Ruhe, A. , Fejer, R. and Walker, B.F. (2010) The test retest reliability of centre of pressure measures in bipedal static task conditions: A systematic review of the literature. Gait & Posture, 32 (4). pp. 436-445.

The test-retest reliability of centre of pressure measures in bipedal static task conditions

A systematic review of the literature

Alexander Ruhe[1] MChiro, René Fejer[2] MHSc (Clin Biomechanics), PhD, Bruce Walker[3] DC, MPH, DrPH

1. Corresponding author: Alexander Ruhe, PhD student, Murdoch University, Praxis fuer Chiropraktik Wolfsburg, Porschestrasse 1, 38440 Wolfsburg, Germany. Tel.: +49 5361 8481358; fax +49 5361 8481378. E-mail address: alexander_ruhe@hotmail.com

2. Adjunct Senior Lecturer , Rygmedicinsk Afdeling, Middlefart Sygehus, Denmark

3. Senior Lecturer, School of Chiropractic and Sports Science, Murdoch University, Murdoch, 6150, Western Australia, Australia.

Acknowledgements

The author(s) like to thank Tino Bos DC (TB) for his valuable contribution during the first phase of the literature search and Alexander Steinbrenner DC (AS) for volunteering as a third reviewer.

Abstract:

Summary of Background data: The analysis of center of pressure (COP) excursions is used as an index of postural stability in standing. Conflicting data have been reported over the past 20 years regarding the reliability of COP measures and no standard procedure for COP measure use in study design has been established.

Search methods: Six online databases (January 1980 to February 2009) were systematically searched followed by a manual search of retrieved papers.

Results: Thirty-two papers met the inclusion criteria. The majority of the papers (26/32, 81.3%) demonstrated acceptable reliability. While COP mean velocity (mVel) demonstrated variable but generally good reliability throughout the different studies (r=0.32-0.94), no single measurement of COP appeared significantly more reliable than the others. Regarding data acquisition duration, a minimum of 90sec is required to reach acceptable reliability for most COP parameters. This review further suggests that while eyes closed readings may show slightly higher reliability coefficients, both eyes open and closed setups allow acceptable readings under the described conditions (r≥0.75). Also averaging the results of three to five repetitions on firm surface is necessary to obtain acceptable reliability. A sampling frequency of 100Hz with a cut-off frequency of 10Hz is also recommended. No final conclusion regarding the feet position could be reached.

Conclusions: The studies reviewed show that bipedal static COP measures may be used as a reliable tool for investigating general postural stability and balance performance under specific conditions. Recommendations for maximizing the reliability of COP data are provided.

Keywords: Balance, center of pressure; force-plate; reproducibility; reliability; systematic review

Background

Postural stability is an important component in maintaining an upright position and in maintaining balance during normal daily movements and activities. Postural stability is also an important factor in the elderly where balance disability may increase the risk of falls and subsequent injury. In sport, problems with balance may lead to serious injuries [1]. Thus, postural stability has important implications in sports and rehabilitation. Many different methods exist today for assessing postural stability. The evaluation of parameters describing COP excursions is a frequently used method of measuring this stability and any associated pathological mechanisms. This is possible as the COP signal is proportional to ankle torque, a combination of descending motor commands as well as mechanical properties of the surrounding musculature [2]. Measurements are most commonly evaluated by using spatial measures such as sway distance, velocity and area traversed based upon sequential locations of the COP in the plane of the force platform.

Many factors contributing to postural control have been identified. This postural control system depends on the unimpaired ability to correctly perceive the environment through peripheral sensory systems, as well as to process and integrate vestibular, visual and proprioceptive inputs at the central nervous system (CNS) level. Depending on whether the task at hand is static or dynamic in nature, the CNS employs different strategies to form appropriate muscle synergies needed to maintain equilibrium [3]. In addition to individual perceptual and motor skills, the area of support in terms of foot position, musculoskeletal characteristics and task constraints play an important role in postural stability.

The methods of measurement of human standing posture can be broadly classified into three main groups: 1) Body segment displacement during standing posture, 2) muscle activity for maintaining postural equilibrium, and 3) measurement of the movement and patterns of the center of mass (COM) or center of pressure (COP) [3].

1) Body segment displacement refers to the change in position of body segments such head or trunk during adaptive movements in order to maintain balance [4].

2) During balance control, the muscle action appears to be an anticipatory feed-forward mechanism that is determined by an internal model of the inverted pendulum and acts in the long-term. It aims at stopping the fall and pushes the body back towards its reference point [2].

In contrast, the intrinsic feedback due to mechanical properties of ankle muscles operates with a zero delay in the short-term in order to slow down the fall of the inverted pendulum. The inverted pendulum model relates the controlled variable (COM) with the controlling variable (COP) [5]. The complementation of this mechanism by the feed-forward control is necessary as the muscle stiffness itself is not sufficient to stabilize the body if the critical level of displacement is reached [2].

3) COP can be defined as the position of the global ground reaction force vector that accommodates the sway of the body. In simple terms, it is the point at which the pressure of the body over the soles of the feet would be if it were concentrated in one spot. This measure, however, is not a true record of body sway but rather a measure of the activity of the motor system in moving the COP. Centre of Mass (COM) is a point equivalent of the total body mass in the global reference system and is commonly accepted to lie around the S2 vertebral level in normal upright posture [6]. Lafond et al. [7] demonstrated the relationship between COP and COM during stance, where COP oscillates on either side of the COM. While COP theoretically completely coincides with COM at low sway frequencies below 1Hz [4], its displacement during sway always exceeds that of the COM [7].

Of these, one of the most commonly used tools to investigate this complex balance system is the stabilogram, which is a measure of the time behaviour of the COP of a person positioned on top of a force platform consisting of a rigid plate supported by force transducers.

Postural sway observed in quiet standing represents the integrated output from the complex interaction between the balance systems mentioned above. As understanding of these balance mechanisms evolved over the last decades, the literature shows a large change in study designs and instruments used to investigate COP.

While the evaluation of COP excursions is a commonly used method for measuring postural stability [21-38] no standardization of this method exists. Further, the reliability of COP needs to be determined if studies using this method are to be considered valid. To our knowledge no systematic literature review has been conducted to investigate the reliability of COP measures.

Aims

The aims of this systematic literature review are 1) to describe and assess the methodological procedures of studies of the most commonly used COP measurements and methods, 2) to determine the reliability of commonly used centre of pressure measures in bipedal static task conditions, and 3) to provide recommendations regarding standardized COP methods for future use in study designs.

Methods

Search strategy

A comprehensive search strategy was developed by identifying all potentially relevant search terms, categorizing these terms into specific search phases and subsequently combining them by using Boolean terms. This search strategy was designed to be used in six different electronic databases. These were PubMed, MEDLINE, EMBASE, Web of Science, ScienceDirect, Digital Dissertations and the Cochrane library. The search strategy is available upon contacting the corresponding author.

Electronic searches

All databases were searched using the search strategy described above. Appropriate minor modifications to the basic search template were made to optimize the strategy in individual databases. Papers were limited to human studies published between January 1980 and February 2009.

