

# Validity of balance measures in cerebellar ataxia: A prospective study with 12-month follow-up

Stanley J. Winser PT PhD<sup>1,2</sup>  |

Anne Y. Y. Chan MB ChB, MRCP (UK), FHKCP, FHKAM<sup>3,4</sup> | Raymond Chung PhD<sup>1</sup> |

Susan Whitney PhD FAPTA<sup>5</sup> | Priya Kannan PT PhD<sup>1</sup>

<sup>1</sup>Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hong Kong

<sup>2</sup>Centre for Eye and Vision Research (CEVR), The Hong Kong Polytechnic University, Kowloon, Hong Kong

<sup>3</sup>Division of Neurology, Prince of Wales Hospital, Hong Kong

<sup>4</sup>Department of Medicine and Therapeutics, Chinese University of Hong Kong, Hong Kong

<sup>5</sup>School of Health and Rehabilitation Sciences, University of Pittsburgh, USA

## Correspondence

Stanley Winser, Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Kowloon, Hong Kong.  
Email: [stanley.j.winsler@polyu.edu.hk](mailto:stanley.j.winsler@polyu.edu.hk)

## Funding information

Start-up fund for early careers, The Hong Kong Polytechnic University (Ref: 1-ZE6J). Early Career Scheme (ECS), RGC, HK (Ref: PP6A)

## Abstract

**Introduction:** Balance deficits are common in cerebellar ataxia. Determining which balance outcome measures are psychometrically strong for this population remains an unmet need.

**Objective:** To evaluate the validity and responsiveness of two clinic-based balance measures (Berg Balance Scale [BBS] and balance sub-component of the Scale for the Assessment and Rating of Ataxia [SARA-bal]) and two laboratory-based balance measures (Sensory Organization Test [SOT] and Limits of Stability [LOS]) in cerebellar ataxia.

**Design:** Prospective cohort study.

**Setting:** Institutional study assessing 40 participants with cerebellar ataxia at baseline, and 6 and 12 months.

**Main outcome measures:** Balance was assessed using the BBS, SARA-bal, SOT, and LOS; disease severity was assessed using the SARA; and Patients' Global Impression of Change (PGIC) was used to estimate responsiveness to disease progress at 6 and 12 months.

**Results:** BBS and SARA-bal (Spearman's correlation coefficient,  $\rho_S = -0.89$ ,  $p < .01$ ) demonstrated strong criterion validity. Convergent validity was moderate to high ( $\rho_S$  range:  $-0.75$  to  $0.92$ ) and external validity was low ( $\rho_S$  range:  $-0.75$  to  $0.11$ ). Composite SOT scores (SOT-COM;  $\rho_S = 0.29$ ,  $p < .01$ ) and maximal excursion (MXE-LOS) in the forward (F) and right (R) directions of the LOS ( $\rho_S = 0.18$ ,  $p < .01$ ) demonstrated moderate to low criterion and convergent validity. The area under the receiver-operating characteristic curve ( $AUC_{ROC}$ ) and its effect size (standard response mean [SRM]) for categorizing "stable" and "worsened" patients at 6 and 12 months were satisfactory for the BBS ( $AUC_{ROC}$ : 0.75; SRM-Stable: 1.06; SRM-Worsened: 1.16), SARA-bal ( $AUC_{ROC}$ : 0.76; SRM-Stable: 0.86; SRM-Worsened: 0.85), and MXE-LOS (R) ( $AUC_{ROC}$ : 0.29; SRM-Stable: 0.41; SRM-Worsened: 1.39).

**Conclusion:** BBS and SARA-bal have moderate to strong criterion and convergent validity and adequate responsiveness to balance changes. Both laboratory-based measures (SOT and LOS) demonstrated a high floor effect. The SOT-COM and MXE-LOS(R) demonstrated moderate to low criterion validity, with only the MXE-LOS(R) displaying adequate responsiveness to balance changes after 6 and 12 months.

## INTRODUCTION

Cerebellar ataxia is not a single disease but refers to a collection of symptoms associated with cerebellar damage or dysfunction. Depending on the underlying cause, cerebellar ataxia can be categorized as genetic, sporadic, or acquired onset.<sup>1</sup> The prevalence of cerebellar ataxia varies across geographical and ethnic groups.<sup>2</sup> Friedreich's ataxia, a genetic disorder, has a prevalence of 2 to 5 per 100,000,<sup>3</sup> and the incidence of spinocerebellar ataxia ranges from 0.9 to 3 per 100,000, depending on the sub-type.<sup>4</sup> Individuals with cerebellar ataxia report a significant decline in quality of life<sup>5</sup> and a greater reduction in the performance activities of daily living.<sup>6</sup> Cerebellar ataxia is associated with high economic costs, with mean per-patient annual costs reported at EUR 18,776 in Spain (2004),<sup>7</sup> USD 12,850 in the United States (2010),<sup>8</sup> and EUR 59,993 in Ireland (2016).<sup>9</sup> The 6-month per-patient annual cost in Hong Kong is reported to be HKD 146,832 (2020).<sup>10</sup>

Balance deficits in cerebellar ataxia are common due to a lack of rhythmic muscle contractions,<sup>11</sup> movement planning and initiation deficits,<sup>12</sup> increased body sway,<sup>13</sup> and poor anticipatory postural adjustments.<sup>14</sup> Up to 93% of patients with cerebellar ataxia report experiencing at least one fall within 12 months,<sup>15</sup> and 87% of patients report experiencing frequent falls.<sup>10</sup> Improving balance to prevent falls and fall-related consequences is a primary focus of rehabilitative interventions in this population. The accurate estimation of balance function requires psychometrically strong and standardized assessment tools, and a wide spectrum of balance outcome measures is currently available. Previous reviews reporting on the currently available balance outcome measures<sup>16</sup> have highlighted the use of specific recommended outcome measures in patients with cerebellar ataxia based on psychometric property testing in this population.<sup>17</sup>

