Routledge
Taylor & Francis Group

Measurement in Physical Education and Exercise Science

ISSN: 1091-367X (Print) 1532-7841 (Online) Journal homepage: www.tandfonline.com/journals/hmpe20

Force-Time Analysis of the Drop Jump: Reliability of Jump Measures and Calculation Methods for Measuring Jump Height

Huseyin Celik, Ahmet Yildirim, Evrim Unver, Caner Mavili, Ekrem Yilmaz, Ferhat Ozturk, Pinar Arpinar-Avsar & Sukru Alpan Cinemre

To cite this article: Huseyin Celik, Ahmet Yildirim, Evrim Unver, Caner Mavili, Ekrem Yilmaz, Ferhat Ozturk, Pinar Arpinar-Avsar & Sukru Alpan Cinemre (2024) Force-Time Analysis of the Drop Jump: Reliability of Jump Measures and Calculation Methods for Measuring Jump Height, Measurement in Physical Education and Exercise Science, 28:2, 119-132, DOI: 10.1080/1091367X.2023.2257169

To link to this article: https://doi.org/10.1080/1091367X.2023.2257169

MEASUREMENT IN PHYSICAL EDUCATION AND EXERCISE SCIENCE
2024, VOL. 28, NO. 2, 119–132
https://doi.org/10.1080/1091367X.2023.2257169

Force-Time Analysis of the Drop Jump: Reliability of Jump Measures and Calculation Methods for Measuring Jump Height

Huseyin Celik[a], Ahmet Yildirim[b], Evrim Unver[c], Caner Mavili[c], Ekrem Yilmaz[c,d], Ferhat Ozturk[e], Pinar Arpinar-Avsar[a], and Sukru Alpan Cinemre[c]

[a]Department of Biomechanics and Motor Control, Faculty of Sport Sciences, Hacettepe University, Ankara, Turkey; [b]Department of Physical Education and Sports, Faculty of Education, Middle East Technical University, Ankara, Turkey; [c]Department of Exercise and Sport Sciences, Faculty of Sport Sciences, Hacettepe University, Ankara, Turkey; [d]Department of Coaching Education, Faculty of Sport Sciences, Kirsehir Ahi Evran University, Kirsehir, Turkey; [e]Department of Physiotherapy and Rehabilitation, Faculty of Physical Therapy and Rehabilitation, Hacettepe University, Ankara, Turkey

ABSTRACT

The purpose was to evaluate reliability of measurements during drop jump (DJ) and magnitude of differences in vertical jump height (VJH) when calculated using two variables: time-in-the-air and center-of-mass velocity. Thirty-seven handball players performed three single-leg and double-leg DJs on a portable force plate during two sessions. Sixteen outcome measures and four sets of reliability metrics, including intraclass correlation coefficients (ICCs), were estimated. All outcome measures, except vertical time-to-stabilization, yielded high (0.70–0.89) to very high (0.90–0.96) interday and intraday ICCs for the single-leg and double-leg DJ tests. These results indicated that the single-leg and double-leg DJ tests can be considered reliable for short-term and long-term monitoring of collegiate handball players of both sexes. In addition, the single-leg DJ VJH calculated using these two variables differed in magnitude (mean difference in test measurements: 0.64 cm, $p < .001$, effect size: 0.165). Therefore, we recommend using the same method to calculate single-leg VJH for long-term monitoring.

KEYWORDS

vertical jump test; vertical jump height; handball; single and double leg drop jump; athlete monitoring

Introduction

Muscle exercises are classified into static and dynamic types (Knuttgen & Komi, 2003). Plyometrics is a dynamic type of exercise included in training programs to improve speed, endurance, reactive strength, and explosive power in athletes (e.g., Chu & Meyer, 2013). Plyometric exercise consists of an eccentric stretch of the muscle-tendon unit (MTU) immediately followed by concentric shortening of the MTU. This combination of MTU undergoing active stretching followed by rapid shortening during the stretch-shortening cycle (SSC) is the basis of plyometric exercise, which seeks to exploit the force-potentiating capabilities of the SSC (Norman & Komi, 1979). Different types of jumps, such as tuck jump, single leg hop, and depth or drop jump, are commonly used in plyometric exercise programs to improve vertical jump performance (Chu & Meyer, 2013).

One of the most popular vertical jump exercises in plyometric training programs is drop jump (DJ) (Bobbert, 1990). The DJ consists of standing on a box, stepping off and dropping to the ground, and immediately performing a vertical jump (concentric shortening or propulsion phase) after landing (eccentric stretching or braking phase) (Komi, 2003). One of the proposed mechanisms responsible for the force potentiation observed in SSC is the recovery of elastic potential energy (e.g., Blazevich, 2011). The elastic potential energy is stored in the MTU during the eccentric stretch phase, and this stored elastic energy is released during the concentric shortening phase, with MTU loading being the dominant factor affecting the energy storage (Blazevich, 2011). Baxter et al. (2021) evaluated the Achilles tendon loading profiles of various exercises and proposed that single and double leg drop jumps are tiers 3 and 4 (the highest tier), respectively. This research suggests that single and double leg drop jumps could be evaluated separately as the load profiles are different, yet little research has been conducted on single leg DJ compared to double leg DJ.

Several previous studies have investigated the double leg DJ and compared it to other vertical jumps using biomechanical instruments such as electromyography systems, motion capture cameras, and force plates (e.g., Bobbert et al., 1987; Healy et al., 2014; Ishikawa et al., 2006). In one of the first studies on the subject, Komi and Bosco (1978)

CONTACT Huseyin Celik huseyincelik@hacettepe.edu.tr Department of Biomechanics and Motor Control, Faculty of Sport Sciences, Hacettepe Universitesi, Spor Bilimleri Fakultesi, Beytepe, Cankaya, Ankara 06800, Turkiye

examined the vertical jump performance of men and women for three types of jumps: squat jump (SJ), countermovement jump (CMJ), and DJ. In many subsequent studies, vertical jump performance has been extensively studied from the perspective of biomechanical outcomes, neuromuscular control, and SSC performance variables using whole-body biomechanics (e.g., Harry et al., 2021; McMahon et al., 2021). To this end, several measures, such as jump height, ground contact time, and reactive strength index, were calculated from the force-time curve variables recorded by a force plate during vertical jumps (McMahon et al., 2021). These measures, if obtained in a reproducible manner, have the potential to be useful for profiling and benchmarking, injury prevention, rehabilitation, and return-to-sports perspectives and can be used to evaluate the effectiveness and impact of vertical jump exercises in training programs.

In order to be able to identify meaningful changes for both short-term and long-term monitoring needs, reliable and repeatable test protocols are critical to the creation of a performance assessment system (French & Torres Ronda, 2022). Exercise scientists evaluating performance must ensure that the procedures they use do not contain errors greater than the differences they are trying to assess and that the outcome measures are reliable enough to be relied upon (Kibele, 1998). Thus, the jump data obtained from these tests must be reproducible and consistent. In other words, the subsequent measurements under identical conditions should yield virtually the same results as the first measurement. Reproducibility is usually quantified using reliability metrics by implementing a test-retest procedure in which the first set of data obtained in the first measurements is compared with the data obtained from the same subjects in subsequent measurements at different times under identical conditions (Vincent & Weir, 2021). Reliability metrics such as the intraclass correlation coefficient (ICC) and standard error of measurement are often used to quantify test-retest reliability in exercise and sports science (Hopkins, 2000; Weir, 2005).

However, the evidence of reliability is often unclear due to common problems in reporting reliability (Morrow & Jackson, 1993). The methodology used to obtain the outcome measures and reliability metrics may not be defined in detail in published studies (Liljequist et al., 2019). In addition, studies aimed at establishing the reliability of a test or measure should have at least 30 participants who are representative of the defined population (e.g., age, sex), as relatively small sample sizes may lead to an overestimation of reliability (Morrow & Jackson, 1993). As noted in a recent study (Harry et al., 2021), most of the data generated on vertical jumps came from studies with only males. This situation is not limited to jumping studies, as it has been demonstrated quantitatively in sports and exercise science and medicine research in general (Cowan et al., 2023; Cowley et al., 2021). Specifically, Cowley et al. (2021) examined the sex distribution of approximately 12 million participants in 5261 articles published in the field of sports and exercise science; 63% of the publications presented data on both male and female participants, 31% were male only, and only 6% were female (approximately 8.2 million male and 4.2 million female participants in total). The results of this study (Cowley et al., 2021) clearly show that women are significantly under-represented in sports and exercise science research. Cowley et al. (2021) discuss that there are two ways to address the gender disparity in sport and exercise science and medicine research: (i) increase research on female-only participants on topics related to female-specific factors that may influence women's health and performance and (ii) increase women's participation in research that applies to both sexes, despite the fact that women's participation rates in sport are lower than men's at both the recreational and elite levels.

