



ELSEVIER

Contents lists available at ScienceDirect

Gait & Posture

journal homepage: www.elsevier.com/locate/gaitpost

Full length article

Validation of the stabilometer balance test: Bridging the gap between clinical and research based balance control assessments for stroke patients

Rafaël Brouwer^{a,b,*}, Elmar Kal^{a,c,d}, John van der Kamp^{c,e}, Han Houdijk^{a,c}

^a Research and Development, Heliomare Rehabilitation Center, Relweg 51, 1949 EC, Wijk aan Zee, the Netherlands

^b Royal Dutch Society for Physical Therapy (KNGF), PO Box 248, 3800 AE, Amersfoort, the Netherlands

^c Department of Human Movement Sciences, Faculty of Behavioural and Movement Sciences, Vrije Universiteit, De Boelelaan 1105, 1081 HV, Amsterdam, the Netherlands

^d Department of Clinical Neuropsychology, Faculty of Behavioural and Movement Sciences, Vrije Universiteit, De Boelelaan 1105, 1081 HV, Amsterdam, the Netherlands

^e Research Centre for Exercise, School and Sport, Windesheim University of Applied Sciences, PO Box 10090, 8000 GB, Zwolle, the Netherlands

ARTICLE INFO

Keywords:

Stroke
Balance assessment
Stabilometer
Validity
Reliability

ABSTRACT

Background: A pronounced discrepancy exists between balance assessments for stroke survivors that are used for clinical purposes and those used for research. Clinical assessments like the Berg Balance Scale generally have stronger ecological validity, whereas research-based assessments like posturography are generally more reliable and precise. We developed a stabilometer balance test (SBT) that aims to couple measurement reliability and precision to clinical meaningfulness by means of a personalized and adaptive test procedure.

Research question: To examine the validity, reliability, and measurement error of the stabilometer balance test in inpatient stroke patients.

Methods: In this cross-sectional study, inpatient stroke patients (FAC > 2) were tested on a stabilometer with adjustable resistance to mediolateral movement. A modified staircase procedure was used to adapt task difficulty (i.e., rotational stiffness) on a trial-by-trial basis. The main outcome was the threshold stiffness at which a patient could just stay balanced. Threshold stiffness was correlated with the Berg Balance Scale and posturography measurements to determine concurrent validity (N = 86). Test-retest reliability (N = 23) was analyzed with the Intraclass Correlation Coefficient (ICC). Floor and ceiling effects were assessed. The minimal detectable change was determined at individual and group level.

Results: Threshold rotational stiffness moderately correlated with the Berg Balance Scale ($r = -0.559$, $p < 0.001$), and the absolute path length of the center of pressure during posturography ($r = 0.348$, $p = 0.006$). Test-retest reliability was good to excellent (ICC = 0.869; 95%CI = 0.696–0.944). There were no floor or ceiling effects. The minimal detectable change was sufficiently small to detect relevant changes in balance control both on individual and group level.

Relevance: The SBT is both a valid and reliable balance assessment in stroke patients. It is at least as precise as current clinically preferred measures and does not suffer from ceiling effects. Therefore, it is suitable for use in clinical practice as well as research.

1. Introduction

Reduced balance is one of the most frequent impairments people experience after a stroke. Impaired balance adversely affects sitting, standing, and walking and causes an increased risk of falling [1]. Therefore, improving balance is a main goal of stroke rehabilitation [2].

Balance is a complex phenomenon. Pollock et al. define it as “a multidimensional concept, referring to the ability of a person not to fall” [3]. Balance *control* is closely related to postural control, which is defined as “the act of maintaining, achieving or restoring a state of

balance during any posture or stability” [3]. In order to maintain balanced, one generally has to keep the center of mass (CoM) within the base of support [4]. Research shows that this can be achieved by: 1) moving the center of pressure with respect to the vertical projection of the CoM; 2) counter-rotating segments around the CoM; and 3) applying an external force. To clinicians these strategies might be familiar as the ankle, hip, and stepping (and/or seeking manual support) strategy, respectively [5].

There is no agreement on a ‘golden standard’ for evaluating balance control after stroke [6]. Different categories of assessment can be

* Corresponding author at: Research and Development, Heliomare Rehabilitation Center, Relweg 51, 1949 EC, Wijk aan Zee, the Netherlands.

E-mail address: r.brouwer@kngf.nl (R. Brouwer).

<https://doi.org/10.1016/j.gaitpost.2018.09.025>

Received 9 June 2018; Received in revised form 24 September 2018; Accepted 25 September 2018

0966-6362/ © 2018 Elsevier B.V. All rights reserved.

distinguished based on the aspect of balance control that is evaluated and/or the setting in which the assessment is performed. With respect to the latter, we can distinguish between assessments used primarily for clinical practice and for research purposes. Frequently used clinical assessments are the Berg Balance Scale (BBS), the Trunk Impairment Scale (TIS), the Functional Reach Test (FRT), and the Postural Assessment Scale for Stroke (PASS) [7–11]. Assessments used for research often involve posturography and analysis of gait kinematics [4,12]. For clinical evaluation, the costs, time investment, (clinical) feasibility, and the close relationship with functional activities of daily life are important aspects for selecting an assessment. Typically, however, these assessments suffer from disadvantages such as floor and ceiling effects, limited reliability and a lack of a precise outcome measure [13]. By contrast, assessments used in research are often more reliable, specific and precise, but also more expensive and time consuming. Moreover, the outcomes are less easily translated to activities of daily living and thus often have limited clinical and ecological validity [14].