Based on the findings of our previous systematic review<sup>17</sup> and a Delphi survey on choosing an appropriate outcome measure for testing balance conducted among physiotherapists and neurologists specializing in cerebellar ataxia,<sup>18</sup> we identified and tested the reliability and validity of the Berg Balance Scale (BBS)<sup>19</sup> and the balance sub-components of Scale for the Assessment and Rating of Ataxia (SARA-bal) among a population of people with cerebellar ataxia secondary to multiple sclerosis. We found these two measures to be both reliable and valid for assessing balance at the activity level in this population.<sup>20</sup> Genetic and sporadic diseases that result in cerebellar ataxia are often chronic and progressive; therefore, outcome measures used to assess symptoms in these populations must be responsive to the change in disease status due to natural progression. However, evaluations of the responsiveness of the BBS in this population are lacking. Two laboratory-based assessments, the Sensory

Organization Test (SOT) and the Limits of Stability (LOS) test, are accurate and objective measures of balance at the body, structure and function, and activity levels.<sup>16</sup> However, the psychometric validation of these two measures among people with cerebellar ataxia has not yet been established. Therefore, the current study aimed to test the validity (criterion, convergent and external validity) and responsiveness of four balance measures, including two clinic-based measures (BBS and SARA-bal) and two laboratory-based measures (SOT and LOS), among individuals with genetic and sporadic cerebellar ataxia.<sup>1</sup> Based on the findings of our previous work, we hypothesize a moderate to strong correlation between all four balance measures. We also hypothesize that the ataxia-specific balance outcome measure (SARA-bal) will be the most responsive measure to changes in balance after 6 and 12 months.

## METHODS

### Participants

Forty participants with a confirmed diagnosis of cerebellar ataxia were enrolled from the Hong Kong Spinocerebellar Ataxia Association (HKSCAA) between April 2018 and February 2020. Ethical approval for this study was obtained from the Human Subjects Ethics Application Review System of Hong Kong Polytechnic University (Ref: HSEARS20160714003), and the study was registered with the Australian New Zealand Clinical Trials Registry (Ref: ACTRN12618000605291). Volunteers were included in the study if they had a confirmed diagnosis of cerebellar ataxia secondary to a genetic or sporadic cause and were ambulant either with or without an assistive walking device. We excluded patients who were (1) unwilling to reveal their personal information, (2) bed-ridden, (3) having difficulty understanding simple commands, and (4) diagnosed with cerebellar ataxia due to acquired lesions, such as cerebellar stroke, cerebral palsy, trauma, or multiple sclerosis.<sup>1</sup>

### Procedure

All participants were assessed on three occasions, and each assessment lasted for ~90 minutes. The timeline included assessments at baseline, and after 6 and 12 months. To ensure the repeatability of all assessments, the same assessor performed the assessments at all three assessment time points. The assessments were conducted at the neurorehabilitation research laboratory of Hong Kong Polytechnic University. The venue, lighting, types of equipment used, and instructions provided during the assessments were standardized across all three assessments.

## Instruments

### Berg Balance Scale (BBS)

The BBS is a generic measure of balance<sup>19</sup> that includes 14 items relevant to static and dynamic balance, with each item scored between 0 and 4. Higher scores indicate better balance function. The BBS has been shown to be a reliable and valid measure for balance assessments in multiple neurological disorders. For patients with cerebellar ataxia, our previous study reported high intra-rater reliability (intraclass correlation coefficient [ICC] = 0.99), inter-rater reliability (ICC = 0.97), internal consistency (Cronbach's  $\alpha$  = 0.94), and construct and criterion validity, with a minimum detectable change value of 3 points.<sup>20</sup>

### Scale for the Assessment and Rating of Ataxia (SARA) and its balance sub-component (SARA-bal)

The severity of cerebellar ataxia was assessed using the SARA.<sup>21</sup> A higher score indicates increased cerebellar ataxia severity. The "gait," "standing," and "sitting" sub-components of the SARA can be scored independently and together are referred to as the SARA-bal.<sup>20</sup> The SARA has been reported to have high inter-rater reliability (ICC = 0.98),<sup>21</sup> test-retest reliability (ICC = 0.90),<sup>21</sup> internal consistency (Cronbach's  $\alpha$  = 0.94), and construct validity<sup>20</sup> for estimating disease severity in patients with cerebellar ataxia.

### Sensory Organization Test (SOT)

The SOT is an objective laboratory-based assessment of dynamic balance. We used a Bertec Balance Advantage Dynamic Posturography (CDP) System with a dual-balance force plate and an immersive virtual reality visual surround to assess the SOT. Participants were instructed to stand on the motion-sensing dual-balance force plate with both medial malleoli aligned with the foot markings. To ensure safety, we used a dynamic balance harness for all participants. To protect against fatigue, we offered one practice trial and one test assessment. The participants were free to take any number of breaks during the assessment. Assessments were considered incomplete if participants lost their balance in the middle of the assessment or informed the assessors of any inability to proceed further. The methods used to replace missing values are discussed in the statistics section. The assessor used a touchscreen monitor to conduct each assessment and record assessment results. The participants were instructed to look at the immersive virtual reality visual surround screen in front of them and to follow instructions provided by the assessor. The SOT estimates the influences of visual, vestibular, and somatosensory inputs for balance maintenance.<sup>22</sup> Each participant's ability to stand

unsupported was tested across six sensory conditions, using combinations of occluded vision, visual reference, and a swaying surface. Each condition lasted for 20 seconds, with three repeated measurements. In this study, we assessed the Bertec-generated visual (SOT-VIS), vestibular (SOT-VES), somatosensory (SOT-SOM), and composite scores (SOT-COM). Appendix 1A illustrates the assessment set-up for the SOT. The SOT has been reported to have good test-retest reliability (ICC = 0.68) among older adults<sup>23</sup> and is valid for assessing balance among patients with vestibular disorders.<sup>24</sup>

