Validity and Reliability of Two Methods of Calculating Leg Power

- Force plate only vs combined force plate and motion analysis

Anja Zoellner

THE SWEDISH SCHOOL OF SPORT
AND HEALTH SCIENCES
Master Degree Project: 32:2015
Master Program in Sport Science: 2014-2016
Supervisor: Maria Ekblom
Examiner: Karin Söderlund

Abstract

Aim

The aim of this study was to compare the reliability and describe the concurrent validity of two methods of measuring leg power; the use of a force plate only and the combination of force plate and motion analysis data.

Methods

Eight female second division volleyball players (Mean±SD: Age: 26.5±3.93 years, height: 1.70±0.053m, weight: 65.9±11.1kg) participated in two testing sessions, separated by one to six weeks. In each testing session participants performed three squat jumps (SJ) and three countermovement jumps (CMJ), with leg power for all jumps calculated both from force plate data only, and using force plate data combined with displacement data from motion analysis. Statistics calculated were t-test, intra-class correlation (ICC), and Bland-Atman plots.

Results

While the ICC of the calculated peak power using the two methods was moderate to high (ranging 0.715-0.847) for all jumps, the peak power was always significantly lower using the combined method indicating poor concurrent validity. Leg power decreased from test to retest in both CMJ and SJ, using both methods of calculation, although this decrease was only significant in the SJ using the combined method. The force plate only method showed no significant difference between test and retest, with a small bias (indicating a small systematic error) and a strong test retest correlation in both SJ and CMJ. However, both of these jumps also had quite a large range in the limits of agreements in the Bland-Altman plots. Peak power calculated using the combined method of force plate and motion analysis data showed a greater difference between test and retest values, significant in SJ and tending towards significance in CMJ. This indicates a greater systematic error than in the force plate only method. However, it still had a strong correlation and the range of the limits of agreement was slightly smaller than it was when using the force plate only, indicating a smaller random error.

Conclusions

The systematic differences between test and retest for the two tests indicates that the true peak power achieved by the subjects of the subjects in this study was lower in the retest. The ICC values still indicate results from this study suggest that reliability is acceptable for both methods. Limits of agreement and CI of the mean differences are presented to guide interpretation of individual and group data respectively. Either method could be used to test leg power, but the results from different methods should not be used interchangeably.

Contents

Tables

Figures

1. Introduction

1.1. Background

Power, in particular leg power, is vital for success in many sports, and is one of the most common performance variables reported in sport (McMaster, et al., 2014). For this reason, it is an area which has been the focus of a number of studies (Luebbers, et al., 2003, Marques, et al., 2008, McBride, et al., 2002, Newton, et al., 2006, Wilson, et al., 1993), looking at different training methods to increase leg power. Due to the importance of leg power to sports performance, it is important to find testing procedures with the greatest validity and reliability possible in order to monitor training effects or to assess talent or performance factors (Markovic, et al., 2004). By doing this, it will ensure that changes or differences really are due to the training or the factor that is being changed and not just due to errors or inaccuracies in measurement, and it will allow conclusions to be confidently drawn from results. There are a number of common methods of measuring leg power, including the use of a force plate, either by itself or in combination with a vertical displacement measure. These are the two methods which will be included in this study.

Using a force plate by itself is a common method of measuring leg power during jumps, and it is possible to use the information obtained from the force plate to calculate power production. However, this involves a number of calculations and assumptions to be made, which can have adverse effects on the results. With every calculation, any errors that may have been made are amplified, thus decreasing both the validity and the reliability of the results. In order to reduce this risk, and to increase validity and reliability, it is common to include a displacement measurement as well, most commonly a linear position transducer. Another method that could be used, and which has not been studied so much, is the use of motion analysis in order to determine the displacement and from that derive the velocity of the movement. Using this method it is then possible to calculate power as the product of the vertical reaction force (from the force plate) and the velocity (calculated from the motion analysis). Using this combined method decreases the number of steps (calculations) required to determine power, which is one way to decrease the risk of error, as methods involving more steps have a greater risk of error due to calculations or increased noise (McMaster, et al., 2014).

As the use of motion analysis is not very common and there is limited information available on the reliability or validity of this method, this will be one method that is used in this study. The motion analysis will replace the use of the more common linear position transducer as the displacement measure to be used in combination with the force plate data to calculate leg power. These results will then be compared to those of a much more common method, the use of a force plate alone, which has been shown to have reasonable validity and reliability (Hansen, et al., 2011).