Despite the paucity of data on women, scientific papers, technical notes, and technological improvements in the measurement, training, and analysis methods for vertical jumps are accumulating (e.g., Comfort et al., 2019). With new technological improvements, portable instruments, such as portable force plates, have been used in jump testing with a field-testing approach (e.g., Uzelac-Sciran et al., 2020). However, there is limited research investigating the reliability of outcome measures of DJ tests measured with portable force plates and even less for single leg DJs. Previous research has mostly focused on double leg DJs measured using a stationary force plate in fewer than 30 male subjects. For example, Flanagan et al. (2008) used a stationary force plate to examine the intraday reliability of the outcome measures of the double leg DJ test in 22 subjects. The authors found that the jump height, ground contact time, and reactive strength index were highly reliable from trial to trial, as evidenced by the very high ICC values (0.977–0.989). Conversely, the time to stabilization was found to be unreliable from trial to trial, as evidenced by the moderate ICC value (0.687). However, arm position was not controlled throughout the jumping and landing movements, which could explain this result. In addition, the sex of the participants was not explicitly reported. In a more recent study, Byrne et al. (2017) examined the interday reliability of the reactive strength index and optimal drop height using an optical system for measurement.

The subjects were 19 male hurling players. The reactive strength index and optimal drop height measurements showed high reliability with ICC values of 0.87 and 0.80 and coefficient of variation values of 4.20% and 2.98%, respectively. However, this study was conducted only on male participants. In addition, an optical system rather than a force plate system was used for measurements. Furthermore, the ICC calculations and models were not reported in detail, as in many other studies (Liljequist et al., 2019).

The choice of method, parameters, and variables used in the calculations has the potential to affect the reliability and magnitude of DJ outcome measures. For example, a recent study showed that changing the thresholds for detecting the takeoff instance affected the reliability and magnitude of countermovement jump outcome measures (Pérez-Castilla et al., 2021). In one of the earliest studies of vertical jump height calculation, Linthorne (2001) showed that vertical jump height could be calculated in two ways using a force plate and two different variables: time in the air or vertical velocity of the center of mass at takeoff. Then, Moir (2008) calculated the vertical jump height during the countermovement jump test using these two variables and pointed out that the results, although consistent within themselves, may differ from each other. However, previous studies have mostly focused on countermovement jumps or double leg DJs (e.g., Moir, 2008; Wank & Coenning, 2019; Xu et al., 2023 for a systematic review on this topic). Therefore, the effects of the method, parameter, and variable choices in the calculation of DJ outcome measures, especially for single leg DJs, are understudied.

As vertical jumping is considered a central and fundamental element of many sports, vertical jumping ability is one of the factors that determine performance in a variety of sports (Aragón, 2000; Bobbert, 1990). Handball, a contact team sport that involves high-speed and short-duration activities, such as jumping, is no exception. Due to the nature of handball, players' biomechanical performance, neuromuscular control, and SSC function have been emphasized. Therefore, monitoring variables related to these aspects using the outcome measures of the DJ tests could be beneficial. In addition, the similarity of single-leg jumps to some fundamental movements in handball suggests that it may be worthwhile to study both single and double leg DJ.

The purpose of this experimental study was to test the hypothesis that outcome measures of single and double leg DJ tests, such as ground contact time, vertical stiffness, and reactive strength index, obtained from 37 male and female handball players using a portable force plate, would provide reliable measurements from trial to trial and from day to day. To this end, a repeated-measures experimental design was implemented in which participants performed three single leg and three double leg DJs from a fixed height in a non-fatigued state. The experimental measurements were repeated after one week, and 16 outcome measures were obtained. The intraday and interday reliability of each outcome measure were statistically assessed using reliability metrics: ICC, standard error of measurement, coefficient of variation, and minimal metrically detectable change. The secondary objective of this study was to evaluate the reliability and magnitude differences in DJ vertical jump height when it was calculated using two methods based on these two variables, namely, the time in the air and the vertical velocity of the center of mass at takeoff. We hypothesized that the DJ vertical jump height measures calculated using the two methods based on these two variables would not differ in terms of magnitude and level of reliability.

Materials and methods

Subjects

Thirty-seven volunteers (22 males and 15 females) participated in the experiments (age:20.6 ± 2.2 years, height:177.0 ± 8.5 cm, weight:69.9 ± 13.5 kg) with no neuromusculoskeletal problems that would affect their jumping performance at the time of the measurements. Participants were members of the university handball team and in season at the time of data collection. Although we could not achieve an equal number of male and female subjects due to the lower participation of females in the university handball team, we combined the data as a cohort because the research questions of our study are not gender specific but apply to both sexes.

This study was conducted at Ankara in December 2022. All participants had experience in plyometric exercises but did not perform such exercises at the time of the study. Participants had not performed any strength training 48 h prior to data collection. All participants provided written informed consent to participate in the study. This study was reviewed and approved by the Human Research Ethics Committee of METU. The authors did not have access to any information that could identify individual participants during or after data collection.

For sample size estimation, we conducted an a priori statistical power analysis for correlation based on the correlation coefficient of jump height magnitudes between trials and between measurements. Portney and Watkins (2015) suggested that since the correlation coefficient is a unitless measure, the effect size measure does not require adjustment and is simply the value of correlation coefficient. Based on previously published and obtained jump height data in the DJ experiments, we expected a correlation coefficient of $r = 0.8$ between the trials and between the measurements. With a correlation coefficient of $r = 0.80$ and power = 0.90, the predicted sample size required was 11 participants (Portney & Watkins, 2015). However, since a minimum of 30 participants is required to establish the reliability of a test or measure (Morrow & Jackson, 1993), 37 members of the university handball team were recruited. Previous studies on the reliability of the outcome measures of the drop jump test had 22, 18, 13, and 19 participants in the studies by Flanagan et al. (2008), Lloyd et al. (2009), Markwick et al. (2015), and Byrne et al. (2017), respectively.

Instrumentation

During the DJs, the ground reaction force in the vertical direction (VGRF) was recorded at 2000 Hz with a portable single force plate (Kistler 9260AA, Switzerland)) for the sample period (10 s). The computer software interface used was Kistler's MARS software. The same force plate has previously been used in the evaluation of vertical jumps with three jumps: SJ, CMJ, and single-leg CMJ (Kozinc & Šarabon, 2022). The raw data of each trial were stored on the same laptop and exported as delimiter-separated values to the MATLAB environment for the calculation of outcome measures and reliability metrics.

Procedures

One week prior to testing, the athletes completed a familiarization session, in which they had the opportunity to practice single and double DJs. The recorded DJ tests were then administered by the same raters in two sessions, one week apart, at the same time of day (from 18:00 h–20:00 h), at the same indoor venue where the collegiate athletes train on a level sports court with a synthetic hard floor. Before each measurement session, the athletes were informed of the study procedures. Participants wore the same shoes and clothing for testing and were instructed to refrain from drinking (except water), eating, and exercising for one hour prior to testing. The DJ tests were performed in two different conditions: double leg and single leg; first, three trials of double leg, then three trials of single leg to provide participants with a gradual increase in neuromuscular load (Lloyd et al., 2009).

Prior to the measurements, the subjects were given a verbal description and visual demonstration of the single and double leg DJs. The instructions were as follows: (i) step forward from the box of 30 cm without stepping down and drop onto the force plate, (ii) when first landing, perform a rebound jump as high and as fast as possible, think of the force plate as a hot plate, (iii) when touching down after the rebound jump (i.e., the second landing), stick your landing and stabilize as quickly as possible while looking straight ahead and remaining motionless by simply looking at a cross sign on the wall approximately 4 m in front of you, (iv) keep your hands on your hips throughout the jump, and (v) do not tuck your legs in the air (Byrne et al., 2017; Dalleau et al., 2004; Flanagan et al., 2008). In both DJ tests, participants stepped forward from the box with their preferred foot, which was established in the familiarization session, and remained the same in subsequent sessions (Maloney et al., 2016).

The warm-up procedure consisted of five minutes of jogging followed by dynamic stretching of each major muscle group of the lower limbs (Flanagan et al., 2008). After the warm-up, participants were given the opportunity to practice both jumps. The participants were then allowed to rest for five minutes before the DJ tests. The participants performed six maximal DJs (three double legs and three single legs). The rest period between jumps was approximately one minute. If the participant failed to stabilize during single leg DJ, the trial was repeated. Data were successfully recorded for both jumps at a drop height of 30 cm for all participants.

The rationale for using a 30 cm box was to accommodate the administration of the tests. Peng (2011) found that the intensity of double leg DJ increased with increasing drop height and drop heights greater than 40 cm did not provide any benefits in terms of mechanical efficiency and leg stiffness. In addition, the mechanical load on the Achilles tendon during single leg DJ is higher than that of double leg DJ (Baxter et al., 2021). We wanted to select a drop height that would accommodate each subject in the cohort for both single leg and double leg DJ, yet at a height that would be challenging to assess reactive strength while trying to minimize the risk of injury. In the familiarization session, subjects tried different drop heights, i.e., 20, 30, and 40 cm. Some subjects experienced balance problems after landing on one leg at 40 cm. At a drop height of 30 cm, all subjects successfully performed single and double leg DJ. It is also common to use 30 cm in the literature (Flanagan et al., 2008; Fransz et al., 2015; Healy et al., 2014).