### Limits of Stability (LOS)

The LOS is an objective laboratory-based assessment of dynamic balance that assesses the ability to shift one's body weight in eight different directions without moving the feet.<sup>25</sup> Body sway in all eight directions is guided by visual feedback, provided by the Bertec system described earlier, on the immersive virtual reality visual surround screen. In this study, we estimated the reaction time (RT-LOS) and the maximal excursion (MXE-LOS) in four directions, including forward, right, back, and left. RT refers to the time delay between the command to move and movement onset and is expressed in milliseconds. The MXE is the maximum displacement of the center of pressure in a given direction. The MXE is expressed as a percentage, with a higher percentage indicating better dynamic control during weight shift in the tested direction. Appendix 1B illustrates the assessment set-up for the LOS.

### Other measures

Independence in activities of daily living was assessed using the Barthel Index.<sup>26,27</sup> The EuroQol visual analog scale (EQ-VAS) was used to assess quality of life.<sup>28,29</sup> We used the Patients' Global Impression of Change (PGIC) scale<sup>30</sup> to categorize all included participants into either stable and/or improved (stable) or worsened groups, based on the assessment of disease progression at each follow-up. This classification system was used to anchor the responsiveness of the outcome measures of interest. The PGIC is a self-reported scale that requires the individual to indicate their perceptions of disease progress. In this study, we required the participants to select either "stable and/or improved," if they perceived no change or reduced cerebellar ataxia symptoms, or "worsened," if they perceived any decline in their health status compared with the baseline assessment.

### Sample size and psychometric properties tested

COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) recommends a

sample size of 30 to 49 to provide “fair” quality of evidence for psychometric property estimation.<sup>31</sup> Considering the limited number of patients with cerebellar ataxia in Hong Kong,<sup>10</sup> we targeted the recruitment of at least 30 participants for this study. The criterion validity is defined as the degree to which the scores of the assessment tool being investigated correlate well with a gold standard.<sup>32</sup> Based on the findings of previous studies supporting the reliability, validity,<sup>20</sup> and appropriateness<sup>33</sup> of the BBS and SARA-bal for assessing balance in people with cerebellar ataxia, we consider these assessments to be gold standards. To estimate criterion validity, we correlated the four balance outcome measures against each other, which is in line with a previous study.<sup>20</sup> For convergent validity, we correlated the measures of balance against disease severity and functional independence, and the measures of balance were correlated against disease duration and quality of life for the estimation of external validity. We define responsiveness as the degree to which the outcome measure accurately distinguishes important changes due to disease progression during the prospective assessment. In line with a similar study, we estimated responsiveness at 6 and 12 months after baseline.<sup>34</sup>

## Statistical analysis

Descriptive data are reported as the mean and standard deviation (SD) or the number and percentage. Spearman's correlation coefficient ( $\rho_S$ ), a bivariate analysis for non-parametric samples, was used to establish criterion and convergent validity. In this study, we interpreted correlation coefficients  $<0.5$  as low, those ranging between 0.5 and 0.79 as moderate, and those  $\leq 0.8$  as high.<sup>35,36</sup> Responsiveness was estimated using the area under the receiver-operating characteristic (ROC) curve ( $AUC_{ROC}$ ), which was plotted for each change in outcome measure score at 6 and 12 months, using the PGIC-based classification (stable vs. worsened) as the external criterion. An AUC value of  $\leq 0.70$  was considered to indicate satisfactory discrimination ability between the stable and/or improved group and the worsened group. This method is in line with that of a similar previous study on responsiveness.<sup>37</sup> The standard response mean (SRM), an indication of the effect size, was estimated for both groups to gauge responsiveness using the formula  $SRM = \text{Mean score change} / \text{SD of the change}$ . Due to the prospective nature of the study design, we anticipated the need to replace missing data. Missing data were replaced using the series mean imputation method,<sup>38</sup> and all analyses were conducted using the complete data set containing imputations for missing values. Sensitivity analysis was performed to increase the robustness of the findings. A secondary analysis was conducted for the psychometrically strong measures by excluding participants with missing data (complete case analysis).