One sport which has been the focus of a number of training intervention studies involving training for leg power is volleyball (Marques, et al., 2008, Newton, et al., 1999, & Newton, et al., 2006). The methods used to measure leg power in these studies vary and therefore the changes found in different studies are difficult to compare, as they may have found different values or results had they used different methods to calculate leg power. The methods used to measure leg power in these studies include combined use of force plate and linear position transducer (Newton, et al., 2006), Plyometric Power System (Newton, et al., 1999), and force plate only (Vaverka, et al., 2013). As leg power is very important to both team and individual performance in volleyball players (Marques, et al., 2008), it is important to be able to confidently, and with as much accuracy and reliability as possible, compare training techniques. There is not much information on the reliability and validity of methods for determining leg power specifically in volleyball players, so this is an area which needs to be addressed.

1.2. Existing research

There are a number of methods commonly used for measuring and calculating power production during jumping. Differences between these methods are, at least partially, the reason for variations in power values in current literature (Dugan, et al., 2004). All methods seem to have high reliability (McMaster, et al., 2014) when using the same method. However there is greater variation when different methods are used, due to differences between methods, so they should not be used interchangeably and values can only be reliably compared if they were obtained using the same methods (Hansen, et al., 2011 & McMaster, et al., 2014). Common methods of determining power output include measuring displacement only, vertical ground reaction force, combining vertical ground reaction force measurement methods with displacement methods, accelerometers, and using video to track anatomical landmarks.

Measurement of displacement only most commonly uses a linear position transducer (Harris, et al., 2007 & Newton, et al., 2006), but it can also be measured using a rotary encoder, V-scope (Dugan, et al., 2004) or something similar. In this study motion analysis will be used in order to analyse displacement. This method gives displacement and time values throughout the jump which can be differentiated to find instantaneous velocity, which in turn can be differentiated again to give instantaneous acceleration, which, together with mass (body plus any added weights) and acceleration due to gravity can be used to calculate force and power (Dugan, et al., 2004). Using a linear position transducer makes the assumption that whatever the linear position transducer is attached to (e.g. barbell) is moving at the same velocity as the person (Dugan, et al., 2004), an assumption which is often incorrect (Hansen, et al., 2011). This method has been shown to have poor validity, as has been suggested by the significant difference ($p<0.05$) from the combination of force plate and vertical displacement (Cormie, et al., 2007a). However, Cormie, et al., (2007a) suggest that by using two linear position transducers instead of one, the validity and reliability can be improved.

Tracking anatomical landmarks and centre of mass on video (McMaster, et al., 2014) can track both vertical and horizontal movement and the coordinates can be used to digitally calculate jump height, then differentiated to give velocity and acceleration. These can, in turn, be used to calculate force and power. This method eliminates the assumption of the linear position transducer; that the object it is attached to is moving at the same speed as the person; as the markers are actually attached to the person. This could result in more accurate data. The assumption is made, however, that the markers being tracked are moving at the same velocity as the centre of mass. However, this can be managed by using multiple markers, including some closer to the centre of mass, which allows the displacement and velocity of the centre of mass to be calculated.

Vertical ground reaction force is measured using a force plate (Kawamori, et al., 2005 & Dugan, et al., 2004). This method calculates power only on the vertical axis, using the impulse-momentum approach and the ground reaction force (McMaster, et al., 2014). For this to be accurate the initial velocity needs to be zero (Dugan, et al., 2004). Hansen, et al. (2011) suggest that this method has the greatest validity, due to its greater absolute consistency, with an inter-day coefficient of variation of 4.6%. This was much better than the results for the linear position transducer (CV=8.0%), and minimally better than the combined force plate and linear position transducer methods (CV=4.8%).

