Validity and test-retest reliability of the jumpo app for jump performance measurement

Amilton Vieira
Anthony J. Blazevich
Edith Cowan University
Alexandre Souza da Costa
James J. Tufano
Martim Bottaro

Follow this and additional works at: https://ro.ecu.edu.au/ecuworkspost2013

Part of the Sports Sciences Commons

This Journal Article is posted at Research Online. https://ro.ecu.edu.au/ecuworkspost2013/10447
Validity and Test-retest Reliability of the Jumpo App for Jump Performance Measurement

AMILTON VIEIRA†1, ANTHONY J. BLAZEVICH‡2, ALEXANDRE SOUZA DA COSTA*1, JAMES J. TUFANO‡3, and MARTIM BOTTARO†1

1Strength and Conditioning Laboratory, Faculty of Physical Education, University of Brasilia (UnB), Brasilia, DF, Brazil; 2Centre for Exercise and Sports Science Research (CESSR), School of Medical and Health Sciences, Edith Cowan University, Joondalup, WA, Australia; 3Faculty of Physical Education and Sport, Charles University in Prague, Prague, Czech Republic.

*Denotes undergraduate student author, †Denotes professional author

ABSTRACT

International Journal of Exercise Science 14(7): 677-686, 2021. The vertical jump test is one of the simplest and most prevalent physical tests used in practice and research. This study investigated the validity and reliability of a new mobile application (Jumpo) for measuring jump performance on Android devices. University-aged students (n = 10; 20 ± 3 years; 176 ± 6 cm; 68 ± 9 kg) reported to the laboratory on three occasions (2-7 days apart): to be familiarized with the jump performance measurements and then for test-retest reliability assessments. Participants performed countermovement jumps (CMJ), squat jumps (SJ), and right and left single-legged jumps in random order on a force platform while being recorded by a smartphone’s slow-motion camera. Flight time was selected as the criterion variable. Strong positive correlations between the Jumpo and force platform were observed for each jump type tested (r ≥ 0.93), although the flight times obtained with the Jumpo App were systematically shorter than those provided by the force platform by 3-6% (p < 0.001). The Jumpo App demonstrated a high test-retest reliability (ICC ≥ 0.94, CV ≤ 3.7%) with no differences between the coefficients of variation obtained from the Jumpo App and force platform (p ≥ 0.25). With respect to jump type, data from double-legged jumps (CMJ and SJ) were more accurately measured than data from single-legged jumps. The Jumpo App provides a valid and reliable measurement of jump performance, but the following equation should be used to calibrate its flight time results, allowing comparisons to be made to force platform data: Force platform flight time = 0.948 × Jumpo flight time + 41.515. Future studies should cross-validate the calibration equation in a different sample of individuals.

KEY WORDS: Android, mobile application, countermovement jump, squat jump, single-legged jump.

INTRODUCTION

The vertical jump test is one of the simplest and most prevalent physical tests used in practice and research, providing valuable information relating to neuromuscular capabilities (20). From vertical jump performance, several useful variables can be obtained by mathematical
calculations (1, 2, 4, 17, 22), but flight time is the simplest and most straightforward variable that can be directly obtained during a vertical jump using various equipment (2, 17).

Force platforms are considered to be the ‘gold-standard’ equipment for assessing jump performance and are among the most frequently used biomechanical apparatuses in sport science and in research (20). However, the high cost of laboratory-grade force platforms (upwards of $10,000 to $30,000 USD) (17) has led coaches and practitioners to seek out less expensive technologies. Many of these technologies, including contact mats (5, 18), accelerometer-based systems (16), infrared photocell technologies (6), and smartphone applications (apps) (3), measure flight time and then compute other variables such as jump height, force, velocity, and power (22). Since these devices are designed to be used in both practice and research, studies have investigated their validity and reliability. For instance, previous researchers have reported an almost perfect relationship ($r^2 = 0.99$) between jumps assessed using a force platform and a commercial contact mat (18), whereas others have noted that when using an accelerometer-based system, measurements obtained during jumping should not be used due their poor agreement with force platform-derived measurements (16). These results highlight the need for new technologies to undergo rigorous validity and reliability testing prior to use.