Searching other resources

The hand search included analyzing references cited in studies selected from the original online search. Citation searches of relevant studies were conducted using the PubMed, MEDLINE and ScienceDirect databases.

Selection Criteria

Articles were limited to peer-reviewed journals and dissertations without restrictions regarding language. Wide inclusion and exclusion criteria for study designs were used in order to avoid limitation of potentially relevant papers.

The inclusion criteria were: Articles that were fully or partially concerned with the intra- and intersession reliability of COP data derived from bipedal static tasks on a force plate. For this systematic review, all COP measures, experimental setups and statistical models fitting these criteria were considered. No limitations of the type of patient demographics or health status applied.

We excluded studies with insufficient documentation of patient demographics or experimental setup. In addition, papers that were anecdotal, speculative or editorial in nature or studies that employed dynamic task conditions such as one-leg hopping, walking or some form of translation of the force platform were excluded.

If any title and abstract did not provide enough information to decide whether or not the inclusion criteria were met, the full text of the article was obtained.

Data extraction and management

For the purpose of this review AR acted as the principal reviewer. A colleague (TB) was involved independently in the process of identifying relevant studies and did not participate in further analysis of the finally included papers. A third reviewer (AS) was used for a majority decision in case discrepancies between AR and TB were not reconciled by discussion. To standardize the procedure between the reviewers, the principal reviewer developed a detailed protocol sheet for critical appraisal by which general information on objectives, design, participant's demographics and outcomes were extracted. Each reviewer retrieved the information independently. A test was conducted with two articles similar but unrelated to the review question and the procedures discussed.

Assessment of methodological procedures

The reviewers specifically assessed the application, documentation and association of six individual items with regards to test- retest reliability. These were 1) subject demographics and morphology, 2) sample duration, 3) number of trial repetitions, 4) visual condition (eyes open or eyes closed), 5) foot position, and 6) type of platform surface. Papers not describing the items need to be considered with caution as these are necessary for full understanding and appreciation of a reliability study. The rationale for choosing these factors was based on the fact that they were considered particularly relevant for reliability outcomes by the available literature [e.g. 30-38].

Results

Literature search results

Initially, the online search strategy identified 215 study abstracts which were screened individually by the reviewers. The application of inclusion/exclusion criteria by the reviewers on the titles and abstracts eliminated a further 162 papers. The most common reason for exclusion was not meeting the selection criteria like static or bipedal tasks. From the titles and abstracts of papers selected (n=53), full articles were reviewed and the same two reviewers (AR and TB) applied the inclusion criteria to the full text. Of these, 32 papers met the inclusion criteria and were included in this review. Eleven of these articles were added after the hand search of reference lists of included papers.

The selection process of suitable studies identified only minor variance between the reviewers. AR and TB initially disagreed on the inclusion of two papers, giving an overall agreement of 97%. The differences were documented and consensus reached after discussion.

Study results

Characteristics of participants and methods

About 30% of the studies (9/32) provided either insufficient description of the selection criteria for participants or none at all. No study described blinding of the examiners to the subject's health status.

While about half the authors described the baseline demographics in appropriate detail (18/32, 56%), only one study included a physical examination in order to validate their health status prior to study enrollment [8]. The other authors relied only on self-reports or did not provide any description at all. Only four studies reported calibration procedures of the force-plate, mostly by means of a calibrated static load [9-12]. With regard to patient demographics, most studies (83%) enrolled mixed gender groups of healthy participants between 21-40 years of age. Subject demographics and health status for all studies is shown in Table 1.

[Table 1]

In order to challenge postural control by modifying the integration of visual, vestibular or proprioceptive input, the included studies variously applied a selection or a combination of all three conditions (eyes open/eyes closed, firm/compliant surface, narrow/ normal stance). About 78% of the trials were performed under both eyes closed (EC) and eyes open (EO) conditions. Most authors conducted between 2-5 repetitions of postural sway recordings (14/32, 44%). In addition, the majority of trials were conducted on the firm surface (26/32, 81%) of a force platform.

[Table 2]

3. The statistics

As with the general experimental setups, an equally heterogeneous selection of statistics for describing the reliability was used, including the coefficient of variation (CV), generalizability coefficient (GC) as well as Pearson's correlation coefficient (PCC). The most commonly applied statistic however, were the different forms of the intraclass correlation coefficient (ICC). While most used models described originally by Shrout and Fleiss [13], others again employed modified versions [14]. About 30% (6/22) of the studies using the ICC failed to state the exact model used. The corresponding authors of these studies were contacted in order to gather the missing information but only two replies [14,15] were received. Where ICC models were reported, the two-way random effect model ($ICC_{2,1}$) was employed most often. Two studies [16,17] used the related generalizability theory.

4. Relationships between methods and reliability

While various studies have investigated the same COP parameters such as mean velocity or area of sway, an inter-study comparison of each parameter's individual reliability is often problematic because of differences in study designs. Only a few studies offer similar experimental procedures that allow for comparing the effect of various factors on the reliability of COP measures (Table 3 a, b, c).

[Table 3 a]

[Table 3 b]

[Table 3 c]

Discussion

General considerations

Due to the heterogeneous study designs and statistical models used there remains little common ground for combining the reliability of all data presented. Only a few papers allowed for direct inter-study comparison of results and most of the conclusions had to be drawn from those studies. No quantitative pooling of results from the studies was possible, but we were nevertheless able to extract enough information to make recommendations regarding reliable experimental setups for COP measurements.

Many trials on the reliability of COP measures were conducted as a complimentary part of papers concerned with postural control and as such COP did not appear in the title or keyword lists. Our search strategy aimed to address that problem; in addition, selected hand searches of reference lists were necessary to identify some of the relevant papers. However, of those studies none contributed any new information to the discussion. It may be therefore safely assumed that as the vast majority of papers were included, no implications regarding the overall conclusions arise.

With regards to differences between within-day and between-day reliabilities, it has been shown that trials run on the same day yield higher values [10,14]. While intra-trial and inter-trial reliability needs to be discussed, inter-rater reliability is unlikely to be of concern due to the simplicity of the apparatus, task and instructions. It appears, however, that it was this simplicity that has led to a lack of standardization in operation.

When considering potential sources of variability affecting the reliability of COP measures one may distinguish between effects of the measurement procedures themselves that can be controlled (e.g.

sampling duration, signal processing) and sources of variability that may not (e.g. joint/muscle function). Generally, the inter-subject variability may be at least partially explained by the learning effect observed [12], leading to an optimization of energy expenditure by progressively reducing body sway over the course of repetitions.

Choice of statistics

The choice of statistics has a profound effect on the reliability results of identical data sets - with subsequent consequences for the interpretation. The most commonly applied statistical tests were different models of the intra-class correlation coefficient (ICC) [13] and the coefficient of variation (CV). Two studies [18,19] employed Pearson's correlation coefficients (PCC) although its application in test-retest reliability studies is often discouraged for its inability to detect systematic error [20].

