Time-dependency of test-retest reliability and measurement error of center-of-pressure synchronization and symmetry during quiet standing within specific frequencies in early sub-acute stroke

Peer-reviewed author version

van Hinsberg, Amber; Saeys, Wim; Hallemans, Ann; Kwakkel, Gert; Yperzeele, Laetitia; Truijen, Steven & SCHRODER, Jonas (2025) Time-dependency of test-retest reliability and measurement error of center-of-pressure synchronization and symmetry during quiet standing within specific frequencies in early sub-acute stroke. In: Gait & posture, 119 , p. 110 -117.

DOI: 10.1016/j.gaitpost.2025.02.026
Handle: http://hdl.handle.net/1942/45783

Time-dependency of test-retest reliability and measurement error of center-of-pressure synchronization and symmetry during quiet standing within specific frequencies in early sub-acute stroke.

Authors: Amber van Hinsberg MSc[1], Wim Saeys PhD[1], Ann Hallemans PhD[1], Gert Kwakkel PhD[2,3,4], Laetitia Yperzeele PhD[5,6], Steven Truijen PhD[1], Jonas Schröder PhD[7]

1. *Research group MOVANT, Department of Rehabilitation Sciences & Physiotherapy, Faculty of Medicine and Health Sciences, University of Antwerp, Wilrijk, Belgium*
2. *Department of Rehabilitation Medicine and Amsterdam Neuroscience, Amsterdam Movement Sciences, Amsterdam University Medical Centre, Amsterdam, Netherlands*
3. *Department of Physical Therapy and Human Movement Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States*
4. *Department of Neurorehabilitation, Amsterdam Rehabilitation Research Centre, Reade, Amsterdam, Netherlands*
5. *Department of Neurology, Stroke Unit and NeuroVascular Center, Antwerp University Hospital (UZA), Edegem, Belgium*
6. *Research group Translational Neurosciences, Faculty of Medicine and Health Sciences, University of Antwerp, Wilrijk, Belgium*
7. *REVAL Rehabilitation Research Center, Faculty of Rehabilitation Sciences, Hasselt University, Diepenbeek, Belgium*

Correspondence: Jonas Schröder, PhD; REVAL Rehabilitation Research Center, Faculty of Rehabilitation Sciences, Hasselt University; Postal address: Campus Diepenbeek, Wetenschapspark 7, 3590 Diepenbeek, Belgium; E-mail: jonas.schroder@uhasselt.be

Conflict of interest: The authors declare that the research was conducted in the absence of any potential conflict of interest.

Acknowledgments: The authors would like to thank the health care professionals at the cooperating hospitals for helping with recruitment and scheduling of the measurements sessions, and the patients for participating.

Financial support: JS received a doctoral (PhD) fellowship by the Research Foundation Flanders (FWO), Belgium (application no.1S64819N) for conducting the TARGEt studies; AvH received a doctoral (PhD) fellowship from the DOCPRO, University of Antwerp, Belgium (application no.44607) for the analyses set forth here.

Abstract:

Background: People with unilateral sensorimotor impairments after a stroke exhibit poor between-limb synchronization and asymmetries in balance control by relying on their less-affected side to maintain stability during standing. Therefore, there is a growing consensus to include center-of-pressure metrics as between-limb synchronization and dynamic control asymmetry when investigating balance improvements poststroke. However, the test-retest reliability of these metrics remains under-investigated, hindering uptake of these assessments in future stroke rehabilitation and recovery studies.

Research question: First, are three immediate test-repetitions necessary for obtaining reliable synchronization and asymmetry scores in individuals with sub-acute stroke, or can fewer repetitions suffice? Second, does timing of assessments at 3-, 5-, 8-, and 12-weeks poststroke affect these estimates' test-retest reliability and measurement error?

Methods: Thirty stroke survivors with moderate-to-severe motor impairments were tested at 3-, 5-, 8-, and 12-weeks poststroke. At each timepoint, they completed three 40-second quiet standing trials on a dual force plate which measured center-of-pressure signals separately on each side. We calculated between-limb synchronization (i.e., cross-correlation) and dynamic control asymmetry (i.e., symmetry index) using the original center-of-pressure signal, and after decomposition into high and low (cut-off 0.4 Hz) frequency bands. Intraclass correlation coefficients (ICC) and Bland-Altman plots were used to assess test-retest reliability and measurement error. A cut-off was used to determine acceptable reliability (ICC>0.75).

Results: Between-limb synchronization and dynamic control asymmetry showed good-to-excellent reliability (ICCs=0.80-0.97) across three repetitions at all timepoints. Reducing to two repetitions yielded acceptable reliability (ICC=0.77-0.95) for dynamic control asymmetry and high-frequency measures, yet insufficient reliability for between-limb synchronization. At the 3-weeks timepoint, ICCs were generally lower, with more error, compared to later timepoints.

Conclusion: Novel performance metrics quantifying synchronization and asymmetry of balance control can be reliably obtained during the early sub-acute phase if three test-repetitions are administered per assessment.

Key Words: Stroke, Balance, Reliability, Force plate, Measurement error, Center of Pressure.

List of abbreviations

ADL: Activities of daily living

AP: Anteroposterior

BLS: Between-limb synchronization

CoP: Center-of-pressure

DCA: Dynamic control asymmetry

ICC: Intraclass correlation coefficient

LOA: Limits of agreement

MDC: Minimal detectable change

MI-LE: Motricity index lower extremity

MRI: Magnetic Resonance Imaging

NIHSS: National Institutes of Health Stroke Scale

SEM: Standard error of measurement

SRRR: Stroke Recovery and Rehabilitation Roundtable

WHO: World Health Organization

1. Introduction

When standing still, steady-state balance is controlled mainly by shifting the center-of-pressure (CoP) within the base of support through ankle torques to move the center-of-mass back to equilibrium ("ankle strategy") [1]. Stroke usually impairs this strategy, causing postural instability reflected by enlarged CoP displacements in anteroposterior (AP) and mediolateral directions compared to healthy age-matched controls [2,3]. These limitations affect the safe performance of more complex activities of daily living (ADL), such as transfers and walking [4]. Therefore, improving steady-state balance, is an important early step in the rehabilitation of stroke survivors.

