Concurrent validity, inter-, and intrarater reliabilities of smart device based application for measuring vertical jump performance

Dong-Sik Oh
Department of Physical Therapy, Hanseo University, Republic of Korea, mmlee@dju.kr

Yoon-Hee Choi
Department of Physical Therapy, Graduate School, Daejeon University, Daejeon Republic of Korea, mmlee@dju.kr

Yu-Jin Shim
Department of Physical Therapy, Graduate School, Daejeon University, Daejeon Republic of Korea, mmlee@dju.kr

Sam-Ho Park
Department of Physical Therapy, Graduate School, Daejeon University, Daejeon Republic of Korea, mmlee@dju.kr

Myung-Mo Lee
Department of Physical Therapy, Daejeon University, Daejeon, Republic of Korea, mmlee@dju.kr

Follow this and additional works at: https://dcgdansk.bepress.com/journal

Part of the Health and Physical Education Commons, Sports Medicine Commons, Sports Sciences Commons, and the Sports Studies Commons

Recommended Citation

This Article is brought to you for free and open access by Baltic Journal of Health and Physical Activity. It has been accepted for inclusion in Baltic Journal of Health and Physical Activity by an authorized editor of Baltic Journal of Health and Physical Activity.
Concurrent validity, inter-, and intra-rater reliabilities of smart device based application for measuring vertical jump performance

Dong-Sik Oh1 AC, Yoon-Hee Choi2 BCD, Yu-Jin Shim2 CE, Sam-Ho Park2 CDF, Myung-Mo Lee3 ACD

1 Department of Physical Therapy, Hanseo University, Republic of Korea
2 Department of Physical Therapy, Graduate School, Daejeon University, Daejeon Republic of Korea
3 Department of Physical Therapy, Daejeon University, Daejeon, Republic of Korea

abstract

Background: The aim of this study was to examine the concurrent validity and inter-and intra-rater reliabilities of smart device based application, against force platform-based portable device Wii balance board (WBB) measurements for the flight time and jump height during the vertical jump.

Material and methods: Thirty participants (23.8 ±3.41 years) completed three vertical jumps, which were evaluated using WBB and application for smart device. To assess the concurrent validity, jump height and flight times were obtained from each device. Inter-and intra-rater reliabilities were determined by replicating data analysis of smart device based application recordings.

Results: Flight time and jump height collected from smart device based application showed excellent agreement level with WBB (flight time and jump height: ICC [2,1]=0.972). However mean flight time and jump height from the smart device based application was significantly higher than WBB (mean difference: 0.006 sec, 0.745 cm, p < 0.05). Intra-rater sessions showed good level of agreement (flight time: ICC [2,1] = 0.967, jump height: ICC [2,1] = 0.974), and inter-rater session showed almost perfect reliability (flight time: ICC [2,1] = 0.985, jump height: ICC [2,1] = 0.987).

Conclusions: Smart device-based applications could be used to replace pressure-based portable devices for clinical evaluations in post-injury rehabilitation as well as evaluating sports performance.

Key words: computing methodologies, validity and reliability, vertical jump performance, spatio-temporal analysis.
INTRODUCTION

Vertical jump is one of the most widely used methods of predicting motor performance in various domains [1]. It is used as a notable index to identify lower body muscle power and fatigue after strengthening lower body extensor muscles in athletes [2–4]. In clinical settings, vertical jump is used as a method to enhance bone intensity and bone density in adults [5]. In addition, it is also used as a health promotion screening tool for predicting bone strength in young adults [6]. Therefore, appropriate validity and reliability for measurement of vertical jumps must be established, since it is an index that describes not only training, but also shows physical performance abilities and physical health [7].

The traditional method of measuring the vertical jump is to jump with an ink at the tip of the fingers and touch the paper posted on the wall. Another way is using 'Vertec', a vertical jump measurement device, but nevertheless, these two methods both require direct measurement, analysis, and recording manually by the examiner [8]. Later on, contact mats, belt mats, video equipment, photoelectric cells, etc. were utilized to measure jump heights, but validation tests of measurement methods remained very limited [1, 2, 9].