## RESULTS

Seventy-four potential participants were approached through the HKSCAA, and 42 volunteers responded to our invitation. Among these 42 volunteers, 2 were excluded, one of whom was bound to a wheelchair/bed ( $n = 1$ ) and one who was diagnosed with cerebellar stroke ( $n = 1$ ). During the 6-month follow-up assessments, we lost four participants, three of whom withdrew and one who relocated. During the 12-month follow-up, we lost two additional participants, one of whom we were unable to reach ( $n = 1$ ) and one who did not attend the scheduled assessment ( $n = 1$ ). Among the 40 enrolled participants, 6 participants were unable to complete all six conditions of the SOT or all directional evaluations in the LOS, indicating a high floor effect (15%) for both measures. The missing values were substituted using the imputation method. All participants were ambulant, either with or without the use of an assistive device. For safety, 22 participants used a powered wheelchair for mobility

**TABLE 1** Demographic data of the included participants ( $n = 40$ )

Age (years)	Mean (SD)	50.1 (10.3)
	Range	22–62
Gender	Male	16 (45%)
	Female	20 (55%)
Ethnicity	Chinese	36 (100%)
Occupation	Employed	0 (0%)
	Unemployed	32 (88.9%)
	Retired	4 (11.1%)
Diagnosis	Spinocerebellar ataxia	
	SCA-1	10 (27.8%)
	SCA-2	2 (5.6%)
	SCA-3	8 (22.2%)
	SCA-6	4 (11.1%)
	SCA-11	2 (5.6%)
	Unknown type	2 (5.6%)
Cerebellar degeneration		
	Unknown cause for ataxia	6 (16.7%)
Age at disease onset	Mean (SD) in years	39.3 (11.7)
	Range	18–57
Disease duration	Mean (SD) in years	10.9 (6.9)
	Range	1–26
Screening for cerebellar ataxia	Gait ataxia	34 (94%)
	Limb ataxia	36 (100%)
	Dysarthria	36 (100%)
	Nystagmus	30 (83%)
Use of assistive walking device	Yes	31 (86.1%)
	No	5 (13.9%)
Disease severity (SARA)		21.19 (7.22)

Abbreviations: SD, standard deviation; SARA, Scale for the Assessment and Rating of Ataxia; SCA, spinocerebellar ataxia.

within the community. The demographic data for the included participants are presented in Table 1. The mean and SD for all outcome measures, assessed at baseline and at 6 and 12 months, are reported in Appendix S2.

## Validity

We found a significant negative correlation between the two clinic-based balance outcome measures ( $\rho_S = -0.89$ ), indicating strong criterion validity for both the BBS and the SARA-bal. The correlations between the clinic-based measures and the SOT-COM and MXE-LOS in the forward and right directions were weak to moderate ( $\rho_S$  range 0.11 to  $-0.46$ ), indicating weak to

moderate criterion validity. Table 2 reports the criterion validity estimated for the four balance outcome measures. We found significant correlations between the clinic-based outcome measures (BBS and SARA-bal) and both functional independence (Barthel index:  $\rho_S$  0.74 and  $-0.75$ , respectively) and disease severity (SARA:  $\rho_S = -0.85$  and 0.92, respectively), indicating moderate to good convergent validity. The correlations between the clinic-based outcome measures and both quality of life (EQ-VAS) and disease duration were insignificant, indicating poor external validity. Among the laboratory-based outcome measures, the correlation between the SOT-COM and the SARA ( $\rho_S = 0.50$ ) was significant and moderate, and the correlation between MXE-LOS in the forward direction and the

**TABLE 2** Criterion validity estimated by correlating the balance outcome measures with each other

Spearman's rho $p$ value ( $n = 40$ )	BBS	SARA-bal	SOT-SOM	SOT-VIS	SOT-VES	SOT-COM
BBS	1	-0.89	-0.09	-0.88	-0.15	0.29
		<0.001*	0.95	0.61	0.35	0.03*
SARA-bal		1	-0.45	0.02	0.03	-0.18
			0.79	0.89	0.82	0.02*

Spearman's rho $p$ value ( $n = 40$ )	RT-LOS (F)	MXE-LOS (F)	RT-LOS (R)	MXE-LOS (R)	RT-LOS (B)	MXE-LOS (B)	RT-LOS (L)	MXE-LOS (L)
BBS	-0.19	0.11	0.04	0.50	-0.96	0.30	-0.46	0.18
	0.25	0.004*	0.81	0.002*	0.57	0.70	0.40	0.27
SARA-bal	0.10	-0.17	0.16	-0.46	0.11	-0.31	0.34	-0.94
	0.54	0.03*	0.34	0.004*	0.51	0.61	0.47	0.25

Abbreviations: BBS, Berg Balance Scale; SARA-bal, balance sub-component of the Scale for the Assessment and Rating of Ataxia; SOT-SOM, Somatosensory score of the Sensory Organization Test; SOT-VIS, visual score of the SOT; SOT-VES, vestibular score of the SOT; SOT-COM, composite score of the SOT; R, right; F, front; L, left; B, back; RT, Reaction time; MXE, maximal excursion; LOS, Limits of Stability.

\*Significant difference.

**TABLE 3** Convergent and external validity of selective measures of balance with significant findings

Spearman's rho $p$ value	Barthel Index	EQ-VAS	SARA	Disease duration
BBS	0.74*	0.31	-0.85*	-0.14
	<0.001	0.16	<0.001	0.38
SARA-bal	-0.75*	0.11	0.92*	-0.80
	<0.001	0.63	<0.001	0.64
SOT-COM	0.10	0.59	0.50*	0.17
	0.60	0.08	0.02	0.29
MXE-LOS (F)	0.32	-0.14	0.59*	0.12
	0.05	0.54	0.007	0.46
MXE-LOS (R)	0.24	0.17	0.23	0.10
	0.31	0.46	0.34	0.53
MXE-LOS (B)	0.19	-0.07	-0.17	0.39*
	0.41	0.77	0.46	0.01
MXE-LOS (L)	0.01	0.18	-0.21	0.35*
	0.94	0.94	0.37	0.03