Currently, balance recovery is mainly assessed through capacity tests, e.g., the Berg Balance Scale [5]. While these tools show some improvement during the early sub-acute phase (i.e., the first 3 months poststroke) [4,6], they exhibit ceiling effects as the test contains a reduced range of changes in less-affected patients [7]. Furthermore, the test merely rates task completion without identifying underlying compensatory strategies, thereby reflecting mainly improvements due to compensation instead of behavioral restitution [3,5]. An inferior motor control on the affected side necessitates stroke survivors to rely on their less-affected side's ankle strategy for maintaining steady-state balance, causing persistent asymmetries [2,3,8,9]. Therefore, the Stroke Recovery and Rehabilitation Roundtable (SRRR) [5] recommends CoP measurements from each limb independently via a dual force plate to quantify between-limb synchronization (BLS) and dynamic control asymmetry (DCA), based on a slightly adjusted protocol of Mansfield et al. [10]. These performance metrics elucidate whether observed capacity improvements result from the "normalization" of quality of movement through restitution, with BLS and DCA becoming increasingly similar to normative references, or, from compensatory reliance on the less-affected side [5,11]. Likewise, it enables the examination of *if* and *how* therapies interact with movement quality.

Besides the SRRR recommendations, frequency decomposition analysis of CoP signals has been proposed to advance measurements, as steady-state balance is guaranteed by two parallel-acting control mechanisms [12,13]. More specifically, low frequencies below 0.4 Hz are assumed to reflect proactive control of sway, whereas high frequencies above 0.4 Hz are thought to reflect faster, reactive responses to internal perturbations [12,13]. Accordingly, separating CoP signals into low and high frequencies when calculating BLS and DCA could explain the preferred use of each leg for pro- and reactive balance.

To promote uptake in future research, the psychometrics underlying the metrics—including test-retest reliability and measurement error—must be established. However, inconsistencies in existing CoP reliability analyses in stroke populations [14–18] limit comparability of results and coming towards consensus. For example, existing studies investigated test-retest reliability at single, non-standardized timepoints, which frequently fall outside the critical recovery window [14–18], or they mix entirely different recovery stages [15,17]. Given the dynamic and non-linear course of poststroke recovery, it would be ideal to assess reliability at several fixed timepoints. Furthermore, studies should adhere to existing technical recommendations, including using validated instrumentation with a sampling frequency of ≥100 Hz [5,19] or, at the very least, 50 Hz as proposed by the International Society for Posture and Gait Research [20]. The incorporation of three test-repetitions of minimal 25 s is also advised to obtain consistent outcomes in terms of 'traditional' CoP metrics [19], but this is not always followed [17,18] to investigate generalizability of these recommendations for measuring movement quality poststroke. This also requires moving from composite balance measures, as investigated exclusively by some [14,15], towards BLS and DCA. Therefore, standardized recommendations, such as the SRRR initiative [5], can serve as a point of orientation to promote methodological consistency. By investigating test-retest reliability of their proposed protocol and metrics throughout the early sub-acute phase, we aim to strengthen these recommendations (or propose adjustments, if needed) to, eventually, facilitate the construction of large-scale qualitative datasets.

Our first objective was to investigate the test-retest reliability and measurement error of a single, two, or three immediate test-repetitions to ascertain whether three repetitions are necessary to obtain robust BLS and DCA scores in early sub-acute stroke. Our second objective was to determine whether timing of assessment affects the test-retest reliability when assessed at 3-, 5-, 8-, and 12-weeks poststroke. In general, we hypothesized to find sufficient reliability, defined as an intraclass correlation coefficient (ICC)>0.75, for measuring BLS and DCA with a laboratory-grade dual force plate. This would align with reliability analyses conducted in later, not standardized recovery stages [16–18]. Furthermore, we hypothesized that one-to-two test-repetitions, as applied in other recovery studies [2,8], would be insufficiently reliable (Objective 1). Lastly, we hypothesized that these estimates' reliability is time-dependent, with ICCs exhibiting relatively lower reliability and larger errors at 3-weeks, compared to later timepoints, as patients are still searching for their preferred stance strategy (Objective 2) [2,8,9].

2. Methods

This was a secondary analysis of data collected during the TARGEt project. Specifically, data from a prospective cohort study (TARGEt-1) [9,21] and a pilot randomized trial (TARGEt-2) [22] were considered, given their identical serial measurements. TARGEt, including this analysis, was approved by the Antwerp University Hospital and University of Antwerp's Medical Ethics Committee (No.18/25/305; Belgian trial registration No.B300201837010). All procedures were conducted according to the Declaration of Helsinki.

2.1. Participants

The entire TARGEt sample was recruited from acute hospitals or rehabilitation centers between 5 to 14 days poststroke (i.e., baseline) following these criteria: 1) first-ever ischemic or hemorrhagic unilateral supratentorial stroke in the cortical, subcortical, or midbrain regions, confirmed with CT/MRI [23]; 2) age 18-90 years; 3) reduced unilateral leg strength (i.e., >0 points on 6a/b of National Institutes of Health Stroke Scale <72 h poststroke and Motricity Index-lower extremity score (MI-LE) <91 at inclusion) and walking difficulties (i.e., Functional Ambulation Category (FAC) ≤4 at inclusion); 4) being admitted to inpatient rehabilitation; 5) ADL premorbid independence; 6) no severe orthopedic or another neurological condition; and 7) no severe cognitive and communication impairment. For TARGEt-2, criteria were slightly more restrictive and only included those with MI-LE <75 and FAC ≤1 [22].

From the TARGEt sample, participants (N=30) who underwent CoP assessments with force plates were included, as another measurement plate was also used, which still lacks validation [24]. Of the included participants, eight, seven, and 14 were assessed on two, three, and all four timepoints, respectively. A single participant underwent only one measurement. A detailed flowchart of follow-up numbers is presented as Figure 1.

2.2. Assessment procedure

Participants stood quietly for 40 s with eyes open and feet positioned in a standardized way, similar to other testing protocols [2,21], with an 8.4 cm heel-to-heel distance and 9° toe-out angle. They were instructed to stand still and look at a visual target displayed at approximate 2 m distance, with their arms alongside their trunk, following standardizations [10]. Three repetitions were evaluated with seated breaks in between of maximum 2 min. Ground reaction forces under each limb were sampled at 1,000 Hz with two side-by-side, floor-mounted force plates (Type OR6-7 AMTI, MA, US) in the M^{2}OCEAN gait-laboratory (University of Antwerp).

2.3. Data processing

The last 30 s of each trial were processed and low-pass filtered with a 10 Hz second-order Butterworth filter. The individual-limb CoP was aligned with the anatomical ankle position by rotating the coordinate system. Therefore, the exact foot position was determined by motion capturing (Vicon Motion system Ltd, Oxford, UK) with the AP-axis defined as a line drawn between the head of the second metatarsal bone and heel.