While the use of a force plate by itself may have shown greater absolute consistency, Hansen, et al. (2011) also found that the relative consistency was slightly greater when using a combination of force plate and linear position transducer. This was shown by ICC values of 0.95 for the combined method and 0.94 for the force plate only. There was, however a greater difference with the linear position transducer only, which had an ICC of 0.87. In order to use a combination of ground reaction force and displacement measures, a force plate is used together with another method which measures vertical displacement, most commonly a linear position transducer (Nibali, et al., 2013, Cormie, et al., 2007a, Cormie, et al., 2007b, Newton, et al., 2006, Dugan, et al., 2004, Hansen, et al., 2011, & McMaster, et al., 2014). This allows power to be calculated using force from the force plate and the velocity calculated from the values of the linear position transducer, meaning that there is less calculation required and thus decreasing the risk of error and increasing the reliability (Dugan, et al., 2004 & Hansen, et al., 2011). As mentioned earlier, the most common method of measuring displacement is with a linear position transducer. However, Hansen, et al. (2011) suggest that the use of a linear position transducer can decrease the validity and reliability of the study. This is seen in their study, where the coefficient of variance if almost double (8.0% compared with 4.6%) that of the force plate alone, and the ICC is also lower (0.87 compared with 0.94). This decrease in validity and reliability is due to the assumption that the bar (or whatever the linear position transducer is attached to) moves at the same speed as the centre of mass of the body (Hansen, et al., 2011). This study will be using motion analysis software with anatomical markers to measure vertical displacement. The hope is that this may make the results more valid and reliable, due to markers being able to be placed close to the centre of mass, limiting the error due to the assumption of the velocity of the centre of mass. This method is also being used as there is only limited information available on this particular method. For this combined method, as with the ground reaction force only method, the velocity of the system needs to be zero at least once during measurement (McMaster, et al., 2014).

Integration, differentiation and manipulation of data can lead to calculation error, as well as increasing the effect of any measurement errors. In order to decrease the risk and effect of errors and to increase reliability and validity, it is better to use methods of measurement which obtain data closer to that which is desired, thus minimising the number of calculations needed, such as the combination of force plate and displacement data, than one method by itself (Dugan, et al., 2004, Cormie, et al., 2007a, Hansen, et al., 2011, & McMaster, et al., 2014).

There is some disagreement in previous research over whether the use of a force plate by itself is more valid (Hansen, et al., 2011), or whether a combination of force plate and vertical displacement has greater validity (Dugan, et al., 2004). This means that there is currently no gold standard method of measuring leg power, so this study will be looking at the concurrent validity between the two methods, as the actual validity cannot be calculated by comparison of these methods.

The tests used in this study will be the squat jump (SJ) and countermovement jump (CMJ). These two tests are popular for assessing the lower body, including leg power, and both have been shown to have good reliability (ICC of 0.97 for SJ and 0.98 for CMJ) (Markovic, et al., 2004). These exercises are also common in volleyball training (Newton, et al., 2006) and for training leg power (Ackland, et al., 2009), so the participants may already be familiar with them.

1.3. Aim, research question and hypothesis

The aim of this study was to describe the test retest reliability and the concurrent validity of two methods of measuring leg power; the use of a force plate only and the combined use of a force plate with motion analysis.

The research questions are:

What is the concurrent validity of leg power calculated with “the force plate alone method” versus “the combined force plate and motion analysis method”?

What is the test retest (between session) reliability of each respective method?

There are some studies that find the force plate only method to have greater validity and reliability (Hansen, et al., 2011) and some that find the combination of force plate and linear position transducer to have greater validity and reliability (Dugan, et al., 2004). However, this study will be using motion analysis in combination with the force plate, instead of the linear position transducer. This has less assumptions made than the linear position transducer, and less calculations than the force plate alone. For these reasons, the hypothesis is that the combination of force plate and motion analysis will have greater reliability than the use of the force plate alone.

2. Method

2.1. Participants

This study consisted of 8 female, second division volleyball players. (Mean±SD: Age: 26.5±3.93 years, height: 1.70±0.053m, weight: 65.9±11.1kg). Participants took part in two sessions of jump testing, which were between one and six weeks apart (1 week x 1, 2 weeks x 1, 3 weeks x 3, 4 weeks x 2, and 6 weeks x 1). Participants were asked not to change anything within their training in the time between testing sessions. They were also asked before retesting began whether they had changed anything in their training or health since the first testing session. All participants said there had been no change.

Participants were recruited by advertising on various public Facebook pages asking for volleyball players who would be interested in taking part. Interested people were then asked if other members in their team would also be able to take part. From this, two teams (15 people) were originally recruited, but one team (7 people) had to withdraw prior to the first testing session due to other commitments, leaving 8 participants, all from the same team. Exclusion criteria were any current knee, hip or ankle injuries, and anybody under the age of 18 years.

