Assessments of the Reliability of the Iranian Version of the Berg Balance Scale in Patients with Multiple Sclerosis

Akram Azad, Ghorban Taghizadeh, Arash Khaneghini

Abstract-

Purpose: Because of the balance limitations in many patients, balance assessment is necessary for multiple sclerosis patients in rehabilitation settings. The aim of this study was to investigate the Interrater reliability and the internal consistency of the Iranian version of the Berg Balance Scale (BBS) when applied to patients with multiple sclerosis (MS) in Tehran.

Methods: Fifty MS patients (with mean age of 36.6±9.5 years) from Hospitals of the Iran University of Medical Sciences and MS Society of Iran were included. Interrater reliability was measured with the Kappa statistics and Intraclass Correlation Coefficients (ICCs).

Results: The mean values of the BBS scored by the 2 evaluators were 37.7 ±12.9 and 38.1 ± 12.3, respectively. Kappa scores for BBS varied from 0.7 to 1.0 Intraclass correlation coefficient for the BBS's sum score was excellent (ICC=0.99 with 95% confidence interval, 0.98-0.99). An excellent internal consistency was found within the BBS's sum score (Cronbach Alpha =0.9). The item -to -total correlations for all items were higher than 0.6.

Conclusion: The Iranian version of the BBS has excellent interrater reliability and internal consistency for the assessment of MS patients when applied in clinics.

Key Words: Berg balance scale, multiple sclerosis, outcome assessment

Acta Neurol Taiwan 2011;20:22-28

INTRODUCTION

Multiple sclerosis (MS) is an autoimmune human disease without any fully effective treatment and well known pathogenesis. Extensive demyelization is seen in the neuronal lesions [1-2]. Disability increases steadily over time. Because of the balance limitations in many patients [3], balance assessment is necessary for multiple sclerosis subjects in rehabilitation settings in order to help establish appropriate treatment goals, to increase awareness of fall risk and to assign appropriate assistive devices [4]. The ability to maintain balance or postural control is important for the correct execution of all daily tasks ranging from standing and walking to sitting and rising from a chair [5].

The Berg Balance Scale (BBS) is commonly used as

From the Department of Occupational Therapy, Iran University of Medical Sciences, Tehran, Iran.
Received October 13, 2009. Revised December 9, 2009. Accepted December 22, 2010.
Correspondence to: Akram Azad, P.O.Box: 15875-4391, Tehran, Iran.
E-mail: a_azad@tums.ac.ir

a clinical measure for testing balance abilities in MS patients [6]. Thus it is a usual used laboratory measure of balance [7]. However, the BBS consists of items designed to detect static and dynamic balance in terms of postural maintenance and execution of voluntary movements [8-9]. Several studies have shown high levels of inter and intratester agreement for the test as a whole and for the individual items in different residents and patients [10-14].

During rehabilitation period, more than one occupational therapist or physiotherapist may assess a multiple sclerosis patient, therefore high interrater reliability is essential. Because errors may occur in testing, high reliability is required when repeated measures are used to monitor the clinical status of patients or to evaluate the effectiveness of the treatment. The BBS has been translated into Farsi by Keivan Davatgharan et al [15], but the reliability of the translated version has not been evaluated in MS patients. One reason for this study on Iranians is the possibility of participating in international clinical trails that use this instrument. Another reason is that it can safely be assumed that studies using the English language version could be applicable to MS adults in Iran.

The purposes of this study were to assess the interrater reliability of the Iranian version of the BBS in multiple sclerosis patients when applied to assess the internal consistency, and to investigate how different scoring levels of the 14 items were used.

METHODS

Participants

Fifty over 25 year old MS patients were identified during the study; 12 were inpatients in hospitals of the Iran University of Medical Sciences, whereas 38 were admitted to MS Society of Iran.

The subject inclusion criteria were : (1) to have clinically definite multiple sclerosis [16], (2) not to have a clinically apparent relapse within three months before entry (to be in a stable clinical state), (3) to have impaired balance due to MS disease, (4) not to have any other pathology (orthopedic or neurologic) that affects their ability to balance, (5) to be willing to participate in the study, (6) to be able to communicate sufficiently to participate i.e., to be able to follow five steps of verbal instructions (7), to be independent in ambulation with or without an assistive device (8), not to have major cognitive involvements (Mini Mental Status Examination, MMSE>23).

