Test-retest reliability and minimal detectable change for center of pressure measures of postural stability in elderly subjects

Reza Salehi, MSc,1 Ismaeil Ebrahim-Takamjani, PhD,2 Ali Esteki, PhD,3 Nader Maroufi, PhD4, Mohammad Parnianpour, PhD5

Faculty of Rehabilitation Science, Iran University of Medical Science, Tehran, Iran.

Abstract

Background: Postural instability has been identified as a potential precursor of falls in elderly subjects. Postural stability in quiet stance is commonly assessed with center of pressure (COP) measures. The purpose of this study was to determine test-retest reliability and minimal detectable change (MDC) for the center of pressure (COP) measures in the elderly subjects.

Methods: Eighteen healthy elderly people over the age of 60 years participated in this study. For each subject the COP was recorded during quiet upright stance on different levels of postural difficulty (eyes open versus eyes closed, firm surface versus foam surface) and lean condition (forward and backward). All measurements were done on two sessions with 7 days interval. These indices: mean velocity, standard deviation of amplitude, standard deviation of velocity, phase plane parameter and area (95% confidence ellipse). Intraclass correlation coefficient (ICC), standard error of measurement (SEM) and coefficient of variation (CV) were used to quantify test-retest reliability. The MDC for each measure was calculated to quantify intervention effects.

Results: In general, test-retest reliability of COP measures in the elder subjects was increased whenever postural difficulty of task increased in quiet standing. In standing conditions, mean velocity and phase plane parameter were the most sensitive and the most reliable measures. The lean range was the most sensitive and the most reliable measure, in the lean conditions.

Conclusion: Center of pressure measures in the quiet standing especially in difficult postural conditions demonstrated high sensitivity in the older subjects. These results may be useful in quantification and assessment of balance performance and treatment efficacy.

Keywords: reliability, minimal detectable change, center of pressure, balance, elder People

Introduction

Falls and their consequences are common problems associated with aging. Twenty-five to 35 percent of people over the age of 65 years fall at least once a year and that number rising to 50 percent for people over the 80 years of age [1,2].

The postural instability has been identified as a potential precursor of falls. Adequate postural control depends on temporal and spatial integration of vestibular, as well as visual and so-
matosensory information. The postural instability increases with aging due to decline in function of sensory and motor systems [2,3]. Clinical or/and functional balance tests and laboratory-based methods can be used to estimate the risk of falling [3]. At this viewpoint, reliability represents a required key for all measuring devices to insure that any observed difference in measures between test sessions reflect the real changes in postural stability, rather than producing systematic or random error in the measurement protocol. Reliability refers to the extent by which the measurement is inherently reproducible, or the degree to which the measurement influenced by measurement errors [4,5].

The postural stability in quiet stance is commonly assessed with center of pressure (COP) measures [3]. Like many biological measurements, the COP has an intrinsic variability affecting the reliability of postural control outcomes [6,7]. Therefore, the reliability of postural stability measures should first be established before they used to monitor balance problem and over the course of a clinical intervention [5].

In the older subjects, few studies have investigated the test-retest reliability of COP measures in different sensory conditions (table1). However, most of them did not considered different sensory environments in postural control measurements. Since the results of one testing condition (e.g. standing on force platform with eyes open) can not be generalized to other situations, incorporating different levels of postural difficulty by manipulating sensory information is crucial in reliability studies.

Although, reliability and consistency of outcome measures is important, Haley and Fragala-Pinkham suggested that "Measures to detect important effects of treatment must be valid (measure what is intended), responsive (able to detect an important change, even if that change is small), and interpretable (the intended audience must understand the magnitude of effect)" [12]. Ability to assess longitudinal changes in health status is critical for outcome measures used in the study of treatment efficacy. This aspect of measurement is termed responsiveness or sensitivity to change. Responsiveness has been defined as the ability of an instrument to detect a small but important change in health status over time. Recent authors have emphasized the importance of responsiveness at the level of the individual patient [12,13]. The Minimum Detectable Change (MDC) also known as reliable change or smallest real difference reflects true change rather than measurement error. The MDC is the smallest threshold of change scores that are detectable and beyond random error at a certain level of confidence (usually 95%). Both clinicians and researchers can use the MDC as a threshold to determine whether the changed score on a measure of an individual patient has reached a real improvement (or deterioration) or is due to the measurement error [12-14]. Thus, the MDC of a measure is critical for interpretation of data in clinical or research settings. Nevertheless, few studies in the elderly reported the MDC for postural sway measures [13].