There are numerous versions of the ICC described in the literature some of which were employed in the presented studies. The ICC is a ratio of variances deriving from ANOVA that is unitless and theoretically varies between 0.0 and 1.0. For the purpose of this review, we used the definition stated in the classic Shrout and Fleiss [13] paper, regarding an ICC≥0.75 as indicative of good reliability.

The issue with the described heterogeneity of the chosen ICC models is that, depending on the data, different models are likely to yield varying results [20]. This, in conjunction with the heterogeneous experimental setups, renders a broader direct comparison of results even more difficult. Five studies failed to provide information on the ICC version selected.

As it will be seen later on in the discussion, the magnitude of the ICC is dependent on the variability of the COP data. The heterogeneity of the participants therefore needs to be carefully considered, as high ICC values may mask poor test-retest consistency if there is a large variability between the participants as it would be expected, for example, in the elderly. Conversely, even in the presence of low inter-participant variability, small test-retest variations may cause low ICC value [20,21].

Tables 3a.b,c show that results of the related models $ICC_{2,1}$ and $ICC_{2,3}$ are very similar. This also accounts for many of the values derived from unknown ICC types, which suggests the application of the same statistics as the experimental setups are similar. It may be argued that despite the unknown ICC models, the different results allow at least a limited comparison of results. As the error term of the ANOVA reflects the interaction between trials and subjects, this error term is small if the subjects' readings change in a similar fashion across a recording session. This would be expected as the baseline demographics of the participants in the studies are homogeneous. If the systematic error is small, ICC results derived from different formulas will be similar. This can be observed in Table 3.b when comparing the values reported by Lafond et al. ($ICC_{2,1}$) [22] and Carpenter et al. ($ICC_{3,k}$) [15].

In conclusion, it needs to be kept in mind that while the variations resulting from different statistics may be marginal under the described conditions, only studies employing the same formulae can be directly compared with confidence. Results derived from similar or identical experimental setups may nevertheless offer a limited comparability. Trends like higher reliability with increasing trial numbers or under visual deprivation are present irrespective of the ICC model used, the overall conclusions therefore remain unaffected.

Subject demographics and morphology

While most articles provided basic details on the baseline demographics, only few articles addressed the effect of intrinsic physical differences between subjects such as body mass index (BMI), height or weight on the reliability of COP measures [23,24]. This should be included in all COP studies as it has been demonstrated that selected temporal-distance COP parameters such as mean velocity or range are strongly dependent on the subject's height [23] and weight [25].

All but one of the studies reviewed relied on self-reported health information from the subjects without conducting some form of physical examination prior to the study. It remains questionable whether the participants in all cases remembered to report relevant previous injuries. Best practice would suggest conducting thorough physical examinations to rule out or identify biomechanical problems that may influence the readings.

A linear increase of COP velocity with increasing body weight, accounting for more than 50% of the observed variance, has also been demonstrated. As with increasing BMI (obesity) the centre of mass is located more anteriorly of the base of support and the foot mechanoreceptor afferents may be de-sensitized [25], the resulting postural instability may affect the reliability of COP measures. Another study argued that these effect are minimal when averaging at least three trials [24]. Until further evidence is established we nevertheless suggest normalizing the acquired data to these factors as originally described by O'Malley [26] and recently employed by Chiari [23] and Pinsault [27].

Age and gender

It is difficult to reach a conclusion regarding the effect of age and gender on the reliability of COP measures as only four studies offer direct comparability. Most studies enrolled mixed-gender groups which have shown high correlation coefficients [8]. In addition, even though it has been shown that COP measures differ between age groups [8,28,29], the reliability of these measures is not influenced by gender.

Demura et al. showed excellent reliability for a selection of different COP measures in both young and elderly subjects [29]. Lin et al., however, found higher inter-class correlation coefficients in groups of healthy elderly participants [14]. As discussed before, the higher ICCs reported in the elderly may be at least partially attributed to a higher variability of measures due to the expected age-related deficits in vision, proprioception or muscle strength.

The possible effect of fatigue, especially in a population of balance impaired or otherwise pathologically affected elderly subjects, has to be considered when increasing the trial number or duration on a single day. Finding the best ratio between trial duration and number of repetitions is of special interest. For example, it may prove impossible for such a group to perform multiple recordings of 180sec duration [10,30].

COP parameters

Recent studies suggest that the COP time series may represent the dynamics of a nonlinear (chaotic) system [31] that may be characterized using fractal dimension [13,19,30] and Stabilogram Diffusion Analysis (SDA) [30,32]. SDA assumes that COP can be modeled as a system of correlated, random walks, thereby addressing the dynamic nature of COP motion, its analysis is based on the random selection of two pairs of COP data [30]. Doyle et al. [11] noted that reliability coefficients for traditional measures such as mean velocity (mVel) or area were low ($ICC_{2,1}$ 0.05-0.71) while fractal dimension showed high values ($ICC_{2,1}$ 0.62-0.90) with low coefficients of variation (CV%) (1.8-6.7). It was therefore concluded that fractal analysis is a superior tool for COP investigations. In a later study, Santos et al. [17] did not support this conclusion. Their results show that fractal dimension data sets have comparable reliability values to traditional measures. In addition to different GC formulas, it is possible that the differences may be explained by the study design. Santos et al. used 60sec sampling duration, while Doyle et al. recorded data for only 10sec, which is surprising as previous research quoted in their own study [22] indicated that this is an insufficient time frame to gain reliable data. Amoud et al. [32] compared the reliability of Stabilogram Diffusion Analysis (SDA) and Detrended Fluctuation analysis (DFA) over three time intervals (2.5, 2 and 10sec) and showed that only AP motion of elderly subjects at 10sec duration could be assessed with a satisfactory reliability ($ICC_{3,1}$ ≥0.75). Limitations of their study include that no instructions regarding the foot placement were given as well as the short sampling durations. As it will be shown later on, longer durations may have yielded higher reliability coefficients.

Traditional parameters that employ minimal, maximal or peak-to-peak readings such as maximal amplitude should be avoided as they use only one or two data points among the entire recorded data and are therefore subject to great variances with subsequent low reliability. As averaging data may decrease the extreme effects of individual extreme readings, COP summary measures such as COP mean velocity should be used instead. Considering the low number of participants throughout the available studies, extreme values will nevertheless influence these means, as the great spectrum of some confidence intervals suggest.

The data available shows that mean velocity (mVel) is one of the most commonly used COP parameters. While the overall limitations described earlier have to be considered, it also shows consistently acceptable reliability values (Table 3) and can be considered the most reliable traditional COP parameter.

The results of this review suggest that with sufficient repetitions and sampling duration, all COP parameters will gain acceptable reliability (r≥0.75). Depending on the specific research purpose, the selection should include both distance (e.g. area) as well as time-distance (e.g. mVel) based parameters to gain a diverse description of the COP excursion.

