Available online at www.sciencedirect.com
ScienceDirect
Gait & Posture 28 (2008) 337–342
www.elsevier.com/locate/gaitpost

Reliability of COP-based postural sway measures and age-related differences

Dingding Lin [a], Hyang Seol [a], Maury A. Nussbaum [a,*], Michael L. Madigan [b]

[a] *Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA, USA*
[b] *Department of Engineering Science and Mechanics, Virginia Tech, Blacksburg, VA, USA*

Received 15 March 2007; received in revised form 17 January 2008; accepted 19 January 2008

Abstract

The objectives of this study were to assess the within-day and between-day reliability of several center of pressure (COP)-based measures of postural sway and identify whether there were age-related differences in reliability. Thirty-two healthy individuals (16 younger and 16 older) participated. COP was recorded during quiet upright stance on 4 different days, and a variety of measures determined: mean velocity, median power frequency, RMS distance, sway area, and two fractal measures derived from Hurst rescaled range analysis ($H_{R/S}$) and detrended fluctuation analysis (DFA). Intraclass correlation coefficient (ICC) and standard error of measurement (SEM) were used to quantify reliability. Mean velocity was the most reliable measure. DFA exponents had relatively better reliability than $H_{R/S}$ exponents. In general, within-day reliability was better than between-day. In comparison with younger participants, older participants exhibited better relative reliability (ICC) for all COP measures and comparable absolute reliability (SEM) except for mean velocity and sway area. These results may be useful in guiding the future selection and interpretation of COP-based measures.

Keywords: Center of pressure; Aging; Postural control; Reliability

1. Introduction

Assessment of human postural (or balance) control is of frequent interest to researchers and clinicians. Decrements in balance control have been considered an important factor in the probability of a fall [1,2], and falls are often a proximal cause of injury and disability in the general population as well as in contemporary industry [3,4]. Performance of the balance control system is commonly assessed in the laboratory by quantifying sway during quiet upright stance. The most commonly used measures of sway are various parameters derived from temporal patterns of center of pressure (COP). Traditional COP-based measures (e.g. ellipse area, mean COP velocity, etc.) have demonstrated differences with age [5,6], sensory conditions [7], and pathology [8,9], and have been related to risk of falls [1,2]. However, it has also been suggested that these measures ignore the dynamic characteristics of COP movement and do not represent more subtle aspects of postural control [10]. To overcome these limitations, some modern approaches have been described, including Hurst rescaled range analysis ($H_{R/S}$) and detrended fluctuation analysis (DFA), which are intended to reveal the underlying fractal properties of the COP time series. Some of these modern methods have been reported as being more sensitive to factors thought to affect balance (e.g. fractal COP measures are more sensitive to age-related changes [11]).

In general, selection of a sway measure for research or clinical use should be motivated by several factors, including the measure reliability. COP mean velocity, for example, has been reported as the most reliable among traditional measures [12–14], and within-day results appear to have better reliability than between-day results [15]. Despite these valuable studies, important questions remain regarding the reliability of COP-based measures of postural control. No study to our knowledge has investigated the reliability of

* Corresponding author at: Department of Industrial and Systems Engineering, Virginia Tech, 250 Durham Hall (0118), Blacksburg, VA 24061, USA. Tel.: +1 540 231 6053; fax: +1 540 231 3322.
E-mail address: nussbaum@vt.edu (M.A. Nussbaum).

0966-6362/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.gaitpost.2008.01.005

the $H_{R/S}$ exponent. Only one study assessed within-day reliability of Hurst exponent derived from DFA method and age-related differences [16]. Whether age-related differences in reliability are present for other COP measures remains unknown. In the present study, we assessed within- and between-day reliability of both traditional COP measures and two more recent measures. Results from this work were intended to facilitate better selection of COP-based sway measures in future studies.

2. Methods

2.1. Participants

Thirty-two healthy individuals (16 younger and 16 older, gender balanced in each group) from the University and local community volunteered for the study (Table 1). The two age groups were intended to represent individuals near the typical beginning and end of working life. They had no self-reported injuries, illnesses, musculoskeletal disorders, or occurrences of falls in the past year. All completed an informed consent procedure approved by the Virginia Tech Institutional Review Board.