Data analysis

The data analysis consisted of two stages: (i) detection of key time points, such as landing and touchdown instants and (ii) calculation of outcome measures of DJs. Both the single and double leg DJ performances of the participants were analyzed using the very same code based on the methods described by Baca (1999). Figure 1 shows an example of a VGRF recording of a double-leg DJ with key time points of the jump. We used the same numbering scheme as Baca (1999). Specifically, t2 was identified as the instant of landing on the force plate by locating the data point where the VGRF value first exceeded the 10 N threshold (Lloyd et al., 2009; Pérez-Castilla et al., 2021) after the start of the recording. Similarly, t5 was identified as the instant of landing on the force plate by locating the data point where the VGRF value first exceeded the 10 N threshold after the maximal vertical jump. The instant of takeoff, t4, was identified as the data point where the VGRF value first fell below the 10 N threshold after instant t2. t6 is the last point of the force recording at which the participant stood still on the force plate. The vertical velocity of the total body center of mass (BCOM) at the landing instant (t2) was estimated by numerically integrating the vertical force minus the body weight (BW = m × g, where m is the body mass, and g is the gravitational acceleration, 9.80665 m/s2) from t6 to t2 using the trapezoidal method (Baca, 1999).

$$V_L = -\int_{t2}^{t6} (F_z(t) - mg)dt/m \tag{1}$$

The t3 was the instant at which the vertical velocity of the BCOM is zero and was located by numerically solving the following equation (Baca, 1999).

$$V_L + \int_{t2}^{t3} (F_z(t) - mg)dt/m = 0 \tag{2}$$

The t3 instant divides the ground contact phase (from t2 to t4) into two phases: eccentric or braking (from t2 to t3) and concentric or propulsion (from t3 to t4) (Baca, 1999; Komi, 2003). The t1 instant was when the participant took off from the box, yet we did not locate that instant, as Baca (1999) used a second force plate for that purpose.

After locating the key time points, we calculated several outcome measures for DJs. Kinetic analysis of vertical jumping using force plates enables extensive assessment of neuromuscular function in the eccentric (braking) and concentric (propulsion) phases of SSC separately and jointly, allowing researchers and practitioners to go beyond testing only simple measures of jump height (Bishop et al., 2021). Recently, Bishop et al. (2021) published a framework for selecting outcome measures during the DJ tests. The authors suggested reactive strength index and its individual components, jump height, ground contact time, and vertical stiffness. Other outcome measures, such as time to stabilization, peak maximal VGRF, peak BCOM displacement, rate of force development, mean braking and propulsion force, braking, and propulsion phase time, have also been reported in previous studies (Flanagan et al., 2008; Maloney et al., 2017; McMahon et al., 2021; Šarabon

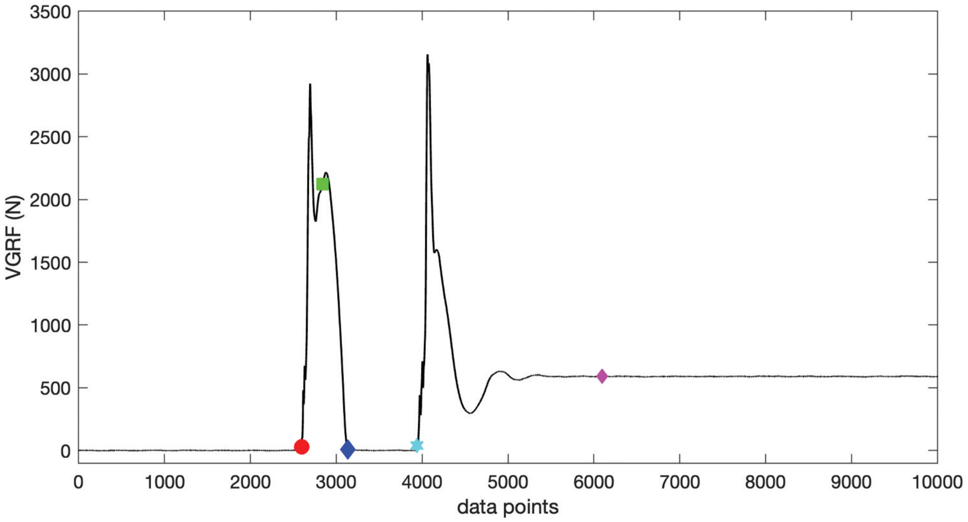

Figure 1. Exemplary force-time curve of vertical ground reaction force (VGRF) during a double leg drop jump. t2 (red circle): landing onto the portable force plate, t3 (green square): the instant where the vertical velocity of the total body center of mass is zero; t4 (blue diamond): the instant of take-off, t5 (cyan hexagram): the instant of touch-down, t6 (black circle): the last point of force record at which the participant stands still on the force plate, vertical time to stabilization (magenta diamond), defined below, from t2 to t3 is the eccentric or braking phase, from t3 to t4 concentric or propulsion phase.

et al., 2020). In his recent book, Cleather (2021) argued that impulse is the most important outcome measure in explosive physical performance. Therefore, we included impulse measures in our analysis.

We calculated the following outcome measures from the force-time curves obtained from the force plate.

- **Jump height** (JH in cm): JH was the peak displacement attained by the BCOM during the flight phase (from t4 to t5 in Figure 1) and was calculated using the flight time in the air (JH-TIA) and the vertical velocity of the BCOM at the takeoff instant (JH-TOV) (Linthorne, 2001; Moir, 2008). As filtering had been shown to introduce errors in JH calculations (Harry et al., 2022; Pinto & Callaghan, 2021), unfiltered data were used in the JH calculations.
- **Ground contact time** (GCT in s): GCT was the time spent on the ground after landing until takeoff (from t2 to t4 in Figure 1) and was used to classify SSC as fast (100-250 ms) and slow (>250 ms) (Schmidtbleicher, 1992).
- **Reactive strength index** (RSI in cm/s): RSI was used to quantify SSC performance and was calculated as the ratio of JH to GCT (Flanagan & Comyns, 2008).
- **Braking phase time** (BPT in s): BPT was the time spent on the ground in the eccentric braking phase and was calculated as the time between the landing instant and the following first instant of zero vertical velocity of the BCOM (from t2 to t3 in Figure 1) (McMahon et al., 2021).
- **Propulsion phase time** (PPT in s): PPT was the time spent on the ground in the concentric propulsion phase and was calculated as the time between the first instant of zero vertical velocity of the BCOM after landing and takeoff (from t3 to t4 in Figure 1) (McMahon et al., 2021).
- **Mean braking force** (MBF in BW): The ability to tolerate high braking forces was related to reactive strength and was calculated as the average of each vertical force data point during the braking phase (from t2 to t3 in Figure 1) (McMahon et al., 2021).
- **Mean propulsion force** (MPF in BW): The ability to exert high propulsive forces was associated with vertical jump performance (Garhammer & Gregor, 1992) and was calculated as the average of each vertical force data point during the propulsion phase (from t3 to t4 in Figure 1) (McMahon et al., 2021).
- **Peak maximal ground reaction force** (PGRF in BW): PGRF was the largest VGRF encountered during the jump and was calculated as the global maximum value of the VGRF data points in the contact phase (braking + propulsion, from t2 to t4 in Figure 1) (Lloyd et al., 2009).
- **Peak center of mass displacement** (PCOMD in cm): PCOMD was the maximum vertical displacement during the jump and was calculated as the maximum displacement value of the vertical displacement data points during the contact phase (from t2 to t4 in Figure 1) after the instant of landing onto the force plate. The vertical displacement during the contact phase was calculated from the double integration of the vertical acceleration of BCOM with respect to time (from t2 to t4 in Figure 1) with integration constants for velocity as $V(t=t2) = V_L$ and for position as $Y(t=t2) = 0$ (arbitrary) (Baca, 1999).
- **Rate force development** (RFD in BW/s): RFD has been used to assess explosive strength in athletes and has been proposed to be determined by the ability to exert maximal voluntary effort in the early phase of an explosive movement (Maffiuletti et al., 2016). RFD was calculated as the peak value of the time derivative of the filtered vertical force-time curve in the braking phase (from t2 to t3 in Figure 1). The vertical force signal was filtered only for the RFD analysis using a second-order low-pass zero-lag Butterworth filter with a cutoff frequency of 5 Hz (Sarabon et al., 2020).
- **Vertical jump impulse** (VJI in BW*s): The vertical impulse applied during the jump was equal to the change in momentum (the momentum of an object is equal to its mass times its velocity). The impulse-momentum relationship requires that the impulse or total VGRF applied to the BCOM be directly proportional to the change in vertical velocity (Cleather, 2021). VJI has been used to assess performance (Cleather, 2021) and deficits prior to return to sport following injury (Costley et al., 2021). The VJI was calculated by integrating the vertical force-time curve during the contact phase of the jump (from t2 to t4 in Figure 1) and dividing the result by the BW.
- **Absolute vertical stiffness** (Abs Kvert in kN/m): Abs Kvert describes the vertical displacement of BCOM in response to VGRF during a movement (Latash & Zatsiorsky, 1993) and has been proposed as a representative outcome measure of summative lower limb stiffness (Maloney & Fletcher, 2021). Absolute Kvert was calculated as the ratio of the peak VGRF during the contact phase of the jump (from t2 to t4 in Figure 1) and the displacement of the BCOM at the time of peak VGRF (McMahon & Cheng, 1990).