Subjects were excluded if they had per- morbid or co -morbid neurological problems unrelated to MS (n=2) or were currently receiving medications known to affect balance (n=5). Subjects with sever pain (Visual Analog Scale >70 mm) were excluded (n=6) because we thought that pain would affect their performance. Ethical approval was granted by Ethics Committees in Rehabilitation Faculty of Iran University of Medical Sciences. Informed consent was obtained from all patients before enrollment.

The mean age was 36.6 years (25-55 year). 33 subjects were women and 17 were men. The range of Expanded Disability Status Scale (EDSS) score was 3.5 to 6 (moderate to sever level of disability). All the subjects were ambulatory. Data on demographic characteristics and co- morbidity were collected from medical records. Twenty one subjects did not require walking aids, 15 used a cane, and 14 used walking frames.

Procedure

The BBS was administered by 2 occupational therapists simultaneously on the same patients. They alternated between instructing and scoring as well as observing and scoring. The therapists, who administered these tests, were blind to each other's results. All the tests were performed in the same room. The patients were tested once only. This model was chosen because all the patients were undergoing rehabilitation and their conditions could have been improved if they had been tested on 2 days to prevent a habituation bias. Two experienced observers were accustomed to using the standardized of administering the test. Before administering the study, they had trained and practiced 3 weeks with the Iranian version of the BBS including discussing and comparing results of testing in order to be quite sure how details concerning the patients' performances should be scored. A digital chronometer, a 30- cm ruler, a 20- cm high stool, a 42- cm high chair with a backrest and no armrest, and a 42- cm high chair with a backrest and armrest

were used for the assessment.

Instrument

The BBS [17] evaluates a person's performance on 14 items (5 items being static and 9 items dynamic) related to balance function that are frequently encountered in everyday life. The scoring method is based on a 5- point ordinal scale of 0 to 4, with the total score ranging from 0 to 56 [17] and greater scores indicating better balance. The BBS was originally developed for screening the elderly at risk of falling [17], but the psychometric properties at the BBS used in MS patients have also been examined by various researchers with supportive results[18-19].

Cross - cultural translation

The method used to produce the Iranian version of the BBS by Davatgharan was the forward-backward translation method [23], including the following steps.

Step one: The original version of BBS was translated into Iranian version in this step. English- Iranian translators, native Iranian speakers were involved. Each translator independently translated the BBS and then compared and discussed the result with that of the other, until a common version was reached.

Step two: Back - translation of the Iranian version of the BBS into English was administered in this step: The preliminary version was given to 2 native English people who were experienced translators, each producing a translation into English. These translators were unaware of either the methodology or the aims of the study.

Statistical analysis

The interrater agreement on individual items of the BBS was analyzed with the weighted Kappa Statistic (k). The weighted Kappa score measures the agreement among raters adjusted for the amount of agreement expected by chance and the magnitude of disagreements[20]. A Kappa value >0.7 indicates excellent agreement, 0.4 to 0.7 indicates fair to good agreement, and <0.4 indicates poor agreement [21].

The interrater reliability of the total score of the BBS was analyzed with the intraclass correlation coefficient (ICC) statistic. An ICC of 0.8 or higher reflects high reliability, 0.6 to 0.8 moderate reliability, and less than 0.6 indicates that reliability is poor [20-21]. The floor and ceiling effects of the sum score reflect the extent to which scores cluster at the bottom and top of the scale range. Floor and ceiling effects of more than 20% are considered to be significant [11].

Table 1. *Distribution of Sum Scores of 2 Raters within each of the14 items of the BBS (N=50)*

	Scoring Values					
Item	0	1	2	3	4	mean
1. Sitting to standing	3	2	4	21	70	3.5
2. Standing unsupported	12	0	8	2	78	3.3
3. Sitting unsupported	0	0	0	0	100	4.0
4. Standing to sitting	1	1	3	25	70	3.6
5. Transfers	0	8	4	26	62	3.4
6.Standing with eyes closed	15	3	10	14	58	2.9
7.Standing with feet together	40	6	1	9	44	2.1
8.Reaching forward with outstretched arm	8	6	8	37	41	2.9
9. Retrieving an object from floor	8	3	1	16	72	3.4
10.Turning to look behind	14	7	18	4	57	2.8
11.Turning 360	31	11	22	12	24	1.8
12.Placing alternate foot on stool	68	5	3	11	13	0.9
13. Standing with 1 foot in front	29	8	16	46	1	1.8
14. Standing on 1 foot	28	49	17	4	2	1.0
Total	257	1.9	115	227	692	-

Factor analysis with varimax rotation was performed to test the construct validity and dimensionality of the BBS. Internal consistency of the BBS was tested by item - to - total correlation and by calculating the Cronbach for each tester's scorings. The Cronbach is regarded as high if it is at least 0.8 [20]. An item - to - total correlation shows the degree of relationship between each item and the total score of the other items in the scale. An item - to - total correlation is considered adequate if it is above 0.4 [22].