Therefore, the objectives of this study were to determine test-retest reliability and MDC of some COP measures in a group of elderly subjects at different levels of postural difficulty over two sessions.

Methods

Subjects: Eighteen older healthy adults were (6 males and 12 females, age: 67.07±2.78 years, height: 159.78±6.33 cm, body mass: 66.71±8.48 kg) voluntarily participated in the experiment. Subjects gave their informed consent to the experimental procedure, and the study was approved by the Iran Medical Science University Ethic Committee.

The inclusion criteria for this study were age of 60 years or older, ability to walk independently without assistant devices and reported to have normal or corrected to normal vision. Ex-
<table>
<thead>
<tr>
<th>Study</th>
<th>Population (N)</th>
<th>Condition</th>
<th>Variable</th>
<th>Statistic</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lafond et al. [7]</td>
<td>Healthy older adult (7)</td>
<td>EO</td>
<td>Sway area; COP range; RMSa; COP mean velocity</td>
<td>ICC</td>
<td>ICC = 0.22 (Area); 0.44 (RangeML); 0.29 (RangeAP); 0.35 (RMSaM); 0.39 (RMSaAP); 0.87 (ML Mean velocity); 0.73 (AP Mean velocity)</td>
</tr>
<tr>
<td>Benventi et al. [8]</td>
<td>Healthy older adult (36)</td>
<td>Firm; Foam; EC; EO</td>
<td>MLpos; APpos; COP mean velocity</td>
<td>ICC</td>
<td>ICC = 0.65 (MLpos); 0.83 (APpos); 0.58 (Mean velocity)</td>
</tr>
<tr>
<td>Bauer et al. [9]</td>
<td>Healthy older adult (63)</td>
<td>EC; EO</td>
<td>Total path; sway area; frontal displacement</td>
<td>ICC</td>
<td>For EC ICC = 0.710-0.946 For EO ICC = 0.841-0.945</td>
</tr>
<tr>
<td>Swanenberg et al. [10]</td>
<td>Faller &amp; nonfaller older (37)</td>
<td>EC; EO Dual task</td>
<td>Mean velocity; sway area; maximal amplitude in AP and ML direction</td>
<td>ICC; SSD</td>
<td>For Max-ML; RMS-ML; MV and Sway area: ICC = 0.70-0.89, For Max-AP' and 'RMS-AP' ICC = 0.52-0.74 SSD = 0.37 cm (MaxML); 0.83 cm (MaxAP) 0.48-1.2 cm/s (Mean velocity); 1.48-3.75 cm² (Area)</td>
</tr>
<tr>
<td>Lin et al. [11]</td>
<td>Healthy young &amp; older adult (32)</td>
<td>EC</td>
<td>Mean velocity; RMS distance; sway area</td>
<td>ICC; SEM</td>
<td>For elderly subjects ICC = 0.91-0.92 (Mean velocity); 0.80 (RMS distance); 0.90 (sway area)</td>
</tr>
</tbody>
</table>

Abbreviations: N: sample size; EO: eyes open; EC: eyes closed; ML: mediolateral; AP: anteroposterior; MLpos: COP ML mean position; APpos: COP AP mean position; RMSa: root mean square amplitude; SSD: Smallest significant difference; Max: maximum. Note: Results of all variables for all conditions are not reported here.

Table 1. Summary of related literatures for reliability of the COP measures used to assess postural stability in elderly subjects.
clusion criteria were acute or subacute disease of the cardiovascular or respiratory systems, neurological or musculoskeletal impairments (such as history of CVA or Parkinson or joint replacement) and cognitive impairment that was quantified with the Mini-Mental State Examination (MMSE score of <24).

**Instrumentation:** The COP data for subjects were obtained using strain gauge Bertec 4060-10 force platform and Bertec AM-6701 amplifier. The amount of noise presented in the force platform measurements determined by applying a calibrated static load between 10 to 40 kg on the center of force platform [15]. The calculated COP displacement of force platform was less than 1 mm, which represented an acceptable amount of noise [15]. All data were collected with the sample frequency of 100 Hz.