The use of force plates is a relatively objective and quantitative measurement method for vertical jumps. However, the downside is that it could not be efficiently utilized in the field due to expensive costs, limited space, examiner’s competence, and inconvenient portability [10, 11]. To compensate for these limitations, portable devices are introduced for easy and efficient use. Wii balance board (WBB) is a supplementary device for Wii, and it is designed to measure balance [12]. Through the load cells located at each corners of the device, measuring both the center of pressure (CoP) and the vertical force applied to the surface is possible [13]. Various software which collects data measured with WBB has been developed, allowing broad usage as a measurement device with similar properties of a force plate. Recently, tracking of CoP transfers and measurement of change in surface repulsive force applied vertically have been showing high validity and reliability [14]. Low cost and portability allow WBB to be an alternative device for the force plate in the clinical field [15].

Recently, there is an increase in clinical attempts to measure and monitor motor performance abilities using smart devices [7, 16]. Movements are measured with acceleration sensors and gyroscope sensors of smart devices, and various applications are developed to display, save, and transmit measured data [17]. Movement performance evaluation using the application is rated to be a very useful measurement tool not only in a sports related field, but also in both clinical and everyday settings [18].

Therefore, this study will first investigate the concurrent validity between the measured flight time during vertical jumps and the height of the jump by simultaneously running WBB, which may be used more efficiently in the clinical field as an alternative to force plates, and smart device applications, thanks to which measurement and monitoring is easy. Secondly, this study will also suggest a clinical possibility by deriving test-retest reliability and intrarater reliability of measurement methods using smart device applications.
MATERIAL AND METHODS

PARTICIPANTS

This study is a cross-sectional study design. At least 30 participants for the sample size were required based on Walter et al. and considering acceptable reliability ICC = 0.900 and expected reliability ICC = 0.960 of the values observed 3 times per participant with 5% of drop out [19]. Significant level was set to \( \alpha = 0.05 \), power = 0.8. A total of 30 healthy adult university students (15 females and 15 males, age: 23.8 ±3.41 years, body mass index: 23.24 ±4.19 kg/m²) were recruited. None of the participants had any musculoskeletal, cardiovascular, neurological diseases that might affect the study outcomes. They were fully informed and understood the objectives of the study, and agreed to participate voluntarily. This study was approved by the Ethical Committee of Daejeon University.

INSTRUMENTS

Vertical jump was conducted on WBB (Nintendo Co. Ltd., Kyoto, Japan). WBB is a rectangular board with the size of 20.5 × 13.2 × 3.2 inches. Continuous data collection of ground reaction force (GRF) applied vertically and the shifting of CoP is possible through the load cells at each corner of the board. Continuously shifting values of GRF on WBB during vertical jumps are transmitted to the PC via Bluetooth, and the data are saved in 100 Hz using Balancia software (Balancia software ver. 2.0, Mintosys Inc., Korea). Vertical jumping motions were video recorded with the camera in the iOS based smart device (iPad Air 2, Apple Inc., USA) taking 240 frames per second. In order to obtain accurate times for take-off and landing of both feet, the camera of the smart device was placed evenly with height of the participants' feet, and the recording was made from frontal view in frontal plane with 3 meters distance. A smart device cradle was used to fix the camera firmly to prevent any shaking caused from manipulation of the device and jumping movements (Figure 1). Recorded video data were analyzed in same sampling rate using application for iPad Air 2 (Whatsmyvert iOS App ver 1.0, ©Andreas Rauh 2017).

Fig. 1. Procedures for performing vertical jumps and data collection methods
Prior to vertical jumps, the participants warmed up for 5 minutes by walking on a treadmill or lightly jogging. The participants placed both hands on the side of their waist, both feet in comfortable distance apart, and with the verbal cue of 'jump', they quickly bent their knees to a squat position in about 90 degrees and jumped up as strongly as possible. After take-off, both legs remained fully extended, and the participants were asked to have both feet land at the same time. In order to make safe landing on top of WBB, the participants were given 3 practice trials, and then 3 actual vertical jump performances were measured. Two examiners each operated WBB and application for iPad Air 2 to collect and record the data.