- **Normalized vertical stiffness** (Norm Kvert in N/m/kg): Norm Kvert has been associated with injury risk and athletic performance (Brazier et al., 2019; Butler et al., 2003) and was calculated by dividing the absolute Kvert value by body mass (Maloney et al., 2016). The logic behind this normalization is that vertical stiffness has been suggested to be influenced by body size (Farley et al., 1993).
- **Peak propulsion power** (PPP in W/kg): PPP was defined as the maximal power generated during the propulsion phase of the jump (Bishop et al., 2020). The ability to generate PPP was associated with success in explosive activities such as jumping (Haff & Stone, 2015). PPP was calculated by first obtaining the power-time curve as the dot product of vertical force and velocity (e.g., Laurson et al., 2022), and then locating the global maximum during the propulsion phase (from t3 to t4 in Figure 1). PPP values were reported after normalization to body mass.
- **Vertical time to stabilization** (VTTS in s): VTSS is an outcome measure of jumps and has been linked to neuromuscular control, which involves sensory and mechanical systems during landing from a jump (Flanagan et al., 2008; Wikstrom et al., 2004). Several methods can be used to calculate VTTS (Fransz et al., 2015). In this study, a method known as sequential estimation was used (Colby et al., 1999). Specifically, the sequential average was calculated as the cumulative average of VGRF data points after touchdown (t5 in Figure 1) by successively adding one data point at a time, then calculating the average of the first two data points, then the average of the first three data points, and so on. The VGRF was considered stable, and the VTSS was located when the sequential average remained within one quarter of the standard deviation of the VGRF in the three-second post-touchdown interval (Colby et al., 1999).

Statistical analysis

For day-to-day, test-retest, or interday reliability analyses, the means of three trials of the DJ outcome measures in single and double leg tests were used. We performed a paired t-test on the difference between the means of the outcome measures obtained in the test and retest sessions to test for the absence of systematic bias (Atkinson & Nevill, 1998). The effect sizes (ES) were also estimated (the calculation details are given below). The alpha level was set at 0.05 for all statistical analyses. For trial-to-trial or intraday reliability analyses, only three trials of the DJ outcome measures from the retest session were used.

To estimate relative reliability, we used the two-way random effects model of ICC, which in simple terms compared the within-subject variability of the measurement data with the between-subject variability of the measurement data (McGraw & Wong, 1996; Shrout & Fleiss, 1979). The test-retest and trial-to-trial correlations were quantified using two forms of ICC, denoted by C and A, referring to the consistency and absolute agreement models: consistency ICC(C,k) and absolute agreement ICC(A,k) (Liljequist et al., 2019; McGraw & Wong, 1996).

$$ICC(C,k) = \frac{MS_R - MS_E}{MS_R} \tag{3}$$

$$ICC(A,k) = \frac{MS_R - MS_E}{MS_R + \frac{MS_C - MS_E}{n}} \tag{4}$$

where MS_R, MS_E, MS_C, and n are the mean square of the rows (i.e., subjects), the mean square error, the mean square of the columns (i.e., trials or measurements), and the number of subjects, respectively. ICC (C,k) reflects the degree of consistency for outcome measures that are averages of k independent trials or measurements, whereas ICC(A,k) reflects the degree of absolute agreement for outcome measures that are averages based on k independent trials or measurements. The conceptual difference between the two models lies in the definition of the ICC denominator (as seen in Equations (3) and (4)), i.e. the column variance (MS_C) is excluded from the denominator variance in the ICC(C,k) model, but is included in the ICC (A,k) model. Thus, ICC(C,k) accounts for random error, while ICC(A,k) accounts for both random and systematic error. Then, if the two models of ICC give similar values to each other, the systematic differences between trials and days would be considered absent (Liljequist et al., 2019; McGraw & Wong, 1996).

For each ICC value, the lower and upper bounds of the 95% confidence intervals (CIs) were also calculated using the equations presented in Table 7 of McGraw and Wong (1996). ICC is a real number, usually between 0 (no reliability) and 1 (perfect reliability) (Liljequist et al., 2019). According to Munro's classification, the strength of the correlation coefficients was interpreted as follows to describe the degree of relative reliability, 0.00–0.25: little or no correlation, 0.26–0.49: low correlation, 0.50–0.69: moderate correlation, 0.70–0.89: high correlation, and 0.90–1.00: very high correlation (Carter & Lubinsky, 2016).

To estimate the absolute reliability, we calculated three reliability metrics. The first was the standard error of measurement (SEM), which was calculated as the square root of the mean square error term from repeated-measures analysis of variance (Atkinson & Nevill, 1998). The SEM was expressed in the actual units of the outcome measures; therefore, it is practical to interpret it as the smaller the SEM, the more reliable the measurement (Atkinson & Nevill, 1998). The second absolute measure of reliability was the minimal metrically detectable change (MMDC), the amount of change that could be considered different between two measurements, calculated as the 95% CI of the SEM of the outcome measures (i.e., ±1.96 SEM). The coefficient of variation (CV), which is the ratio of the standard deviation (SD) to the mean of the data, was also calculated to assess the absolute reliability of the outcome measures. Calculating the CV allows us to compare the absolute reliability between variables that have different units of measurement. This was performed by calculating the mean CV value from the individual CV values (Atkinson & Nevill, 1998). For trial-to-trial CV calculations, the second and third trials were included. The following equations were used to calculate the CV metric (n=number of subjects; $x1$ and $x2$ are measurements for the ith subject) (Shechtman, 2013).

$$SD_i = \sqrt{(x_1 - x_2)^2/2} \tag{5}$$

$$\bar{x}_i = (x_1 + x_2)/2 \tag{6}$$

$$CV_i = \frac{SD_i}{\bar{x}_i} \tag{7}$$

$$CV(\%) = 100 \times \frac{\sum CV_i}{n} \tag{8}$$

Additional statistical analyses were performed to compare the magnitude of the differences in DJ vertical jump height when calculated using the two methods based on two variables (i.e., time in the air and vertical velocity of the center of mass at takeoff). Specifically, a dependent samples t-test was performed on the vertical jump height data by comparing the means of JH-TIA and JH-TOV. Separate statistical analyses were performed for test and retest measurements. In addition, ES were calculated using Hedges' g with correction for small sample sizes and interpreted as minimal effect (0.20), small effect (0.20–0.50), moderate effect (0.50–0.80), and large effect (>0.80) (Cohen, 1988). Statistical analyses were performed using MATLAB Statistics and Machine Learning Toolbox version 12.4 (R2022b). The significance level was set at $p < .05$.

Results

Table 1 presents the means and SDs of 16 outcome measures of the single and double leg DJ tests for test and retest measurements. Tables 2 and 3 present the absolute agreement and consistency ICCs and their 95% CIs (in parentheses in Tables 2 and 3), SEM, MMDC, and CV metrics of the outcome measures of the single and double leg DJ tests, respectively.

Table 1. Descriptive data for outcome measures during the single and double leg drop jump tests in the test and retest measurements separated by one week.

	Double leg		Single leg	
	Test Mean (SD)	Retest mean (SD)	Test Mean (SD)	Retest mean (SD)
JH-TIA	20.68 (4.69)	20.20 (5.42)	9.98 (3.70)	9.97 (3.40)
JH-TOV	20.66 (4.89)	20.21 (5.62)	9.34 (4.06)	9.53 (3.45)
GCT	0.27 (0.04)	0.27 (0.04)	0.34 (0.05)	0.34 (0.05)
RSI	0.80 (0.25)	0.77 (0.24)	0.31 (0.16)	0.30 (0.11)
BPT	0.12 (0.02)	0.13 (0.02)	0.15 (0.02)	0.15 (0.02)
PPT	0.14 (0.02)	0.15 (0.02)	0.19 (0.03)	0.19 (0.03)
MBF	2.83 (0.37)	2.78 (0.35)	2.29 (0.27)	2.24 (0.25)
MPF	2.48 (0.31)	2.43 (0.32)	1.75 (0.27)	1.75 (0.19)
PGRF	4.72 (0.86)	4.70 (0.85)	3.73 (0.55)	3.65 (0.55)
PCOMD	1.74 (2.41)	18.03 (2.47)	16.51 (1.91)	16.47 (1.86)
RFD	36.91 (5.11)	36.63 (4.79)	29.98 (3.97)	29.41 (3.70)
VJI	0.69 (0.04)	0.69 (0.05)	0.65 (0.04)	0.65 (0.05)
Abs Kvert	24.67 (8.88)	24.52 (8.73)	24.16 (6.90)	23.80 (7.15)
Norm Kvert	350.01 (115.81)	344.45 (103.31)	345.53 (103.29)	339.93 (108.34)
PPP	51.00 (10.38)	49.87 (10.54)	27.81 (8.18)	27.60 (5.51)
VTTS	1.06 (0.09)	1.07 (0.08)	1.13 (0.08)	1.12 (0.10)

SD: standard deviation (in parenthesis), JH-TIA(cm): jump height from flight time in the air, JH-TOV(cm): jump height from takeoff velocity, GCT(s): ground contact time, RSI(cm/s): reactive strength index, BPT (s): braking phase time, PPT(s): propulsion phase time, MBF (body weight(BW)): mean braking force, MPF(BW): mean propulsion force, PGRF(BW): peak maximal ground reaction force, PCOMD(cm): peak center of mass displacement, RFD(BW/s): rate force development, VJI(BW*s): vertical jump impulse, Abs Kvert(kN/m): absolute vertical stiffness, Norm Kvert(N/m/kg): normalized vertical stiffness, PPP(W/kg): peak propulsion power, VTTS(s): vertical time to stabilization.