**Procedure:** The test and retest measurements of COP signal were performed with 7 days interval by the same tester. The conditions were randomly ordered in both sessions. Postural sway was assessed in six different conditions and levels of difficulty. Quiet stance conditions included: 1) standing with eyes open on a firm surface, 2) standing with eyes closed on a firm surface, 3) standing with eyes open on a foam surface and 4) standing with eyes closed on a foam surface. Lean conditions included: 1) standing with eyes open on a firm surface and maximum forward lean and 2) standing with eyes open on a firm surface and maximum backward lean.

Subjects stood barefoot with their heels separated by 50% of pelvic width and their arms resting at their sides. Subjects were instructed to "stand as still as possible" for data collection period [5]. In eyes-open condition, subjects were instructed to look at a fixed target approximately 2m in front of their faces. In eyes-closed condition, subjects were a blindfold to eliminate visual input. In the foam surface condition, the force platform was covered with a 10cm thick piece of foam. Each condition was performed for 30 seconds in random order and repeated three times.
Table 3. Reliability analysis of the COP measures in different level of postural difficulty (N=18). Bold items are significant at p < 0.05.

<table>
<thead>
<tr>
<th>Measure</th>
<th>AP S.D. of Amplitude (cm)</th>
<th>ICC (95% CI)</th>
<th>SEM</th>
<th>CV (%)</th>
<th>MDC</th>
<th>ICC (95% CI)</th>
<th>SEM</th>
<th>CV (%)</th>
<th>MDC</th>
<th>ICC (95% CI)</th>
<th>SEM</th>
<th>CV (%)</th>
<th>MDC</th>
<th>ICC (95% CI)</th>
<th>SEM</th>
<th>CV (%)</th>
<th>MDC</th>
</tr>
</thead>
<tbody>
<tr>
<td>S.D. of Velocity (cm/s)</td>
<td>0.81 (0.39-0.94)</td>
<td>0.05</td>
<td>16.80 ±0.15</td>
<td>0.61 (0.60-0.87)</td>
<td>0.06</td>
<td>13.49 ±0.14</td>
<td>0.84 (0.63-0.95)</td>
<td>0.07</td>
<td>12.09 ±0.19</td>
<td>0.94 (0.84-0.98)</td>
<td>0.05</td>
<td>7.61 ±0.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phase Plane parameter</td>
<td>0.78 (0.26-0.95)</td>
<td>0.05</td>
<td>3.88 ±0.15</td>
<td>0.80 (0.36-0.93)</td>
<td>0.1</td>
<td>5.16 ±0.27</td>
<td>0.94 (0.83-0.98)</td>
<td>0.07</td>
<td>4.62 ±0.21</td>
<td>0.97 (0.93-0.99)</td>
<td>0.14</td>
<td>5.52 ±0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ML S.D. of Amplitude (cm)</td>
<td>0.88 (0.64-0.96)</td>
<td>0.02</td>
<td>11.59 ±0.05</td>
<td>0.80 (0.36-0.93)</td>
<td>0.03</td>
<td>18.38 ±0.07</td>
<td>0.86 (0.59-0.95)</td>
<td>0.03</td>
<td>8.63 ±0.1</td>
<td>0.95 (0.86-0.98)</td>
<td>0.02</td>
<td>6.24 ±0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S.D. of Velocity (cm/s)</td>
<td>0.81 (0.38-0.95)</td>
<td>0.03</td>
<td>3.31 ±0.09</td>
<td>0.80 (0.36-0.93)</td>
<td>0.06</td>
<td>4.57 ±0.15</td>
<td>0.93 (0.79-0.97)</td>
<td>0.08</td>
<td>6.39 ±0.23</td>
<td>0.98 (0.94-0.99)</td>
<td>0.15</td>
<td>8.07 ±0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phase Plane parameter</td>
<td>0.83 (0.43-0.94)</td>
<td>0.03</td>
<td>3.64 ±0.1</td>
<td>0.82 (0.44-0.94)</td>
<td>0.06</td>
<td>4.51 ±0.17</td>
<td>0.93 (0.79-0.97)</td>
<td>0.08</td>
<td>6.34 ±0.24</td>
<td>0.98 (0.94-0.99)</td>
<td>0.23</td>
<td>7.28 ±0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Velocity (cm/s)</td>
<td>0.81 (0.49-0.94)</td>
<td>0.05</td>
<td>3.65 ±0.14</td>
<td>0.82 (0.43-0.94)</td>
<td>0.09</td>
<td>4.31 ±0.25</td>
<td>0.94 (0.83-0.98)</td>
<td>0.09</td>
<td>4.84 ±0.25</td>
<td>0.98 (0.95-0.99)</td>
<td>0.15</td>
<td>5.05 ±0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phase Plane parameter Total</td>
<td>0.81 (0.42-0.94)</td>
<td>0.06</td>
<td>3.61 ±0.16</td>
<td>0.80 (0.38-0.95)</td>
<td>0.12</td>
<td>4.85 ±0.32</td>
<td>0.94 (0.83-0.98)</td>
<td>0.11</td>
<td>4.86 ±0.3</td>
<td>0.98 (0.95-0.99)</td>
<td>0.18</td>
<td>4.89 ±0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Area (95% ellipse) (cm²)</td>
<td>0.98 (0.65-0.96)</td>
<td>0.29</td>
<td>22.69 ±0.81</td>
<td>0.93 (0.49-0.94)</td>
<td>0.4</td>
<td>23.75 ±0.94</td>
<td>0.98 (0.63-0.96)</td>
<td>0.87</td>
<td>16.36 ±2.43</td>
<td>0.97 (0.92-0.99)</td>
<td>1</td>
<td>8.81 ±2.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
For the maximum lean trials, a feedback system (a monitor at an approximately 2m distance) for the subjects, was used to indicate how far the subject leaned. The subjects executed the leaning movement at their own pace until they reached their maximal range [6]. The subjects received continuous verbal feedback from the experimenter until the appropriate position was reached. Maximal leaning position was maintained until the end of the trial. During this time, subjects were instructed not to raise their heels or fingers while performing the task. To avoid the fatigue effect, a rest period of 1 minute between each test was given.