We developed a balance control assessment which aims to combine the benefits of existing clinical and research assessments. This assessment uses a stabilometer with adjustable resistance (stiffness) to mediolateral movement, allowing for different and incremental levels of task difficulty (Fig. 1). Stabilometers are low-cost, easy to use, and can provide precise and quantifiable outcome measures (i.e., stiffness and sway). Furthermore, stabilometers have been used extensively both in research and in physiotherapy practice [15]. This stabilometer balance test (SBT) uses a modified staircase procedure, an approach which we adopted from psychophysics [16]. This procedure involves a series of trials in which, unlike most known assessments, the difficulty level of a trial depends on the patient's performance in the previous trial. This allows a quantification of the patient's capacity for balance control by determining the maximum difficulty level an individual can handle [17]. The proposed SBT procedure also allows the patient to exploit all three mechanisms for balance control [5].

In the current study, we examine the concurrent validity, test-retest reliability and minimal detectable change (MDC) of the SBT in ambulatory (FAC > 2) stroke patients.



Fig. 1. Experimental set-up of the Stabilometer Balance Test.

2. Methods

This cross-sectional validation study was part of a larger project into motor learning after stroke (Dutch CCMO-register ID NL54560.029.15). The study protocol was approved by the medical ethics committee of the VU University Medical Center Amsterdam (the Netherlands, protocol ID: 2015/354).

2.1. Participants

Power analysis using nQuery software (Statistical Solutions Ltd, Ireland) showed that a sample size of $N = 52$ was sufficient for construct validity analysis ($r = 0.8$, $\alpha = 0.05$, 95%CI: ± 0.10) and a sample of $N = 24$ to be sufficient for test-retest reliability analysis (ICC = 0.8, $\alpha = 0.05$, 95%CI: ± 0.15).

All eligible adult inpatient stroke patients recovering in a rehabilitation unit between March 2016 and June 2017 were informed about the study in writing and verbally, and invited to participate. Inclusion criteria were: 1) First-ever or recurrent stroke < 6 months ago; 2) Functional Ambulation Categories (FAC) > 2; 3) able to stand independently > 1 min; 4) able to understand instructions and cooperate with neuropsychological assessment; and 5) no other central nervous system, orthopedic, and uncorrected visual/hearing impairments. Patients who agreed to participate and signed informed consent were scheduled for measurements.

The data described in this study come from two groups of patients. First, we used baseline data from stroke patients who participated in a RCT of Kal et al. (recruited March 2016–February 2017) [18]. These data were used for the validation analyses only. The second group of patients (February 2017–June 2017) performed measurements for the purpose of both validity and reliability analyses. Inclusion of the second group followed directly after the inclusion of the first group, with the same inclusion criteria and measurement protocol. For this latter group, the only difference in measurement protocol was that an extra SBT assessment (T1) was planned one day after the initial assessment (T0).

2.1.1. Patient characteristics

We collected background characteristics from patients' medical files. Specifically, we collected general/demographic information, stroke characteristics, comorbidities, mobility, cognition (attention), executive function, and working memory. Also, we administered the Conscious Motor Processing subscale of the Movement-Specific Reinvestment Scale (MSRS-CMP) [19].

2.2. Baseline (T0) and retest (T1) measurement

At baseline, patients performed the SBT, the Berg Balance Scale and posturography measurements. Patients who were included for the test-retest reliability sample were scheduled for a retest (T1) during which only the SBT was repeated with procedures identical to T0 (Fig. 2). We aimed to perform T1 one day after T0, to minimize the influence of patients' recovery on their performance.

2.2.1. Stabilometer balance test

Patients performed the SBT on a custom-made stabilometer with a surface of 75×75 cm that could rotate to a maximum of 30° to either side along an axis in the frontal plane (Fig. 1). Therefore, the task specifically assessed mediolateral balance control, which is often impaired after stroke [20]. Patients were secured with a safety harness and were instructed to stand as still as possible with their feet 11 cm outside the board's midline. Patients were instructed to only use the handrail if they completely lost balance. The inclination angle of the stabilometer was sampled by a potentiometer at 100 Hz. The rotational stiffness of the stabilometer could be adjusted (0–220 N m/rad) by changing a set of bilateral preloaded springs (800 vs. 390 N/m springs), by using either one or two parallel springs, or by altering their moment

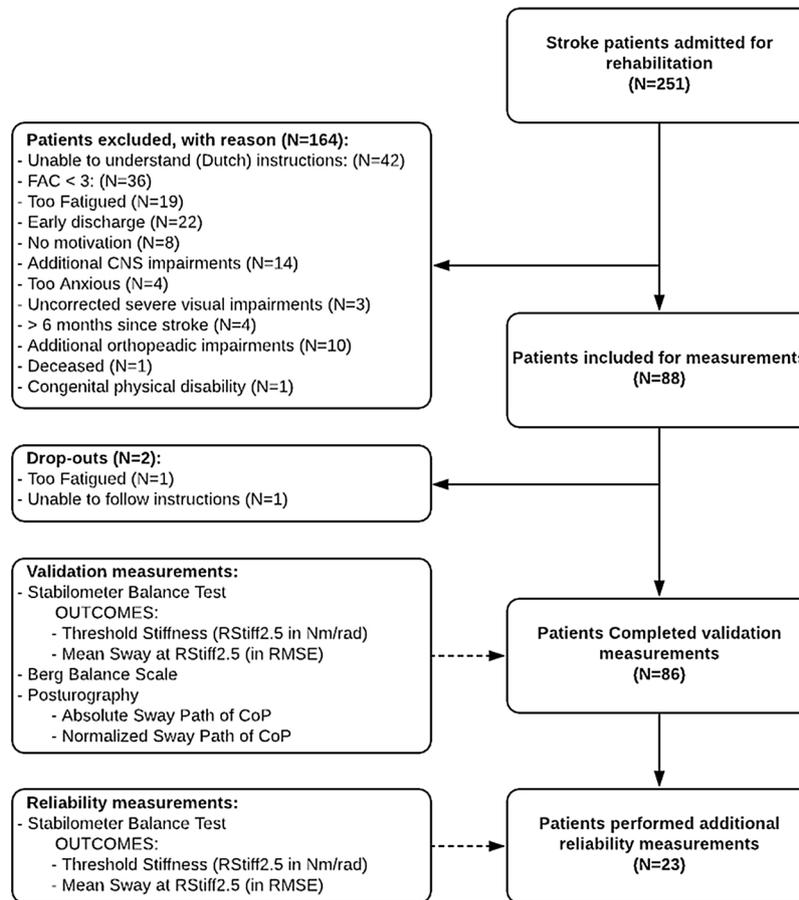


Fig. 2. Flowchart of patient inclusion.

arm.