Experimental Setup

About 28% (9/32) of the studies reviewed failed to state the instructions given to participants for the experiment. The two most commonly used instruction in the studies reviewed were “stand quietly” and “stand as still as possible”. In their study, Zok et al. [33] showed that the instructions issued to the participants during posturography may have a significant impact on the results. Most COP parameters investigated showed variations of 8% to 71% depending on which one of the instructions was given. Results obtained when the subjects were asked to “stand as still as possible” showed narrower confidence intervals indicating a higher consistency. We therefore recommend explicit instructions be given to participants in COP measurement studies. These instructions should be “stand as still as possible” while looking straight ahead.

Just a few studies reported some form of standardization of the environment such as lighting, temperature or time of day for the follow-ups [9,34]. Another potential limitation was varying foot positions when stepping off and back on the force platform during breaks. Only one study avoided this effect by having the participants sit down during breaks while maintaining the original foot position [35]. The arms at sides position was most commonly used position (60%). From a biomechanical point of view, this is more likely to keep the COP in a natural position than a position with hands in front or on the back. Accordingly, we recommend to remove shoes and have the arms at sides when data is being recorded.

Sampling and cut-off frequency

It has been shown that COP measures and its reliabilities vary depending on both the acquisition and cut-off frequency chosen [30,36]. In the literature, sampling frequencies ranging from 10-200Hz have been reported [9,16,17,22,27,37-41] and it seems that the reported variations in COP reliability across similar experimental setups are at least partially due to the different frequencies chosen.

Filtering of any signal is aimed at the selective rejection, or attenuation, of certain frequencies. The effect on parameters defined on the basis of frequency distribution of data such as mean power frequency is marked, whereas measures of mean displacement such as mean velocity or mean amplitude are far less sensitive to different sampling frequencies [36]. It has been shown that COP mean displacement velocity and path length were 26.1% greater when sampling frequencies of 50Hz were used compared to 10Hz [30] as it would be expected with more data points describing the shape of the COP. This however, did not significantly affect reliability as mean velocity showed generally consistent reliabilities (r=0.82-0.89) across different frequencies ranging from 64-200Hz [27,38,39] (Table 3).

Depending on the parameter selected, the choice of the cut-off frequencies has a significant effect on the reliability of COP data. The results for mean velocity for example showed low variation from $ICC_{2,1}$ 0.75 at 0.8Hz to 0.71 at 10Hz, while the reliability values of mean power frequency dropped from 0.21 to 0.13 under the same condition. A cut-off frequency of 10Hz has been suggested as the best compromise to reject noise power [36].

Depending on the COP parameter chosen, care should be taken with regards to the sampling frequency. Although further research is necessary, a sampling frequency of 100Hz with a cut-off level of 10Hz appears advisable for traditional COP measures.

Sampling duration

The test-retest results suggest that the number of trial recordings and duration appears to be a critical factor for obtaining reliable data sets. There have been few attempts to provide recommendations on both the length and number of trials that should be used when assessing balance. While earlier studies suggest that reliable data may be obtained with sample durations of 10 to 60sec [18,36,42,43]. This has later been disputed by studies investigating multiple time intervals of up to 120sec. They concluded that between 90 and 120sec are necessary to reach correlation coefficients of ≥0.75 for most COP parameters with confidence [15,22,38], further lengthening trial duration once an acceptable level is reached did not significantly reduce variability [9].

When similar studies are compared, the results confirm a trend towards increased reliability values with longer sampling durations. While the data presented includes only a limited selection of parameters from few studies and deriving from different statistical models, the values for mVel and RMS (AP/ML) show a positive relationship between sampling duration and reliability coefficient. This is also true for COP area, although the results for the different time intervals show a greater variation. Similar results can be observed with similar ICC models (Table 3).

Overall a sampling duration of 90sec can be expected to yield good reliability for all traditional COP parameters.

Number of repetitions

In addition to trial duration, the number of repetitions needed to gain acceptable reliability (≥0.75) also varies with the COP parameter under investigation and conflicting results have been reported.

For COP mean velocity for example, just two 120sec trials were required to reach an $ICC_{2,1}$ >0.90, whereas COP range and RMS needed four 120sec trials to reach similar reliability levels [22]. Furthermore, it has been stated that averaging two [35], three [24], four [38] or seven [17] trials yields acceptable reliability for the majority of COP parameters.

When comparing results of similar setups, the trend for increased trial numbers to yield more reliable data is apparent (Table 3). In a clinical setting, however, it may be argued that setups involving 10 trials in elderly people are impractical. Given the heterogeneous study designs in this review we conclude that averaging 3-5 trials of sufficient duration over one day is appropriate under most conditions.

Visual condition

Loss of vision does not affect COP measures of a young population during quite standing, while the effect was more marked in the elderly [44]. Under eyes closed conditions the reliability is lower for short sampling durations and rises as the individual adapts [43], leading to higher overall reliability values under eyes closed condition compared to eyes open [8,9,11,17,22,37,38,40,45,46].

It has also been shown that while both conditions showed high reliability values, the overall eyes closed data was more reliable than eyes open even in elderly subjects [9]. This appears a bit surprising as postural stability in the eyes closed position would be expected to be harder to maintain due to the reduced effectiveness of peripheral proprioception with increasing age. While loss of vision leads to increased muscle stiffness [47], the higher variances of measures caused by the decreased postural stability under this visual condition would be expected to result in higher ICC values, as described earlier. In addition, the trend by recent papers to report higher reliability estimates under eyes closed conditions may at least partially be attributed to improved technical equipment, a more rigorous scientific procedures in conducting the studies or a higher true score variability. For best practice we recommend that data be collected under eyes closed conditions

Foot position

It has been shown that widening of the foot position increases the passive stability of the musculoskeletal system and decreases active neural control [23,48]. A wide foot position acts to strengthen the coupling between hips and ankles and would be expected to yield higher reliability coefficients under eyes closed conditions (especially in the elderly).

Only one study by Hill et al. [41] directly compared narrow and normal stance. It showed that narrow stance measurements lead to lower overall reliability than feet apart ($ICC_{2,1}$ 0.27 compared to 0.55). The sampling duration, however, was short (25sec). Comparing selected data of similar studies indicates that while the correlation coefficients for seven repetitions after 60sec were significantly higher during normal stance (GC 0.96) compared to narrow foot position (GC 0.75) [17], both reached acceptable reliability. When data from a single 30sec trial were compared, narrow stance reached higher reliability values than a normal foot position [37] (Table 3).

No conclusion regarding the more reliable foot position can be reached with the current data available; accordingly best practice suggests that the position of the feet should be standardized. This may depend on the specific purpose of research and whether the participant's physical condition allows for a more challenging position for the proprioceptive system or not.

Surface condition

Three studies investigated data obtained from both firm (F) and compliant surfaces (C). All of them enrolled subjects with various conditions ranging from vestibular impairment [19] and LBP to lower limb injuries [16,40]. Without testing with open eyes, Salavati et al. [24] reported lower $ICC_{2,3}$ values with comparatively high standard error of measurement and coefficient of variation values for trials run on compliant surfaces with closed eyes. Benvenuti et al. [10] agree with this trend but added that the parameter COP distance antero-posterior tested on a compliant surface may be as reliable as on a firm base. This was the only study using elderly subjects (74.5 years), while the others enrolled young participants (14.9-38.4 years). In contrast, Harriage et al. [40] found generally lower correlation coefficients ($ICC_{2,1}$) during eyes closed and eyes open trials for both 60 and 120sec sampling duration on firm surface.