We calculated BLS and DCA using the original CoP signal, and separately for high- and low-frequency bands. A Fast Fourier Transformation has been applied using Matlab (MathWorks Inc., Natick, US) [25] with 0.4 Hz cut-off [13], as was chosen in prior stroke studies [12,13].

- Between-limb synchronization (BLS) was calculated as a cross-correlation between the CoP displacement from each side in AP-direction, at zero time-lag [1,26]. Higher values (closer to 1) reflect 'better' synchronization.

- Dynamic control asymmetry (DCA) was calculated as a symmetry index using CoP root mean square (rms) velocities in AP-direction and represents each leg's contribution ratio. Positive DCA values reflect a relatively greater contribution by the less-affected leg, whereas negative DCA values reflect greater control by the most-affected leg [2,3,9].

DCA = 100%

2.4. Statistical analysis

Statistical analysis was conducted using SPSS (v29.0.2.0 IBM, Armonk, US).

Test-retest reliability was defined as "the proportion of the total variance in the measurements, which is due to 'true' differences *between* patients" and is presented by ICC point-estimates [27]. Therefore, an ICC for each metric and timepoint was computed with 95% confidence intervals (95%CI) using a two-way mixed-effects model measuring absolute agreement of *k* measurements ($ICC_{3,k}$) [28]. An average-measure outcome was used to reflect the protocol's use of averaging multiple test-repetitions [19]. An $ICC_{3,k}$ was calculated for all three test-repetitions ($ICC_{3,3}$), the first two ($ICC_{3,2}$), and a single ($ICC_{3,1}$) repetition to compare the necessities to achieve ICCs ≥0.75. This cut-off has been used by others for defining good reliability of measurement techniques for research [19,29]. Reliability was interpreted as: ICC 0-0.5: poor; 0.5-0.75: moderate; 0.75-0.9: good. >0.9: excellent [28].

Measurement error, defined as "the systematic and random error, independent of the true changes in a repeated measurement within a subject" is presented by the standard error of measurement (SEM) [27,30]. The SEM was calculated as follows to estimate the distribution around a "true" score of multiple measures *within* a patient [27].

Subsequently, the minimal detectable change (MDC) was measured to determine the smallest difference in test results that is not explained by random error [31].

Bland-Altman plots were created to investigate systematic bias between immediate test-repetitions. These plots include the mean difference line with its standard error and limits of agreement (LOA) [32]. A narrower LOA encompassing zero reflects better test-retest agreement, or lower error, whereas the distribution of difference scores was visually analyzed for bias. A significant difference by a constant amount was interpreted as fixed bias and a slope pattern as proportional bias [32].

2.5. Sample size justification

The sample size of N=30, with 20-25 observations per assessment (Figure 1), was based on data availability. Using Zou's lower bound CI method [33], this sample approximates the N=23 required to detect a 95%CI lower bound of 0.5 (null hypothesis) with 80% power, assuming an ICC=0.75 for BLS based on prior results [16–18], and considering three test-repetitions. While this analytical approach is supported by COSMIN [27,31], it was developed for a one-way ICC, making the estimate conservative for our two-way model [31].

3. Results

As shown in Table 1, the 30 participants had a mean±standard deviation age of 63.67±15.37 years. Twelve (40%) were female, thirteen (43.33%) had left-sided weakness, and 25 (83.33%) experienced an ischemic stroke. Clinical baseline scores are provided, including MI-LE of 47.7±25.93 and seven participants (23.3%) were independent walkers (FAC =4). Follow-up assessment scores and 'traditional' CoP metrics are included in Table 1 as well.

Overall, as shown in Table 2, we found regarding test-retest reliability of BLS and DCA, if measured with all three available test-repetitions $ICC_{3,3}$=0.80-0.97 throughout the early sub-acute phase, thereby exceeding our threshold of >0.75. The lower bound of the 95%CI did not exceed <0.50.

For our specific objectives, Figures 2AB show tendencies of change in ICC scores when considering fewer test-repetitions (Objective 1) and across different timepoints (Objective 2). Summarized, when considering only two repetitions, DCA, DCA_{high}, $DCA_{low,}$ and BLS_{high} yielded consistently $ICC_{3,2}$ >0.75and the 95%CI did not exceed <0.50, except for DCA_{high} at 3-weeks (95%CI[0.41;0.91]). Regarding BLS, at 3-weeks we found an $ICC_{3,2}$=0.65 with 95%CI[0.10;0.86]. At the 5-, 8-, and 12-weeks assessments, $ICCs_{3.2}$ were >0.75, while at 8-weeks the 95%CI[0.45;0.90] exceeded <0.50. BLS_{low} yielded generally insufficient reliability with $ICC_{3.2}$ <0.75 and/or the 95%CI lower bound <0.50. When considering the single-measurement condition, $ICC_{3,1}$ <0.75 were found.

Regarding test-retest measurement error, the SEM estimates appear to follow the same tendencies as ICCs ; all metrics across all timepoints decrease their measurement error by including more test-repetitions. Higher SEMs are found at 3-weeks compared to the last 12-week assessment (Table 2).

Bland-Altman plots (Figure 3 and supplementary material) did not yield any bias between immediate test-repetitions for DCA. BLS yielded some proportional bias due to a decrease in change, or less measurement error, with a higher synchronization score.

4. Discussion

To our knowledge, this is the first study to investigate the test-retest reliability and measurement error of BLS and DCA at various timepoints throughout the early sub-acute phase, using a standardized protocol that employs a laboratory-grade dual force plate [5,10]. The timepoints 3-, 5-, 8-, and 12-weeks poststroke were chosen in line with SRRR recommendations for monitoring recovery courses, with consideration that most standing balance recovery appears to occur within the initial eight weeks poststroke [5,9]. Overall, when all three available test-repetitions were considered, we found sufficient reliability ($ICC_{3,3}$=0.84-0.95) throughout the early sub-acute phase. These ICC values were similar for

BLS and DCA metrics, with and without frequency decomposition. Regarding our first objective, the first two test-repetitions yielded sufficient reliability for DCA, DCA_{high}, DCA_{low}, and BLS_{high}, yet it tended to decrease for BLS and BLS_{low}. Single-test conditions were generally unreliable. Regarding our second objective, which concerns the timing of assessment, we found the test-retest reliability to be lower at 3-weeks poststroke, with relatively larger measurement errors, compared to later assessments.