2.2. Procedures

Each participant completed four test sessions (sessions 1–4) on 4 different days. All sessions were done in the same laboratory environment, by the same experimenters with a minimum of 2 days between sessions. Each session consisted of three successive trials of quiet upright stance, with a rest period of approximately 1 min between trials. During each trial, participants stood as still as possible while barefoot, feet together, eyes closed, head facing straight ahead, and arms at their sides. Trials were conducted in a closed room to minimize any noise or other disturbances. Repeatability of foot placement between trials was maintained by outlining the feet on poster board placed on top of the force platform. Each trial lasted 75 s, with the initial 10 s and last 5 s removed to avoid initial transients and anticipation effects, respectively. Triaxial ground reaction forces and moments were sampled at 100 Hz using a force platform (AMTI OR6-7-1000, Watertown, MA, USA), low-pass filtered (2nd order, zero-phase-lag, Butterworth, 5 Hz cut-off frequency), and transformed to obtain COP values [17].

2.3. Data analysis

Dependent variables were COP-based sway measures derived from the stabilograms (temporal patterns of COP). The following traditional COP-based measures were determined according to the recommendations of Prieto et al. [6]: (1) median power frequency; (2) mean velocity; (3) RMS distance; and (4) sway area. In addition, $H_{R/S}$ exponent [18] and DFA exponent [11] were calculated. Median power frequency, mean velocity, $H_{R/S}$ exponent, and DFA exponent were determined in both the antero-posterior (AP) and medial–lateral (ML) directions. Increases in the traditional COP-based measures are typically interpreted as an overall deterioration of postural control [6,11,19]. $H_{R/S}$ and DFA exponents are measures of long-range dependence (persistence or anti-persistence) in a time series [11,20,21]. $H_{R/S}$ ranges from 0 to 1. $H_{R/S} > 0.5$ represents persistence (the greater $H_{R/S}$ the more persistent) and $H_{R/S} < 0.5$ represents anti-persistence (smaller $H_{R/S}$ = more anti-persistent), though this only applies to fractional Gaussian noise (fGn). For fractional Brownian motion (fBm), $H_{R/S}$ analysis can only be applied to anti-persistent data series and $H_{R/S} > 0.5$ represents anti-persistence (smaller $H_{R/S}$ = more anti-persistent) [11,21]. DFA exponents vary between 0 and 2. If the signal behaves as fGn, DFA exponents range from 0 to 1. If the signal behaves as fBm, DFA exponents range from 1 to 2. When DFA < 0.5 or $1 <$ DFA < 1.5, the signal is anti-persistent (smaller DFA = more anti-persistent). When $0.5 <$ DFA < 1 or $1.5 <$ DFA < 2, the signal is persistent (larger DFA = more persistent). Some diversity exists regarding how these exponents are to be practically interpreted. However, Collins et al. [10] have suggested that greater persistence of the COP series is correlated with increased muscle activity and decline in postural stability. Greater anti-persistence has been suggested to reflect a more tightly controlled postural system [16].

Preliminary analysis of the data was focused on systematic variations associated with day, trial, and age. Since the time of the experiment in each day and the interval between sessions were not tightly controlled, these two factors were considered as potential covariates to minimize the variability associated with them. Thus, an initial three-way MANCOVA with time of day and session interval as covariates was performed to test the effects of age, trial, and day on all dependent variables. A significant main effect was followed up with univariate ANCOVAs for each COP measure.

Reliability of each COP-based measure was quantified using ICC and SEM [22]. ICC was chosen to assess relative reliability [23–25], and is determined as between-subject variance versus total variance. ICC ranges from 0 (no reliability) to 1 (perfect reliability), and was interpreted using the following criteria: 0.00–0.39 poor, 0.40–0.59 fair, 0.60–0.74 good and 0.75–1.00 excellent [26]. SEM was selected to test absolute reliability [23–25], and provides an estimate of the precision of measurement [22]. It can help distinguish low ICCs caused by large within-subjects variability from low ICCs caused by a narrow measurement range [27]. A high SEM indicates a high level of error and implies non-reproducibility of tested values.