Even considering the differences in patient demographics and health condition, it may be concluded that data obtained on a firm surface tends to be more reliable, although no similar setups allow for a specific inter-study comparison of results. If the study purpose allows, we recommend using a firm surface although further research is required.

Conclusion

The overall results indicate that the reliability of traditional COP parameters is acceptable if our recommendations are followed in the study design. The test-retest reliability depends primarily on factors such as the number of trial recordings and duration rather than the selection of particular COP parameters. Care should be taken to thoroughly assess the subject’s physical status and anthropometric properties prior to the measurements. The primary finding of this systematic review is there is relatively little consistency in the methods employed and measurements selected for COP analysis when using a force-platform.

We recommend the following methods should be employed: Regarding the data acquisition duration, the results suggest that a minimum of 90sec is required to reach acceptable reliability for all traditional COP parameters in healthy subjects. A sampling frequency of 100Hz with a cut-off frequency of 10Hz is advisable. In addition, measurements should be conducted under eyes closed condition on a firm surface. Averaging the results of three to five repetitions can be expected to yield reliable data. Although the specific effect on the reliability remains unclear, the current evidence suggests that “stand as still as possible” should be the instruction issued prior to the recording. No final recommendation regarding the foot position is possible at this point.

References

1. McGuine TA, Greene JJ, Best T, Leverson G. Balance as a predictor of ankle injuries in high school basketball players. Clin J Sport Med 2000;10:239-244
2. Baratto L, Morasso PG, Re C, Spada G. A new look at posturographic analysis in the clinical context: sway-density versus other parameterization techniques. Motor Control 2002;6:246-270
3. Balasubramaniam R, Wing AM. The dynamics of standing balance. Trends Cogn Sci 2002;6:531-536
4. Winter DA. Human balance and posture control during standing and walking Gait & Posture 1995;3:193-214
5. Gatev P, Thomas S, Kepple T, Hallett M. Feedforward ankle strategy of balance during quiet stance in adults. Journal of Physiology-London 1999;514:915-928
6. Gard SA, Miff SC, Kuo AD. Comparison of kinematic and kinetic methods for computing the vertical motion of the body center of mass during walking. Hum Mov Sci 2004;22:597-610
7. Lafond D, Duarte M, Prince F. Comparison of three methods to estimate the center of mass during balance assessment. J Biomech 2004;37:1421-1426
8. Hageman PA, Leibowitz JM, Blanke D. Age and gender effects on postural control measures. Arch Phys Med Rehabil 1995;76:961-965
9. Bauer C, Groger I, Rupprecht R, Gassmann KG. Intrasession reliability of force platform parameters in community-dwelling older adults. Arch Phys Med Rehabil 2008;89:1977-1982
10. Benvenuti F, Mecacci R, Gineprari I, et al. Kinematic characteristics of standing disequilibrium: reliability and validity of a posturographic protocol. Arch Phys Med Rehabil 1999;80:278-287
11. 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;86:2034-2040

12. Nordahl SH, Aasen T, Dyrkorn BM, Eidsvik S, Molvaer OI. Static stabilometry and repeated testing in a normal population. Aviat Space Environ Med 2000;71:889-893

13. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86:420-428

14. Lin D, Seol H, Nussbaum MA, Madigan ML. Reliability of COP-based postural sway measures and age-related differences. Gait Posture 2008;28:337-342

15. Carpenter MG, Frank JS, Winter DA, Peysar GW. Sampling duration effects on centre of pressure summary measures. Gait Posture 2001;13:35-40

16. Doyle RJ, Ragan BG, Rajendran K, Rosengren KS, Hsiao-Wecksler ET. Generalizability of Stabilogram Diffusion Analysis of center of pressure measures. Gait Posture 2008;27:223-230

17. Santos BR, Delisle A, Lariviere C, Plamondon A, Imbeau D. Reliability of centre of pressure summary measures of postural steadiness in healthy young adults. Gait Posture 2008;27:408-415

18. Letz R, Gerr F. Standing steadiness measurements: empirical selection of testing protocol and outcome measures. Neurotoxicol Teratol 1995;17:611-616

19. Riley PO, Benda BJ, Gill-Body KM, Krebs DE. Phase plane analysis of stability in quiet standing. Journal of Rehabilitation Research and Development 1995;32:227-235

20. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res 2005;19:231-240

21. Keating J, Matyas T. Unreliable inferences from reliable measurements. Aust J Physiother 1998;44:5-10

22. Lafond D, Corriveau H, Hebert R, Prince F. Intrasession reliability of center of pressure measures of postural steadiness in healthy elderly people. Archives of Physical Medicine and Rehabilitation 2004;85:896-901

23. Chiari L, Rocchi L, Cappello A. Stabilometric parameters are affected by anthropometry and foot placement. Clin Biomech (Bristol, Avon) 2002;17:666-677

24. Salavati M, Hadian MR, Mazaheri M, et al. Test-retest reliability of center of pressure measures of postural stability during quiet standing in a group with musculoskeletal disorders consisting of low back pain, anterior cruciate ligament injury and functional ankle instability. Gait Posture 2009;29:460-464

25. Hue O, Simoneau M, Marcotte J, et al. Body weight is a strong predictor of postural stability. Gait Posture 2007;26:32-38

26. O'Malley MJ. Normalization of temporal-distance parameters in pediatric gait. J Biomech 1996;29:619-625

27. Pinsault N, Vuillerme N. Test-retest reliability of centre of foot pressure measures to assess postural control during unperturbed stance. Med Eng Phys 2009;31:276-286

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

29. Demura S, Kitabayashi T, Noda M, Aoki H. Age-stage differences in body sway during a static upright posture based on sway factors and relative accumulation of power frequency. Percept Mot Skills 2008;107:89-98

30. Raymakers JA, Samson MM, Verhaar HJ. The assessment of body sway and the choice of the stability parameter(s). Gait Posture 2005;21:48-58

31. Myklebust JB, Prieto T, Myklebust B. Evaluation of nonlinear dynamics in postural steadiness time series. Ann Biomed Eng 1995;23:711-719

32. Amoud H, Abadi M, Hewson DJ, et al. Fractal time series analysis of postural stability in elderly and control subjects. J Neuroeng Rehabil 2007;4:12

33. Zok M, Mazzà C, Cappozzo A. Should the instructions issued to the subject in traditional static posturography be standardised? Medical Engineering and Physics 2008;30:913-916

34. Rogind H, Simonsen H, Era P, Bliddal H. Comparison of Kistler 9861A force platform and Chattecx Balance System for measurement of postural sway: correlation and test-retest reliability. Scand J Med Sci Sports 2003;13:106-114

35. Kitabayashi T, Demura S, Noda M. Examination of the factor structure of center of foot pressure movement and cross-validity. J Physiol Anthropol Appl Human Sci 2003;22:265-272