RESULTS

Participants were 25 to 55 year old with a mean of 36.6 years (standard deviation 9.5). There were 17 males and 33 females. The mean values of the BBS scored by the 2 evaluators were 37.7±12.9 and 38.1± 12.3, respectively. The mean value of the disease duration was 0.1±0.3 year. The items are presented in Table 1. There were no relationships between age, disease duration and gender with total score of BBS.

Distribution

Table 1 shows the distributions of the BBS from o to 4 scores at all of the 14 items and the total score. Some rating categories were not used at all, and others were used very sparingly. On the whole, each tester completed 700 scores. The score values 0, 1, 2, 3 and 4 were used in 18.4%, 7.8%, 8.2%, 16.2%, and 49.4% of the times, respectively. The items placing alternated foot on stool and standing on 1 foot had the lowest mean score, indicating a greater degree of difficulty. The item sitting unsupported had mean score 4.

Reliability and Construct

Kappa scores for BBS varied from 0.7 to 1.00 and the mean Kappa was 0.8%. The extent of agreement (Kappa) between scores for each of the items obtained by both evaluators was excellent (Table 2). With the exception of sitting to standing, turning to look behind, and placing the alternate foot on stool with fair to good agreement (Table 2).

The evaluators scored differently on only 75 occasions out of the total 700. Interclass correlation coefficient for inter-rater rating for the BBS's sum score was excellent (ICC = 0.9 with 95% confidence interval, 0.9 - 0.9).

The 3 and 2 scores difference were 3 and l0 times, respectively, which were related to standing to sitting, placing alternate foot on stool, and standing with 1 foot in front for 3 score difference and standing with feet together, turning to look behind , turning 360, and standing on 1 (foot for 2 score difference. The 1 score difference was 62 occasions.

Factor analysis on the 14 items of the BBS gave 2 factors. Together, the 2 factors accounted for 70.2% of the matrix variance (58.7%, and 11.4%, respectively). The first and second factors are shown in Table 3. (Insert table 3 about here)

An excellent internal consistency was found within the BBS'S sum score (cronbach = 0.9). The correlation

Table 2. *Intra-rater Reliability Coefficient (Kappa) for each Item of the BBS (N=50)*

Item	Kappa	Strength of Agreement
1. Sitting to standing	0.7	fair to good
2. Standing unsupported	1.0	excellent
3. Sitting unsupported	*	*
4. Stonding to sitting	0.8[a]	excellent
5. Transfers	0.8	excellent
6. Standing with eyes closed	0.	
7. Standing with feet together	0.9[b]	excellent
8. Reaching forward with outstretched arm	0.9	excellent
9. Retrieving an object from floor	0.8	excellent
10. Turning to look behind	0.7	fair to good
11. Turning 360	0.8	excellent
12. Placing alternate foot on stool	0.9	fair to good
13. Standing with 1 foot in front	0.7[c]	fair to good
14. Standing on 1 foot	0.7	fair to good

* Everyone scored 4.

a. Rating categories 0 and 1 are merged because the 2 score levels were not used by both raters.

b. Rating category 2 is merged because the1 score level was not used by 2 raters.

c. Rating category 4 is merged because the 1 score level was not used by 2 rates.

Table 3. *Result of Factor Analysis of each Iranian BBS Item*

Item	Name of the component	
	Component 1 (Dynamic)	Component 2 (Static)
1. Sitting to standing	0.9	†
2. Standing unsupported	0.8	†
3*. Sitting unsupported	†	†
4. Standing to sitting	0.8	†
5. Transfers	0.8	†
6. Standing with eyes closed	0.6	†
7. Standing with feet together	†	0.7
8. Reaching forward with outstretched arm	0.7	†
9. Retrieving an object from floor	0.9	†
10. Turning to look behind	0.6	0.4
11. Turning 360	†	0.8
12. Placing alternate foot on stool	†	0.9
13. Standing with 1 foot in front	0.4	0.4
14. Standing on 1 foot	†	0.7

* Everyone scored 4 (not entered in the factor analysis).