ICCs and SEMs were calculated using a statistical model reported by Eliasziw et al. [28] with the day effect as random. This model allows concurrent assessment of within- and between-day reliability. Both within- and between-day reliability coefficients were derived from multiple observations, which can effectively increase the sample size available for estimating each coefficient and thereby improve the precision of the results [28]. Significant interaction effects of age × time of day and age × session interval were found from the initial MANCOVA. Thus, these two factors (time of day and session interval) were added into the ANOVA model of Eliasziw et al. [28] as covariates to remove the variability associated with them. A

Table 1
Participant characteristics (mean ± S.D.)

Age group	Gender	Age	Stature (cm)	Body mass (kg)
Younger	Male ($N = 8$)	20.3 ± 1.4	176.1 ± 4.6	74.7 ± 12.1
	Female ($N = 8$)	21.5 ± 2.0	166.1 ± 5.2	59.6 ± 5.1
Older	Male ($N = 8$)	65.4 ± 3.7	175.5 ± 8.1	88.9 ± 13.3
	Female ($N = 8$)	60.8 ± 6.4	160.2 ± 7.5	66.2 ± 15.8

Table 2
COP-based measures obtained from younger and older participants

	Younger		Older	
	Mean	S.D.	Mean	S.D.
Median power frequency (ML) (Hz)	0.36	0.09	0.48	0.14
Median power frequency (AP) (Hz)	0.35	0.07	0.43	0.12
Mean velocity (ML) (mm/s)	11.4	4.7	18.0	9.9
Mean velocity (AP) (mm/s)	8.7	2.6	13.7	8.8
RMS distance (mm)	8.8	2.9	8.7	2.3
Sway area (mm^2/s)	41.7	23.2	68.4	68.9
$H_{R/S}$ exponent (ML)	0.92	0.03	0.90	0.03
$H_{R/S}$ exponent (AP)	0.93	0.02	0.92	0.03
DFA exponent (ML)	1.36	0.10	1.25	0.10
DFA exponent (AP)	1.41	0.08	1.32	0.11

separate ANCOVA was performed for each age group. For each ICC obtained, the 95% one-sided lower-limit confidence interval (CI) was calculated to provide a range of values that is likely to cover the true population value [28]. The level of significance was set at $p = 0.05$ for all statistical tests.

3. Results

Descriptive summaries of the COP-based measures for both age groups are presented in Table 2. Preliminary MANCOVA results indicated significant effects of age ($p < 0.0001$) and day ($p < 0.0001$). There was no significant trial effect ($p = 0.41$) or day × trial interaction ($p = 0.81$). Time of day and session interval as covariates also had no significant effects ($p = 0.28$ and $p = 0.12$, respectively). Subsequent univariate ANCOVAs indicated significant age effects on all COP measures except for RMS distance ($p = 0.97$) and sway area ($p = 0.17$). Significant day effects were found for ML median power frequency, AP median power frequency, AP $H_{R/S}$ exponent, and AP DFA exponent. No clear pattern was evident with respect to day for any of the COP measures, except for median power frequency. In both the AP and ML directions, median power frequency was ∼8% lower in days 2–4 versus day 1.

3.1. Reliability of COP-based sway measures

ICCs, the 95% one-sided lower-limit CIs, and SEMs of younger and older participants are summarized in Tables 3 and 4 for within-day and between-day, respectively. Several trends were evident. Mean velocity (AP and ML) was the most reliable measure and had excellent ICC values. SEM values for mean velocity were slightly lower in the AP than the ML direction. Median power frequency (AP and ML) showed a relatively wide range of ICC values, indicating poor–excellent reliability. SEM values for median power were comparable in the AP and ML directions. RMS distance and sway area showed good-excellent ICC values. ICC values for the fractal measures were fair to excellent. DFA exponents (AP and ML) had higher ICC values compared with $H_{R/S}$ exponents. SEM values for fractal measures were slightly lower in the AP than the ML direction.