Data analysis: The COP output signals were filtered with MATLAB program (second-order, zero phase, Butterworth, low-pass digital filter with cut-off frequency of 15 Hz) and then exported to the Microsoft Excel to calculate the COP parameters. The anteroposterior (AP) and mediolateral (ML) displacement of COP were measured along the y-axis and x-axis, respectively. Parameters calculated from the COP data were: (1) mean velocity, (2) standard deviation (S.D.) of amplitude in AP and ML directions, (3) standard deviation (S.D.) of velocity in AP and ML directions, (4) phase plane parameter in AP-ML (combined), AP and ML directions, (5) sway area (95% confidence ellipse) [16], (6) lean range for the lean conditions (position of COP between the maximal and minimal position in A/P direction) [17].

Statistical analysis: The mean of functional tests three trials, ABC questionnaire and the
COP parameters in each condition was used for statistical analysis. Paired t-test between test and retest scores was calculated to verify the absence of systematic bias and alpha level set at 0.05 for all statistical analyses.

Relative reliability was assessed using two-way random model of intraclass correlation coefficient (ICC_{2,1}). According to Fleiss' classification: ICC ≥ 0.75 which indicate as "excellent" reliability; ICC value contained between 0.41 and 0.74 "fair to good" reliability; and ICC value < 0.40 "poor" reliability [18]. For each ICC, 95% confidence interval (CI) was calculated to take the sampling distribution into account. To assess absolute reliability, the standard error of measurement (SEM) was calculated as the square root of the mean square error term derived from the analysis of variance table [4]. In addition, the coefficient of variation (CV) was determined for comparison of absolute reliability between the COP measure of test and retest sessions (S.D./mean * 100). This was achieved by calculating mean of CV from individual CVs [4]. The SEM is useful for computing the minimal detectable change (MDC) or change that could be considered clinical difference between two measurements. The MDC was defined as 95% CI of SEM of the COP measure (1.96 *√2 *SEM). The multiplier of√2 was used to account for the additional uncertainty introduced by using difference scores from measurements at 2 points in time [14].

**Results**

Table 2 shows the descriptive summaries of the COP measures for test and retest sessions. Table 4 and 5 demonstrate ICC and its 95% CI, SEM, CV, and MDC.

Overall, results showed that the ICC generally tend to increase as levels of difficulty increase. There was no significant difference between test and retest mean scores for any COP measures in all conditions, which indicates absence of any systematic bias (p > 0.05). In the lean conditions, excellent reliability was found for lean range with ICC levels of 0.90 and 0.83 for forward and backward lean, respectively.