The SBT procedure consisted of 16 trials of 30 s each. A 2-down-1-up modified staircase procedure [16] was used to determine the threshold rotational stiffness at which patients were just able to stay balanced. Based on pilot-testing, we defined this as the stiffness at which the root mean squared error (RMSE) of the board inclination was smaller than 2.5° for at least 70% of trial duration [16,21]. Patients performed one familiarization trial of 30 s (at 150 Nm/rad). The first trial was always performed at a rotational stiffness of 150 Nm/rad. The rotational stiffness on all other consecutive trials was determined by the performance on the previous trial, using the criteria outlined in Table 1. In case the patient was successful on two consecutive trials, the stiffness was reduced with 50 Nm/rad. However, if a patient failed once, we immediately increased the rotational stiffness with 40 Nm/rad. In accordance with Taylor and Creelman (1967) (16), these step sizes were halved with every reversal (down to a minimum of -3.125 Nm/rad and +2.5 Nm/rad). Also, step sizes were doubled in case of four consecutive successful or two consecutive failed trials (up to a maximum of -50 and +40 Nm/rad). This procedure was followed for a fixed number of 16 trials. Using non-linear regression, a line was then fitted through the 16 data points obtained to establish the threshold rotational stiffness at which the patient successfully managed to keep the board's deviation below 2.5° for 70% of the trial. This threshold stiffness, so-called RStiff_{2.5}, is the main outcome of the SBT. For detailed description of the RStiff_{2.5} calculation see Appendix A.

Each patient performed two additional 30-second performance trials on the stabilometer at their individually determined RStiff_{2.5}.

The total test procedure resulted in three outcomes:

- 1) RStiff_{2.5}: The minimum threshold rotational stiffness at which the patient is able to keep the stabilometer's inclination below 2.5° for

Table 1

Criteria for evaluating success during stabilometer balance test.

Performance criteria for test sessions			
% of trial duration that board deviates < 2.5 degrees	Number of times patient grabbed handrail for support	Outcome	Next trial ^a
> 70%	0	Success	1 st time: Maintain stiffness 2 nd time in a row: Reduce Stiffness
> 70%	1 or 2	Inconclusive ^b	Maintain Stiffness
< 70%	any number	Failure	Increase Stiffness
> 70%	> 2 ^c	Failure	Increase Stiffness

^a The first trial was always performed at a rotational stiffness of 150 Nm/rad. In case of success, the stiffness was reduced with 50 Nm/rad. In case of failure, the stiffness was increased with 40 Nm/rad. These step sizes were halved with every reversal (down to a minimum of -3.125 Nm/rad and +2.5 Nm/rad).

^b If a patient scored "Inconclusive" on two consecutive trials, this counted as a failure.

^c When a patient grabbed the handrail more than twice, we multiplied the number of times that patients grabbed the rail with -10%. This was subtracted from the patient's score (i.e., scoring 95% while grabbing the handrail 4 times would result in a corrected score of 55%). Handrail support was scored online by the experimenter.

- 70% of the trial;
- 2) RMSE: The average RMSE (in degrees) during the performance trials at RStiff_{2.5};
- 3) Percentage score: The average percentage of the duration of the performance trials during which RMSE was < 2.5° (corrected for

handrail support; Table 1).

2.2.2. Posturography

A force plate (P6000, BTS Bioengineering Corp., New York, USA) recorded patients' center of pressure (CoP) at a frequency of 800 Hz. Patients were instructed to stand as still as possible for 30 s, with their hands alongside their body and their feet 22 cm apart. CoP data were low-pass filtered (cut-off frequency: 6 Hz) using a 2nd order bidirectional Butterworth filter. The 'absolute path length of CoP' (CoP-Abs) and 'normalized path length of CoP' (CoP-Norm) were determined for the middle 24 s of the trial [12]. CoP-Abs quantifies the magnitude of body sway while CoP-Norm is scale independent and quantifies the amount of postural corrections. Longer sway path length and more curviness are generally regarded to indicate worse balance performance [22]. Therefore, these two measures reflect different, but complementary aspects of balance performance [12].

2.2.3. Berg balance scale

The BBS measures different aspects of balance (e.g., sit-to-stand, standing on one leg). It comprises 14 items which are scored on a five-point ordinal scale. The BBS is validated for elderly and stroke patients [11].

2.2.4. Movement-specific reinvestment scale – conscious motor processing subscale (MSRS-CMP)

The MSRS-CMP quantifies one's tendency to engage in conscious motor control. It consists out of five items which are scored on a six-point ordinal scale, ranging from 'Strongly Disagree' to 'Strongly Agree'. The test includes items like 'I try to think about my movements when I carry them out' and 'I try to figure out why my actions failed' [19]. The MSRS-CMP was collected as a descriptive measure since the level of conscious motor control possibly influences the level of balance control [23].

2.3. Statistical analysis

Statistical analysis was conducted using SPSS version 22 (IBM corporation, New York, USA).