Regarding Objective 1, test-retest reliability of DCA remained robust even when only two test-repetitions were considered, suggesting that a minimal protocol could be acceptable if asymmetric limb contributions to steady-state balance are of primary interest. In contrast, reliability of BLS was lower, not surpassing our threshold of ICC >0.75. This agrees with observations of Aryan et al. [18] and Jagroop et al. [17], who found similarly inferior reliability for BLS, or ICCs of 0.70 (95%[0.48;0.84]) [18] and 0.83 (95%[0.57;0.93]) [17], when compared to symmetry indexes as DCA, using a two-test protocol. However, our results complement their findings, as BLS reached sufficient reliability when three test-repetitions were administered. It should be noted that despite this difference, ICC and SEM estimates were generally 'poorer' in their studies when compared with the current study. Although speculative, this may include variation in the selected population and/or protocol. Regarding the latter, they [18] permitted habitual shoes and employed a standardized axis transformation for defining AP orientation. Conversely, we measured participants barefooted, which could have caused observed differences [34]. Furthermore, we also tracked their exact foot position through motion-capturing trial-by-trial. However, systematic investigations would be needed to clarify these factors' influence.

The lower reliability of BLS relative to DCA could be attributed to the way these estimates are calculated. BLS is calculated based on point-by-point correlations, showing non-linear progression [1,26], while DCA as a symmetry index of composite velocity scores using the entire CoP signal on either side [2,3,9]. This likely reduces variability by smoothing out of noise and the impact of outliers [35]. Bland-Altman plots also revealed some proportional bias regarding BLS, with worse synchronization linked to greater variability in scores (Figure 3). This might limit monitoring changes in BLS in more-affected patients as they show greater measurement error. A subgroup analysis by impairment severity is warranted but not feasible due to a limited sample.

Following our expectations regarding Objective 2, reliability tended to be lower early poststroke. As emphasized, this could be attributed to patients still discovering the most effective strategy to adapt to regain stability while standing. In line with this assumption, DCA tended to increase over time poststroke (Table 2), which somewhat agrees with previous studies showing no change in DCA [2,9]. This suggests an enhanced reliance on the less-affected side's ankle strategy with ongoing recovery, which is similar when only considering cases were tested on all occasions in our dataset (Table S1). Once this compensatory strategy is established, patients could be more consistent in their performance, thereby reducing the test-retest variability.

An additional final observation, beyond our pre-specified objectives, was a higher reliability for high-frequency compared to low-frequency BLS outcomes. Technically speaking, capturing low frequencies requires longer measurements [12,36] and one may hypothesize that our measurements were too short. However, prolonging measurements would propose challenges in pathological populations. Furthermore, this observation could be explained by the "throw and catch" model of postural control [37], where the low-frequent signal reflects random, thereby less consistent, sway and high-frequent signals reflect reactive corrections that are more regular [36].

In terms of implications, we suggest that when assessing patients early, particularly around 3-weeks poststroke, a timepoint that likely coincides with the initiation of early rehabilitation trials, using three test-repetitions is advisable. This decreases the random error and, thereby, increases the power to detect recovery of movement quality *within* subjects. This is particularly important for BLS and frequency-decomposed metrics which showed greater measurement error. Nevertheless, the MDC values for BLS (~0.1, Table 2) are lower than approximate differences between stroke (mean ~0.6, Table 1) and healthy populations (mean ~0.8-0.9 [9,38]). Thus, the provided MDC values (Table 2) can be used as benchmarks to distinguish "true" change in the direction of normative ranges, from random error when interpreting findings in future studies.

4.1. Future perspective

Our findings facilitate longitudinal studies monitoring qualitative changes in standing balance performance poststroke using frequency-decomposed BLS and DCA, during ongoing recovery, or in response to therapies. However, replication is warranted using a clinically-feasible test protocol (i.e., portable technology) and in a broader clinically-representative sample, including recurrent strokes and the mildly affected. Then, when translating our findings for clinical practice, we recommend higher ICC cut-offs as more precision is needed for clinical decision-making in individual patients. Additionally, to enhance interpretation of DCA and BLS, large-scale open-access normative datasets are needed to provide age-matched healthy reference values and normative thresholds.

4.2. Limitations

First, our sample size was arbitrary and rather small. At 3-weeks, some participants' balance was still severely reduced, resulting in fewer data than the prerequisite N=23, as determined based on power analysis. Therefore, despite similar findings to prior research [16–18], large confidence intervals at 3-weeks and when fewer test-repetitions are considered warrant cautious interpretation. Second, our sample was limited to inpatients and first-ever major strokes. Third, measurements before 3-weeks poststroke could not be systematically administered, leaving out very early changes in recovery. Fourth, rest breaks were not standardized, potentially influencing the measurements' reliability due to varying fatigue levels. Fifth, alternative approaches to correct for the non-linear nature of BLS were not investigated. Last, laboratory-based protocols were used due to the lack of validated portable systems [5], and protocols were limited to quiet standing.

5. Conclusion

This study examined the test-retest reliability and measurement error of BLS and DCA in early sub-acute stroke, following a standardized protocol of the SRRR for measuring steady-state standing balance with dual force plates [5,10]. The findings indicate reliable outcomes when three 40-s test-repetitions were administered across timepoints, thereby confirming CoP reliability recommendations in other populations [19]. Further, DCA was found sufficiently reliable even after two test-repetitions, offering an opportunity to collect outcomes with less strenuous activity for patients. Future research should include a larger clinically-representative population and test validity of portable technology to improve the feasibility of conducting repeated measurements as part of stroke recovery and rehabilitation studies.

6. References

[1] Winter DA. Human balance and posture control during standing and walking. Gait Posture 1995;3:193–214. https://doi.org/10.1016/0966-6362(96)82849-9.

[2] De Haart M, Geurts AC, Huidekoper SC, Fasotti L, Van Limbeek J. Recovery of standing balance in postacute stroke patients: A rehabilitation cohort study. Arch Phys Med Rehabil 2004;85:886–95. https://doi.org/10.1016/j.apmr.2003.05.012.

[3] Roelofs JMB, van Heugten K, de Kam D, Weerdesteyn V, Geurts ACH. Relationships Between Affected-Leg Motor Impairment, Postural Asymmetry, and Impaired Body Sway Control After Unilateral Supratentorial Stroke. Neurorehabil Neural Repair 2018;32:953–60. https://doi.org/10.1177/1545968318804405.

[4] Kollen B, Van De Port I, Lindeman E, Twisk J, Kwakkel G. Predicting improvement in gait after stroke a longitudinal prospective study. Stroke 2005;36:2676–80. https://doi.org/10.1161/01.STR.0000190839.29234.50.

[5] Van Criekinge T, Heremans C, Burridge J, Deutsch JE, Hammerbeck U, Hollands K, et al. Standardized measurement of balance and mobility post-stroke: Consensus-based core recommendations from the third Stroke Recovery and Rehabilitation Roundtable. International Journal of Stroke 2024;19:158–68. https://doi.org/10.1177/17474930231205207/ASSET/IMAGES/LARGE/10.1177_17474930231205207-FIG5.JPEG.