Due to the increase in smartphone popularity and affordability, mobile apps are becoming more commonplace in sport science research and practice. For example, the inexpensive ($10 USD) MyJump mobile app has been successfully implemented in sport science research and practice, and its validity and reliability has been repeatedly confirmed for use on iOS devices (3, 10). However, the vast majority of devices worldwide (~80% in 2018) use an Android operating system (11) and, to date, no study has been conducted using an Android operating device. To address this level of inaccessibility, a new mobile app for Android systems that can assess jump performance across a variety of jumps has been recently developed at the University of Brasilia, but its reliability and validity have not been assessed. Therefore, the aim of the present study was to investigate the concurrent validity and test-retest reliability of the new mobile application (called Jumpo) for measuring jump performance. Given that a number of vertical jump variations are commonly used in practice and in research, the validity and reliability was assessed across single- and double-leg countermovement (CMJ) and squat jumps (SJ).

**METHODS**

**Participants**

A sample of 10 university-aged men (20 ± 3 years, 176 ± 6 cm, 68 ± 9 kg) participated in this study since results from a pilot testing revealed that a high coefficient of validity (i.e. $r > 0.9$) could be reached (14). They were classified as physically active by IPAQ questionnaire (i.e. 2,340 ± 1,132; ranging from 815 to 3,879 MET-min·week$^{-1}$) (9). Participants were free from any surgery in past 12 months, free from any lower extremity injury or pain that could compromise jump performance, and reported as being right-limb dominant for kicking a ball. The study protocol was approved by the UDF – University Center Ethical Committee (number 2.878.364), and a written informed consent was obtained from each participant before participation. This research
was carried out fully in accordance with the ethical standards of the International Journal of Exercise Science (21).

Protocol

The participants visited the laboratory on three occasions (2-7 days apart). In the first visit, they signed the informed consent form and completed the IPAQ questionnaire. Anthropometric measurements were then taken, and the participants were familiarized with the jumping procedures. In the following two visits, they completed ~10 min of a standardized warm-up composed of traditional and ballistic overloaded back squats (50% body mass) in a Smith Machine followed by five vertical jumps of progressive effort (i.e. 20, 40, 60, 80 and 100% of the perceived maximum). Participants completed four repetitions of four different jump types: bilateral countermovement jump (CMJ), bilateral squat jump (SJ), and right (SLR) and left (SLL) single-leg countermovement jumps. The order of jumps for each participant were selected at random using an online number generator (www.randomization.com). Jumps were performed on a force platform (AMTI, Accupower Portable Force Plate, Watertown, MA, EUA) while being recorded by a slow-motion camera (240 fps) on a Samsung Galaxy S7 phone (model SM-G930F). Each jump (16 in total) was separated by 1 min of rest and the same order of the jumps was kept for the re-test session for each participant. Flight time was measured during CMJ, SJ, SLR, and SLL as it is a reliable metric of jump performance (13) that can be obtained from several feasible devices (6).

Participants performed each jump with hands akimbo and were instructed to jump as fast and as high as possible (i.e. to minimize total movement time and optimize jump height in an explosive manner). They were asked to keep their legs straight during the flight phase of the jump and to land in the same place keeping the same posture as at takeoff (i.e. knees extended with the ankles plantarflexed). For CMJ, participants were instructed to perform a downward (counter) movement followed by a vigorous hip, knee extension, and ankle plantar flexion with both lower limbs. To maintain ecological validity, they were free to determine the countermovement depth for each CMJ. For SJ, they were required to achieve a squat position with ~90º of knee and hip flexion (to determine participant reliability, this angle was not directly measured, as to maintain the ecological validity of SJ testing) and hold this position for 2 s before jumping. After each SJ, the force-time curve was inspected and if a counter-movement was present, the SJ trial was repeated. For SLR and SLL, participants were instructed to keep their trunk in an upright posture and the contralateral knee bent while performing a downward movement followed by a vigorous upward triple-extensions movement using a single lower limb (right or left), much like the CMJ procedure.