**Discussion**

The purpose of present investigation was to assess the reliability and MDC of the COP measures in the elderly subjects. Reliability of the COP measures has previously been addressed. Previous research into reliability of COP measures has not been as conclusive and comprehensive as the present study due incorporation of different levels of postural difficulty manipulated by sensory information. It is difficult to compare our results specifically with the previous studies due to different setting such as; sample duration, feet stance, eyes open versus eyes closed and firm surface versus foam surface.

Our results showed that whenever level of task difficulty increased, test-retest reliability increased. All parameters demonstrated excellent reliability when subjects were asked to close their eyes while standing on the foam surface. In general, a high ICC value indicated that most of the observed variance is attributed to difference between subject measurements and proportionately little variance generated due to error related to repeated measurement. Bauer et al. found that the reliability of the COP measures increased when subjects were asked to close their eyes while maintaining the narrow stance position [9]. Our results suggested that foam surface could be used in assessment of postural stability in the elderly subjects.

Our findings demonstrated good to excellent reliability in quiet standing conditions for elderly subjects. Demura et al. found that the reliability in a static upright posture is also considered to be high in the elderly [19]. Lin et al. reported that older participants exhibit better relative reliability in comparison with young participants [11].

The mean velocity showed excellent relative reliability and high absolute reliability in all
standing conditions consistent with previous studies. Salavati et al. reported high to very high relative reliability of mean velocity in all conditions of postural difficulty [20]. Lafond et al. reported excellent ICC values for COP velocity [7]. Samson and Crowe used CV to assess the reliability of mean velocity across 3 days and reported mean CVs of 8.5% and 10.6% for closed eyes and open eyes conditions, respectively [21]. However, those results were based on the coefficient of variation (CV), not on measurement error. Our results were consistently better in the ML direction compared to AP direction with the similar previous works. This results shows agreement with the results of Swanenburg et al. [10] but not with Corriveau et al reports. [22]. This difference could be due variation in protocol such as feet distance.

As Horak suggested, the primary purposes of the clinical balance assessment are to identify whether or not a balance problem exists in order to predict risk of a fall or to determine whether treatment is needed or has been effective and identify the underlying cause of the balance problem in order to manage or treat it effectively [23]. Reliable measurements are not necessarily sensitive to detect differences. Minimal detectable change values are useful to therapists in rehabilitation and intervention programs in determining whether change during or after intervention is clinically significant. [12]. A more reliable measure with lower MDC monitors the results of intervention with more sensitivity [13]. Our findings indicates that, the mean velocity and phase plane parameter were the most sensitive and the most reliable measures, in standing conditions.

In the lean conditions only one COP parameter showed excellent reliability (lean range with ICC 0.90 and 0.83 for forward and backward lean condition, respectively). For sway area in forward and backward lean the ICC values of 0.73 and 0.67, and CV values of 17.25% and 30.56% were found, respectively. In the lean conditions, lean range was the most sensitive and the most reliable measure. Juras et al. reported the ICC above 0.85 for the range of COP excursion [17]. Van Wegan et al. reported that in the lean condition older individuals increased the COP variability [6].

The lack of reliability of postural stability measures in lean condition decreases the power of a study to detect differences between groups in the same test condition because of the random measurement errors that may cause an increase in the variance. Two potential sources of error contribute to the lack of reliability: The lack of precision of the instrument and the variability of the phenomena measured [24]. The accuracy of the COP measured in this study was 1 mm. Such precision is satisfactory, but good precision dose not necessarily mean that the measure is reliable. The second source was variability of the outcome variable being measured. This is related to the elected procedure chosen and the phenomena being measured [24]. Whereas sex differences have been found in COP parameters [25], the results of the present study could be more applicable to female subjects, who constituted the majority of included participants.

**Conclusion**

The center of pressure measures in the quiet standing (especially in difficult postural conditions) demonstrated high sensitivity in the older subjects. This suggests that these measurements can be used to assessment balance among older subjects and to evaluate the effectiveness of intervention program.

**Acknowledgments**

This research was supported by grants from the Iran University of Medical Sciences. The experiment was conducted in Biomechanics Lab., Rehabilitation Research Center, Iran University of Medical Sciences.
References