Table 2. Reliability metrics of outcome measures during single leg drop jump test.

	Interday					Intraday				
	ICC (A,k)	ICC (C,k)	SEM	MMDC	CV%	ICC (A,k)	ICC (C,k)	SEM	MMDC	CV%
JH-TIA	0.76 (0.53,0.88)	0.75 (0.52,0.87)	2.23	4.38	14.82	0.86 (0.71,0.94)	0.86 (0.71,0.94)	1.65	3.23	9.96
JH-TOV	0.77 (0.56,0.88)	0.77 (0.55,0.88)	2.30	4.51	16.06	0.85 (0.69,0.94)	0.85 (0.69,0.94)	1.71	3.36	10.32
GCT	0.90 (0.80,0.95)	0.90 (0.80,0.95)	0.02	0.04	5.40	0.94 (0.88,0.98)	0.94 (0.87,0.97)	0.02	0.04	4.84
RSI	0.70 (0.40,0.84)	0.70 (0.39,0.84)	0.09	0.18	16.07	0.86 (0.72,0.94)	0.87 (0.72,0.94)	0.05	0.10	10.63
BPT	0.83 (0.68,0.91)	0.83 (0.67,0.91)	0.01	0.02	6.51	0.96 (0.91,0.98)	0.96 (0.91,0.98)	0.01	0.02	5.38
PPT	0.90 (0.80,0.95)	0.90 (0.80,0.95)	0.01	0.03	5.68	0.90 (0.80,0.96)	0.90 (0.79,0.96)	0.01	0.02	5.45
MBF	0.82 (0.66,0.91)	0.83 (0.67,0.91)	0.14	0.27	4.14	0.94 (0.88,0.98)	0.94 (0.88,0.98)	0.11	0.21	3.08
MPF	0.78 (0.57,0.89)	0.77 (0.56,0.88)	0.14	0.27	4.84	0.90 (0.79,0.96)	0.90 (0.79,0.96)	0.08	0.15	2.55
PGRF	0.79 (0.59,0.89)	0.79 (0.59,0.89)	0.33	0.64	6.39	0.90 (0.80,0.96)	0.90 (0.79,0.96)	0.33	0.64	5.76
PCOMD	0.79 (0.58,0.89)	0.78 (0.58,0.89)	1.13	2.21	5.35	0.83 (0.64,0.93)	0.83 (0.63,0.93)	1.36	2.66	6.91
RFD	0.86 (0.72,0.93)	0.86 (0.73,0.93)	1.91	3.73	4.39	0.93 (0.85,0.97)	0.93 (0.85,0.97)	1.81	3.54	4.29
VJI	0.82 (0.65,0.91)	0.82 (0.65,0.91)	0.03	0.05	3.20	0.92 (0.83,0.97)	0.91 (0.82 0.96)	0.03	0.05	3.23
Abs Kvert	0.92 (0.84,0.96)	0.91 (0.83,0.96)	2.78	5.46	9.57	0.94 (0.88,0.98)	0.94 (0.88,0.97)	3.51	6.88	11.76
Norm Kvert	0.91 (0.83,0.96)	0.91 (0.83,0.96)	42.34	83.00	9.60	0.94 (0.88,0.98)	0.94 (0.88,0.98)	52.16	102.23	11.69
PPP	0.74 (0.49,0.87)	0.73 (0.48,0.86)	4.53	8.88	7.68	0.89 (0.78,0.95)	0.89 (0.78,0.95)	2.40	4.70	5.42
VTTS	0.66 (0.34,0.83)	0.66 (0.33,0.82)	0.06	0.13	4.53	0.88 (0.75,0.95)	0.89 (0.77,0.95)	0.05	0.09	3.09

ICC(A,k): absolute agreement ICC, ICC(C,k): consistency ICC, LB: lower bound (in parenthesis), UB: upper bound (in parenthesis), SEM: standard error of measurement, MMDC: minimal metrically detectable change, CV: coefficient of variation, JH-TIA(cm): jump height from flight time in the air, JH-TOV(cm): jump height from takeoff velocity, GCT(s): ground contact time, RSI(cm/s): reactive strength index, BPT (s): braking phase time, PPT(s): propulsion phase time, MBF (body weight(BW)): mean braking force, MPF(BW): mean propulsion force, PGRF(BW): peak maximal ground reaction force, PCOMD(cm): peak center of mass displacement, RFD(BW/s): rate force development, VJI(BW*s): vertical jump impulse, Abs Kvert(kN/m): absolute vertical stiffness, Norm Kvert(N/m/kg): normalized vertical stiffness, PPP(W/kg): peak propulsion power, VTTS(s): vertical time to stabilization.

Table 3. Reliability metrics of outcome measures during double leg drop jump test.

	Interday					Intraday				
	ICC (A,k)	ICC (C,k)	SEM	MMDC	CV%	ICC (A,k)	ICC (C,k)	SEM	MMDC	CV%
JH-TIA	0.87 (0.74,0.93)	0.87 (0.74,0.93)	2.47	4.84	10.40	0.89 (0.77,0.95)	0.89 (0.77,0.95)	2.22	4.36	7.60
JH-TOV	0.87 (0.75,0.93)	0.87 (0.75,0.93)	2.52	4.95	10.90	0.88 (0.75,0.95)	0.88 (0.74,0.95)	2.48	4.86	8.45
GCT	0.80 (0.62,0.90)	0.81 (0.63,0.90)	0.02	0.05	6.78	0.84 (0.65,0.93)	0.83 (0.64,0.93)	0.03	0.05	7.28
RSI	0.88 (0.77,0.94)	0.89 (0.78,0.94)	0.11	0.21	11.69	0.89 (0.76,0.95)	0.88 (0.75,0.95)	0.13	0.25	13.41
BPT	0.81 (0.63,0.90)	0.81 (0.63,0.90)	0.01	0.02	7.71	0.83 (0.65,0.93)	0.83 (0.65,0.93)	0.01	0.03	9.32
PPT	0.80 (0.61,0.90)	0.80 (0.61,0.90)	0.01	0.02	6.34	0.84 (0.67,0.93)	0.84 (0.66,0.93)	0.01	0.03	6.08
MBF	0.83 (0.67,0.91)	0.83 (0.67,0.91)	0.20	0.38	5.54	0.83 (0.65,0.93)	0.83 (0.65,0.93)	0.23	0.46	7.18
MPF	0.87 (0.75,0.93)	0.88 (0.76,0.94)	0.15	0.29	4.92	0.88 (0.76,0.95)	0.88 (0.75,0.95)	0.18	0.35	5.36
PGRF	0.87 (0.76,0.94)	0.87 (0.75,0.93)	0.41	0.80	6.81	0.78 (0.55,0.91)	0.81 (0.60,0.92)	0.59	1.16	9.42
PCOMD	0.76 (0.53,0.88)	0.76 (0.53,0.88)	1.52	2.99	6.73	0.82 (0.62,0.92)	0.82 (0.61,0.92)	1.48	2.91	6.61
RFD	0.88 (0.77,0.94)	0.88 (0.77,0.94)	2.28	4.47	5.11	0.82 (0.63,0.92)	0.83 (0.64,0.93)	2.98	5.83	7.29
VJI	0.72 (0.46,0.86)	0.72 (0.45,0.85)	0.03	0.06	3.34	0.84 (0.67,0.93)	0.84 (0.66,0.93)	0.03	0.05	2.48
Abs Kvert	0.82 (0.65,0.91)	0.82 (0.65,0.91)	4.88	9.56	12.81	0.89 (0.76,0.95)	0.89 (0.77,0.95)	5.11	10.02	11.99
Norm Kvert	0.83 (0.67,0.91)	0.83 (0.66,0.91)	59.64	116.89	13.01	0.83 (0.64,0.93)	0.84 (0.66,0.93)	72.63	142.36	12.01
PPP	0.90 (0.81,0.95)	0.90 (0.81,0.95)	4.45	8.73	7.16	0.89 (0.77,0.95)	0.89 (0.77,0.95)	5.49	10.76	7.85
VTTS	0.72 (0.45,0.86)	0.72 (0.45,0.85)	0.06	0.11	3.70	0.89 (0.77,0.95)	0.89 (0.78,0.95)	0.05	0.10	4.35

ICC(A,k): absolute agreement ICC, ICC(C,k): consistency ICC, LB: lower bound (in parenthesis), UB: upper bound (in parenthesis), SEM: standard error of measurement, MMDC: minimal metrically detectable change, CV: coefficient of variation, JH-TIA(cm): jump height from flight time in the air, JH-TOV(cm): jump height from takeoff velocity, GCT(s): ground contact time, RSI(cm/s): reactive strength index, BPT (s): braking phase time, PPT(s): propulsion phase time, MBF (body weight(BW)): mean braking force, MPF(BW): mean propulsion force, PGRF(BW): peak maximal ground reaction force, PCOMD(cm): peak center of mass displacement, RFD(BW/s): rate force development, VJI(BW*s): vertical jump impulse, Abs Kvert(kN/m): absolute vertical stiffness, Norm Kvert(N/m/kg): normalized vertical stiffness, PPP(W/kg): peak propulsion power, VTTS(s): vertical time to stabilization.