In case of missing data we planned to impute the median of the respective variable(s), but only if the data were considered missing not at random.

Validity was assessed in three ways. First, we assessed concurrent validity. This was deemed to be satisfactory if RStiff_{2.5} was significantly positively correlated with posturography measures and negatively correlated to the BBS score. Second, divergent validity was tested. Given that the resulting RStiff_{2.5} should be equally difficult for each individual patient, their performance at this stiffness (i.e., mean RMSE) should not significantly correlate with posturography, BBS or RStiff_{2.5}. Spearman correlations were used in all of these analyses, as variables were non-parametric. Third, we used a one-sample *t*-test to assess whether patients indeed achieved 70% scores within the performance trials at their determined RStiff_{2.5} – an internal validity test. Floor and ceiling effects were deemed to be absent if less than 15% of the patients reached the minimum (0 Nm/rad) or maximum (220 Nm/rad) RStiff_{2.5} score [24]. For all validity analyses we used the results from T0 only.

Test-retest reliability of the outcomes RStiff_{2.5} and mean RMSE was determined with an Intraclass Correlation Coefficient (ICC; 2-way random, absolute agreement) [25]. These outcome measures were deemed reliable when ICC \geq 0.75 [26]. An ANOVA was used to analyze whether there was no significant difference between patients' RStiff_{2.5} at T0 and T1. The MDC was calculated for RStiff_{2.5} and RMSE at individual and group level (95%CI) in the following way: SEM = SD \times $\sqrt{1-ICC}$, MDC_{individual} = SEM \times 1.96 \times $\sqrt{2}$ and MDC_{group} = SEM \times 1.96 \times $\sqrt{2} / \sqrt{N}$ [25,27].

Table 2
Patient characteristics.

Characteristics	Validation Sample (N = 86)	Reliability Sample (N = 23)
Male/female (n)	56 / 30	13 / 10
Age (y) ^a	60.9 \pm 10.5 (30–82)	65.7 \pm 8.6 (44–82)
Weight (kg) ^a	78.8 \pm 14.5 (50–129)	73.7 \pm 13.5 (50–104)
Height (cm)	175.5 \pm 8.6 (155–193)	173.3 \pm 7.8 (160–186)
Time since stroke (days) ^a	32.7 \pm 17.5 (9–111)	38.1 \pm 18.5 (17–83)
Time since admission (days) ^a	17.3 \pm 12.6 (3–71)	23.0 \pm 16.3 (8–71)
Stroke type		
Hemorrhagic (n, %)	22 (25.6%)	7 (30.4%)
Infarction (n, %)	64 (74.4%)	16 (69.6%)
Bamford Stroke Classification		
TACS (n, %)	3 (3.5%)	0 (0.0%)
PACS (n, %)	38 (44.2%)	10 (43.5%)
POCS (n, %)	17 (19.8%)	4 (17.4%)
LACS (n, %)	28 (32.6%)	9 (39.1%)
Recurrent stroke (n, %)	12 (14.0%)	5 (21.7%)
Comorbidities (CCI) ^a	0.73 \pm 0.94 (0–5)	0.96 \pm 0.83 (0–3)
BBS Score ^a	47.6 \pm 9.2 (24–56)	47.6 \pm 10.2 (24–56)
FAC Score		
3 (n, %)	23 (26.7%)	5 (21.7%)
4 (n, %)	32 (37.2%)	7 (30.4%)
5 (n, %)	31 (36.0%)	11 (47.8%)
CoP-Abs (cm) ^{a,c}	42.82 \pm 22.39 (12.43–15.74)	42.11 \pm 21.82 (12.43–85.19)
CoP-Norm ^{a,c}	98.96 \pm 31.80 (38.70–69.38)	111.40 \pm 31.10 (60.18–69.38)
Education level [30] ^b	5 \pm 2 (2–7)	5 \pm 2 (2–7)
Sustained Attention (D2) [31] ^a	118.5 \pm 43.0 (10–213)	115.2 \pm 38.0 (41–175)
Working Memory (DSST) ^a	44.0 \pm 16.6 (1–78)	40.0 \pm 12.8 (12–63)
Executive Function (CTT) ^a	0.96 \pm 0.52 (0.07–2.21)	1.01 \pm 0.57 (0.19–2.21)
MSRS-CMP ^a [19]	20.8 \pm 6.4 (6–30)	19.1 \pm 6.9 (7–30)
Baseline Stabilometer Balance Test		
RStiff _{2.5} (Nm) ^a	38.94 \pm 29.44 (0.07–132.97)	27.63 \pm 20.87 (0.07–73.27)
RMSE (degrees) ^a	2.13 \pm 2.08 (0.00–10.54)	1.15 \pm 1.03 (0.04–3.17)
Handrail support ^b	1, 2, 125 (0.0–5.5)	1, 2, 5 (0.0–3.5)

Abbreviations: BBS = Berg Balance Scale; CCI = Charlson Comorbidities Index; CoP-Abs = Absolute Path Length of Center of Pressure; CoP-Norm = Normalized Path Length of Center of Pressure CTT = Color Trails Test; DSST = Digit-Symbol Substitution Test; FAC = Functional Ambulation Categories; LACS = Lacunar Stroke; MSRS-CMP = Movement Specific Reinvestment Scale – Conscious Motor Processing subscale; PACS = Partial Anterior Circulation Stroke; POCS = Posterior Circulation Stroke; RMSE = Root Mean Squared Error; RStiff_{2.5} = Rotational Stiffness of the balance board (stabilometer balance test outcome); TACS = Total Anterior Circulation Stroke.

^a Values presented as $\bar{X} \pm$ SD (range).

^b Values presented as median \pm IQR (range).

^c Missing values within Posturography measurements N = 24.