Jumbo is an app available in the Play Store (Google Inc., USA). The app was developed (Android Studio, version 2.3) to calculate the flight time as an indicator of jump performance by identifying the take-off and the landing frames from the video; these moments can be determined by moving the recorded video frame-by-frame. To record the jumps, a tripod (~30 cm high) with the smartphone was positioned on the ground 1.5 m from the anterior aspect of the participant’s feet. An observer with no previous experience of video analysis was asked to use the Jumbo App and select the last frame in which at least one foot was on the ground (i.e.
take-off) and subsequently the first frame where at least one foot contacted the ground (i.e. landing). This procedure demonstrated good consistency in our pilot testing. The force platform (101 × 76 cm) recorded data at a sampling frequency of 1000 Hz and was used to measure the flight time of the jumps at the same time as they were being recorded by the smartphone. The force platform was connected to a PC equipped with the software to analyze the force data (AccuPower 2.0.3 Dickinson, ND, USA). The commercial software has a default detection threshold of 54.3 N to identify take-off and landing times and the raw data (i.e. non-filtered) were used in the further analysis.

**Statistical Analysis**

Several analyses were conducted to determine the concurrent validity and test-retest reliability of the Jumpo App and force platform. To calculate the concurrent validity, paired student t-test, Cohen’s d effect size ($d$), and the bivariate Pearson product moment correlation coefficient ($r$) were applied to the data. In addition, Bland–Altman plots were created to represent the agreement between the two instruments (19). A regression analysis was performed to determine a simple linear equation for calibration using the flight times from the Jumpo App and from the force plate, inclusive of all jump types (15). To analyze the test-retest reliability of both devices, a paired student t-test, intraclass correlation coefficient (ICC 2,k), change in percentage ($\Delta\%$), coefficient of variation (CV), and Cohen’s $d$ effect size were used. For Cohen’s $d$, values of < 0.20, 0.20-0.49, 0.50-0.79, and > 0.80 were classified as trivial, small, moderate, and large, respectively. The level of statistical significance was set at $p < 0.05$. Calculations were performed using IBM SPSS Statistics 22 for Windows (IBM Co., USA) and a custom-made spreadsheet (Microsoft Excel).

**RESULTS**

**Validity:** A Pearson's product-moment correlation was run to assess the relationship between flight time obtained from the Jumpo App and force platform. Three hundred and twenty-two jumps were compared (10 individuals × 16 jumps × 2 days + 2 extra jumps). Preliminary analyses showed the relationship to be linear with both variables normally distributed with no outliers, as assessed by the Shapiro-Wilk test ($p > 0.05$). A strong positive correlation was observed between the Jumpo App and force platform for each of the vertical jump types tested ($r = 0.93-0.96$), and a higher value was subsequently observed when data from all 322 jumps were considered ($r = 0.98; \ p < 0.001$; Figure 1). This strong relationship between both devices allows this study to present a calibration equation that can be used to more accurately transfer the flight time values from the app to flight time values that would have occurred on the force plate for the same jump: force platform flight time = 0.948 × Jumpo flight time + 41.515.
Figure 1. Relationship between the flight times obtained using the Jumbo app and force platform. Correlations for each jump are presented. Numbers between parentheses describe the number of jumps that were compared for each jump. CMJ = countermovement jump, SJ = squat jump, SLR = right single-legged jump, and SLL = left single-legged jump.

Bland-Altman plots depicted the amount of bias and respective 95% limits of agreement for each of the vertical jump tested (Figure 2). Although systematic error (i.e. bias) can be observed in all the tested jumps, they were relatively low (ranging from 3 to 6%). Bland-Altman plots also reveal heteroscedasticity of the errors between the Jumbo App and force platform for SLL only ($r^2 = 0.428$, Figure 2D), but not for the remaining jump types ($r^2 \leq 0.015$; Figure 2A, B and C).

Figure 2. Bland-Altman plots depicting the bias and 95% limits of agreement for each jump between the Jumbo App and force platform. Heteroscedasticity is depicted by the continuous trend line. (A) Countermovement jump, (B) Squat jump, (C) Right single-legged jump, and (D) Left single-legged jump.
Table 1. Comparison of flight times (ms) between the Jumpo app and force platform.