Based on the results of paired t-tests, there were no significant differences between the test and retest means for any of the outcome measures of the single and double leg DJ tests ($p > .05$ for each comparison, ES: 0.003 (JH-TIA) – 0.206 (MBF) for double leg, ES: 0.016 (Abs Kvert) – 0.169 (PPT) for single leg). This result would demonstrate the absence of any systematic bias in the DJ tests from measurement to measurement.

For single leg DJ tests (Table 2), the interday absolute agreement ICC values of four outcome measures, namely, GCT, PPT, Abs Kvert, and Norm Kvert, could be considered a very high level of reliability according to Munro's classification. The interday absolute agreement ICC values of the 11 outcome measures, that is, JH-TIA, JH-TOV, RSI, BPT, MBF, MPF, PGRF, PCOMD, RFD, VJI, and PPP, ranged from 0.70 to 0.89, which could be considered a high level of reliability. In addition, the interday absolute agreement ICC values of only one outcome measure, VTTS, ranged from 0.50 to 0.69, which could be considered a moderate level of reliability. In terms of absolute reliability, the lowest CV level (i.e., the most absolute reliable) was observed for VJI of 3.20%.

For double leg DJ tests (Table 3), the interday absolute agreement ICC value of one outcome measure, PPP, was 0.90, which could be considered a very high level of reliability according to Munro's classification. The interday absolute agreement ICC values of the other 15 outcome measures, for example, RSI, GCT, and Abs Kvert, ranged from 0.70 to 0.89, which could be considered a high level of reliability. In terms of absolute reliability, the lowest value was again for the VJI, based on a CV level of 3.34%.

In terms of the intraday absolute agreement ICC and single leg DJ (Table 2), very high reliability was evidenced for GCT, BPT, PPT, MBF, MPF, PGRF, RFD, VJI, Abs Kvert, and Norm Kvert, with absolute agreement ICC values of 0.90–0.96 and CV levels of 2.55–11.76%. For all other measures of both single and double leg DJ, the absolute agreement ICC values were at the level of high reliability.

For the vertical jump height comparisons between the two methods, the DJ jump height outcome measures JH-TIA and JH-TOV were not significantly different for the double leg test (mean difference = 0.02 cm, $p = .901$, ES = 0.004 for test; and mean difference = 0.01 cm, $p = .944$, ES = 0.002 for retest measurements). In contrast, the DJ jump height measures JH-TIA and JH-TOV were significantly different from each other for the single leg test (mean difference = 0.64 cm, $p < .001$, ES = 0.165 for test; and mean difference = 0.44 cm, $p < .001$, ES = 0.129 for retest measurements).

Discussion

This reliability study was conducted to present the reliability metrics for the outcome measures of the single and double leg DJ tests. For the double leg DJ test, interday absolute agreement ICC values of all 16 outcome measures calculated in this study could be considered high to very high reliability in a group of collegiate athletes of both sexes with more than 30 participants. Similarly, for the single leg DJ test, the interday ICC values of 15 outcome measures (except VTTS) could be considered to have a high to very high level of reliability. Furthermore, for intraday reliability, all outcome measures for both single and double leg DJs had a high to very high level of reliability. According to our first hypothesis, we expected the outcome measures of the single and double leg DJ tests to be reliable from trial to trial and day to day. Our hypothesis was supported by the present results, as the ICC values for absolute agreement and consistency between and within days were high to very high (except for VTTS). Almost equal absolute agreement and consistency interday and intraday ICC values were observed for all 16 outcome measures calculated in this study, indicating the absence of bias or systematic error in DJ measurements with a portable force plate (Liljequist et al., 2019).

In general, the interday absolute agreement ICC values of outcome variables in the double leg DJ test were higher than those in the single leg DJ test (9 of 16). Conversely, the intraday absolute agreement ICC values of the outcome variables in the single leg DJ test were mostly greater than those in the double leg DJ test (11 of 16). The results of Baxter et al. (2021) suggest that single and double leg DJs are not identical exercises in terms of the Achilles tendon loading profiles. This implies that single and double leg DJs can be evaluated separately in terms of SSC function, as the load profiles are different. The results of our study showed that the interday and intraday ICC values of outcome measures could be different in single and double leg DJs, and it should not be assumed that the reliability of the double leg DJ is always superior to that of the single leg DJ (Maloney et al., 2018), which is more challenging in terms of Achilles tendon loading.

Among the outcome measures of the single and double leg DJ tests, the VTTS had the lowest interday ICC values. The VTTS or time to stabilization measures in any direction have been used to analyze dynamic postural stability (Ross et al., 2003). The assessment of dynamic stability using dynamic tests, such as single and double leg DJ, is particularly important to detect postural instability in athletes (Ross et al., 2003). Many methods have been developed to estimate the time required for stabilization (Fransz et al., 2015). Flanagan et al. (2008) assessed the reliability of the VTSS using a method in which the VTSS was located as the first instance when the VGRF reached and remained within a threshold for one second, and the analysis yielded an intraday ICC value that could be considered a moderate level of intraday reliability. To estimate the VTTS, we used the sequential estimation method, which was shown to be able to identify postural instability (Colby et al., 1999). For the sequential estimation method, the results of the current study with collegiate athletes showed ICC values that could be considered moderate interday reliability, high intraday reliability, and rather low CV values (CV < 10%). With such reliability metrics, the sequential estimation method can be considered an alternative to assess dynamic postural stability in collegiate athletes.

According to our second hypothesis, we expected that the DJ vertical jump height measures calculated by the two methods based on two variables – the time in the air and the vertical velocity of the center of mass at takeoff – would not differ in magnitude and level of reliability. When a jump mat is used to measure jumping performance (Markwick et al., 2015), the time-in-air method (i.e., JH-TIA) is used to calculate JH. However, when a force plate is used, either method can be used in calculations. Moir (2008) suggested that both methods are valid and consistent; however, when a force plate is available, practitioners or experimenters should use the JH-TOV for the countermovement jump. We wanted to determine if there was a difference in the magnitude and level of reliability that would allow one method to be preferred over the other for DJ testing. Our results partially supported our hypothesis, as the comparisons between the JH-TIA and JH-TOV showed no significant differences, with minimal effect sizes for the double leg DJ. In addition, the interday and intraday ICC values of vertical jump height calculated using the JH-TIA and JH-TOV were almost identical and at the same level (0.70–0.89, high correlation) for single and double leg DJ. In contrast, JH-TIA and JH-TOV for single leg DJ showed significant differences in magnitude, although the ES were minimal. This finding suggests that, although the results for JH-TIA and JH-TOV are reliable at the same level, they may differ in magnitude for single leg DJ. This finding also highlights the importance of using the same method without interchanging for the long-term monitoring of athletes. A similar recommendation was made by McMaster et al. (2021) when the researchers compared CMJ JH obtained from three measurement systems (force plate as the criterion method), and even almost perfect linear correlations between them, significant differences in JH magnitude were observed between measurement systems.

Our study has several limitations. We examined 16 outcome measures; however, there is an opportunity to extract more information from the DJ test, an instrumental jump test that has the potential to generate reliable and rich datasets to study biomechanical performance, neuromuscular control, and SSC function. Future studies may consider adding more outcome measures, possibly combining data from other sources, such as electromyography, three-dimensional motion capture, and modeling and simulation. In addition, the choice of method, parameters, and variables used in the calculations have the potential to affect the reliability and magnitude of DJ outcome measures. We examined the effects of the choice of method for calculating vertical jump height, but there are other choices that may affect the reliability and magnitude of the outcome measures. For example, a recent study showed that changing the thresholds used to detect the takeoff instance affects the reliability and magnitude of countermovement jump outcome measures (Pérez-Castilla et al., 2021). Additional reliability studies could examine this issue using single and double leg DJ tests.

This study presents a comprehensive assessment of reliability statistics for single and double leg DJ tests and provides several implications for future research and practice. Exercise and sports scientists and practitioners could implement methods similar to those presented in this study to better understand measurement quality in terms of reliability specific to their own environment and setup, thereby improving the application and explanation of vertical jump test results. The statistics presented in this paper will also inform exercise and sports scientists and practitioners when selecting outcome measures for DJ tests to assess athlete performance and neuromuscular status. Finally, the reliability statistics would guide the expected amount of typical noise in DJ testing and the amount of minimal detectable change that could be considered meaningful seasonal variability in the long-term monitoring of athletes in cycles of competition and training seasons.

Conclusion

In conclusion, the unique contribution of this study is the comprehensive assessment of the reliability of measurements during single and double leg DJ and the magnitude of differences in VJH when calculated using two variables: time in the air and center of mass velocity with a cohort of 37 participants of both sexes. Overall, our force-time analysis in this study showed that the outcome measures of the single and double leg DJ tests can be considered reliable from trial to trial and day to day. The DJ outcome measures proposed by Bishop et al. (2021), namely RSI and its individual components JH and GCT and normalized Kvert, yielded high to very high interday and intraday ICC values for the DJ tests. The other 12 outcome measures could also be used in various combinations, depending on the needs of the assessment. Single leg DJ vertical jump height measures calculated by the two methods based on two variables, time in the air and vertical velocity of the center of mass at takeoff, would differ in magnitude. Based on this, we recommend using the same method without interchanging to calculate vertical jump height from measurement to measurement. To conclude, the data provide evidence that DJ tests can be considered reliable for short- and long-term monitoring of biomechanical performance, neuromuscular control, and SSC function in collegiate handball players.