2.2. Study design

All testing took place in the late afternoon. Participants were asked to warm up on a cycle ergometer (Nibali, et al., 2013) for five minutes (Samozino, et al., 2008) at a low (10-12 RPE on the Borg scale) intensity. After warming up, two different types of jumps were explained and demonstrated to the participants. They were told they would be performing three countermovement jumps and three squat jumps. Both methods were used at the same time, so only one set of each jump had to be carried out. Once participants were clear on how to perform the jumps, two reflective markers were attached to the lower back, one each on the left and the right posterior superior iliac spine. The positioning of the markers was determined by palpating the posterior superior iliac spine. This was carried out by the same person each time to ensure consistency. Any clothes that might interfere with the visibility of these markers, by covering them, were taped out of the way. The participants were then asked to do one practice jump to ensure the markers could be seen throughout the movement. Once the markers were fixed in place, the participants were asked to stand still on the force plate, facing the same way they would for the jumps, to take a reference measurement. The reference measurement lasted for 3 seconds, during which time the participants were asked to keep as still as possible. After the reference measurement the jumping measurements could start (approximately 5 minutes after the warm up), beginning with the countermovement jump (CMJ), followed by the squat jump (SJ). The order of the jumps was kept the same for everyone, during both test and retest.

2.2.1. Countermovement jumps

Countermovement jump measurements were carried out starting with the participants standing in front of the force plate. They were then asked to step onto the force plate and stand as still as possible. As they were asked this, the measurement was started. Force measurements were visually monitored during this time, and once it could be seen that they were standing still, giving a stable baseline, they were told to jump, at which point a digital marker was inserted into the data. Once they completed the jump they could step off the force plate ready for the next measurement. Measurements were recorded for 15 seconds, to ensure enough time was allowed to get a stable baseline and complete the jump before the measurement ended.

Participants were instructed before jumping to jump as high as they could, to jump straight away, without holding the squat at the bottom, and to keep their hands on their hips throughout the jump. If they did not keep their hands on their hips they were asked to repeat the jump. When they jumped they squatted to a self-selected depth and jumped straight away, without holding the squat at the bottom.

The three countermovement jumps were separated by a one minute rest, during which time the squat jump was explained and demonstrated again.

2.2.1. Squat jumps

Squat jumps started the same as countermovement jumps, with the participants standing behind the force plate. They were then asked to step onto the force plate, at which point the measurement was started. Once on the force plate, a box was placed behind their knees, touching their calves, and they were asked to squat down so that both their calves and their thighs were touching the box, with the edge right in against the back of their knees, ensuring the knee angle was 90 degrees. Once they were in this position they were asked to hold it (1-2 seconds) as the box was removed. The force measurement was monitored again, and once this showed a stable baseline, they were instructed to jump, with a digital marker inserted into the data. Once they landed they could step off the force plate.

Participants were instructed to jump as high as they could, and to keep their heels on the ground when squatting and hands on their hips throughout the jump. They were also told not to drop down any lower before jumping, this was visually monitored. If they did not keep their heels on the ground while squatting, if their hands did not stay on their hips throughout the jump or if they dropped lower than the angle set by the box, they were asked to repeat the jump.

This was carried out twice more, or until three jumps had been completed following the criteria above.

2.3. Measurement methods

2.3.1. Power calculation from force plate only

This method for calculating power used force data only. Force was measured using a Kistler force plate with a sample rate of 1500Hz and a low pass filter set at 10Hz. Measurements started before the participant stepped onto the force plate and lasted for 15 seconds, with a digital marker indicating the point where the participants were asked to jump. Data was processed and calculations made using Visual 3D v5 Professional software.

Calculations:

First, body mass was calculated. This was calculated by rearranging the equation:

F=ma to give ***m=F/a***

Where ***F*** was the mean normal force recorded by the force plate in the half second before the marker indicating the participant had been instructed to jump, and ***a*** being acceleration due to gravity ($9.81ms^{-2}$).

Once body mass had been calculated, total acceleration during the jump was calculated by rearranging the same equation to give:

a=F/m

As this acceleration still involved acceleration due to gravity, this was then subtracted, to give acceleration due to the jump only.

Acceleration was then integrated to give velocity.

Using this, power could be calculated using the equation:

P=Fv

Where ***F*** is measured by the force plate and ***v*** the velocity as calculated above.

Peak power of the movement was then defined as the maximum value on the power curve, throughout the jump.