† Item bad scored below 0/403 on the factors.

matrix, determined for the 13 items and items - to - total correlation, is shown in table 4. A correlation coefficient could not be calculated for item 3 (sitting unsupported) because the scores did not vary. The all item - to - item correlations were significant (r ranged from 0.2 to 0.8). The item - to - total correlations for all items were higher than 0.6 (r ranged from 0.6 to 0.8).

DISCUSSION

Individual items which comprise the Berg Balance Scale score were examined using kappa statistic. Strength of agreement was generally excellent agreement with Kappa value of 0.7 or higher. However 3 items had fair to good agreement (Kappa> 0.7). Our study shows good and excellent intertester reliability when using Iranian version of BBS to evaluate balance of patients under the multiple sclerosis rehabilitation. Although it is not possible to compare the results with other studies on MS patients, the results support the studies that have been reported from different populations and countries [10,13,23]. The generalization of the findings is not strengthened by the lack of control of the test conditions.

The ICC for the sum score of BBS for inter-rater reliability is similar to those from other studies [24,25]. In short the finding of "Fair" to "excellent" agreement of individual BBS items and high intraclass correlation coefficient of total BBS scores -obtained by testing with different observers - is reassuring. However, there are several caveats that apply to this conclusion. Firstly, the BBS may be less reliable in patients with cognitive impairment. Because in this study, the subjects were without any major cognitive involvements (MMSE > 23). The effect on the results cannot be estimated. Secondly, reliability of the BBS is influenced by the degree of disability in the population examined and is higher in subjects who are independent. The range of EDSS score was from moderate to severe in our study. The effect on the findings cannot be detected. Lastly, the demonstration of reliability is dependent on the choice of statistical methods [26,27], but kappa and ICC are acceptable for measuring agreement [27]. Therefore, we cannot be determining the effect of the results. In agreement with many previous studies [25, 26], our findings indicated that the internal consistency of BBS was equally high. All of the item - to -total correlation coefficients are above 0.6 (Table 4). The very high internal consistency of the BBS showed that the items of this instrument measured the same concept: balance. However, the extremely high internal consistency of BBS might indicate the possibility of item redundancy, which needs further study.

Factor analysis in our study showed that 2 factors have emerged (Table 3). The first factor relates to the ability to maintain postural control when changing position or condition (dynamic aspect). The second factor assesses the static aspects of balance. These results are supported by finding of other studies which have shown that only 1 or 2 factors have emerged [13, 27].

We found no variability between patients in the sitting unsupported item, which fits well with the results in the other studies [12,13], which reported that more than 90% had a top score on this item, indicating a very low degree

of difficulty. By condensing item - rating categories, we could eliminate underused categories and construct categories that separated people of differing abilities better. The score values 0, 3, and 4 were significantly more frequently used than the score values 1 and 2, indicating that 3 or 4 levels might be better than 5 levels in our study. This is supported by the findings of Wang [28] and colleagues.

In conclusion, the Iranian version of the BBS has good interrater reliability and internal consistency for the assessment of the MS patients. We think that it should be considered for use in clinical practice and in research.

ACKNOWLEDGEMENTS

Participating facilities include hospitals of the Iran University of Medical Sciences and the MS Society of Iran. We thank all patients for their co-operation, the Occupational Therapy Department at Rehabilitation Faculty of Iran University of Medical Sciences and Ms. Jazaeari N. for their help. We take this opportunity to express our sincere appreciation to experts in Farzan Institute for Research and Technology for their technical assistance.