3.2. Within- and between-day reliability

In general, within-day reliability was superior to between-day reliability for both age groups. All but one of the COP-based measures showed higher ICC values for within-day than between-day; the exception was median power frequency (AP, older group), which had equal ICC values for within-day and between-day. SEM values of within-day were lower than or equal to those of between-day for all COP-based measures.

3.3. Younger versus older participants

For all COP-based measures, higher ICCs were found in the older group for both within-day and between-day measures. Less consistent differences were found for SEMs. The older group had higher SEM values for median power frequency (ML and AP), mean velocity (ML and AP), and sway area, but lower or equal SEM values for RMS distance, $H_{R/S}$ exponents (ML and AP), and DFA exponents (ML and AP).

Table 3
Within-day reliability

	Younger			Older		
	ICC	95% one-sided lower limit CI	SEM	ICC	95% one-sided lower limit CI	SEM
Median power frequency (ML) (Hz)	0.61	0.31	0.06	0.77	0.52	0.07
Median power frequency (AP) (Hz)	0.41	0.01	0.06	0.57	0.27	0.08
Mean velocity (ML) (mm/s)	0.91	0.81	1.4	0.95	0.90	2.3
Mean velocity (AP) (mm/s)	0.86	0.72	1.0	0.95	0.91	1.9
RMS distance (mm)	0.80	0.61	1.3	0.82	0.66	1.0
Sway area (mm^2/s)	0.79	0.60	10.5	0.92	0.84	19.4
$H_{R/S}$ exponent (ML)	0.65	0.30	0.02	0.70	0.44	0.02
$H_{R/S}$ exponent (AP)	0.48	0.12	0.01	0.76	0.53	0.01
DFA exponent (ML)	0.76	0.51	0.05	0.82	0.64	0.04
DFA exponent (AP)	0.70	0.44	0.04	0.85	0.71	0.04

Table 4
Between-day reliability

	Younger			Older		
	ICC	95% one-sided lower limit CI	SEM	ICC	95% one-sided lower limit CI	SEM
Median power frequency (ML) (Hz)	0.55	0.40	0.06	0.56	0.39	0.09
Median power frequency (AP) (Hz)	0.34	0.20	0.06	0.57	0.42	0.08
Mean velocity (ML) (mm/s)	0.79	0.67	2.1	0.91	0.85	2.9
Mean velocity (AP) (mm/s)	0.77	0.65	1.2	0.92	0.87	2.4
RMS distance (mm)	0.71	0.57	1.6	0.80	0.70	1.0
Sway area (mm^2/s)	0.72	0.59	12.2	0.90	0.83	22.1
$H_{R/S}$ exponent (ML)	0.46	0.29	0.02	0.61	0.45	0.02
$H_{R/S}$ exponent (AP)	0.44	0.30	0.02	0.67	0.53	0.02
DFA (ML)	0.61	0.45	0.06	0.74	0.61	0.05
DFA (AP)	0.61	0.46	0.05	0.82	0.72	0.04

4. Discussion

The main objective of this study was to assess within- and between-day test–retest reliability of several postural sway measures and to investigate age-related differences in reliability. Overall, the results indicate that COP-based measures have quite diverse levels of reliability. Among the measures obtained, mean velocity was the most reliable, in accordance with previous studies [12–14]. Using seven healthy and older participants, Lafond et al. [12] reported within-day ICC values (AP and ML) of 0.77 and 0.90 for mean velocity, 0.02 and 0.24 for median power frequency, 0.52 and 0.61 for RMS distance, and an ICC = 0.47 for sway area. In the present study, a similar participant group and the same condition (eyes closed, hard surface, 60 s trial) were used, with higher ICC values obtained for all these measures. This discrepancy may be due to the different statistical model used by this study in determining mean square terms in ICC calculations. By using this model, both within-day and between-day reliability coefficients are derived from multiple observations [28]. It has been suggested that increasing the number of repetitions can decrease the weight of the error variance compared with the true score [29], which leads to higher ICC values. Another possible explanation is the larger number of participants recruited in the current study.