Data collected with WBB and Balancia software were calculated to obtain flight time after coding in Microsoft Excel for conversion to GRF data. The flight time was set to time taken from 'take-off' to 'landing'. In concordance with the research processes of prior studies, take-off time was identified as the point where GRF drops below 10N, and landing time was identified as the point where GRF increases over 10N [20] (Figure 2-A). Flight time ($t_{flight}$) calculated from each data were substituted in the following formula to derive jump height ($h_{peak}$).

$$h_{peak} = \frac{v_{take-off}^2}{2g} \quad (1)$$

$$v_{take-off} = \frac{gt_{uftake-off}}{2} \quad (2)$$

Gravitational acceleration ($g$) = 9.799

After the vertical jump, recordings were imported to a jump analysis application, and video frames were searched to set take-off frames (when both feet go away from WBB) and landing frames (when both or one foot touches WBB).

The relationship between jump height and take-off velocity is calculated with the law of conservation of mechanical energy (the sum of potential energy and kinetic energy at the take-off is equal to the sum of potential energy and kinetic energy at its peak).

$$\frac{1}{2}mv_{take-off}^2 + mgh_{take-off} = \frac{1}{2}mv_{peak}^2 + mgh_{peak} \quad (3)$$

During the vertical jump, ignoring air resistance, potential energy at take-off ($h_{take-off} = 0$) and kinetic energy at peak ($v_{peak} = 0$), therefore

$$\frac{1}{2}mv_{take-off}^2 = mgh_{peak} \quad (4)$$

$$h_{peak} = \frac{v_{take-off}^2}{2g} \quad (5)$$

The time between take-off and landing is the flight time; instantaneous velocities of take-off and landing are equal values but in different directions, and since gravity velocity applied onto the participants during flight time is consistent,

$$\Delta v/\Delta t = -g \quad (6-a)$$

$$\frac{(v_{peak} - v_{take-off})}{(t_{peak} - t_{take-off})} = -g \quad (6-b)$$

$v_{take-off}$ calculated from the formula above may be substituted in (5) to calculate jump height.
The iOS based art device used in this study captured motion in 240 Hz per second, and the application ‘Whatsmyvert’ was set to import the video to search 240 frames per second. When take-off and landing points are set after searching the frames, the number of frames in between the two points was calculated to yield flight time. These values were then substituted in the physics formula above to calculate the jump height.

Based on the number of frames between the two points and sampling rate, take-off time and landing time were each displayed on the screen of the smart device, and the time between the two points were recorded as flight time. When take-off frame and landing frame were set by the examiner, application for iPad Air 2 automatically substitutes the flight time into the formula to compute jump height (Figure 2-B).

\[ sm(0 - v_{\text{take-off}}) / \frac{t_{\text{flight}}}{2} = -g \]  
\[ -v_{\text{take-off}} = -g \times \frac{t_{\text{flight}}}{2} = -1 \]  
\[ v_{\text{take-off}} = g \times \frac{t_{\text{flight}}}{2} \]

(a): take-off point (where the GRF is less than 10N), (b): flight time, (c): landing point (where the GRF is over than 10N), (d): height of center of mass at take-off point, (e): height of center of mass at landing point.

Figure 2. (A) Changes in the ground reaction force and major landmarks during a vertical jump, and (B) ‘Whatsmyvert’ jump analysis application.
The concurrent validity was evaluated by comparing the flight time and the jump height collected through WBB and jump analysis application. The same examiner searched the recording to re-evaluate flight time and jump height for comparison after 3 days to check from intra-rater reliability of jump analysis application. Another trained examiner evaluated the same recording to determine inter-rater reliability.