<table>
<thead>
<tr>
<th>Vertical Jump</th>
<th>Jumbo App</th>
<th>Jumbo App</th>
<th>Force Platform</th>
<th>p-value</th>
<th>d (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMJ</td>
<td>530 ± 47</td>
<td>554 ± 45</td>
<td>547 ± 45</td>
<td>0.343</td>
<td>0.16 (-0.29, 0.60)</td>
</tr>
<tr>
<td>SJ</td>
<td>479 ± 53</td>
<td>504 ± 51</td>
<td>496 ± 55</td>
<td>0.386</td>
<td>0.15 (-0.29, 0.59)</td>
</tr>
<tr>
<td>SLR</td>
<td>348 ± 46</td>
<td>378 ± 44</td>
<td>370 ± 44</td>
<td>0.289</td>
<td>0.18 (-0.25, 0.62)</td>
</tr>
<tr>
<td>SLL</td>
<td>343 ± 45</td>
<td>373 ± 43</td>
<td>365 ± 36</td>
<td>0.226</td>
<td>0.20 (-0.23, 0.64)</td>
</tr>
</tbody>
</table>

Note: Flight time data in milliseconds are presented as mean ± standard deviation. CMJ = countermovement jump, SJ = squat jump, SLR = right single-legged jump, SLL = left single-legged jump. d = effect size, 95% CI, 95% confidence interval.

Test-retest reliability: Table 2 shows the test-retest reliability of flight time measurements from the Jumbo App and force platform. Both devices demonstrated high reliability (i.e. ICC ≥ 0.94, CV ≤ 3.7%), and there were no differences between CVs (p ≥ 0.25). Only the force platform value obtained from SLR was less on day 2 compared to day 1 (p = 0.02), however the magnitude of the difference was small (-3.3%, d = -0.28).

Table 2. Test-retest reliability of flight time measurements from the Jumbo app and force platform.

<table>
<thead>
<tr>
<th>Vertical Jump</th>
<th>Device</th>
<th>Day 1 (ms)</th>
<th>Day 2 (ms)</th>
<th>Δ% (95% CI)</th>
<th>ICC (95% CI)</th>
<th>CV% (95% CI)</th>
<th>p-value</th>
<th>d (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMJ</td>
<td>Jumbo app</td>
<td>530</td>
<td>527</td>
<td>-0.4</td>
<td>0.99</td>
<td>1.2</td>
<td>0.51</td>
<td>-0.05</td>
</tr>
<tr>
<td></td>
<td>Force platform</td>
<td>± 49</td>
<td>± 45</td>
<td>(-2.1, 1.2)</td>
<td>(0.94, 1.00)</td>
<td>(0.3, 2.0)</td>
<td></td>
<td>(-0.93, 0.82)</td>
</tr>
<tr>
<td>SJ</td>
<td>Jumbo app</td>
<td>476</td>
<td>489</td>
<td>2.7</td>
<td>0.95</td>
<td>2.6</td>
<td>0.47</td>
<td>-0.95, 0.81</td>
</tr>
<tr>
<td></td>
<td>Force platform</td>
<td>547</td>
<td>544</td>
<td>-0.6</td>
<td>0.98</td>
<td>1.4</td>
<td></td>
<td>-0.07</td>
</tr>
<tr>
<td>SLR</td>
<td>Jumbo app</td>
<td>497</td>
<td>501</td>
<td>-1.3, 2.9</td>
<td>(0.93, 1.00)</td>
<td>(0.6, 2.5)</td>
<td>0.44</td>
<td>-0.81, 0.95</td>
</tr>
<tr>
<td></td>
<td>Force platform</td>
<td>± 56</td>
<td>± 55</td>
<td>-2.0</td>
<td>0.96</td>
<td>3.0</td>
<td></td>
<td>-0.16</td>
</tr>
<tr>
<td>SLL</td>
<td>Jumbo app</td>
<td>353</td>
<td>346</td>
<td>-0.8</td>
<td>0.94</td>
<td>3.7</td>
<td>0.66</td>
<td>-0.95, 0.81</td>
</tr>
<tr>
<td></td>
<td>Force platform</td>
<td>± 47</td>
<td>± 44</td>
<td>-0.2</td>
<td>0.97</td>
<td>2.3</td>
<td></td>
<td>-0.10</td>
</tr>
<tr>
<td></td>
<td>± 49</td>
<td>± 42</td>
<td>(-6.0, 4.3)</td>
<td>(0.78, 0.99)</td>
<td>(1.1, 6.3)</td>
<td>0.42</td>
<td>(-0.97, 0.78)</td>
<td></td>
</tr>
</tbody>
</table>