[6] Buvarp D, Rafsten L, Abzhandadze T, Sunnerhagen KS. A cohort study on longitudinal changes in postural balance during the first year after stroke. BMC Neurol 2022;22. https://doi.org/10.1186/S12883-022-02851-7.

[7] Knorr S, Brouwer B, Garland SJ. Validity of the Community Balance and Mobility Scale in Community-Dwelling Persons After Stroke. Arch Phys Med Rehabil 2010;91:890–6. https://doi.org/10.1016/j.apmr.2010.02.010.

[8] Roerdink M, Geurts ACH, De Haart M, Beek PJ. On the relative contribution of the paretic leg to the control of posture after stroke. Neurorehabil Neural Repair 2009;23:267–74. https://doi.org/10.1177/1545968308323928.

[9] Schröder J, Saeys W, Embrechts E, Hallemans A, Yperzeele L, Truijen S, et al. Recovery of Quiet Standing Balance and Lower Limb Motor Impairment Early Poststroke: How Are They Related? Neurorehabil Neural Repair 2023;37:530–44. https://doi.org/10.1177/15459683231186983.

[10] Mansfield A, Inness EL. Force Plate Assessment of Quiet Standing Balance Control: Perspectives on Clinical Application within Stroke Rehabilitation. Rehabilitation Process and Outcome 2015;4:RPO.S20363. https://doi.org/10.4137/rpo.s20363.

[11] Kwakkel G, Van Wegen EEH, Burridge JH, Winstein CJ, van Dokkum LEH, Alt Murphy M, et al. Standardized measurement of quality of upper limb movement after stroke: Consensus-based core recommendations from the Second Stroke Recovery and Rehabilitation Roundtable. International Journal of Stroke 2019;14:783–91.

https://doi.org/10.1177/1747493019873519/ASSET/IMAGES/LARGE/10.1177_1747493019873519-FIG3.JPEG.

[12] Schinkel-Ivy A, Singer JC, Inness EL, Mansfield A. Do quiet standing centre of pressure measures within specific frequencies differ based on ability to recover balance in individuals with stroke? Clinical Neurophysiology 2016;127:2463–71. https://doi.org/10.1016/j.clinph.2016.02.021.

[13] Winter DA, Patla AE, Prince F, Ishac M, Gielo-perczak K. Stiffness control of balance in quiet standing. J Neurophysiol 1998;80:1211–21. https://doi.org/10.1152/JN.1998.80.3.1211/ASSET/IMAGES/LARGE/JNP.AU16F8.JPEG.

[14] Gasq D, Labrunée M, Amarantini D, Dupui P, Montoya R, Marque P. Between-day reliability of centre of pressure measures for balance assessment in hemiplegic stroke patients. J Neuroeng Rehabil 2014;11:39. https://doi.org/10.1186/1743-0003-11-39.

[15] Gray VL, Ivanova TD, Garland SJ. Reliability of center of pressure measures within and between sessions in individuals post-stroke and healthy controls. Gait Posture 2014;40:198–203. https://doi.org/10.1016/j.gaitpost.2014.03.191.

[16] Martello SK, Boumer TC, de Almeida JC, Correa KP, Devetak GF, Faucz R, et al. Reliability and minimal detectable change of between-limb synchronization, weight-bearing symmetry, and amplitude of postural sway in individuals with stroke. Research on Biomedical Engineering 2017;33:113–20. https://doi.org/10.1590/2446-4740.06816.

[17] Jagroop D, Aryan R, Schinkel-Ivy A, Mansfield A. Reliability of unconventional centre of pressure measures of quiet standing balance in people with chronic stroke. Gait Posture 2023;102:159–63. https://doi.org/10.1016/J.GAITPOST.2023.03.021.

[18] Aryan R, Inness E, Patterson KK, Mochizuki G, Mansfield A. Reliability of force plate-based measures of standing balance in the sub-acute stage of post-stroke recovery. Heliyon 2023;9. https://doi.org/10.1016/j.heliyon.2023.e21046.

[19] Ruhe A, Fejer R, Walker B. The test–retest reliability of centre of pressure measures in bipedal static task conditions – A systematic review of the literature. Gait Posture 2010;32:436–45. https://doi.org/10.1016/J.GAITPOST.2010.09.012.

[20] Scoppa F, Capra R, Gallamini M, Shiffer R. Clinical stabilometry standardization. Basic definitions - Acquisition interval - Sampling frequency. Gait Posture 2013;37:290–2. https://doi.org/10.1016/j.gaitpost.2012.07.009.

[21] Schröder J, Saeys W, Yperzeele L, Kwakkel G, Truijen S. Time Course and Mechanisms Underlying Standing Balance Recovery Early After Stroke: Design of a Prospective Cohort Study With Repeated Measurements. Front Neurol 2022;13. https://doi.org/10.3389/FNEUR.2022.781416/FULL.

[22] Schröder J, Yperzeele L, Embrechts E, Loureiro-Chaves R, Hallemans A, Lafosse C, et al. Exoskeleton-assisted training to enhance lower limb motor recovery in subacute stroke: does timing matter? A pilot randomized trial. Frontiers in Stroke 2024;3:1379083. https://doi.org/10.3389/FSTRO.2024.1379083.

[23] Sacco RL, Kasner SE, Broderick JP, Caplan LR, Connors JJ, Culebras A, et al. An updated definition of stroke for the 21st century: A statement for healthcare professionals from the American heart association/American stroke association. Stroke 2013;44:2064–89. https://doi.org/10.1161/STR.0B013E318296AECA.

[24] Schröder J, Hallemans A, Saeys W, Yperzeele L, Kwakkel G, Truijen S. Is a portable pressure plate an alternative to force plates for measuring postural stability and interlimb coordination of quiet standing balance control? J Rehabil Assist Technol Eng 2024;11:1–10. https://doi.org/10.1177/20556683241234858.

[25] Vieira TMM, Oliveira LF, Nadal J. Estimation procedures affect the center of pressure frequency analysis. Brazilian Journal of Medical and Biological Research 2009;42:665–73. https://doi.org/10.1590/s0100-879x2009000700012.

[26] Mansfield A, Mochizuki G, Inness EL, McIlroy WE. Clinical correlates of between-limb synchronization of standing balance control and falls during inpatient stroke rehabilitation. Neurorehabil Neural Repair 2012;26:627–35. https://doi.org/10.1177/1545968311429688.