36. Schmid M, Conforto S, Camomilla V, Cappozzo A, D'Alessio T. The sensitivity of posturographic parameters to acquisition settings. Med Eng Phys 2002;24:623-631

37. Chiari L, Cappello A, Lenzi D, Della Croce U. An improved technique for the extraction of stochastic parameters from stabilograms. Gait Posture 2000;12:225-234

38. Doyle RJ, Hsiao-Wecksler ET, Ragan BG, Rosengren KS. Generalizability of center of pressure measures of quiet standing. Gait Posture 2007;25:166-171

39. Haidan MR, Negahban H, Talebian S, et al. Reliability of Center of Pressure Measures of Postural Stability in Patients With Unilateral Anterior Cruciate Ligament Injury. Journal of Applied Sciences 2008;8:3019-3025

40. Harringe ML, Halvorsen K, Renstrom P, Werner S. Postural control measured as the center of pressure excursion in young female gymnasts with low back pain or lower extremity injury. Gait Posture 2008;28:38-45

41. Hill KD, G. CS, Kalogeropoulos A, Schwarz JA. Retest reliability of centre of pressure measures of standing balance in healthy older women. Australian Journal on Ageing 1995;14:76-80

42. Goldie PA, Bach TM, Evans OM. Force platform measures for evaluating postural control: reliability and validity. Arch Phys Med Rehabil 1989;70:510-517

43. Le Cliar K, Riach C. Postural stability measures: what to measure and for how long. Clinical Biomechanics 1996;11:176-178

44. Blaszczyk JW, Prince F, Raiche M, Hebert R. Effect of ageing and vision on limb load asymmetry during quiet stance. J Biomech 2000;33:1243-1248

45. Geurts AC, Nienhuis B, Mulder TW. Intrasubject variability of selected force-platform parameters in the quantification of postural control. Arch Phys Med Rehabil 1993;74:1144-1150

46. Mientjes MI, Frank JS. Balance in chronic low back pain patients compared to healthy people under various conditions in upright standing. Clin Biomech (Bristol, Avon) 1999;14:710-716

47. Redfern MS, Yardley L, Bronstein AM. Visual influences on balance. J Anxiety Disord 2001;15:81-94

48. Henry SM, Fung J, Horak FB. Effect of stance width on multidirectional postural responses. J Neurophysiol 2001;85:559-570

Table 1: Participant demographics and health status

Study	Number	Gender		Age	Weight (kg)	Height (cm)	Health status
		female	male	plus range			
Goldie et al., 1989 [44]	28	14	14	28.1±8	-	-	healthy
Hageman et al., 1995 [10]	A: 24	12	12	20-35	-	-	healthy
	B: 24	12	12	60-75	-	-	healthy
Hill et al., 1995 [42]	17	17	0	69.5±7	-	-	healthy
Le Cliar et al., 1995 [45]	25	13	12	19-32	-	-	healthy
Letz et al., 1995 [21]	A: 8	4	4	20-40	-	-	healthy
	B: 30	15	15	23-60	-	-	healthy
Mattacola et al. 1995 [51]	12	10	2	24.7±3	62.2±7.5	164.8±7	healthy
Riley et al., 1995 [22]	A: 11	4	7	50.3	-	-	Healthy
	B: 15	11	4	"	-	-	BVH
Samson et al., 1995 [52]	15	8	7	20-60	-	-	healthy
Takala et al., 1997 [53]	18	9	9	38.7	69.5	173	healthy
Moe-Nilssen, 1998 [54]	19	15	4	22.9	-	-	healthy
Benvenuti et al., 1999 [12]	A: 12	6	6	74.5	72.5	162	healthy
	B: 12	6	6	"	"	"	mod. disequilibrium
	C: 12	6	6	"	"	"	severe disequilibrium
Geurts et al., 1999 [47]	A: 8	4	4	44.3±20	-	-	helthy
	B: 8	4	4	24.9±2.4	-	-	healthy
Mientjes et al., 1999 [48]	8	3	5	38.4	-	-	CLBP
Carpenter et al., 2000 [17]	49	29	20	19-34	-	-	healthy
Chiari et al., 2000 [38]	12	6	6	26-40	-	-	healthy
Schmid et al., 2002 [37]	8	4	4	24-32	-	-	healthy
Kitabayashi et al., 2003 [36]	220	112	108	20	60.7	167	healthy
Rogind et al., 2003 [35]	12	12	0	25.8	60.0	166	healthy
Lafond et al., 2004 [25]	7	4	3	67±4	65±17.5	161±12	healthy
Doyle et al., 2005 [13]	30	10	20	23±5	71±12	175±9	healthy
Raymakers et al., 2005 [33]	A : 45	unclear		21-45	unclear	unclear	healthy
	B : 38	unclear		61-78	unclear	unclear	healthy
	C : 10	unclear		75-89	unclear	unclear	geriatric
	D : 21	unclear		65-87	unclear	unclear	Parkinson's
Amoud et al., 2007 [32]	A: 90	33	57	19.7	65.3	175.0	healthy
	B: 10	6	4	80.4	75.0	166.8	healthy
Doyle et al., 2007 [39]	15	8	7	19,9±1	72.2±12.5	169±4	healthy
Harringe et al., 2007 [41]	A: 9	9	0	14.9	50.4	161	healthy
	B: 7	7	0	"	"	"	LBP
	C: 8	8	0	"	"	"	LEI
Bauer et al., 2008 [11]	63	42	21	78.74±6.65	-	161.±11	healthy
Demura et al., 2008 [55]	A: 50	33	17	73	57	155	healthy
	B: 50	25	25	21	60	167	healthy
Doyle et al., 2008 [18]	15	8	7	19,9±1	72.2±12.5	169±4	healthy
Haidan et al., 2008 [40]	12	0	12	27.5±7	74.9±13.1	175±7	healthy
Lin et al., 2008 [16]	A: 16	8	8	20.9	67.2	171.1	healthy
	B: 16	8	8	63.1	77.6	167.9	healthy
Pinsault et al., 2008 [43]	10	5	5	24.6±3	68.9±14.2	175.±10	healthy
Santos et al., 2008 [19]	12	0	12	26.9±1	74.9±13.1	175±7	healthy
Salavati et al., 2009 [27]	A: 11	2	9	26.1±7	76.4±13	175±1	LBP
	B: 12	0	12	"	"	"	ACL injury
	C: 10	1	9	"	"	"	ankle instability

ACL: anterior cruciate ligament, BVH: bilateral vestibular hypofunction, CLBP: chronic low back pain, LBP: low back pain, LEI: lower extremity injury.