REFERENCES

1. Bauer HJ, Hanefeld FA. Multiple sclerosis: Its impact from childhood to old age London: WB. Saunders; 1993.
2. Boyden KM. The path physiology of demyelization and the ionic basis of nerve conduction in multiple sclerosis: An overview. J Neurosis Nurs 2000; 32: 49 - 53.
3. Vaney C, Vaney S, Wade DT. SaGAS, the short and graphic ability score: an alternative scoring method for the motor components of the multiple sclerosis functions composite. Mul Scl 2004; l0: 231 - 247.
4. Soyuer F, Mirza M, Erkorkmaz V. Balance performance in three forms of multiple sclerosis. Neural Res 2006; 28 : 555 - 562.
5. Schultz AB, Alexander NB, Ashton - Miller JA. Biomechanical analyses of rising from a chair. J Biomechanics 1992; 25: 1381-1391.
6. Lord SE, Wade DT, Halligan PW. A comparison of two physiotherapy treatment approaches to improve walking in multiple sclerosis: a pilot randomized controlled study. Clin Rehabil 1998; 12: 477 - 486.
7. Frzovic D, Morris ME, Vowels L. Clinical tests of standing balance: Performance of persons with multiple sclerosis. Arch Phys Med Rehabil 2000; 81: 215 - 221.
8. Dleiden S. Weight shifting as a treatment for balance deficits: A literature review. Physiother Can. 1990; 42: 81 - 87.
9. Shepherd RB. Adaptive motor behavior in response to perturbations of balance. Physiother Theory Pract 1992; 8: 137 - 143.
10. Berg KO, Wood - Dauphinee SL, Williams JI. The balance scale: reliability assessment with elderly residents and patients with acute stroke. Scand J Rehabil Med 1995; 27: 27-36.
11. Mao HF, Hsueh IP, Tang PF, Sheu CF, Hsieh CL. Analysis and comparison of the psychometric properties of three balance measures for stroke patients. Stroke 2002; 33: 1022 - 1027.
12. Norén AM, Bogren U, Bolin J, Stenström C. Balance assessment in patients with peripheral arthritis: Applicability and reliability of some clinical assessments. Physiother Res Int 2001; 6: 193 - 204.
13. Ottonello M, Ferriero G, Benevolo E, Sessarego P, Dughi D. Psychometric evaluation of the Italian Version of the Berg Balance Scale in rehabilitation inpatients. Eur Med Phys 2003; 39: 181 - 189.
14. Miyamoto ST, Lombardi Junior I, Berg KO, Ramos LR, Natour J. Brazilian Version of the Berg Balance Scale. Braz J Med Biol Res 2004; 37: 1411 - 1421.
15. Davatgharan K. Translation and normalization of Iranian Version of the Berg Balance Scale in older. Tesis of Physical Therapy Master Sciences, Rehabilitation Welfare University of Tehran, 2004.
16. Poser CM, Paty DW, Scheinberg L, McDonald WI, Davis FA, Ebers GC, Johnson KP, Sibley WA, Silberberg DH, Tourtellotte WW. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. Ann Neurol 1983; 13: 227 - 231.
17. Berg KO, Wood-Dauphinee SL, Williams JI, Maki B. Measuring balance in the elderly: preliminary development of an instrument. Can J Public Health 1989; 41: 304 - 311.

18. Cattaneo D, Regola A, Meotti M. Validity of six balance disorders scales in persons with multiple sclerosis. Disabil Rehabil 2006; 28: 789 - 795.
19. Paltamaa J, West H, Sarasoja T, Wikström J, Mälkiä E. Reliability of physical functioning measures in ambulatory subjects with MS. Physiother Res Int 2005; 10: 93 -109.
20. Domholdt E. Rehabilitation research: principles and applications. Philadelphia: WB Saunders; 2005.
21. McCluggage WG, Bharucha H, Caughley LM, Date A, Hamilton PW, Thornton CM, Walsh MY. Hamilton PW, Thornton CM, Walsh MY. Interobserver variation in the reporting of cervical colposcopie biopsy specimens: comparison of grading systems. J Clin Pathol 1996; 49: 833 - 835.
22. Altman DG. Practical statistics for medical research. London: Chapman and Hall; 1991.
23. Tyson SF, DeSouza LH. Reliability and validity of functional balance tests post stroke. Clin Rehabil 2004; 18: 916 - 923.
24. Harada N, Chiu V, Damron-Rodriguez J, Fowler E, Siu A, Reuben DB. Screening for balance and mobility impairment in elderly individuals living in residential care facilities. Phys Ther 1995; 75: 462 - 469.
25. Wang CY, Hsieh CL, Olson SL, Wang CH, Sheu CF, Liang CC. Psychometric properties of the Berg Balance Scale in a community - dwelling elderly resident population in Taiwan. J Formos Med Assoc 2006; 105: 992 - 1000.
26. Berg K, Norman KE. Functional assessment of balance and gait. Clin Geriatr Med 1996; 12: 705 - 723.
27. Berg KO, Maki BE, Williams JI, Holliday PJ, Wood-Dauphinee SL. Clinical and laboratory measures of postural balance in an elderly population. Arch Phys Med Rehabil 1992; 73:1073 - 1080.
28. Wang CH, Hsueh IP, Sheu CF, Yao G, Hsieh CL. Psychometric properties of 2 simplified 3-level balance scales used for patients with stroke. Phys Ther 2004; 84: 430 - 480.