3. Results

3.1. Recruitment/response

A total of eighty-eight patients were included (Fig. 2). Eighty-six patients completed validity measurements, of which twenty-three performed additional test-retest reliability measurements.

3.2. Sample characteristics

Detailed sample characteristics are presented in Table 2. Within the validation sample (N = 86), mean age was 60.9 years (\pm 10.5). Mean time since stroke was 32.7 days (\pm 17.5). Mean BBS score was 47.6 (\pm 9.2). Within the reliability sample (N = 23) mean age was 65.7

Table 3
validity and reliability measurements.

Measurement	Value	Hypothesis confirmed
VALIDITY		
<i>Concurrent (r)</i>		
RStiff _{2.5} – CoP-Abs	0.348 (p = 0.006)	YES
RStiff _{2.5} – CoP-Norm	0.203 (p = 0.117)	NO
RStiff _{2.5} – BBS	–0.559 (p < 0.001)	YES
<i>Divergent (r)</i>		
RMSE – CoP-Abs	0.174 (p = 0.176)	YES
RMSE – CoP-Norm	0.081 (p = 0.531)	YES
RMSE – BBS	–0.077 (p = 0.482)	YES
RMSE – RStiff _{2.5}	0.232 (p = 0.032)	NO
<i>Internal</i>		
Percentage score $\bar{X} \pm$ (sd)	73.21% (± 28.61)	
Percentage score ≈ 70% (t-score, p-value)	1.040 (p = 0.302)	YES
Number of minimal scores (N, %)	0 (0%)	YES
Number of maximal scores (N, %)	0 (0%)	YES
RELIABILITY		
<i>Test - Retest</i>		
RStiff _{2.5} (ICC)	0.869 (95%CI = 0.696–0.944)	YES
RMSE (ICC)	0.281 (95%CI = –0.678–0.697)	NO
RStiff _{2.5} (ANOVA, F)	2.135 (p = 0.158)	YES
MEASUREMENT ERROR		
<i>Standard Error of the Mean</i>		
RStiff _{2.5} (Nm)	7.754	
RMSE (Nm)	1.336	
<i>Minimal Detectable Change</i>		
RStiff _{2.5} (individual) (Nm)	21.494	
RStiff _{2.5} (group) (Nm)	3.204	
RMSE (individual) (Nm)	3.704	
RMSE (group) (Nm)	0.552	

Abbreviations: BBS = Berg Balance Scale, CoP-Abs = Absolute Path Length of Center Of Pressure, CoP-Norm = Normalized Path Length of Center Of Pressure; ICC = Intraclass Correlation Coefficient; RStiff_{2.5} = Rotational Stiffness of the balance board (stabilometer balance test outcome); RMSE = Root Mean Square Error.

years (± 8.6). Mean time since stroke was 38.1 days (± 18.5). Mean BBS score was 47.6 (± 10.2).

3.3. Missing data

Missing data was present in twenty-four of the posturography measurements. This was caused by a temporary malfunctioning of the force plate. Missing data occurred completely at random, and therefore was not imputed.

3.4. Validity

The validity and reliability results are presented in Table 3. Mean RStiff_{2.5} of the validation sample at baseline was 38.94Nm (± 29.44). Since there was a significant correlation between RStiff_{2.5} and patients' weight (r = 0.422, P < 0.001), all concurrent validity correlations were corrected for weight. A significant correlation was found between RStiff_{2.5} and the measurements CoP-Abs (r = 0.348 p = 0.006) and BBS (r = –0.559 p < 0.001), as hypothesized. The correlation between RStiff_{2.5} and CoP-Norm was non-significant (r = 0.203 p = 0.117). Within the analysis for divergent validity all analyses were non-significant as hypothesized, except for the correlation between RMSE and RStiff_{2.5} (r = 0.279 p = 0.009). On average, patients could keep the RMSE below 2.5° for 73.21% (± 28.61) of trial duration. This value did not differ significantly from 70% (t = 1.040, df = 85, p = 0.302). No floor or ceiling effects were found.

3.5. Reliability

Nineteen participants performed T1 one day after T0, as planned. Two participants performed T1 two days later, one participant performed T1 five days later, and one participant seven days later (median_{T0-T1} = 1, IQR = 0, range = 1–7). Mean RStiff_{2.5} for the reliability sample was 27.63Nm (± 20.87) at baseline and 23.25Nm (± 22.21) at retest (F = 2.135, df = 22, p = 0.158). We found a high ICC for the RStiff_{2.5} analysis (ICC = 0.869, 95%CI = 0.696–0.944). However, the ICC of the RMSE was lower than hypothesized (ICC = 0.281, 95%CI = –0.678 to 0.697). The MDC on individual and group level was 21.49 Nm and 3.20 Nm for RStiff_{2.5}, respectively. For RMSE this was 3.70° and 0.55° at individual and group level respectively.

4. Discussion

The aim of our study was to validate a stabilometer balance test with an adaptive test procedure in order to couple measurement reliability and precision to clinical meaningfulness. Our results show that the SBT is valid, reliable, and sufficiently precise to use in ambulatory (FAC > 2) patients after stroke. These findings will be discussed below.

4.1. Validity

Different aspects of validity were comprehensively assessed in a relatively large patient sample. The moderate negative correlation between RStiff_{2.5} and the BBS and the moderate positive correlation between RStiff_{2.5} and CoP-Abs confirm that the RStiff_{2.5} quantifies a similar, but not identical construct as other balance measures. The correlation between RStiff_{2.5} and CoP-Norm did not reach significance, contrary to our hypothesis. However, CoP-Norm also correlated poorly with the BBS (r = –0.208, p = 0.104) while the correlation between BBS and CoP-Abs did yield a significant result (r = –0.314, p = 0.013). This suggests that the amount of postural corrections (i.e., CoP-Norm) is more related to a patient's balance strategy. The balance strategy that a person uses might not necessarily determine the level of balance control itself. However, it might affect the amount of postural activity needed for balance control [28].

