-Meyer Assessment-Upper Extremity with patients following stroke. *J Rehabil Res Dev* 2011; 48: 1211–1222. DOI: 10.1682/jrrd.2010.10.0203
17. PatientProgress.org. Keyform ability maps. (accessed July 27, 2022). Available from: <https://www.patientprogress.org/keyform-recovery-maps>
18. Goljar N, Giordano A, Vrbancic TSL, Rudolf M, Banicek-Sosa I, Albensi C, et al. Rasch validation and comparison of Slovenian, Croatian, and Italian versions of the Mini-BESTest in patients with subacute stroke. *Int J Rehabil Res* 2017; 40: 232–239. DOI: 10.1097/MRR.0000000000000233
19. Franchignoni F, Godi M, Guglielmetti S, Giordano A. Enhancing the usefulness of the Mini-BESTest for measuring dynamic balance: a Rasch validation study. *Eur J Phys Rehabil Med* 2015; 51: 429–437.
20. Mallinson T, Kozlowski AJ, Johnston MV, Weaver J, Terhorst L, Grampurohit N, et al. Rasch Reporting Guideline for Rehabilitation Research (RULER): the RULER Statement. *Arch Phys Med Rehabil* 2022; 103: 1477–1486. DOI: 10.1016/j.apmr.2022.03.013
21. de Winckel AV, Kozlowski AJ, Johnston MV, Weaver J, Grampurohit N, Terhorst L, et al. Reporting Guideline for RULER: Rasch Reporting Guideline for Rehabilitation Research: Explanation and Elaboration. *Arch Phys Med Rehabil* 2022; 103: 1487–1498. DOI: 10.1016/j.apmr.2022.03.019
22. Linacre JM. Sample size and item calibration stability. *Rasch Meas Trans* 1994; 7: 328.
23. Linacre JM. Optimizing rating scale category effectiveness. *J Appl Meas* 2002; 3: 85–106.
24. Grattan ES, Velozo CA, Skidmore ER, Page SJ, Woodbury ML. Interpreting action research arm test assessment scores to plan treatment. *OTJR* 2019; 39: 64–73. DOI: 10.1177/1539449218757740
25. Alghadir AH, Al-Eisa ES, Anwer S, Sarkar B. Reliability, validity, and responsiveness of three scales for measuring balance in patients with chronic stroke. *BMC Neurol* 2018; 18: 141. DOI: 10.1186/s12883-018-1146-9

26. Hiengkaew V, Jitaree K, Chaiyawat P. Minimal detectable changes of the berg balance scale, Fugl-Meyer assessment scale, timed "up & go" test, gait speed, and 2-minute walk test in individuals with chronic stroke with different degrees of ankle plantarflexor tone. *Arch Phys Med Rehabil* 2012; 93: 1201–1208. DOI: 10.1016/j.apmr.2012.01.014
27. Joa KL, Kwon SY, Choi JW, Hong SE, Kim CH, Jung HY. Classification of walking ability of household walker versus community walkers based on K-BBS, gait velocity and upright motor control. *Eur J Phys Rehabil Med* 2015; 51: 619–625.
28. Sahin IE, Guclu-Gunduz A, Yazici G, Ozkul C, Volkan-Yazici M, Nazliel B, et al. The sensitivity and specificity of the balance evaluation systems test-BESTest in determining risk of fall in stroke patients. *NeuroRehabilitation* 2019; 44: 67–77. DOI: 10.3233/NRE-182558
29. Tsang CSL, Liao LR, Chung RCK, Pang MYC. Psychometric properties of the Mini-Balance Evaluation Systems Test (Mini-BESTest) in community-dwelling individuals with chronic stroke. *Phys Ther* 2013; 93: 1102–1115. DOI: 10.2522/ptj.20120454
30. Caronni A, Picardi M, Redaelli V, Antoniotti P, Pintavalle G, Aristidou E, et al. The Falls Efficacy Scale International is a valid measure to assess the concern about falling and its changes induced by treatments. *Clin Rehabil* 2022; 36: 558–570. DOI: 10.1177/02692155211062110
31. Draak TH, Vanhoutte EK, Van Nes SI, Gorson KC, der Pol WLV, Notermans NC, et al. Changing outcome in inflammatory neuropathies: Rasch – comparative responsiveness. *Neurology* 2014; 83: 2124–2132. DOI: 10.1212/WNL.0000000000001044
32. Caronni A, Picardi M, Zaina F, Negrini S. The Italian spine youth quality of life questionnaire measures health-related quality of life of adolescents with spinal deformities better than the reference standard, the scoliosis research society 22 questionnaire. *Clin Rehabil* 2019; 33: 1404–1415. DOI: 10.1177/0269215519842246