[27] Mokkink LB, Eekhout I, Boers M, van der Vleuten CP, de Vet HC. Studies on Reliability and Measurement Error of Measurements in Medicine – From Design to Statistics Explained for Medical Researchers. Patient Relat Outcome Meas 2023;Volume 14:193–212. https://doi.org/10.2147/prom.s398886.

[28] Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med 2016;15:155–63. https://doi.org/10.1016/j.jcm.2016.02.012.

[29] Portney LG. Foundations of Clinical Research: Applications to Evidence-Based Practice. F.A. Davis; 2020.

[30] Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 2010;63:737–45. https://doi.org/10.1016/J.JCLINEPI.2010.02.006.

[31] Mokkink LB, de Vet H, Diemeer S, Eekhout I. Sample size recommendations for studies on reliability and measurement error: an online application based on simulation studies. Health Serv Outcomes Res Methodol 2023;23:241–65. https://doi.org/10.1007/s10742-022-00293-9.

[32] Martin Bland J, Altman DG. STATISTICAL METHODS FOR ASSESSING AGREEMENT BETWEEN TWO METHODS OF CLINICAL MEASUREMENT. The Lancet 1986;327:307–10. https://doi.org/10.1016/S0140-6736(86)90837-8.

[33] Zou GY. Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. Stat Med 2012;31:3972–81. https://doi.org/10.1002/sim.5466.

[34] Reutimann S, Hill-Strathy MJ, Krewer C, Bergmann J, Müller F, Jahn K, et al. Influence of footwear on postural sway: A systematic review and meta-analysis on barefoot and shod bipedal static posturography in patients and healthy subjects. Gait Posture 2022;92:302–14. https://doi.org/10.1016/J.GAITPOST.2021.11.022.

[35] Cohen J, Cohen P, West SG, Aiken LS. Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. 3rd edition. New York: Routledge; 2002.

[36] Zatsiorsky VM, Duarte M. Rambling and Trembling in Quiet Standing. Motor Control 2000;4:185–200. https://doi.org/10.1123/MCJ.4.2.185.

[37] Loram ID, Lakie M. Human balancing of an inverted pendulum: Position control by small, ballistic-like, throw and catch movements. Volume 540, Issue 3, Pages 1111 - 1124 2002;540:1111–24. https://doi.org/10.1113/jphysiol.2001.013077.

[38] Mansfield A, Danells CJ, Inness E, Mochizuki G, McIlroy WE. Between-limb synchronization for control of standing balance in individuals with stroke. Clinical Biomechanics 2011;26:312–7. https://doi.org/10.1016/J.CLINBIOMECH.2010.10.001.

Figure captions

Figure 1. Flowchart of follow-up data with reasons of missing data.

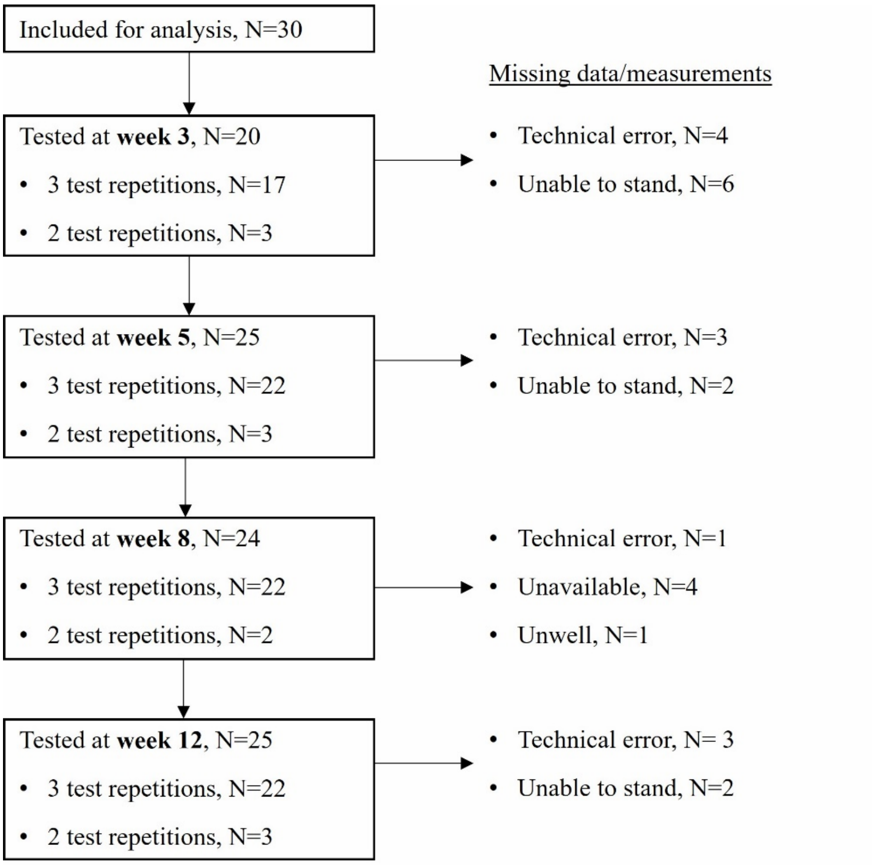


Figure 2AB. Values shown are intraclass correlation coefficient (ICC3,k) point estimates with their 95% confidence interval [95%CI] as error bars for immediate test-retest reliability of dual force plates with and without frequency decomposition (high and low) at fixed assessment timepoints in the early sub-acute phase poststroke.

The difference symbols represent ICC scores for 3 trials (ICC3,3; circle ●), 2 trials (ICC3,2; triangle ▲), and 1 trial (ICC3,1; square ■). Panel A shows ICC outcomes for BLS, and panel B shows ICC outcomes for DCA.