For divergent validity, we hypothesized that a patient's RMSE score would be independent from CoP-Abs, CoP-Norm, BBS, and RStiff_{2.5}. Three of these hypotheses were confirmed. The only anomaly was a weak but significant correlation between RMSE and RStiff_{2.5} (r = 0.232, p = 0.03). Participants with a higher RStiff_{2.5} score showed a relatively high and more variable RMSE score (Fig. 3). This suggests that participants with poorer balance control have more difficulty to perform consistently at their threshold stiffness (RStiff_{2.5}).

The internal validity was confirmed by the fact that patients' mean percentage score did not significantly differ from the 70% they are supposed to score when tested on their individually determined RStiff_{2.5}. Also, floor and ceiling effects were absent. Overall, the validity of the SBT is confirmed.

A relatively large amount of data (N = 24) was missing for the posturography measurements, due to a temporary malfunctioning force plate. Despite these missing data, however, the number of successful posturography measurements still exceeded our required sample size. Missing data did not cause any added risk of bias because it occurred completely at random. Therefore missing data was not imputed.

4.2. Reliability

The test retest reliability of the main outcome, RStiff_{2.5}, was high. However, the reliability of the RMSE score was lower than expected. Therefore, the RMSE seems less reliable as a secondary outcome measure. This may be related to the fact that the RMSE reflects a patient's

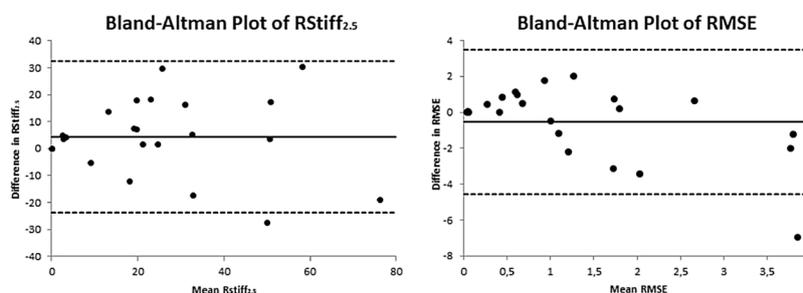


Fig. 3. Bland-Altman plot of T0 and T1 RStiff_{2.5} and RMSE scores. Abbreviations: RMSE = Root Mean Squared Error; RStiff_{2.5} = Rotational Stiffness of the balance board (stabilometer balance test outcome).

performance at the threshold difficulty level at which he or she can just stay balanced. In those conditions, variance in performance between trials may be inevitable.

4.3. Measurement error

At group level the MDC of RStiff_{2.5} and RMSE are small (i.e., 3.2 and 0.6, respectively). At individual level the MDC appears sufficient to detect relevant changes in RStiff_{2.5} (21.5), but not in RMSE (3.7). The SBT seems at least as sensitive as the BBS, which is commonly used in clinical practice. To illustrate this, we analyzed how many of the 51 patients who completed the 3-week training intervention in the study by Kal et al. [18] achieved improvements in RStiff_{2.5} and BBS that exceeded the respective MDCs of these measures. This was the case for twenty-four patients (47%) for the RStiff_{2.5} measure, compared to 17 patients for the BBS (35%). This shows that the SBT is suitable for use in clinical practice with ambulatory (FAC > 2) stroke patients. Nonetheless, the MDC at individual level was larger than anticipated. A possible explanation for this is the increased inconsistency in performance in patients with less balance control (Fig. 3). Potentially, increasing the number of iterations of the staircase procedure might improve this.

4.4. Implications for clinical practice

The SBT is valid and reliable with sufficient sensitivity at individual level and group level. In ambulatory stroke patients (FAC > 2) the SBT does not suffer from floor or ceiling effects that are commonly found in other balance control assessments [13]. An advantage of the SBT is that it is performed on stabilometers, which are frequently used in clinical balance training and are relatively low-cost when compared to balance assessments preferred for research purposes. The SBT offers a new approach in balance assessments, in which task difficulty is continuously adapted to the capacity of individual patients. This is especially relevant for stroke patients, whose balance capacity can vary substantially from patient to patient, but also can improve rapidly from one day to the next (especially early in recovery). The primary outcome (RStiff_{2.5}) is easy to interpret and provides a direct insight to patients' capacity. There is no previous research to our knowledge that used the modified staircase procedure for balance testing or other motor skills. We found only one validated balance assessment that used a

stabilometer, but this assessment did not use adjustable task difficulty nor an adaptive test protocol [29]. The SBT primarily quantifies (dynamic) mediolateral standing balance control. This differs from current clinically preferred assessments like the BBS, which aim to quantify general functional balance control. The SBT may complement these clinical assessments in several ways. For one, it gives a precise estimate for mediolateral balance control, and is thus well-suited for assessment of patients who have particular problems with this aspect of balance performance. Second, the SBT is tailored to the individual patient and does not suffer from ceiling or floor effects, unlike most clinical balance assessments. This makes it a suitable test to track changes in balance performance over time. Although the SBT correlates significantly with measures for general functional balance control (BBS), generalizability might be limited in populations with balance impairments that are not as strongly related to medio-lateral balance control as stroke patients. Therefore, cross-validation of the SBT in other patient populations would be of additional value. Future research may use a similar approach to develop and validate assessments for other aspects of balance, which can be combined into a core-set for general functional balance. The modified staircase procedure used in this study might be a valuable method for other motor skill assessments since it aims to provide a measure that is tailored to the individual capacity of the patient.