2.3.2. Power calculation from combined motion analysis and force plate

This method of power calculation used a combination of force plate and motion analysis data. Force data was taken from the force plate, using the same method as described above, and velocity data was calculated by differentiation of displacement data from motion analysis. Motion analysis was carried out using an 8 camera (Oqus 4, Qualysis) set up of optoelectronic motion capture system, with a sample rate of 150Hz and a low pass filter set at 7Hz. Markers on the left and right posterior superior iliac spine were tracked, with the midpoint of these points calculated and used for analysis. Data from Qualysis was exported to Visual 3D v5 Professional software used for calculations.

Motion capture was carried out from before the participants stepped onto the force plate and lasted for 15 seconds for each jump, as with the force measurements.

Calculations:
Define centre point between markers, using a model based on the reference measurement.
Differentiation of displacement data of midpoint to give velocity.
Power is then calculated using the formula:

$$P=Fv$$

Where v is the velocity that has just been calculated and F is the normal force from the force plate.

2.4. Statistical analysis

Distribution of the data was checked for normality using a Shapiro-Wilk test in Statistica. The mean value of the three CMJ and SJ attempts was used for statistical analysis. This was carried out using Microsoft Excel for descriptive statistics and paired t-test and Statistica software for intra-class correlation coefficient (ICC2,k), 95% confidence intervals, and Bland-Altman plots. Correlation coefficients were calculated for the Bland-Altman plots. If heteroscedacity was seen, raw data was plotted in a Scatterplot with a regression line and 95% confidence intervals instead of a Bland-Altman plot. P values of less than 0.05 were deemed as significant, with values less than 0.1 deemed as having a tendency, and a correlation greater than 0.7 was deemed as high.

2.5. Ethical considerations

All participants were required to give informed consent (Appendix 2) before the study started. Participants were informed, both orally and in writing, of all possible risks and were assured anonymity. They were given the chance to ask questions and were informed that participation was voluntary and that they could withdraw at any time without explanation.

3. Results

All data was found to be normally distributed. In all tests there was a decrease from test to retest and the values obtained using the combined method were lower in all tests than those obtained using the force plate only.

3.1 Peak power measured using a force plate only

Table 1 shows that in both the CMJ and the SJ, there was a small decrease in peak power (when calculated using force plate data only) from test to retest. However, this difference was not significant in either case.

Table 1: Descriptive statistics and t-test results for peak power values calculated using a force plate only.

Peak Power (W)	**Test (mean ± SD)**	**Retest (mean ± SD)**	**T-test**	**Mean Difference (W) (lower to upper 95% Confidence Interval)**
CMJ (mean ± SD)	2604.9 ± 336.7	2585.6 ± 323.6	p = 0.592	-19.3 (-100.5 to 61.9)
SJ (mean ± SD)	2576.6 ± 373.5	2543.4 ± 353.0	p = 0.494	-33.2 (-142.1 to 75.6)

There was no significant difference in peak power (calculated using the force plate only) between test and retest results for CMJ or SJ, which can also be seen in Figures 1 and 2, where there is only a small bias (Figure 1 = -19.30W, 0.744%, Figure 2 = -33.24W, 1.30%). These also both showed high correlations with values of 0.9797 ($p < 0.05$) for CMJ and 0.9688 ($p < 0.05$) for SJ. The limits of agreement in Figure 1 (CMJ) have a range of 380.8W, from -209.7W to 171.1W. This equates to 7.34% on either side of the bias line. Six out of eight points fall below both the bias and zero, while one of the two remaining points falls outside of the limits of agreement. Figure 2 (SJ) has a greater range in the limits of agreement, of 510.3W (9.97% either side of the bias), from -228.4W to 221.9W, but the points are more evenly distributed above and below the bias and zero lines than with the CMJ (Figure 1).