Note: Flight time data in milliseconds are presented as mean ± standard deviation. CMJ = countermovement jump, SJ = squat jump, SLR = right single-legged jump, SLL = left single-legged jump. Δ% = change in percentage, 95% CI = 95% confidence interval, ICC = intraclass correlation coefficient, CV, coefficient of variation; d = effect size.

DISCUSSION

The aim of the present study was to quantify the concurrent validity and reliability of a new Android-based mobile application for measuring jump performance. The Jumbo App provided a valid measure of jump performance since its measures of flight time were strongly associated with flight time measures obtained from a force platform (i.e. external criteria), although a small systematic error was observed that was similar in magnitude to that reported in studies using other apps (3). The strong relationship between Jumbo and the force platform allowed for the calculation of a calibration equation to account for the small systematic error (15). Thus, although the Jumbo app can be confidently used on its own and without the need of a force
plate, for small inaccuracies (compared to force platform data), the following equation can be used to correct the Jumpo flight time, with 96.5% confidence of determination: Force platform flight time = 0.9484 × Jumpo flight time + 41.515. Data from the Jumpo App were also highly reliable (ICC 0.94 - 0.99 and CV ≤ 3.7%) and comparable to data obtained from the force platform (Table 2). Based on these results, the Jumpo App appears to provide a valid and reliable method for measuring jump performance (i.e. flight time) on Android devices (phone or tablet).

The present data support the validity of the app, however the Jumpo slightly underestimated flight time. The error was considered small for CMJ, SJ, and SLR, but moderate for SLL. This suggests that any direct comparison between Jumpo and force platform data benefits from the use of the calibration equation. Notably, previous studies have also demonstrated statistically significant, although small, errors when comparing CMJ data obtained from another mobile app, the MyJump for iOS, to data obtained using a force platform (3, 7, 8, 23). These studies showed excellent scores of validity for CMJ “height” (e.g. Pearson’s correlations ranging 0.98 to 1.00), but with some systematic bias. Similar to the present study, most of these studies found a negative bias, i.e. flight time and the resultant jump “height” data from MyJump tended to be lower than that obtained using the force platform (7, 8, 23). For instance, one study reported a negative bias of -2.98 cm (95%LoA -7.09 to 1.13) when comparing MyJump to an AMTI AccuPower force platform (7), which is slightly greater than our estimated bias of 2.17 cm (-16.4 ms flight time). Another study also demonstrated a bias (3), but in the opposite direction (+1.2 cm, 95%LoA 0.1 to 2.3). Although there is not a straightforward reason for these discrepancies, factors such as a different detection threshold and/or data filtering might partly explain them. Additionally, the common force plate sampling frequency (1000 Hz) is greater than the frame rate of most mobile phone cameras (240 Hz in the present study). Hence, it is possible that the sensitivity of the timescale of data collection could also be partially responsible for the slight negative bias. Regardless of the direction of bias, it should be noted that bias is likely always present when using a mobile app or any other indirect measuring methods, but the amount of bias in the Jumpo App is small and, if necessary, can be corrected with a simple linear equation, as presented earlier. It would be of future interest to determine whether the systematic errors observed for other apps might also be amenable to calibration correction. In any case, it is also important to mention that there is a large variation in the camera quality among Android devices, and different frame rates may be observed. Considering the aforementioned points, a low frame rate (< 120 fps) should be avoided since it might not be sensitive enough to precisely capture the instance of take-off or landing, likely increasing the degree of measurement error.