Although a number of methods have been employed to analyze fractal properties of COP data series [10,11,16,18,30], few studies have reported the reliability of these fractal measures. Amoud et al. [16] investigated within-day reliability of Hurst exponents derived from DFA for different trial durations (2.5 s, 5 s, and 10 s). ICC values increased with trial duration. Mean ICC values (AP and ML) in the 10 s trials were 0.48 and 0.52 among a younger group, and 0.75 and 0.62 in an older group. DFA exponent was derived similarly in the present study, and the corresponding within-day ICC values (AP and ML) were 0.70 and 0.76 for younger group, and 0.85 and 0.82 (AP and ML) for older group. The higher ICC values found here can be attributed to the substantially longer trial duration (60 s) employed, since reliability is directly related to trial duration [12,16]. In addition, Amoud et al. [16] had no constraints on foot placement. Doyle et al. [31] assessed the within-day reliability of the COP fractal dimension (D) in different vision and surface conditions using 30 healthy young participants. D measures the complexity of a signal and is a local property. In principle, D is independent of $H_{R/S}$ and DFA exponents, since the latter two are measures of global characteristics [32]. However, for a COP time series, which can be considered a fractal process [11,16,18,21], the local properties are reflected in the global ones. Thus, D is related to $H_{R/S}$ and DFA exponents for COP series. Doyle et al. [31] reported ICC values for D of 0.80 and 0.75 (AP and ML) for the condition with eyes closed and a hard surface. Results for within-day ICC of DFA exponents in the present study, from younger participants and using the same testing conditions, were 0.70 and 0.76 (AP and ML), comparable to the findings of Doyle et al. ICCs for $H_{R/S}$ exponents in the current study were lower than those for DFA exponents, and also lower than values of D reported by Doyle et al. [31], though they were comparable with the findings of Amoud et al. [16].

It is unclear why the different fractal measures exhibited different levels of reliability. Differences in the algorithms used to derive the measures may account for some of the discrepancies. However, ICC should be interpreted with caution. ICC is determined as the ratio of the variance between subjects to the total variance [24]. In general, a high ICC value indicates that most of the observed variance is attributed to difference between subject measurements and proportionately little variance is due to error related with repeated measurement. However, the lower ICC values may have resulted from either large inter-session variances or small variance between subjects. With less between-subjects variance, the same test–retest variation will represent proportionately more of the total variance and lead to lower ICC [33]. Here, the smaller variance of $H_{R/S}$ exponent among subjects (as indicated by the SDs in Table 2) may have contributed to the lower ICC values.

Relatively few studies have reported between-day reliability of COP-based measures. However, it is common to assess the effect of an intervention on postural stability over the course of several days in practice. Samson and

Crowe [34] used CV to assess the reliability of mean velocity across 3 days and reported mean CVs of 8.5% and 10.6% for eyes closed and eyes open conditions, respectively. Benvenuti et al. [15] used ICC to assess the short-term reliability (4 h between trials) and long-term reliability (1 week between sessions) of several COP-based measures and found higher short-term ICC values. The present study found that within-day reliability was better than between-day reliability, consistent with the work of Benvenuti et al. [15]. The lower between-day reliability may be attributed to a change in postural control over days, as a significant multivariate day effect was found. It has been suggested that a long-term postural control adaptation can occur between consecutive test days [35]. Such an adaptation can improve balance maintenance by changing the postural strategy. However, except for a decrease in median power frequency, no other obvious trends were found over the repeated experimental sessions. It is likely that as yet unidentified random effects between days were responsible.

The other objective of this study was to identify age-related differences in the reliability of COP-based measures. Higher ICCs were found among healthy older participants. This is consistent with the findings of Amoud et al. [16], who reported higher ICCs for Hurst exponents among older individuals, though significant differences were only observed in the AP and resultant directions. The higher ICC values may have resulted from a higher variation of measures among the older participants (as indicated by the SDs in Table 2). A number of age-related deficits (e.g. decline in visual function, reduced muscle strength, etc.) are possible causes of the decreased postural stability in the elderly [36,37]. These deficits, however, are likely specific for each individual, leading to a higher between-subject variance in the older group.