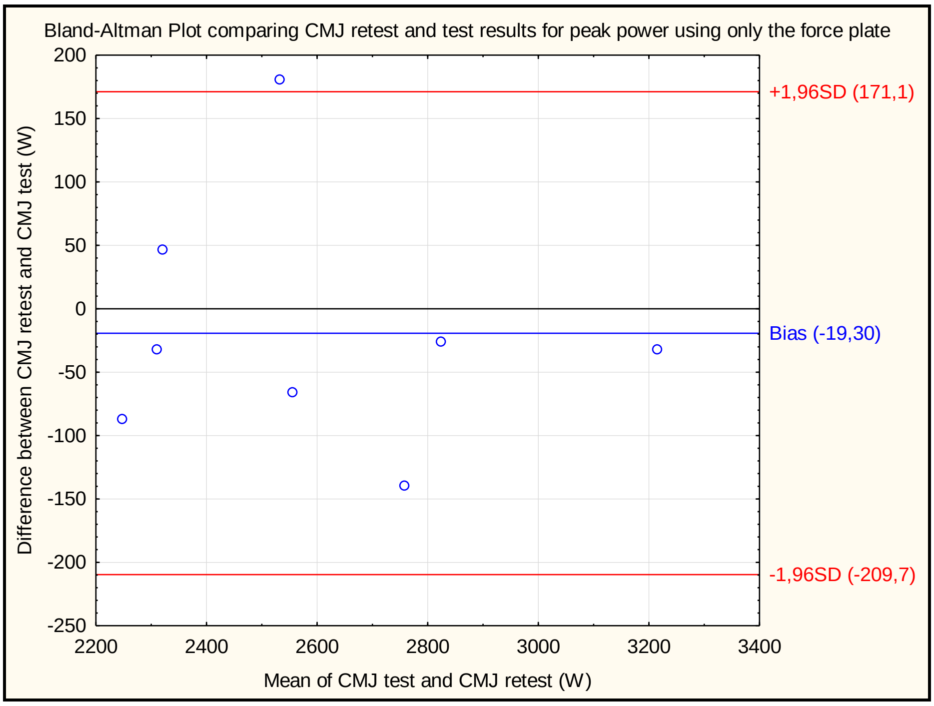

Figure 1: Bland-Altman Plot comparing CMJ retest and test results for peak power calculated using only the force plate, showing the bias (blue line) and upper and lower limits of agreement (red lines).

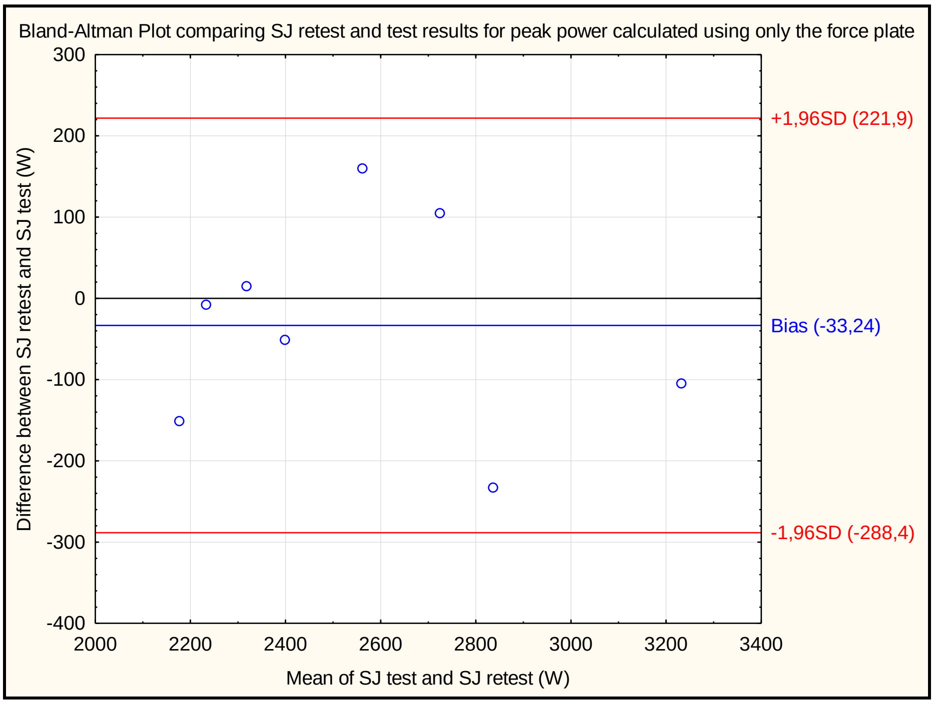

Figure 2: Bland-Altman Plot comparing SJ retest and test results for peak power calculated using only the force plate, showing the bias (blue line) and upper and lower limits of agreement (red lines).

3.2 Peak power measured using a force plate combined with motion analysis

The descriptive statistics in Table 2 show that when peak power is calculated using a combination of force plate and motion analysis data there is also a decrease in peak power from test to retest in both CMJ and SJ. In CMJ this tends towards significance, but in SJ it is a significant difference.