**STATISTICAL ANALYSES**

Statistical analyses were performed using MedCalc for Windows, version 14.8.1 (MedCalc Software, Ostend, Belgium). General characteristics and measurement values are described as mean ± standard deviation. Variables obtained using WBB and an application for iPad Air 2 was used for paired t-test to determine the systematic difference. The level of agreement of measurement results are suggested as (Intra-class correlation coefficient, ICC [2, 1]) and 95% confidence interval (95% CI). Bland-Altman plot graph and 95% limits of agreement (95% LOA) was used for absolute comparison of the measurement results, and regression line and $R^2$ value were suggested for linear regression analysis. Coefficients of variation of method error CV$_{ME}$ were calculated for method of error (ME) using standard deviation of the difference scores (SD), and coefficient of variation (CV) was used for conversion to percentage (CV %).

\[
\text{ME} = \frac{SD}{\sqrt{2}}
\]

\[
\text{CV} \% = \frac{2 \times ME}{(X_1 + X_2)} \times 100\%
\]

In order to calculate range of error, the standard error of measurement (SEM) was calculated. SEM was derived by substituting the higher SD value between the two measurement outcome values in the formula:

\[
\text{SEM} = SD \times \sqrt{1 - \text{ICC}[2, 1]}
\]

and it was converted into percentage of the mean value to display in SEM %.

Intra-rater reliability and inter-rater reliability of application for iPad Air 2 were ICC [2, 1] and 95% CI. Bland-Altman plot graph, 95% LOA, and CV % values were calculated and the repeated measurement values of the same examiner and the measurement values of the two examiners were each categorically compared. In addition, SEM % of the mean values was calculated to measure the range of error of the two sessions. Statistical significance was $\alpha < .05$.

**RESULTS**

Data for 90 vertical jumps were collected, and 85 data were used for statistical analysis after excluding 5 sessions that may affect the data due to unstable landing.

**CONCURRENT VALIDITY**

The results of comparison of the collected data through WBB and application for iPad Air 2 are shown in Table 1. The result of paired t-test, which was conducted to compare the two measurement systems, showed that application for iPad Air 2 measurement was 0.006 sec longer in the flight time and 0.745 cm higher in the jump height than the measurements from WBB, producing a statistically significant difference ($p < 0.05$). However, the consistency level
of the two measurement results were flight time ICC \([2, 1] = 0.972\) and jump height ICC \([2, 1] = 0.972\), showing excellent levels, and CV % (flight time = 2.738, jump height = 5.433) and SEM % (flight time = 2.306, jump height = 4.625) both showed low levels of error. These consistency levels were also reflected in Bland-Altman plot graph, symmetrical distribution of scatter plot graph, and 95% LOA values (flight time = \(-0.045 \sim 0.031\), jump height = \(-5.328 \sim 3.839\)) (Figure 3).

Table 1. Concurrent validity of BB and jump analysis App

<table>
<thead>
<tr>
<th>Variables</th>
<th>WBB</th>
<th>Jump App</th>
<th>t(p)</th>
<th>ICC ([2, 1]) (95% CI)</th>
<th>CV %</th>
<th>SEM %</th>
<th>95% LOA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flight time</td>
<td>0.487 ±0.059</td>
<td>0.493 ±0.062</td>
<td>3.162 (0.002)</td>
<td>0.972 (0.952~0.983)</td>
<td>2.738</td>
<td>2.306</td>
<td>-0.045~0.031</td>
</tr>
<tr>
<td>Jump height</td>
<td>29.457 ±4.141</td>
<td>30.202 ±4.426</td>
<td>2.980 (0.003)</td>
<td>0.972 (0.954~0.983)</td>
<td>5.433</td>
<td>4.625</td>
<td>-5.328~3.839</td>
</tr>
</tbody>
</table>

\(^1\)Mean ±S.D. WBB: Wii balance board; App: iPad Air 2 application; ICC: intra correlation coefficient; CV %: coefficients of variation of method error %; SEM: standard error of measurement; LOA: limits of agreement.