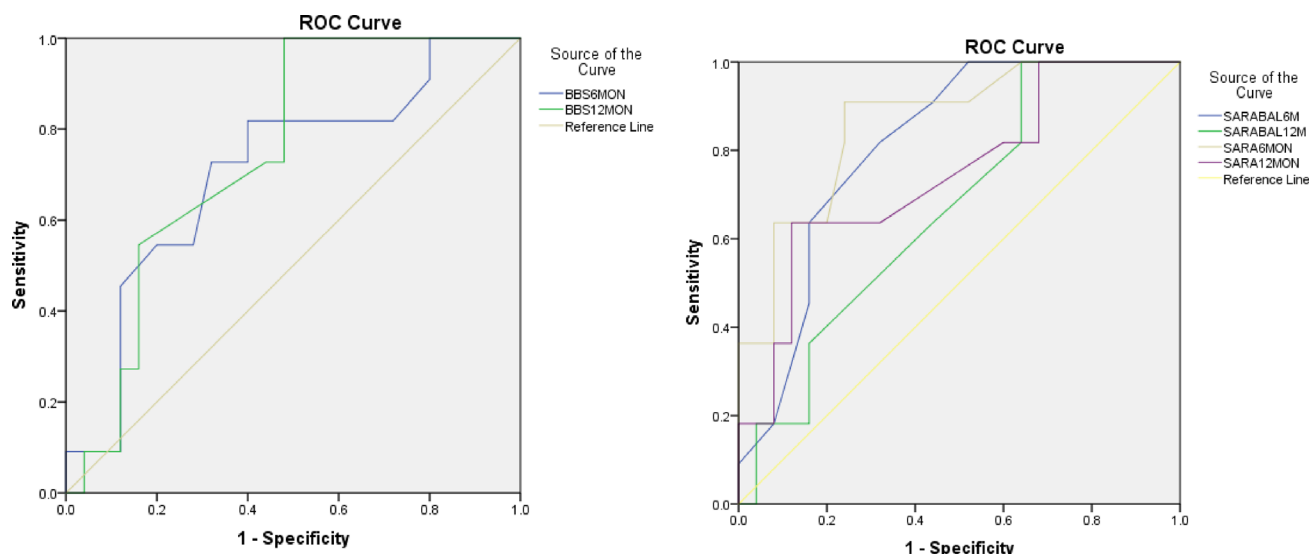
Abbreviations: BBS, Berg Balance Scale; SARA-bal, balance sub-component of the Scale for the Assessment and Rating of Ataxia; SOT-COM, composite score of the SOT; R, right; F, front; L, left; B, back; MXE, maximal excursion; LOS, Limits of Stability; EQ-VAS, visual analogue scale of the Euro Qol questionnaire; SARA, Scale for the Assessment and Rating of Ataxia.

\*Significant difference.

**TABLE 4** Mean baseline change score, standardized response mean (SRM) for the “stable” and “worsened” group and area under the receiver-operating characteristic curve ( $AUC_{ROC}$ ) for selective measures of balance and SARA with significant findings at 6 and 12 months

	6 Months				12 Months			
	Baseline change score (SD) ( $n = 36$ )	SRM-Stable ( $n = 24$ )	SRM-Worsen ( $n = 12$ )	$AUC_{ROC}$	Baseline change score (SD) ( $n = 36$ )	SRM-Stable ( $n = 24$ )	SRM-Worsen ( $n = 12$ )	$AUC_{ROC}$
BBS	6.43 (15.35)	1.06	1.16	0.70*	-1.16 (3.86)	0.27	0.35	0.75*
SARA-Bal	-1.44 (5.47)	0.86	0.85	0.81*	0.85 (2.61)	0.36	0.20	0.76*
MXE-LOS (R)	11.28 (21.03)	0.41	1.39	0.29*	-6.94 (8.74)	0.91	0.91	0.54*
SARA	0.76 (8.66)	0.73	1.10	0.86*	-1.60 (4.39)	0.37	1.50	0.74*

Abbreviations: AUCROCA, AUCROC- area under the receiver-operating characteristic curve; BBS, Berg Balance Scale; SARA-bal, balance sub-component of the Scale for the Assessment and Rating of Ataxia; SARA, Scale for the Assessment and Rating of Ataxia; R, right; MXE, maximal excursion; LOS, Limits of Stability; SARA, SRM-standardized response mean.



**FIGURE 1** (A) ROC curve for BBS at 6 and 12 months. (B) ROC curve for SARA and SARA-bal at 6 and 12 months. ROC, Receiver-operating characteristic; BBS6MON, Berg Balance Scale at 6 months; BBS12MON, Berg Balance Scale at 12 months; SARABAL6MON, balance sub-component of the Scale for the Assessment and Rating of Ataxia at 6 months; SARABAL12MON, balance sub-component of the Scale for the Assessment and Rating of Ataxia at 12 months; SARA6MON, Scale for the Assessment and Rating of Ataxia at 6 months; SARA12MON, Scale for the Assessment and Rating of Ataxia at 12 months.

SARA was ( $\rho_S = 0.59$ ) moderate. The MXE-LOS in the back and left directions were correlated weakly with disease duration ( $\rho_S < 0.40$ ), indicating moderate convergent and weak external validity. The correlations examined for all other variables measured for the SOT and LOS were insignificant. Table 3 reports the convergent and external validity of selected measures with significant findings, and the convergent and external validity for all tested outcome measures can be found in Appendix S3.