Acknowledgments

This research was partially supported by The Scientific and Technological Research Council of Turkey. The datasets generated and analyzed during the current study are not publicly available, but are available from the corresponding author upon request. The authors declare that there is no conflict of interest.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the Türkiye Bilimsel ve Teknolojik Araştırma Kurumu [219S813].

ORCID

Huseyin Celik http://orcid.org/0000-0001-8316-6468

References

Aragón, L. F. (2000). Evaluation of four vertical jump tests: Methodology, reliability, validity, and accuracy. *Measurement in Physical Education and Exercise Science*, *4*(4), 215–228. https://doi.org/10.1207/S15327841MPEE0404_2

Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine*, *26*(4), 217–238. https://doi.org/10.2165/00007256-199826040-00002

Baca, A. (1999). A comparison of methods for analyzing drop jump performance. *Medicine & Science in Sports & Exercise*, *31*(3), 437–442. https://doi.org/10.1097/00005768-199903000-00013

Baxter, J. R., Corrigan, P., Hullfish, T. J., O'Rourke, P., & Silbernagel, K. G. (2021). Exercise progression to incrementally load the achilles tendon. *Medicine and Science in Sports and Exercise*, *53*(1), 124–130. https://doi.org/10.1249/mss.0000000000002459

Bishop, C., Pereira, L. A., Reis, V. P., Read, P., Turner, A. N., & Loturco, I. (2020). Comparing the magnitude and direction of asymmetry during the squat, countermovement and drop jump tests in elite youth female soccer players. *Journal of Sports Sciences*, *38*(11–12), 1296–1303. https://doi.org/10.1080/02640414.2019.1649525

Bishop, C., Turner, A., Jordan, M., Harry, J., Loturco, I., Lake, J., & Comfort, P. (2021). A framework to guide practitioners for selecting metrics during the countermovement and drop jump tests. *Strength & Conditioning Journal*, *44*(4), 95–103. https://doi.org/10.1519/SSC.0000000000000677

Blazevich, A. J. (2011). The stretch–shortening cycle. In M. Cardinale, R. Newton, & K. Nosaka (Eds.), *Strength and conditioning: Biological principles and practical applications* (pp. 209–222). Wiley-Blackwell.

Bobbert, M. F. (1990). Drop jumping as a training method for jumping ability. *Sports Medicine*, *9*(1), 7–22. https://doi.org/10.2165/00007256-199009010-00002

Bobbert, M. F., Huijing, P. A., & van Ingen Schenau, G. J. (1987). Drop jumping. I. The influence of jumping technique on the biomechanics of jumping. *Medicine and Science in Sports and Exercise*, *19*(4), 332–338. https://doi.org/10.1249/00005768-198708000-00003

Brazier, J., Maloney, S., Bishop, C., Read, P. J., & Turner, A. N. (2019). Lower extremity stiffness: Considerations for testing, performance enhancement, and injury risk. *Journal of Strength and Conditioning Research*, *33*(4), 1156–1166. https://doi.org/10.1519/JSC.0000000000002283

Butler, R. J., Crowell, H. P., & Davis, I. M. (2003). Lower extremity stiffness: Implications for performance and injury. *Clinical Biomechanics*, *18*(6), 511–517. https://doi.org/10.1016/S0268-0033(03)00071-8

Byrne, D. J., Browne, D. T., Byrne, P. J., & Richardson, N. (2017). Interday reliability of the reactive strength index and optimal drop height. *Journal of Strength and Conditioning Research*, *31*(3), 721–726. https://doi.org/10.1519/JSC.0000000000001534

Carter, R. E., & Lubinsky, J. (2016). *Rehabilitation research: Principles and applications* (Fifth edition ed.). Elsevier.

Chu, D. A., & Meyer, G. C. (2013). *Plyometrics*. Human kinetics.

Cleather, D. (2021). *Force: The biomechanics of training*. KMA Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed ed.). L. Erlbaum Associates.

Colby, S. M., Hintermeister, R. A., Torry, M. R., & Steadman, J. R. (1999). Lower limb stability with ACL impairment. *Journal of Orthopaedic & Sports Physical Therapy*, *29*(8), 444–454. https://doi.org/10.2519/jospt.1999.29.8.444

Comfort, P., Jones, P., & McMahon, J. J. (Eds.). (2019). *Performance assessment in strength and conditioning*. Routledge, Taylor & Francis Group. https://doi.org/10.4324/9781315222813

Costley, J. A. E., Miles, J. J., King, E., & Daniels, K. A. J. (2021). Vertical jump impulse deficits persist from six to nine months after ACL reconstruction. *Sports Biomechanics*, *22*(1), 123–141. https://doi.org/10.1080/14763141.2021.1945137

Cowan, S. M., Kemp, J. L., Ardern, C. L., Thornton, J. S., Rio, E. K., Bruder, A. M., Mosler, A. B., Patterson, B., Haberfield, M., Roughead, E. A., Hart, H., To, L., Neufeld, S., Mazahir, N., & Crossley, K. M. (2023). Sport and exercise medicine/physiotherapy publishing has a gender/sex equity problem: We need action now! *British Journal of Sports Medicine*, *57*(7), 401–407. https://doi.org/10.1136/bjsports-2022-106055

Cowley, E. S., Olenick, A. A., McNulty, K. L., & Ross, E. Z. (2021). "Invisible sportswomen": The sex data gap in sport and exercise science research. *Women in Sport & Physical Activity Journal*, *1*(aop), 1–6. https://doi.org/10.1123/wspaj.2021-0028

Dalleau, G., Belli, A., Viale, F., Lacour, J.-R., & Bourdin, M. (2004). A simple method for field measurements of leg stiffness in hopping. *International Journal of Sports Medicine*, *25* (3), 170–176. https://doi.org/10.1055/s-2003-45252

Farley, C. T., Glasheen, J., & McMahon, T. A. (1993). Running springs: Speed and animal size. *Journal of Experimental Biology*, *185*(1), 71–86. https://doi.org/10.1242/jeb.185.1.71

Flanagan, E. P., & Comyns, T. M. (2008). The use of contact time and the reactive strength index to optimize fast stretch-shortening cycle training. *Strength & Conditioning Journal*, *30*(5), 32–38. https://doi.org/10.1519/SSC.0b013e318187e25b

Flanagan, E. P., Ebben, W. P., & Jensen, R. L. (2008). Reliability of the reactive strength index and time to stabilization during depth jumps. *The Journal of Strength & Conditioning Research*, *22*(5), 1677–1682. https://doi.org/10.1519/JSC.0b013e318182034b

Fransz, D. P., Huurnink, A., de Boode, V. A., Kingma, I., & van Dieën, J. H. (2015). Time to stabilization in single leg drop jump landings: An examination of calculation methods and assessment of differences in sample rate, filter settings and trial length on outcome values. *Gait & Posture*, *41*(1), 63–69. https://doi.org/10.1016/j.gaitpost.2014.08.018

Garhammer, J., & Gregor, R. (1992). Propulsion forces as a function of intensity for weightlifting and vertical jumping. *The Journal of Strength & Conditioning Research*, *6*(3), 129–134. https://doi.org/10.1519/00124278-199208000-00001

Haff, G. G., & Stone, M. H. (2015). Methods of developing power with special reference to football players. *Strength & Conditioning Journal*, *37*(6), 2–16. https://doi.org/10.1519/SSC.0000000000000153

Harry, J. R., Barker, L. A., Tinsley, G. M., Krzyszkowski, J., Chowning, L. D., McMahon, J. J., & Lake, J. (2021). Relationships among countermovement vertical jump performance metrics, strategy variables, and inter-limb asymmetry in females. *Sports Biomechanics*, *8*(8), 1–19. https://doi.org/10.1080/14763141.2021.1908412

Harry, J. R., Blinch, J., Barker, L. A., Krzyszkowski, J., & Chowning, L. (2022). Low-pass filter effects on metrics of countermovement vertical jump performance. *Journal of Strength and Conditioning Research*, *36*(5), 1459–1467. https://doi.org/10.1519/JSC.0000000000003611

Healy, R., Howard, R. M., Kenny, I. C., & Harrison, A. J. (2014). A comparison of methods to examine double and single leg drop jump performance. *Proceedings of the 2014 International Society of Biomechanics in Sports Conference*, July 12 – July 16, 2014, Johnson City, TN, USA.

Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine*, *30*(1), 1–15. https://doi.org/10.2165/00007256-200030010-00001

Ishikawa, M., Komi, P. V., Finni, T., & Kuitunen, S. (2006). Contribution of the tendinous tissue to force enhancement during stretch–shortening cycle exercise depends on the prestretch and concentric phase intensities. *Journal of Electromyography and Kinesiology*, *16*(5), 423–431. https://doi.org/10.1016/j.jelekin.2005.08.006

Kibele, A. (1998). Possibilities and limitations in the biomechanical analysis of countermovement jumps: A methodological study. *Journal of Applied Biomechanics*, *14*(1), 105–117. https://doi.org/10.1123/jab.14.1.105

Knuttgen, H. G., & Komi, P. V. (2003). Basic considerations for exercise. In P. V. Komi (Ed.), *Strength and power in sport* (2nd ed., Vol. 3, pp. 3–7). Wiley. https://doi.org/10.1002/9780470757215.ch1

Komi, P. V. (2003). Stretch–shortening cycle. In P. V. Komi (Ed.), *Strength and power in sport* (2nd, Vol. 3). Blackwell Science. https://doi.org/10.1002/9780470757215

Komi, P. V., & Bosco, C. (1978). Utilization of stored elastic energy in leg extensor muscles by men and women. *Medicine and Science in Sports*, *10*(4), 261–265. https://journals.lww.com/acsm-msse/toc/1978/10040

Kozinc, Ž., & Šarabon, N. (2022). Measurements of lower-limb isometric single-joint maximal voluntary torque and rate of torque development capacity offer limited insight into vertical jumping performance. *Measurement in Physical Education and Exercise Science*, *26*(1), 15–26. https://doi.org/10.1080/1091367X.2021.1937176

Latash, M. L., & Zatsiorsky, V. M. (1993). Joint stiffness: Myth or reality? *Human Movement Science*, *12*(6), 653–692. https://doi.org/10.1016/0167-9457(93)90010-M

Laurson, K. R., Baptista, F., Mahar, M. T., Welk, G. J., & Janz, K. F. (2022). Long jump, vertical jump, and vertical jump power reference curves for 10-18 year olds. *Measurement in Physical Education and Exercise Science*, *26*(4), 306–314. https://doi.org/10.1080/1091367X.2021.2017291

Liljequist, D., Elfving, B., Skavberg Roaldsen, K., & Chiacchio, F. (2019). Intraclass correlation – a discussion and demonstration of basic features. *PLOS ONE*, *14*(7), e0219854. https://doi.org/10.1371/journal.pone.0219854

Linthorne, N. P. (2001). Analysis of standing vertical jumps using a force platform. *American Journal of Physics*, *69*(11), 1198–1204. https://doi.org/10.1119/1.1397460

Lloyd, R. S., Oliver, J. L., Hughes, M. G., & Williams, C. A. (2009). Reliability and validity of field-based measures of leg stiffness and reactive strength index in youths. *Journal of Sports Sciences*, *27*(14), 1565–1573. https://doi.org/10.1080/02640410903311572

Maffiuletti, N. A., Aagaard, P., Blazevich, A. J., Folland, J., Tillin, N., & Duchateau, J. (2016). Rate of force development: Physiological and methodological considerations. *European Journal of Applied Physiology*, *116*(6), 1091–1116. https://doi.org/10.1007/s00421-016-3346-6

Maloney, S. J., & Fletcher, I. M. (2021). Lower limb stiffness testing in athletic performance: A critical review. *Sports Biomechanics*, *20*(1), 109–130. https://doi.org/10.1080/14763141.2018.1460395

Maloney, S. J., Fletcher, I. M., & Richards, J. (2016). A comparison of methods to determine bilateral asymmetries in vertical leg stiffness. *Journal of Sports Sciences*, *34* (9), 829–835. https://doi.org/10.1080/02640414.2015.1075055

Maloney, S. J., Richards, J., & Fletcher, I. M. (2018). A comparison of bilateral and unilateral drop jumping tasks in the assessment of vertical stiffness. *Journal of Applied Biomechanics*, *34*(3), 199–204. https://doi.org/10.1123/jab.2017-0094

Maloney, S. J., Richards, J., Nixon, D. G. D., Harvey, L. J., & Fletcher, I. M. (2017). Do stiffness and asymmetries predict change of direction performance? *Journal of Sports Sciences*, *35*(6), 547–556. https://doi.org/10.1080/02640414.2016.1179775

Markwick, W. J., Bird, S. P., Tufano, J. J., Seitz, L. B., & Haff, G. G. (2015). The intraday reliability of the reactive strength index calculated from a drop jump in professional men's basketball. *International Journal of Sports Physiology and Performance*, *10*(4), 482–488. https://doi.org/10.1123/ijspp.2014-0265

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*(1), 30–46. https://doi.org/10.1037/1082-989X.1.1.30

McMahon, T. A., & Cheng, G. C. (1990). The mechanics of running: How does stiffness couple with speed? *Journal of Biomechanics*, *23*, 65–78. https://doi.org/10.1016/0021-9290(90)90042-2

McMahon, J. J., Lake, J. P., Stratford, C., & Comfort, P. (2021). A proposed method for evaluating drop jump performance with one force platform. *Biomechanics*, *1*(2), 178–189. Article 2. https://doi.org/10.3390/biomechanics1020015

McMaster, D. T., Tavares, F., O'Donnell, S., & Driller, M. (2021). Validity of vertical jump measurement systems. *Measurement in Physical Education and Exercise Science*, *25*(2), 95–100. https://doi.org/10.1080/1091367X.2020.1835664

Moir, G. L. (2008). Three different methods of calculating vertical jump height from force platform data in men and women. *Measurement in Physical Education and Exercise Science*, *12*(4), 207–218. https://doi.org/10.1080/10913670802349766

Morrow, J. R., & Jackson, A. W. (1993). How "significant" is your reliability? *Research Quarterly for Exercise and Sport*, *64*(3), 352–355. https://doi.org/10.1080/02701367.1993.10608821

French, D. N., & Torres Ronda, L. (Eds.), & National Strength & Conditioning Association (U.S.). (2022). *NSCA's essentials of sport science*. Human Kinetics, Incorporated.

Norman, R. W., & Komi, P. V. (1979). Electromechanical delay in skeletal muscle under normal movement conditions. *Acta Physiologica Scandinavica*, *106*(3), 241–248. https://doi.org/10.1111/j.1748-1716.1979.tb06394.x

Peng, H.-T. (2011). Changes in biomechanical properties during drop jumps of incremental height. *Journal of Strength and Conditioning Research*, *25*(9), 2510–2518. https://doi.org/10.1519/JSC.0b013e318201bcb3

Pérez-Castilla, A., Fernandes, J. F. T., Rojas, F. J., & García-Ramos, A. (2021). Reliability and magnitude of countermovement jump performance variables: Influence of the take-off threshold. *Measurement in Physical Education and Exercise Science*, *25*(3), 227–235. https://doi.org/10.1080/1091367X.2021.1872578

Pinto, B. L., & Callaghan, J. P. (2021). An appropriate criterion reveals that low pass filtering can improve the estimation of counter-movement jump height from force plate data. *Measurement in Physical Education and Exercise Science*, *25*(4), 344–352. https://doi.org/10.1080/1091367X.2021.1906253

Portney, L. G., & Watkins, M. P. (2015). *Foundations of clinical research: Applications to practice* (3rd edition [revised] ed.). Pearson Education Limited.

Ross, S. E., Guskiewicz, K. M., & Guskiewicz, K. M. (2003). Time to stabilization: A method for analyzing dynamic postural stability. *Athletic Therapy Today*, *8*(3), 37–39. https://doi.org/10.1123/att.8.3.37

Šarabon, N., Knezevic, M. O., Mirkov, M. D., & Smajla, D. (2020). Introduction of dynamic rate-of-force development scaling factor in progressive drop jumps. *Journal of Biomechanics*, *110*, 109980. https://doi.org/10.1016/j.jbiomech.2020.109980

Schmidtbleicher, D. (1992). Training for power events. In P. V. Komi (Ed.), *Strength and power in sport* (1st ed., pp. 381–395). Blackwell Scientific Publications.

Shechtman, O. (2013). The coefficient of variation as an index of measurement reliability. In S. A. R. Doi & G. M. Williams (Eds.), *Methods of clinical epidemiology* (pp. 39–49). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-37131-8_4

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Uzelac-Sciran, T., Sarabon, N., & Mikulic, P. (2020). Effects of 8-week jump training program on sprint and jump performance and leg strength in pre- and post-peak height velocity aged boys. *Journal of Sports Science & Medicine*, *19*(3), 547–555. https://www.jssm.org/jssm-19-547.xml%3EFulltext

Vincent, W. J., & Weir, J. P. (2021). *Statistics in kinesiology* (Fifth ed.). Human Kinetics.

Wank, V., & Coenning, C. (2019). On the estimation of centre of gravity height in vertical jumping. *German Journal of Exercise and Sport Research*, *49*(4), 454–462. https://doi.org/10.1007/s12662-019-00581-6

Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *The Journal of Strength and Conditioning Research*, *19*(1), 231. https://doi.org/10.1519/15184.1

Wikstrom, E. A., Powers, M. E., & Tillman, M. D. (2004). Dynamic stabilization time after isokinetic and functional fatigue. *Journal of Athletic Training*, *39*(3), 247–253.

Xu, J., Turner, A., Comfort, P., Harry, J. R., McMahon, J. J., Chavda, S., & Bishop, C. (2023). A systematic review of the different calculation methods for measuring jump height during the countermovement and drop jump tests. *Sports Medicine*, *53*(5), 1055–1072. https://doi.org/10.1007/s40279-023-01828-x