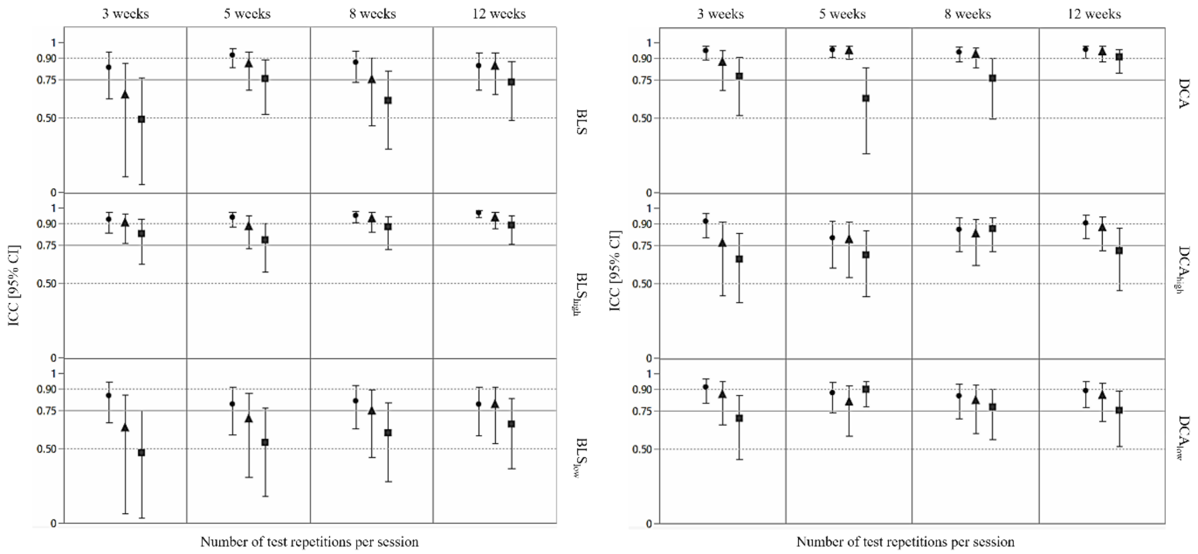
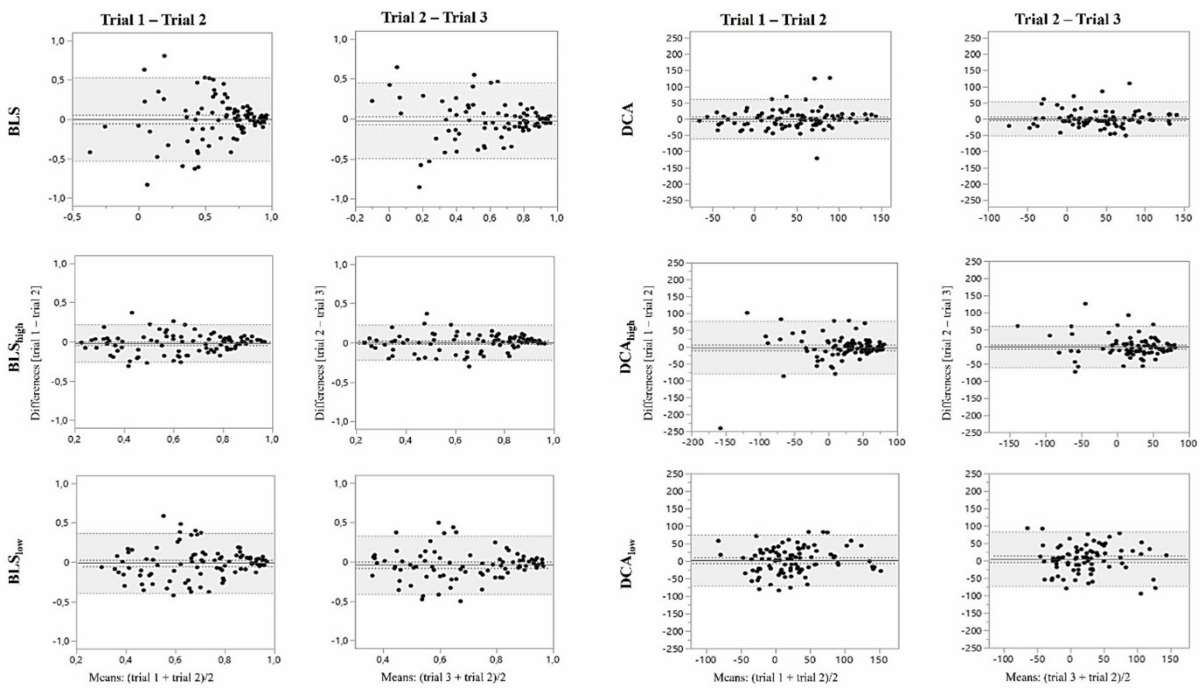


Figure 3. Bland-Altman plots with subject-specific differences and mean scores between trial 1 vs. trial 2, and trial 2 vs. trial 3 at all investigated assessment timepoints.

The Y-axis shows subject-specific difference scores between the two intended trials, and the X-axis represents the mean score per subject. A solid line represents the mean difference, a dashed line indicates the standard error of the mean, and the grey area represents the area between the upper and lower end of the limit of agreement. On the left, the between-limb synchronization (BLS) with and without frequency decomposition is displayed. On the right, the dynamic control asymmetry (DCA) with and without frequency decomposition is shown.



Tables

Table 1. Participants' characteristics at baseline (5-14 days poststroke) and results on clinical scales and force plate assessments at 3-, 5-, 8-, and 12-weeks poststroke.

Values presented with * are means with standard deviations in parenthesis, otherwise, they are numbers (N).

Assessment timepoint	**Baseline**	**3-weeks**	**5-weeks**	**8-weeks**	**12-weeks**
Sex (female/male) (N)	12/18				
Age (years)*	63.67 (15.37)				
Weight (kg)*	72.99 (16.05)				
Height (cm)*	166.28 (8.37)				
Affected side (right/left) (N)	13/22				
Stroke type (ischemic/ hemorrhagic) (N)	25/5				
measurements (N)		20	25	24	25
Additional descriptive outcomes of impairment severity and balance performance					
FM_LE [0-34]*	17.27 (8.53)	20.60 (7.47)	22.45 (7.18)	24.27 (6.56)	24.52 (5.56)
MI_LE [0-100]*	47.7 (25.93)	57.33 (24.34)	61.72 (23.58)	67.5 (20.88)	70.68 (17.17)
FAC [0-5]*	1.03 (1.7)	1.80 (1.85)	2.59 (1.78)	3.12 (1.48)	3.8 (1.35)
WBA*(%)		40.34 (12.07)	42.18 (9.85)	41.35 (9.03)	41.03 (8.17)
CoP ellipse area*(mm^2)		571.52 (383.77)	498.53 (355.95)	410.20 (333.00)	365.97 (280.49)
CoPa_ap (mm)*		35.13 (11.49)	34.29 (10.61)	30.11 (10.05)	28.67 (8.43)
CoPa_ml (mm)*		35.81 (17.46)	31.30 (16.63)	28.86 (14.10)	26.74 (12.54
CoPv_ap (mm/s)*		16.10 (5.79)	15.63 (6.48)	14.62 (6.01)	13.90 (5.67)
CoPv_ml (mm/s)*		14.00 (7.56)	12.35 (7.27)	11.61 (6.22)	11.10 (5.78)
CoP-based performance metrics of interest					
BLS (r)*		0.60 (0.24)	0.57 (0.34)	0.63 (0.257)	0.62 (0.246)
DCA (%)*		28.50 (51.05)	43.15 (52.78)	48.90 (39.73)	51.63 (43.20)
BLS_high (r)*		0.65 (0.19)	0.64 (0.22)	0.65 (0.21)	0.65 (0.22)
BLS_low (r)*		0.66 (0.20)	0.69 (0.18)	0.70 (0.19)	0.70 (0.15)
DCA_high (%)*		6.90 (48.60)	16.92 (52.52)	30.27 (33.74)	35.17 (32.38)
DCA_low (%)*		18.33 (53.80)	27.53 (48.66)	24.19 (40.65)	29.06 (39.43)