## Responsiveness

Among the 36 participants who attended the 6-month assessments, 24 participants were categorized as stable, and the remaining were categorized as worsened,

based on their PGIC responses. Two participants from the stable group were lost to follow-up at 12 months, and the data were replaced using the imputation method. The clinic-based measures for balance (BBS and SARA-bal), and the laboratory-based MXE-LOS in the right direction, demonstrated significant responsiveness for the correct classification of participants as either stable or worsened at both the 6- and 12-month follow-ups. Cerebellar ataxia-specific measures (SARA and SARA-bal) showed higher accuracy at both 6 ( $AUC_{ROC} > 0.8$ ) and 12 months compared ( $AUC_{ROC} > 0.7$ ) with both the generic balance scale (BBS,  $AUC_{ROC} > 0.7-0.8$ ) and the other scales. Table 4 reports the SRM and responsiveness for selected balance measures and SARA with significant results at 6 and 12 months. Appendix S4 reports the SRM and responsiveness for all assessed outcome measures. The ROC curves for the

BBS, SARA, and SARA-bal at 6 and 12 months are illustrated in Figure 1(A),(B). The ROC curves for the SOT and LOS variables are illustrated in Appendices S5a and 5b.

## Sensitivity analysis

No differences were found in the complete case analysis performed for those measures of balance with significant findings in the main analysis, with results comparable to those obtained in the main analysis. Appendices S6 and S7 report the results of criterion validity and responsiveness for the complete case analyses of selected measures at 6 and 12 months.

## DISCUSSION

We aimed to estimate the validity and responsiveness of two clinic-based measures (BBS and SARA-bal) and two laboratory-based measures (SOT and LOS) for assessing balance among individuals with genetic and sporadic cerebellar ataxia. Our findings demonstrated strong criterion validity, moderate to strong convergent validity, weak external validity, and adequate responsiveness for the clinic-based outcome measures, in line with our previous findings showing good reliability and validity for the BBS and SARA-bal in cerebellar ataxia secondary to multiple sclerosis.<sup>20</sup> Among the laboratory-based measures, the SOT-COM and the MXE-LOS in the forward and right directions demonstrated moderate to weak criterion, convergent, and external validity. The prospective change scores measured at 6 and 12 months were represented by the SRM, which showed large effects for both the BBS and SARA-bal at 6 months and moderate effects at 12 months. The MXE-LOS in the right direction demonstrated a weak to moderate ability to discriminate participants between the stable and worsened groups at 6 and 12 months.

Cerebellar ataxia is progressive, depending on the underlying cause. Therefore, identifying outcome measures that are responsive to changes secondary to natural disease progression is crucial. Although previous studies have reported the responsiveness of the SARA<sup>21,34</sup> and the SARA-bal to disease progression,<sup>39</sup> this study is the first to report the responsiveness of the BBS, SOT, and LOS in individuals with cerebellar ataxia. Fitzpatrick et al.<sup>40</sup> provided recommendations for choosing suitable outcome measures, which included the combination of both measure (such as the BBS) and disease-specific (such as the SARA-bal) balance measures. We observed that both clinic-based balance measures were well tolerated by all patients, resulting in minimal missing values, further supporting the acceptability of these scales. By contrast, the laboratory-based measures demonstrated high floor effects. Finally, both of the clinic-

based scales are available free of cost and require no more than 20 minutes to complete, making the outcome measures feasible for all clinical practices. Therefore, the BBS and SARA-bal appear to fulfill most of the recommended criteria, including acceptable psychometric properties. Based on the findings of this study, we recommend the use of the BBS and SARA-bal for the assessment of balance in individuals with cerebellar ataxia, both in clinical practice and research settings.

Except for the SOT-COM and MXE-LOS in the forward and right directions, the variables examined on the laboratory-based tests did not meet the criteria for acceptable validity or responsiveness. The order of assessment may have contributed to the significant findings detected for the MXE-LOS in the forward and right directions. The LOS variables are programmed and assessed in the same order for all participants, starting with the forward direction, followed by the right, back, and left directions. We speculate that the patients' performance was acceptable during the early assessments (forward and right) and may have reduced due to fatigue during the later assessments (back and left). Second, the patients' righting reactions may also have influenced better performance in the forward and right directions. However, restricting the LOS to the forward and right directions is questionable. Future studies may consider randomizing the order directions during LOS assessments to better assess the psychometric strength of the LOS in the back and left directions.

The poor psychometric scores observed for the SOT and LOS are likely due to the patient population tested in this study. Our participants had a mean ataxia severity score of 21, as measured by the SARA, indicating moderate-to-maximal dependency for activities of daily living.<sup>41</sup> We did not restrict the participants in this study based on their balance abilities. Among the 40 enrolled participants, only 34 were able to complete all of the items measured during the SOT and LOS assessments. Both the SOT and LOS require the user to stand unsupported or to counteract unpredictable challenges, such as changes in the visual surroundings during the SOT. Among those participants who were assessed by the SOT and LOS, 15% of values were missing due to the inability of the participants to complete all of the test tasks, indicating significant floor effects for both outcome measures among this population. The participants who were unable to complete all items were either ambulant with an assistive walking device or were unable to stand unsupported for more than 20 seconds, suggesting that these two measures may be more suitable for patients with reasonable balance, such as those who do not use an assistive device for walking. Future studies may consider retesting the psychometric properties of these two laboratory-based scales among cerebellar ataxia patients by restricting the population to those with the ability to stand

unsupported for at least 20 seconds, thereby ensuring that all participants are capable of completing the tasks required for the SOT and LOS with minimal missing values. We speculate that the psychometric properties of these measures are likely to improve in both performance and accuracy in such a population.