Table 2: Study Characteristics

Study	Condition	Parameters	Duration (sec)	Repetitions	Statistics	Results
Goldie et al., 1989 [44]	BP, tandem, EO/EC/F, narrow stance.	COP ML/AP Force AP/ML/ vertical	32	2	LR	EO: ML 0.30, AP=0.11
Hageman et al., 1995 [10]	BP EO/EC/F, normal stance	COP sway area	20	2	$ICC_{3,4}$	EO: 0.91, EC: 0.97
Hill et al., 1995 [42]	BP EO/F, normal, narrow stance+ others	Dispersion Index (DI)	25	9x3	$ICC_{2,1}$ CV	EO: normal ICC 0.55, CV 0.17, narrow ICC 0.27, CV 0.19
Le Cliar et al., 1995 [45]	BP, normal stance EO/EC/F	SD COP ML/ AP, mVel, SD force AP/ML/vertical	10, 20, 30, 40, 50, 60	2	RC	SD ML: 0.81, SD AP: 0.86, mVel: 0.84
Letz et al., 1995 [21]	BP, narrow, EC/EO/F	Vel, SD path, RMS AP/ML, mean excursion AP/ML	60 (2x30)	2	PCC	EO/EC/F 60sec: RMS path AP/ML 0.28-0.79, SD range 0.50-0.83, Vel 0.85-0.92.
Mattacola et al., 1995 [51]	BP, normal stance, EO/EC/F	Sway index	10	10	ICC (unclear), SEM, CI	EO: ICC 0.75, SEM 0.06, 95%CI 0.16-0.40 EC: ICC 0.06, SEM 0.26, 95% CI 0.13-0.87
Riley et al., 1995 [22]	BP, normal stance, tandem, EO/EC/F	Phase plane	7	2	PCC	Healthy subjects: COP ML 0.91, AP 0.78
Samson et al., 1995 [52]	BP, EO/EC/F narrow stance, tandem	Mean velocity	60	10	CV (%)	EO: 9.46% (4.55-29.38), EC: 10.53% (3.68-24.28)
Takala et al., 1997 [53]	BP, EO/EC/F narrow stance	Mean Vel, area, frequency, amplitude	30	2x2	ICC (unclear)	Short term: EO mVel 0.64, EC 0.56, area EO 0.55, EC 0.43. Long term: EO mVel 0.86, EC 0.77, area EO 0.44, EC 0.40
Moe-Nilssen, 1998 [54]	BP, EO/EC/F narrow stance	RMS AP/ML	30	3	$ICC_{1,1; 3,1}$ CV(%)	All parameters ICC <0.60 EO/EC, CV (%) 19.2-25.2
Benvenuti et al., 1999 [12]	BP, LB/NB, F/C, EO/EC	Mean velocity, Quadratic fit AP/ML	40, last 15 recorded	3	$ICC_{3,k}$	Vel 0.51-0.75, ML 0.65-0.77, AP 0.82-0.83.
Geurts et al., 1999 [47]	A:B, EO/EC/BV B: BP, EO normal stance	AP/ML RMS amplitude, RMS velocity	A: 3x20 B: 2x30	5	CV (%)	RMS area: ML 36%, AP 33%, RMS vel: ML 35%, AP 20%, range ML 32%, AP 27%.
Mientjes et al., 1999 [48]	BP, normal stance, EO/EC, F/C	AP/ML RMS, COP mean, MPF	unclear	3	ICC (unclear)	EO: RMS AP 0.14, ML 0.54, EC: RMS AP 0.41, ML 0.89
Carpenter et al., 2000 [17]	BP, narrow stance, EO/F	RMS, MPF, MPOS AP/ML	120 (8x15, 4x 30, 2x60)	3	$ICC_{3,k}$	Pos ML: 0.86-0.91, AP: 0.75-0.85, SD pos ML: 0.32-0.73, AP: 0.32-0.73
Chiari et al., 2000 [38]	BP, normal stance EO/EC/F	mVel, FD, area, Diffusion & Hurst coefficient (H)	50	10	ICC (modified)	mVel EO 0.83, EC 0.87, area EO 0.58, EC 0.70, FD EO 0.53, EC 0.80, SMP 0.20-0.79, NSMP 0.54-0.85
Schmid et al., 2002 [37]	BP, EO/F, normal stance	mVel, area, amplitude, MPF, Hurst	unclear	3	$ICC_{2,1}$	ICC: mVel 0.71-0.75, Ampl 0.36-0.37, area 0.55-0.62, MPF 0.13-0.21, H 0.21-0.39
Kitabayashi et al., 2003 [36]	BP, narrow stance, EO/F	34 parameters (e.g. area, mVel, RMS vel,)	60	3	ICC (unclear)	ICC ≥ 0.70 all parameters, Vel most reliable: mVel AP/ML, RMS vel: 0.96
Rogind et al., 2003 [35]	BP, EO/EC/F, normal/tandem stance	Vel AP/ML, 100% square, Max Ampl., sway index	25	4	CV	CV: 0.13-0.23
Lafond et al., 2004 [25]	BP, 2 platforms Normal stance,	RMS, range, Vel, MPF,	120 (30, 60,120)	9	$ICC_{2,1}$	EO: mVel 2 trials 120s for ICC≥0.90. RMS and range

	EO/F	MedPF AP/ML, area				6-8 trials 120s for ICC≥0.90, mVel ML most reliable
Doyle et al., 2005 [13]	BP, EO/EC/F, normal stance	FD, range, peak vel AP/ML, TEA	10	3	$ICC_{2,1}$ TEM, CV	EO/EC/F AP/ML: ICC FD >0.75, range 0.43-0.71, Vmax 0.12-0.58. EO/EC/C AP/ML: FD 0.62-0.90, range -0.28-0.72, Vmax 0.01-0.14.
Raymakers et al., 2005 [33]	BP, EO/F, narrow stance	Range, mVel, phase plane, area, DC	50	2	CV (%)	CV%: mVel 14, phase plane 18, area 26, DC 30, range AP 28, ML 19.
Amoud et al., 2007 [32]	BP, EO/F, stance unclear	Hurst exponent (SDA, DFA)	up to 30	4	$ICC_{3,?}$	ICC increases with time (10>5>2.5sec), only DFA (elderly) 10sec ICC=0.75.
Doyle et al., 2007 [39]	BP, EC/EO/F, normal stance	SD AP/ML, Vel, Area	90x2	10	GC	GC higher with increased duration, mVel most reliable (0.64-0.95) EO/EC.
Harringe et al., 2007 [41]	BP, EO/EC F/C, normal stance	Path length, SD AP/ML, RMS vel AP/ML/total, area	120	2	$ICC_{2,1}$ MMDC, CV	Healthy: ICC EO/F: 60s 0.34-0.66, 120: 0.40-0.78. EC/F: 60s 0.18-0.82, 120s 0.67-0.91. EO/C: 60s -0.02-0.82, 120s 0.18-0.82, EC/C: 60s 0.14-0.73, 120s 0.47-0.90.
Bauer et al., 2008 [11]	BP, EC/EO/F, narrow stance	Mean area, length, sway	30	3	$ICC_{2,1}$	All parameters ICC >0.75 except area EC (0.71)
Demura et al., 2008 [55]	BP, EO/F, narrow stance	36 parameters (e.g. RMS, area, mVel, RMS Vel)	60	3	ICC (unclear)	All parameters ICC >0.75 (e.g. mVel A: 0.96, B: 0.96, area A: 0.95, B: 0.92)
Doyle et al., 2008 [18]	BP EO/EC/F normal stance	DC AP/ML/ short term/long term	30, 60. 90	10x2	GC	All parameters GC ≥0.70 after 2 trials 30sec.
Haidan et al., 2008 [40]	BP, EC/EO F/C, narrow stance	SD vel, ampl, phase plane,	30	3	$ICC_{2,3}$ CV, MMDC	mVel EC/C 0.89, EC/F 0.87, EO/F 0.80. Area EC/C 0.65, EC/F 0.74, EO/F 0.10
Lin et al., 2008 [16]	BP, EC/F, narrow stance	MPF, mVel, RMS, area, DFA exponent, Hurst exponent (H)	60	2x3	ICC (modified), SEM	Young: mVel, RMS, area, DFA: ICC ≥0.75 same day, only mVel ICC≥0.75 inter-day. Elderly: All parameters ICC >0.75 same day
Pinsault et al., 2008 [43]	BP, EC/F, normal stance	Area, range, vel., Vmax AP/ML	30	10	$ICC_{2,1}$ LOA, SD, SEM	Vel, Vmax, vel AP, Vmax AP >0.75 (one trial). All >0.75 if 3 trials averaged.
Santos et al., 2008 [19]	BP, EO/EC/F narrow stance	FD, mean freq / vel / dist, RMS	60	2	GC	RMS dist: EO 0.43, EC 0.45, mVel EO 0.45, EC 0.36, range EO 0.52, EC 0.28. MPF EO 0.50, EC 0.44.
Salavati et al., 2009 [27]	BP, EO/EC F/C, narrow stance	SD amplitude / velocity, phase plane AP/ML	30	3	$ICC_{2,3}$ SEM, CV, MMDC	SD ampl. AP/ML: EO 0.61-0.64, EC 0.44-0.60. SD Vel AP/ML: EO 0.50-0.77, EC 0.71-0.83, Area: EO 0.33, EC 0.64, mVel EO 0.84, EC 0.91.