Fig. 3. Relationship between the WBB and jump analysis App for the flight time and the jump height in the vertical jump

**INTRA-RATER RELIABILITY**

The comparison results of repeated measurements of the flight time and the jump height by the same examiner using application for iPad Air 2 are shown in Table 2. The results of paired t-test show that there were no significant systematic differences between the measured sessions. The flight time was ICC \([2, 1] = 0.967\), and the jump height was ICC \([2, 1] = 0.974\), showing a high level of consistency. CV % (flight time = 3.220, jump height = 5.679) and SEM %
(flight time = 2.535, jump height = 4.464) both shows low levels of error. The symmetrical distribution of Bland-Altman plot, scatter plot graph, and 95% LOA (flight time = -0.046~0.044, jump height = -4.890~4.818) values also showed high consistency and strong reliability levels (Figure 4).

Table 2. Intra-rater reliability of jump analysis App

<table>
<thead>
<tr>
<th>Variables</th>
<th>Measure1</th>
<th>Measure2</th>
<th>t(p)</th>
<th>ICC [2, 1] (95% CI)</th>
<th>CV %</th>
<th>SEM %</th>
<th>95% LOA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flight time (s)</td>
<td>0.487 ± 0.059(^1)</td>
<td>0.493 ± 0.062</td>
<td>3.162 (0.002)</td>
<td>0.972 (0.952~0.983)</td>
<td>2.738</td>
<td>2.306</td>
<td>-0.045~0.031</td>
</tr>
<tr>
<td>Jump height (cm)</td>
<td>29.457 ± 4.412</td>
<td>30.202 ± 4.426</td>
<td>2.980 (0.003)</td>
<td>0.972 (0.954~0.983)</td>
<td>5.433</td>
<td>4.625</td>
<td>-5.328~3.839</td>
</tr>
</tbody>
</table>

\(^1\)Mean ± S.D, ICC: intra correlation coefficient; CV %: coefficients of variation of method error %; SEM: standard error of measurement; LOA: limits of agreement

**INTER-RATER RELIABILITY**

Inter-rater reliability was ICC [2, 1] = 0.985 for flight time, and ICC [2, 1] = 0.987 for jump height, showing high levels, and there were no statistically significant differences in the systematics. CV\(_{ME}\) (flight time = 2.161, jump height = 4.014) and SEM % (flight time = 1.736, jump height = 3.224) both showed a low level of error (Table 3). Bland-Altman plot graph, symmetrical distribution of scatter plot graph, and 95% LOA (flight time = -0.031 ~ 0.030, jump height = -3.429 ~ 3.429) values showed high consistency levels (Figure 5).
Table 3. Inter-rater reliability of jump analysis App

<table>
<thead>
<tr>
<th>Variables</th>
<th>Rate 1</th>
<th>Rate 2</th>
<th>t(p)</th>
<th>ICC [2, 1] (95% CI)</th>
<th>CV %</th>
<th>SEM %</th>
<th>95% LOA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flight time (s)</td>
<td>0.493 ±0.062¹</td>
<td>0.495 ±0.063</td>
<td>0.437 (0.664)</td>
<td>0.985 (0.977~0.990)</td>
<td>2.161</td>
<td>1.736</td>
<td>-0.031~0.030</td>
</tr>
<tr>
<td>Jump height (cm)</td>
<td>30.202 ±4.426</td>
<td>30.238 ±4.583</td>
<td>0.135 (0.893)</td>
<td>0.987 (0.980~0.992)</td>
<td>4.014</td>
<td>3.224</td>
<td>-3.429~3.429</td>
</tr>
</tbody>
</table>

¹Mean ±S.D, ICC: intra correlation coefficient; CV %: coefficients of variation of method error %; SEM: standard error of measurement; LOA: limits of agreement

![Fig. 5. Agreement for the flight time and the jump height between rater 1 and rater 2 of a jump analysis App](image)

**DISCUSSION**

This study compared iOS based application and force platform based WBB to determine validity of the flight time and the jump height measurements, and intra-rater reliability and inter-rater reliability. As a result, high concurrent validity, intra-rater and inter-rater reliability are shown.

With the existing platform type GRF measuring device, flight time method, impulse-momentum method, and work-energy method are implemented to calculate jump height. Among these methods, the flight time method is the simplest method of measuring the jump height, since it uses take-off moment velocity and jump height using the flight time.