Although jump height can be derived from flight time and is one of the most studied and reported jump-related variables, flight time is also a useful parameter that can be directly measured from force platforms and other devices such as smartphones, contact mats, and photocells. Furthermore, flight time is a direct measurement that is not prone to additional error that might be incurred when the variable is used in subsequent calculations, which can occur when estimating jump height. As such, flight time has been shown to provide more reliable data than jump height estimated by the impulse-momentum method (ICC .99 vs .88 and CV 2.4 vs 5.4%) or by converting flight time to jump height, which can increase data variability (CV from 2.4 to 4.7%) (13). Therefore, considering the purpose of the present study, flight time was the
criterion variable analyzed in order to reduce the probability of introducing additional inaccuracies due to converting data from both devices. It is important to note that estimates of jump height (and other variables) from flight time always carry error because incorrect assumptions are made with respect to the relation between flight time and jump height (e.g. that the body’s center of mass at take-off and landing is equivalent to that during standing). That is, flight “distance” is estimated from flight time, which is very different (~10 – 12 cm) than the true jump height, although this error can be reduced by including additional measurements in more complex equations [e.g. see Wade et al., 2020 (24)].

In addition to being the first study to investigate the validity and reliability of an Android-compatible app, to the best of the authors’ knowledge, this is the first study describing the concurrent validity of single-legged vertical jump performance using any mobile app on any mobile platform. It could be observed that SLR (-22.1 ms, Δ% = 6.0) and SLL (-22.0 ms, Δ% = 6.0) produce greater errors than the double-leg CMJ (-16.4 ms, Δ% = 3.0) and SJ (-17.8 ms, Δ% = 3.6) tests. Furthermore, SLL was the only test that demonstrated non-uniformity of error, where the magnitude of the error increased as SLL performance (i.e. flight time) decreased (Figure 2D). This probably results from a lower signal-to-noise ratio due a poor jump technique applied by individuals, who were not trained jumpers, when using their non-dominant limb. We cannot rule out that some participants might have landed slightly behind or in front of where they took-off, introducing parallax error to the measure. For these reasons, data from the non-dominant limb should be interpreted with caution. Thus, future studies might impose an additional familiarization session and/or allow more than four trials for single-legged jumps; however, the methods used in the present study represent what might most likely be used in many research and applied testing scenarios, and these errors would possibly also occur in other apps, since the same flight time detection method is used. Moreover, it is worth mentioning that the observer had no previous experience of video analysis, as most “everyday users” would likely not be trained in video analysis. This highlights the Jumpo App’s ease of use.

Regardless of the jump type, there was an acceptable level of reliability (ICC ≥ 0.90 and CV < 10%), with CMJ demonstrating the highest scores (ICC = 0.99 and CV = 1.2%) and SLL the lowest (ICC = 0.94 and CV = 3.7). These results are consistent with a previous study showing that CMJs produce higher reliability scores (ICC = 0.95) than other jump types like SJ (ICC = 0.90) and drop jumps (ICC = 0.87) (12). Also, another study reported very high test-retest reliability scores for CMJ “height” using the MyJump app (ICC = 0.97 and CV = 6.7%), while an inertial measurement unit (VERT, USA) might not have reached an acceptable level of reliability (CV = 10.7%) (7). Taken together, these results demonstrate that the Jumpo App provides reliable test-retest scores, which are similar, or superior, to the MyJump and other apps and technologies. However, it is worth mentioning that the present study included only physically active college students and further testing in other populations (e.g. athletes, elderly, and sedentary individuals) is warranted before use. In addition, a future study, conducted by a fully independent laboratory, should cross-validate the calibration equation in a different sample of individuals.
A new Android-based mobile app (Jumpo App) for measuring jump performance (flight time) provided a valid and reliable measurement of jump performance. Furthermore, the following equation allows force platform and Jumpo App results to be used interchangeably: 

\[
\text{force platform}_{\text{flight time}} = 0.948 \times \text{Jumpo}_{\text{flight time}} + 41.515.
\]

Despite the positive results reported herein, future studies are required to cross-validate the present findings, and particularly calibration equation, in a different sample of individuals.

ACKNOWLEDGEMENTS

This study was supported, in part, by a university scholarship for undergraduate students that was provided by PIBEX - DEX/UnB.

REFERENCES