4.5. Conclusion

Our study shows that the SBT provides a valid, reliable, and precise balance measure which is applicable in measurements both at individual level and group level. Future research should focus on further improvement of the sensitivity to change on individual level and on the clinical utility of the SBT, possibly by increasing the number of test, performance and/or familiarization trials. Furthermore, research should study the feasibility to use RStiff_{2.5} as a benchmark for balance training.

Conflict of interest statement

This project was funded by the Revalidatiefonds (number: R2015042; www.revalidatiefonds.nl). The funding party had no involvement in data collection, analysis, or interpretation. It was also not involved in writing the paper or in the decision to submit it for publication.

Appendix A

Non-linear regression to determine threshold rotational stiffness

Non-linear regression was used to model the relation between the board's rotational stiffness (Rst) and balance performance P (i.e., % trial duration for which board deviations < 2.5°; Eq. (A1)). We used the solver add-in functions of Microsoft Excel (Generalized Reduced Gradient Nonlinear method; convergence = 0.000001, central differencing, multistart) to determine which coefficient values would result in optimal model fit (i.e., minimal sum of squares).¹ The resulting coefficients were used to determine the threshold Rst for which P would be 70% (Eq. (A2)). A typical

¹ D.C. Harris, Nonlinear least-squares curve fitting with Microsoft Excel Solver, J Chem Educ. 75 (1998) 119.

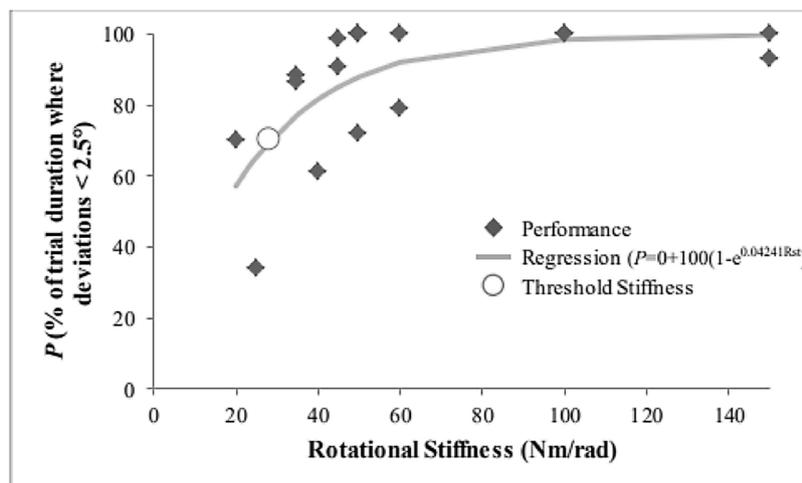


Fig. A1. Example of threshold rotational stiffness assessment.

example is presented in Fig. A1.

$$P = C + A(1 - e^{-k \times Rst}) \quad (A1)$$

$$\text{Threshold } Rst = \frac{\ln\left(\frac{70-C-A}{-A}\right)}{-k} \quad (A2)$$

NB: P = trial duration (%) for which board deviations < 2.5°; C = intercept (0% ≤ C ≤ 100%); A = horizontal asymptote (0% ≤ A ≤ 100%); Rst = rotational stiffness (Nm/rad); k = non-linear constant that describes how P changes for each 1Nm/rad change in Rst; NB: For Eq. (A2), P is set at 70%.

In this example, one patient's performance (P; % of trial duration where board deviations were below 2.5°) is plotted against the board's rotational stiffness (Nm/rad) for each of the sixteen trials performed (dark grey diamonds). The light grey line shows the best-fitting non-linear regression line (C = 0, A = 100, k = 0.04241). The large white circle highlights the threshold rotational stiffness as estimated with Eq. (A2)

$$\frac{\ln\left(\frac{70-0-100}{-100}\right)}{-0.04241} = 28.39 \text{ Nm/Rad.}$$