A slight systematic bias was observed between the two measurement devices that calculated jump height using flight time method. Application for iPad Air 2 gave a measurement mean of 0.7 cm greater than WBB. In the case where accurate take-off and landing points cannot be captured with the application,

www.balticsportscience.com
within the video, the frame with the foot at the closest contact on top of WBB was set as take-off and landing points. With this result in a more lenient flight time calculation, and when the flight time is substituted in the formula, greater jump height differences may have been calculated.

Despite the systematic bias, the flight time ICC [2, 1] = 0.972 and the jump height ICC [2, 1] = 0.972 calculated from the two measurement devices show an excellent consistency level, and a high correlation is shown in Bland-Altman plot. CVME, which shows the variation of the two measurement results, showed a very low level of 2.74% and 5.43% for the flight time and the jump height, respectively. 95% LOA were -0.045 ~ 0.031 and -5.328 ~ 3.839 each, resulting in a distribution within very small area. SEM % value also showed a low level of error between the two measurement results for flight time and jump height, yielding 2.30% and 4.62% each.

There are studies that report jump performance measurement comparisons between sensor-based measurement devices and force plates. Casartelli and Muller [21] compared squat jump using an accelerometric system and a force plate, and reported approximately 5.6cm bias and ICC = 0.74. Glatthorn and Gouge [1] measured jump performance using a photoelectric cell sensor and when compared with force plate, high ICC value of 0.997 ~ 0.998 and a systematic bias of over 1cm were reported. When these results are compared to this study, application for iPad Air 2 ‘Whatsmyvert’ demonstrated a relatively low systematic bias and high correlation to confirm that it is a device that has a high validity for jump performance measurements. This consistency level is similar to the studies that analyzed video recordings of jump performances using smart devices. Gallardo-Fuentes and Gallardo-Fuentes [22] compared contact platforms and jump analysis applications in smart phones to measure jump performance abilities in male and female athletes. The researchers reported approximately 0.1cm low systematic bias and high ICC = 0.99 for the squat jump, ICC = 0.99 for the countermovement jump, and ICC = 0.99 for the drop jump along with high test-retest reliability (r = 0.86 - 0.95). Stanton and Wintour [10] analyzed smart phone-recorded jump performance with the application and when the results were compared with force plates, approximately 0.87 cm systematic bias, high correlation of ICC = 0.991 - 0.933, and high test-retest reliability of ICC = 0.997-0.998 are reported.

This study used a convenient, portable, and low cost platform type WBB and application for iPad Air 2 to measure and compare the flight time and the jump height during jump performances. In the traditional method of movement analyses, movements are video recorded with expensive equipment, restrictions in space, and it took a long time to obtain the results. However, current supply of smart devices and the development of applications that use various sensors allowed low cost, easy to use, and portable analyses methods. The application for iPad Air 2 that is used in this study can measure flight time and jump height easily and accurately with simple drags and clicks if take-off and landing criteria are clearly defined. With these advantages, the application for iPad Air 2 used in this study showed excellent intra-rater reliability (ICC = 0.967 for the flight time, ICC = 0.974 for the jump height) and inter-rater reliability (ICC = 0.985 for the flight time, ICC = 0.987 for the jump height). In addition, Bland-Altman plot and 95% LOA showed very high consistencies and symmetrical distribution within a small range.
CONCLUSIONS

Force platform based WBB records GRF during a vertical jump and it is analyzed for the flight time and the jump height. Vertical jump performance video recorded with a smart device was analyzed with an application for iPad Air 2 for flight time and jump height. The two methods implemented in this study measured the jump height based on the flight time method, and the two measured data showed high consistency when compared with each other. Specifically, jump height analysis that used iPad Air2 showed high consistency in intra-rater sessions and high reliability in inter-rater sessions. Based on these results, a jump analysis application for iPad Air2 may be utilized not only in the field of sports, but also for rehabilitation assessment after injury.

REFERENCES


www.balticsportscience.com