In a previous systematic review, the authors reported the need to identify a classification-based assessment tool for balance.<sup>42</sup> We used the International Classification of Functioning, Disability, and Health model to categorize balance outcome measures into those that assess the body, structure and function, activity, or participatory levels.<sup>43</sup> Based on the findings of this study, the BBS and SARA-bal have adequate psychometric properties to assess balance at the activity level. The SOT and LOS, which both assess balance at the body, structure and function, and activity levels, demonstrated limitations in both validity and responsiveness in the population tested. Future studies testing or developing balance outcome measures for the assessment of balance at the body and structure and function levels for people with cerebellar ataxia remain warranted. Few outcome measures that assess balance at the participatory level are currently available,<sup>16</sup> and no evidence has been reported regarding the psychometric properties for participatory level measures of balance in cerebellar ataxia. Future research and tool development for disease-specific balance measures for participatory level assessments are recommended.

## Study limitations

The findings of this study should be interpreted with caution, and the following limitations must be carefully considered. (1) First, the relatively small sample size ( $n = 40$ ) may have influenced the power of the study. Hong Kong is a small city with a low prevalence of cerebellar ataxia, and the HKSCAA is the only special interest group that exists to support people with ataxia in Hong Kong. However, the association currently includes only 250 registered patients, and our recruited sample represents one-sixth of the total population, which may be considered a representative sample. (2) Second, the order of the assessments was not randomized, and the SOT and LOS were assessed toward the end of each session, which may have resulted in fatigue-induced underperformance for these tasks. In addition, while assessing the LOS, the sequence of directions was standardized, which may have influenced the performance of the participants in the directions tested toward the end of the assessment. Therefore, future studies should consider randomizing the order of assessments to reduce response bias. (3) Finally, the generalizability of the study findings is restricted to patients with either genetic or sporadic underlying causes for cerebellar ataxia and to those who are ambulant with or without a walking assistive

device. Balance assessment among non-ambulant patients is more challenging, and future studies need to consider including such populations.

## CONCLUSION

This study demonstrates moderate to strong criterion and convergent validity and adequate responsiveness for the BBS and SARA-bal for assessing balance among people with cerebellar ataxia. The SOT and LOS have high floor effects among this population. The SOT-COM and the MXE-LOS in the forward and right directions showed moderate to weak criterion, convergent, and external validity and responsiveness. Future studies remain necessary to assess the psychometric properties of SOT and LOS among populations restricted to ambulant people with the ability to stand unsupported for at least 20 seconds. Based on the findings of this study, we recommend the use of the BBS and SARA-bal for clinical practice and research. The SOT and LOS may not be suitable for balance assessments in all people with cerebellar ataxia, and we recommend restricting the use of the SOT and LOS in ambulant patients who either do not use an assistive device for walking or can stand unsupported for 20 seconds or more.


## ACKNOWLEDGMENTS

The authors would like to acknowledge the participants, the Hong Kong Spinocerebellar ataxia Association (HKSCAA) and research assistants Mr. Will and Ms. Joe Pun who assisted us with data collection and data entry.

## DISCLOSURES

The authors have no conflicts of interest to declare.

## ORCID

Stanley J. Winsler  <https://orcid.org/0000-0002-8766-3688>

## REFERENCES

1. Manto M, Marmolino D. Cerebellar ataxias. *Curr Opin Neurol*. 2009;22(4):419-429.
2. Infante J, Combarros O, Volpini V, Corral J, Llorca J, Berciano J. Autosomal dominant cerebellar ataxias in Spain: molecular and clinical correlations, prevalence estimation and survival analysis. *Acta Neurol Scand*. 2005;111(6):391-399.
3. Joo B-E, Lee C-N, Park K-W. Prevalence rate and functional status of cerebellar ataxia in Korea. *Cerebellum*. 2012;11(3):733-738.
4. Klockgether T. The clinical diagnosis of autosomal dominant spinocerebellar ataxias. *Cerebellum*. 2008;7(2):101-105.
5. Pérez-Flores J, Hernández-Torres A, Montón F, Nieto A. Health-related quality of life and depressive symptoms in Friedreich ataxia. *Qual Life Res*. 2020;29(2):413-420.
6. Miyai I. Challenge of neurorehabilitation for cerebellar degenerative diseases. *Cerebellum*. 2012;11(2):436-437.