References

- S.F. Tyson, M. Hanley, J. Chillala, A. Selley, R.C. Tallis, Balance disability after stroke, *Phys. Ther.* 86 (1) (2006) 30–38.
- A. Pollock, G. Baer, V. Pomeroy, P. Langhorne, Physiotherapy treatment approaches for the recovery of postural control and lower limb function following stroke, *Cochrane Database Syst. Rev.* 4 (2007) Art. No.: CD001920.
- A.S. Pollock, B.R. Durward, P.J. Rowe, J.P. Paul, What is balance? *Clin. Rehabil.* 14 (2000) 402.
- S. Bruijn, O. Meijer, P. Beek, J. van Dieën, Assessing the stability of human locomotion: a review of current measures, *J. R. Soc. Interface* 10 (2013) 20120999.
- A.L. Hof, The equations of motion for a standing human reveal three mechanisms for balance, *J. Biomech.* 40 (2007) 451–457.
- H. Mao, I. Hsueh, P. Tang, C.-F. Sheu, C.-L. Hsieh, Analysis and comparison of the psychometric properties of three balance measures for stroke patients, *Stroke* 33 (2002) 1022–1027.
- C.J. Winstein, J. Stein, R. Arena, B. Bates, L.R. Cherney, S.C. Cramer, F. Deruyter, J.J. Eng, B. Fisher, R.L. Harvey, C.E. Lang, M. MacKay-Lyons, K.J. Ottenbacher, S. Pugh, M.J. Reeves, L.G. Richards, W.Z.R. Stiers, Guidelines for adult stroke rehabilitation and recovery: a guideline for healthcare professionals from the American Heart Association/American Stroke Association, *Stroke* 47 (2016).
- C. Benaim, D.A. Perennou, J. Villy, M. Rousseaux, J.Y. Pelissier, Validation of a standardized assessment of postural the postural assessment scale for stroke patients (PASS), *Stroke* 30 (1999) 1862–1868.
- P.W. Duncan, D.K. Weiner, J. Chandler, S. Studenski, Functional reach: a new clinical measure of balance, *J. Gerontol. Med. Sci.* 45 (6) (1990) 192–197.
- G. Verheyden, A. Nieuwboer, J. Mertin, R. Preger, C. Kiekens, W. De Weerd, The trunk impairment scale: a new tool to measure motor impairment of the trunk after stroke, *Clin. Rehabil.* 18 (2004) 326–334.
- K.O. Berg, S.L. Wood-Dauphinee, J.I. Williams, B. Maki, Measuring balance in the elderly: validation of an instrument, *Can. J. Public Heal.* 83 (1992) S7–S11.
- S.F. Donker, M. Roerdink, A.J. Greven, P.J. Beek, Regularity of center-of-pressure trajectories depends on the amount of attention invested in postural control, *Exp. Brain Res.* 181 (2007) 1–11.
- SRALAB, Berg Balance Scale, Clinimetric Properties [Internet], Available from: (2017) <https://www.sralab.org/rehabilitation-measures/berg-balance-scale/#stroke>.
- J.E. Visser, M.G. Carpenter, Kooijh. Van Der, B.R. Bloem, The clinical utility of posturography, *Clin. Neurophysiol.* 119 (2008) 2424–2436.
- A.J. Orrell, F.F. Eves, R.S. Masters, Motor learning of a dynamic balancing task after stroke: implicit implications for stroke rehabilitation, *Phys. Ther.* 86 (2006) 369–380.
- M.M. Taylor, C.D. Creelman, PEST: efficient estimates on probability functions, *J. Acoust. Soc. Am.* 41 (1967) 782–787.
- World Health Organization, How to Use the ICF: a Practical Manual for Using the International Classification of Functioning, Disability and Health (ICF). Exposure Draft for Comment, October WHO, Geneva, 2013.
- E. Kal, J. Houdijk, Kamp J. van der, M. Verhoef, R. Prosée, E. Groet, M. Winters, BennekomC. van, E. Scherder, Are the effects of internal focus instructions different from external focus instructions given during balance training in stroke patients? A double-blind randomized controlled trial, *Clin. Rehabil.* (August) (2018) 1–15.
- E. Kal, H. Houdijk, Wurffp. Van Der, E. Groet, BennekomC. Van, E. Scherder, Kamp J. van der, The inclination for conscious motor control after stroke: validating the movement-specific reinvestment scale for use in inpatient stroke patients, *Disabil. Rehabil.* 38 (11) (2016) 1097–1106.
- P.T. Cheng, S.H. Wu, M.Y. Liaw, A.M.K.T.F. Wong, Symmetrical body-weight distribution training in stroke patients and its effect on fall prevention, *Arch. Phys. Med. Rehabil.* 82 (12) (2001) 1650–1654.
- F. Karmali, S.E. Chaudhuri, Y. Yi, D.M. Merfeld, Determining thresholds using adaptive procedures and psychometric fits: evaluating efficiency using theory, simulations, and human experiments, *Exp. Brain Res.* 234 (2016) 773–789 Springer Berlin Heidelberg.
- A. Hufschmidt, J. Dichgans, K.H. Mauritz, M. Hufschmidt, Some methods and parameters of body sway quantification and their neurological applications*, *Arch. Psychiatr. Nervenkr. Z. Gesamte Neurol. Psychiatr.* 228 (1980) 135–150.
- R.P.M. Denneman, E.C. Kal, H. Houdijk, Kamp J. Van Der, Over-focused? The relation between patients' inclination for conscious control and single- and dual-task motor performance after stroke, *Gait Posture* [Internet] 62 (2018) 206–213, <https://doi.org/10.1016/j.gaitpost.2018.03.008> Elsevier B.V Available from:.
- C.B. Terwee, S.D.M. Bot, Boer M.R. De, WindtA.W.M. Van Der, D.L. Knol, J. Dekker, L.M. Bouter, H.C.W. de Vet, Quality criteria were proposed for measurement properties of health status questionnaires, *J. Clin. Epidemiol.* 60 (2007) 34–42.
- Vet H.C.W. de, C.B. Terwee, D.L. Knol, L.M. Bouter, When to use agreement versus reliability measures, *J. Clin. Epidemiol.* 59 (2006) 1033–1039.
- L.G. Portney, M.P. Watkins, Foundations of clinical research, applications to practice, Foundations of Clinical Research, Applications to Practice, 3rd ed., (2014), pp. 595–628.

- [27] J.P. Weir, Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM, *J. Strength Cond. Res.* 19 (1) (2005) 231–240.
- [28] T.E. Prieto, J.B. Myklebust, R.G. Hoffmann, E.G. Lovett, B.M. Myklebust, Measures of postural steadiness: differences between healthy young and elderly adults, *IEEE Trans. Biomed. Eng.* 43 (1996) 9.
- [29] L. Lei, L.-S. Weng, C.-J. Zhao, D.-J. Han, Stabilometer test for the quantified evaluation of equilibrium function in stroke patients with hemiplegia, *Chin. J. Clin. Rehabil.* 9 (17) (2005) 200–202.
- [30] F. Verhage, Intelligentie en leeftijd: onderzoek bij Nederlanders van twaalf tot zeventig jaar [Intelligence and Age: study with Dutch people aged 12 to 77], Assen van Gorcum, 1964.
- [31] R. Brinckenkamp, OosterveldP. De, d2 Aandachts- En Concentratietest, Nederlandse Bewerking van de” Aufmerksamkeits Belastung Test” van R. Brinckenkamp, Hogrefe Uitgevers, Amsterdam, 2007.