AP: anterior-posterior, BP: Bipedal, BV: blurred vision, C: compliant surface, CV: coefficient of variation, DC: diffusion coefficient, DFA: detrended fluctuation analysis, DC: diffusion constant, EC: eyes closed, EO: eyes open, F: firm surface, FD: fractal dimension, GC: G-coefficient, H: Hurst exponent, ICC: intra-class correlation coefficient, LB: large base, LOA: limits of agreement, LR: linear regression, ML: medial-lateral, MMDC: minimal metrical detectable change, MPF: mean power frequency, MPOS: mean position, mVel: mean velocity, NB: narrow base, PCC: Pearson correlation coefficient, RC: reliability coefficient, RMS: root mean square, SD: Standard deviation, SDA: stabilogram diffusion analysis, SEM: standard error of the mean, SL: single leg, TEA: total excursion area.

Commonly accepted interpretations for reliability coefficients are <0.40="poor", 0.40-0.75= "fair to good", >0.75= "excellent" reliability. CV values ≤0.33 are interpreted as acceptable [15].

Table 3a: Visual Condition

Visual condition	Sampling frequency (Hz)	Cut-off frequency	Parameter	Number of trials	Duration (sec)	Result	Study
Eyes open (EO)	100	5	Mean velocity	3	30	GC 0.83	[39]
	20	10	Mean velocity	3	30	$ICC_{2,1}$ 0.89-0.95	[25]
	200	10	Mean velocity	3	30	$ICC_{2,3}$ 0.80	[40]
Eyes closed (EC)	64	unclear	Mean velocity	3	30	$ICC_{2,1}$ 0.84	[43]
	100	5	Mean velocity	3	30	GC 0.84	[39]
	200	10	Mean velocity	3	30	$ICC_{2,3}$ 0.87	[40]

Commonly accepted interpretations of ICC and GC are <0.40="poor", 0.40-0.75= "fair to good", >0.75= "excellent" reliability [15].

Table 3b: Sampling duration

Duration (sec)	Sampling frequency (Hz)	Cut-off frequency	Root Mean Square (RMS) AP/ML	Mean Velocity	Area (A)	Study
30	20	10	EO/F $ICC_{2,1}$ 0.35-0.39	EO/F $ICC_{2,1}$ 0.73-0.87		[25]
	20	5	EO/F $ICC_{3,k}$ 0.32-0.58			[17]
	100	5		EO/F GC 0.64-0.93	EC/F GC 0.45-0.83	[39]
	64	unclear			EC/F $ICC_{2,1}$ 0.61-0.91	[43]
	200	10		EO/F $ICC_{2,3}$ 0.80	EC/F $ICC_{2,3}$ 0.74	[40]
60	20	10	EO/F $ICC_{2,1}$ 0.52-0.61	EO/F $ICC_{2,1}$ 0.77-0.90		[25]
	20	5	EO/F $ICC_{3,k}$ 0.53-0.65			[17]
	100	unclear	EO/F PCC 0.28-0.69	EO/F PCC 0.85-0.86		[21]
	100	5		EO/F GC 0.67-0.94	EC/F GC 0.52-0.88	[39]
	50	10	EO/F $ICC_{2,1}$ 0.46-0.56		EC/F $ICC_{2,1}$ 0.35	[41]
	100	5			EC/F ICC_{mod} 0.79	[16]
90	100	5		EO/F GC 0.68-0.95	EC/F GC 0.55-0.90	[39]
120	20	10			EC/F $ICC_{2,1}$ 0.69	[25]
	20	5	EO/F $ICC_{3,k}$ 0.58			[17]
	50	10	EO/F $ICC_{2,1}$ 0.68-0.74		EC/F $ICC_{2,1}$ 0.56	[41]

EC: eyes closed, EO: eyes open, F: firm surface, GC: G-coefficient, ICC: intra-class correlation coefficient, PCC: Pearson correlation coefficient, RMS : root mean square.
Commonly accepted interpretations of ICC and GC are <0.40=”poor”, 0.40-0.75= “fair to good”, >0.75= “excellent” reliability [15].

Table 3c: Number of repetitions

Study	Sampling frequency (Hz)	Cut-off frequency	Condition	Duration (sec)	Results					
					<3 repetitions		3-5 repetitions		6-10 repetitions	
[43]	64	unclear	EC/F (mVel)	30	$ICC_{2,1}$	0.82-0.83	$ICC_{2,1}$	0.82-0.88	$ICC_{2,1}$	0.88-0.89
[39]	100	5	EC/F (mVel)	30	GC	0.64-0.79	GC	0.84-0.89	GC	0.91-0.94
[40]	200	10	EC/F (mVel)	30			$ICC_{2,3}$	0.87		

EC: eyes closed, EO: eyes open, F: firm surface, GC: G-coefficient, ICC: intra-class correlation coefficient, mVel: mean velocity.
Commonly accepted interpretations of ICC and GC are <0.40="poor", 0.40-0.75= "fair to good", >0.75= "excellent" reliability [15].