One potential limitation should be noted here. Only healthy participants were included. Since it has been suggested that ICC should only be applied when a fixed population of individuals can be well defined [23], the current reliability results may not generalize to other populations, such as older people with a history of falls, and patients with disorders related to and/or affecting postural control.

In summary, the within-day and between-day reliability of several COP measures has been determined, including relatively new fractal-based measures, and age-related differences in reliability were identified. Mean velocity was the most reliable measure, consistent with earlier reports. Among the two fractal-based measures studied, DFA exponents were relatively more reliable than $H_{R/S}$ exponents. Within-day reliability was generally better than between-day reliability. Older participants showed better relative reliability (ICC) across all COP measures and comparable absolute reliability (SEM) except for mean velocity and sway area. These results can facilitate the selection of reliable postural sway measures and protocols, and the interpretation of COP-based results obtained from studies of postural control.

Conflict of interest statement

None.

Acknowledgments

This work was supported in part by Cooperative Agreement Number R01 OH0078802 from the Centers for Disease Control and Prevention (CDC). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of CDC. The authors thank Mr. Navrag Singh for his assistance in data collection.

References

[1] Fernie G, Gryfe C, Holliday P, Llewellyn A. The relationship of postural sway in standing: the incidence of falls in geriatric subjects. Age Ageing 1982;11:11–6.
[2] Maki BE, Holliday PJ, Fernie GR. Aging and postural control: a comparison of spontaneous-and induced-sway balance tests. J Am Geriatr Soc 1990;38:1–9.
[3] Courtney TK, Webster BS. Antecedent factors and disabling occupational morbidity–insights from the new BLS data. AIHA J 2001;62: 622–32.
[4] Rubenstein LZ. Falls in older people: epidemiology, risk factors and strategies for prevention. Age Ageing 2006;35:ii37–41.
[5] Baloh RW, Fife TD, Zwerling L, Socotch T, Jacobson K, Bell T, Beykirch K. Comparison of static and dynamic posturography in young and older normal people. J Am Geriatr Soc 1994;42:405–12.
[6] Prieto TE, Myklebust JB, Hoffmann RG, Lovett EG, Myklebust BM. Measures of postural steadiness: differences between healthy young and elderly adults. IEEE Trans Biomed Eng 1996;43:956–66.
[7] Redfern MS, Moore PL, Yarsky CM. The influence of flooring on standing balance among older persons. Hum Factors 1997;39: 445–55.
[8] Boucher P, Teasdale N, Courtemanche R, Bard C, Fleury M. Postural stability in diabetic polyneuropathy. Diab Care 1995;18:638–45.
[9] Lafond D, Corriveau H, Prince F. Postural control mechanisms during quiet standing in patients with diabetic sensory neuropathy. Diab Care 2004;27:173–8.
[10] Collins JJ, De Luca CJ, Burrows A, Lipsitz LA. Age-related changes in open-loop and closed-loop postural control mechanisms. Exp Brain Res 1995;104:480–92.
[11] Norris JA, Marsh AP, Smith IJ, Kohut RI, Miller ME. Ability of static and statistical mechanics posturographic measures to distinguish between age and fall risk. J Biomech 2005;38:1263–72.
[12] Lafond D, Corriveau H, Hébert R, Prince F, Raiche M. Intrasession reliability of the center of pressure measures of postural steadiness in healthy elderly people. Arch Phys Med Rehabil 2004;85:896–901.
[13] Cornilleau-Pérès V, Shabana N, Droulez J, Goh JCH, Lee GSM, Chew PTK. Measurement of the visual contribution to postural steadiness from the COP movement: methodology and reliability. Gait Posture 2005;22:96–106.
[14] Raymakers JA, Samson MM, Verhaar HJJ. The assessment of body sway and the choice of stability parameter(s). Gait Posture 2005;21: 48–58.
[15] Benvenuti F, Mecacci R, Gineprari I, Bandinelli S, Benvenuti E, Ferrucci L, Baroni A, Rabufetti M, Hallett M, Dambrosia JM, Stanhope SJ. Kinematic characteristics of standing disequilibrium: reliability and validity of posturographic protocol. Arch Phys Med Rehabil 1999;80:278–87.