FM-LE: Fugl-Meyer motor assessment – Lower extremity; MI-LE: Motricity index – Lower extremity; FAC: Functional Ambulation Category; BLS: Between limb synchronization; DCA: Dynamic control asymmetry; WBA: Weight-bearing asymmetry (ie, weight placed on the most-affected leg); CoPa: Center of Pressure amplitude; ap: Anteroposterior; ml: mediolateral; CoPv: Center of Pressure velocity; mm: millimeter; mm/s: millimeter per second.

Table 2. Test-retest reliability and measurement error results.

3-weeks poststroke									
	ICC3,3 (95%CI)	SEM	MDC	ICC3,2 (95%CI)	SEM	MDC	ICC3,1 (95%CI)	SEM	MDC
BLS	0.84 (0.63;0.94)	0.04	0.11	0.65 (0.10;0.86)	0.05	0.13	0.49 (0.05;0.76)	0.06	0.17
DCA	0.95 (0.88;0.98)	0.07	0.20	0.88 (0.68;0.95)	0.08	0.22	0.78 (0.52;0.91)	0.12	0.34
BLS_{high}	0.93 (0.83;0.97)	0.03	0.08	0.91 (0.77;0.96)	0.03	0.09	0.83 (0.62;0.93)	0.05	0.14
BLS_{low}	0.85 (0.67;0.94)	0.04	0.08	0.64 (0.06;0.86)	0.04	0.10	0.47 (0.03;0.75)	0.05	0.14
DCA_{high}	0.91 (0.81.0.97)	0.08	0.21	0.77 (0.41;0.91)	0.10	0.26	0.63 (0.26;0.84)	0.14	0.38
DCA_{low}	0.91 (0.80;0.87)	0.07	0.19	0.87 (0.66;0.95)	0.08	0.23	0.76 (0.49;0.90)	0.11	0.30

5-weeks poststroke									
	ICC3,3 (95%CI)	SEM	MDC	ICC3,2 (95%CI)	SEM	MDC	ICC3,1 (95%CI)	SEM	MDC
BLS	0.92 (0.83;0.96)	0.04	0.12	0.86 (0.69;0.94)	0.05	0.15	0.76 (0.52;0.89)	0.09	0.24
DCA	0.95 (0.91;0.98)	0.06	0.18	0.95 (0.89;0.98)	0.08	0.22	0.91 (0.80;0.96)	0.12	0.33
BLS_{high}	0.94 (0.88;0.97)	0.03	0.07	0.88 (0.73;0.95)	0.03	0.09	0.79 (0.57;0.90)	0.05	0.13
BLS_{low}	0.80 (0.59;0.91)	0.02	0.07	0.70 (0.31;0.87)	0.02	0.07	0.54 (0.18;0.77)	0.04	0.12
DCA_{high}	0.80 (0.60;0.91)	0.06	0.17	0.80 (0.54;0.91)	0.07	0.20	0.66 (0.37;0.84)	0.11	0.31
DCA_{low}	0.87 (0.74;0.94)	0.07	0.19	0.82 (0.58;0.92)	0.09	0.26	0.69 (0.41;0.85)	0.15	0.43

8-weeks poststroke									
	ICC3,3 (95%CI)	SEM	MDC	ICC3,2 (95%CI)	SEM	MDC	ICC3,1 (95%CI)	SEM	MDC
BLS	0.87 (0.73;0.94	0.03	0.09	0.76 (0.45;0.90)	0.04	0.11	0.61 (0.29;0.81)	0.06	0.17
DCA	0.94 (0.87;0.97)	0.05	0.14	0.93 (0.83;0.97)	0.06	0.17	0.87 (0.71;0.94)	0.08	0.23
BLS_{high}	0.95 (0.90;0.98)	0.03	0.07	0.93 (0.84;0.97)	0.03	0.09	0.88 (0.72;0.95)	0.04	0.12
BLS_{low}	0.82 (0.63;0.92)	0.03	0.07	0.75 (0.44;0.89)	0.03	0.09	0.60 (0.28;0.81)	0.05	0.13
DCA_{high}	0.86 (0.71;0.94)	0.06	0.15	0.83 (0.62;0.93)	0.07	0.20	0.72 (0.45;0.87)	0.11	0.29
DCA_{low}	0.85 (0.70;0.93)	0.05	0.13	0.83 (0.60;0.92)	0.05	0.14	0.70 (0.43;0.86)	0.07	0.18

12-weeks poststroke									
	ICC3,3 (95%CI)	SEM	MDC	ICC3,2 (95%CI)	SEM	MDC	ICC3.1 (95%CI)	SEM	MDC
BLS	0.85 (0.69;0.93)	0.03	0.09	0.85 (0.65;0.93)	0.04	0.11	0.74 (0.48;0.87)	0.07	0.18
DCA	0.96 (0.90;0.98)	0.05	0.15	0.95 (0.88;0.98)	0.06	0.17	0.90 (0.78;0.95)	0.09	0.25
BLS_{high}	0.97 (0.94;0.99)	0.03	0.07	0.94 (0.86;0.97)	0.03	0.08	0.90 (0.78;0.95)	0.04	0.12
BLS_{low}	0.80 (0.58;0.91)	0.02	0.06	0.80 (0.54;0.91)	0.03	0.07	0.66 (0.37;0.84)	0.04	0.08
DCA_{high}	0.90 (0.80;0.96)	0.05	0.14	0.88 (0.72;0.95)	0.06	0.17	0.78 (0.56;0.90)	0.09	0.25
DCA_{low}	0.89 (0.77;0.95)	0.04	0.12	0.86 (0.68;0.94)	0.05	0.13	0.76 (0.52;0.89)	0.07	0.19

ICC: Intraclass correlation coefficient; CI: Confidence interval; SEM: Standard error of measurement; MDC: Minimal detectable change; BLS: between limb synchronization; DCA: dynamic control asymmetry.