7. López-Bastida J, Perestelo-Pérez L, Montón-Alvarez F, Serrano-Aguilar P. Social economic costs and health-related quality of life in patients with degenerative cerebellar ataxia in Spain. *Mov Disord*. 2008;23(2):212-217.
8. Polek B, Roach MJ, Andrews WT, Ehling M, Salek S. Burden of Friedreich's ataxia to the patients and healthcare Systems in the United States and Canada. *Front Pharmacol*. 2013;4:66.
9. Kelly MJ, Bogdanova-Mihaylova P, Skeens J, et al. The cost of living with inherited ataxia in Ireland. *Cerebellum*. 2021;21(2):280-296.
10. Winsler SJ, Kelly CKL, Tung CC, et al. Cost of cerebellar ataxia in Hong Kong: a retrospective cost-of-illness analysis. *Front Neurol*. 2020;11:711.
11. Mariotti C, Fancellu R, Di Donato S. An overview of the patient with ataxia. *J Neurol*. 2005;252(5):511-518.
12. Brazis PW, Masdeu JC, Biller J. *Localization in clinical neurology*. Cerebellum. Philadelphia: Lippincott Williams and Wilkins; 2007:371-386.
13. Morton SM, Bastian AJ. Cerebellar control of balance and locomotion. *Neuroscientist*. 2004, 2004;10(3):247-259.
14. Marsden J, Harris C. Cerebellar ataxia: pathophysiology and rehabilitation. *Clin Rehabil*. 2011;25(3):195-216.
15. van de Warrenburg BP, Steijns JA, Munneke M, Kremer BP, Bloem BR. Falls in degenerative cerebellar ataxias. *Mov Disord*. 2005;20(4):497-500.
16. Winsler SJ, Hale L, Claydon LS, Smith C. Outcome measures for the assessment of balance and posture control in cerebellar ataxia. *Phys Ther Rev*. 2013;18(2):117-133.
17. Winsler SJ, Smith CM, Hale LA, Claydon LS, Whitney SL, Mehta P. Systematic review of the psychometric properties of balance measures for cerebellar ataxia. *Clin Rehabil*. 2015;29(1):69-79.
18. Winsler SJ, Smith C, Hale LA, Claydon LS, Whitney SL. Balance outcome measures in cerebellar ataxia: a Delphi survey. *Disabil Rehabil*. 2015;37(2):165-170.
19. Berg K, Wood-Dauphine S, Williams J, Gayton D. Measuring balance in the elderly: preliminary development of an instrument. *Physiother Can*. 1989;41(6):304-311.
20. Winsler S, Smith CM, Hale LA, et al. Psychometric properties of a Core set of measures of balance for people with cerebellar ataxia secondary to multiple sclerosis. *Arch Phys Med Rehabil*. 2017 Feb;98(2):270-276.
21. Schmitz-Hübsch T, Du Montcel ST, Baliko L, et al. Scale for the Assessment and Rating of Ataxia: development of a new clinical scale. *Neurology*. 2006;66(11):1717-1720.
22. Peterka RJ, Black FO. Age-related changes in human posture control: sensory organization tests. 1989.
23. Ford-Smith CD, Wyman JF, Elswick RK Jr, Fernandez T, Newton RA. Test-retest reliability of the sensory organization test in noninstitutionalized older adults. *Arch Phys Med Rehabil*. 1995;76(1):77-81.
24. Cohen HS, Kimball KT. Usefulness of some current balance tests for identifying individuals with disequilibrium due to vestibular impairments. *J Vestib Res Equilib Orientat*. 2008;18(5-6):295-303.
25. Juras G, Stomka K, Fredyk A, Sobota G, Bacik B. Evaluation of the Limits of Stability (LOS) balance test. *J Hum Kinet*. 2008;19(1):39-52.
26. Mahoney FI, Barthel DW. Functional evaluation: the Barthel index: a simple index of independence useful in scoring improvement in the rehabilitation of the chronically ill. *Md State Med J*. 1965;14:61-65.
27. Hsueh I-P, Wang C-H, Sheu C-F, Hsieh C-L. Comparison of psychometric properties of three mobility measures for patients with stroke. *Stroke*. 2003;34(7):1741-1745.
28. The EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy*. 1990;16(3):199-208.
29. Janssen M, Pickard AS, Golicki D, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Qual Life Res*. 2013;22(7):1717-1727.
30. Hurst H, Bolton J. Assessing the clinical significance of change scores recorded on subjective outcome measures. *J Manipulative Physiol Ther*. 2004;27(1):26-35.
31. Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, de Vet HCW. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res*. 2012;21(4):651-657.
32. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018;27(5):1147-1157.
33. Winsler SJ, Smith CM, Hale LA, Claydon LS, Whitney SL. Clinical assessment of balance using BBS and SARAbal in cerebellar ataxia: synthesis of findings of a psychometric property analysis. *Hong Kong Physiother J*. 2018;38(1):53-61.
34. Schmitz-Hübsch T, Fimmers R, Rakowicz M, et al. Responsiveness of different rating instruments in spinocerebellar ataxia patients. *Neurology*. 2010;74(8):678-684.
35. Portney Leslie Gross WMP. Statistical measures of reliability. In: Alexander JL, ed. *Foundations of Clinical Research Application to Practice*. New Jersey; F.A Davis Company; 2009:585-616.
36. Portney Leslie Gross WMP. Correlation. In: Alexander JL, ed. *Foundations of Clinical Research Application to Practice*. New Jersey; F.S Davis Company; 2009:523-537.
37. Monte TL, Reckziegel ER, Augustin MC, et al. NESSCA validation and responsiveness of several rating scales in spinocerebellar ataxia type 2. *Cerebellum*. 2017;16(4):852-858.
38. Cokluk O, Kayri M. The effects of methods of imputation for missing values on the validity and reliability of scales. *Educ Sci Theory Pract*. 2011;11(1):303-309.
39. Lee Y, Liao Y, Wang P, Lee IH, Lin K, Soong B. Comparison of cerebellar ataxias: a three-year prospective longitudinal assessment. *Mov Disord*. 2011;26(11):2081-2087.
40. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998; 2:1-74
41. Kim B-R, Lim J-H, Lee SA, et al. Usefulness of the Scale for the Assessment and Rating of Ataxia (SARA) in ataxic stroke patients. *Ann Rehabil Med*. 2011;35(6):772-780.
42. Winsler SJ, Kannan P, Bello UM, Whitney SL. Measures of balance and falls risk prediction in people with Parkinson's disease: a systematic review of psychometric properties. *Clin Rehabil*. 2019;33(12):1949-1962.
43. WHO. *International Classification of Functioning, Disability and Health (ICF)*. Geneva, Switzerland: WHO Press - World Health Organization; 2001.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Winsler SJ, Chan AYY, Chung R, Whitney S, Kannan P. Validity of balance measures in cerebellar ataxia: A prospective study with 12-month follow-up. *PM&R*. 2023;15(6):742-750. doi:[10.1002/pmrj.12826](https://doi.org/10.1002/pmrj.12826)