[16] Amoud H, Abadi M, Hewson DJ, Michel-Pellegrino V, Doussot M, Duchene J. Fractal time series analysis of postural stability in elderly and control subjects. J Neuroeng Rehabil 2007;4:12.

[17] Winter DA. Human balance and posture control during standing and walking. Gait Posture 1995;3:193–214.

[18] Duarte M, Zatsiorsky VM. On the fractal properties of natural human standing. Neurosci Lett 2000;283:173–6.

[19] Corbeil P, Blouin JS, Bégin F, Nougier V, Teasdale N. Perturbation of the postural control system induced by muscular fatigue. Gait Posture 2003;18:92–100.

[20] Delignieres D, Ramdani S, Lemoine L, Torre K, Fortes M, Ninot G. Fractal analyses for 'short' time series: a re-assessment of classical methods. J Math Psychol 2006;50:525–44.

[21] Delignieres D, Deschamps T, Legros A, Caillou N. A methodological note on nonlinear time series analysis: is the open- and closed-loop model of Collins and De Luca (1993) a statistical artifact? J Mot Behav 2003;35:86–96.

[22] Deneger CR, Ball DW. Assessing reliability and precision of measurement: an introduction to intraclass correlation and standard error of measurement. J Sport Rehabil 1993;2:35–42.

[23] Atkinson DB, Nevill AM. Statistical methods for assessing measurement effort (reliability) in variables relevant to sports medicine. Sports Med 1998;26:217–38.

[24] Bruton A, Conway JH, Holgate ST. Reliability: What is it, and how is it measured? Physiotherapy 2000;86:94–9.

[25] Weir JP. Quantifying test–retest reliability using the intraclass correlation coefficient and the SEM. J Strength and Cond Res 2005;19:231–40.

[26] Cicchetti DV, Sparrow SS. Developing criteria for establishing interrater reliability of specific items: application to assessment of adaptive behavior. Am J Ment Defic 1981;86:127–37.

[27] Henriksen M, Lund H, Moe-Nilssen R, Bliddal H, Danneskiod-Samsoe B. Test–retest reliability of trunk accelerometric gait analysis. Gait Posture 2004;19:288–97.

[28] Eliasziw M, Young SL, Woodbur MG, Fryday-Field K. Statistical methodology for the concurrent assessment of interrater and inrarater reliability: using goniometric measurements as an example. Phys Ther 1994;74:777–88.

[29] Bravo G, Potvin L. Estimating the reliability of continuous measures with Cronbach's alpha or the intraclass correlation coefficient: toward the integration of two traditions. J Clin Epidemiol 1991; 44:381–90.

[30] Doyle TL, Dugan EL, Humphries B, Newton RU. Discriminating between elderly and young using a fractal dimension analysis of center of pressure. Int J Med Sci 2004;1:11–20.

[31] Doyle TL, Newton RU, Burnett AF. Reliability of traditional and fractal dimension measures of quiet stance center of pressure in young, healthy people. Arch Phys Med Rehabil 2005;86:2034–40.

[32] Gneiting T, Schlather M. Stochastic models that separate fractal dimension and the Hurst effect. SIAM Rev 2004;46:269–82.

[33] Keating J, Matyas T. Unreliable inferences from reliable measurements. Aust J Physiother 1998;44:5–10.

[34] Samson M, Crowe A. Intra-subject inconsistencies in quantitative assessment of body sway. Gait Posture 1996;4:252–7.

[35] Tjernstrom F, Fransson PA, Hafstrom A, Magnusson M. Adaptation of postural control to perturbations—a process that initiates long-term motor memory. Gait Posture 2002;15:75–82.

[36] Pyykko I, Jantti P, Aalto H. Postural control in elderly subjects. Age Aging 1990;19:215–21.

[37] Woollacott MH, Shumway-Cook A. Changes in posture control across the life span—a systems approach. Phys Ther 1990;70:799–807.