Table 2: Descriptive statistics and t-test results for peak power values calculated using a combination of force plate and motion analysis data.

Peak Power (W)	**Test (mean ± SD)**	**Retest (mean ± SD)**	**T-test**	**Mean Difference (W) (lower to upper95% Confidence Interval)**
CMJ (mean ± SD)	2408.8 ± 217.2	2351.3 ± 231.0	p = 0.065	-57.50 (-119.6 to 4.6)
SJ (mean ± SD)	2316.2 ± 303.4	2218.0 ± 259.7	p = 0.031	-98.20 (-184.7 to -11.7)

There was a tendency towards a difference between test and retest values for peak power (calculated using a combination of force plate and motion analysis) in CMJ, with a p-value of 0.065. This can also be seen in Figure 3, which has a relatively larger bias (-57.50W, 2.416%) than using the force plate only method (Figure 1; bias of -19.30W, 0.744%). However, there was a strong correlation of 0.9590 ($p < 0.05$), and the range of the limits of agreement was slightly narrower (291.18W, from -203.1W to 88.08W, 6.12% either side of the bias line) using this method than when using the force plate only method (range 380.8W, 7.34%), however the relative range is only 1.22% less, so this difference is reasonably small. The points in this method are however more evenly distributed around the bias and zero lines than in Figure 1. This decrease in the width of the limits of agreements indicates a decrease in variance between individuals. This decrease also coincides with a narrower confidence interval using the combined method (-119.6 to 4.6) than the force plate only (-100.5 to 61.9), suggesting that there may also have been less variation in the groups using this method than the force plate only method. However the mean difference is still greater using this method.

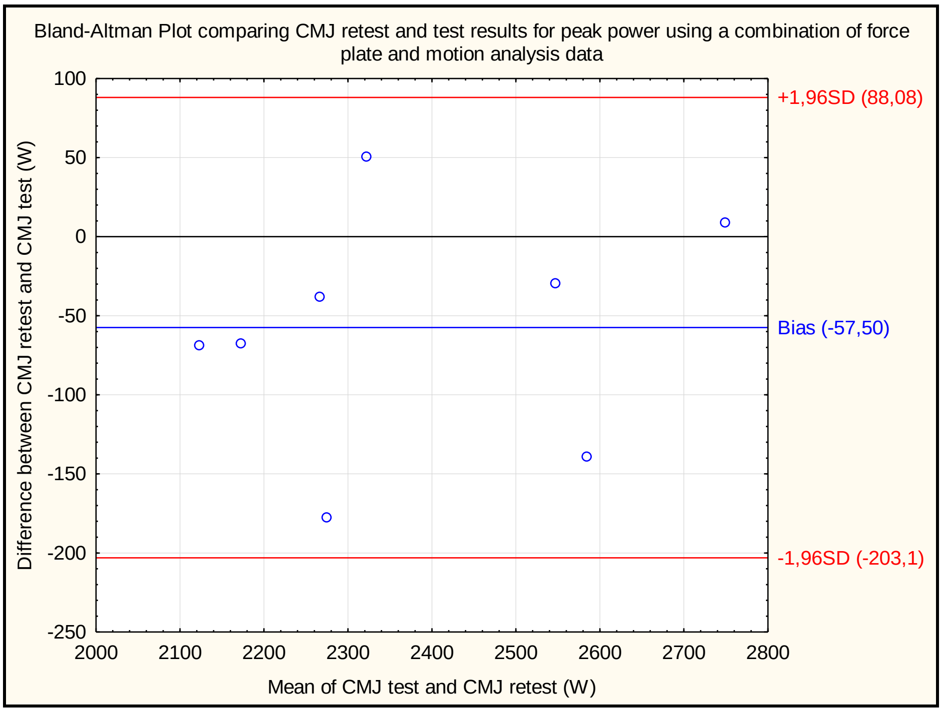

Figure 3: Bland-Altman Plot comparing CMJ retest and test results for peak power calculated using a combination of force plate and motion analysis data, showing the bias (blue line) and the upper and lower limits of agreement (red lines).

There is a significant difference between test and retest values for peak power (calculated using a combination of force plate and motion analysis) in SJ. There is also a larger bias (-98.20W, 4.33%) in Figure 2 than using the force plate method (Figure 2; bias of -33.24W, 1.30%). There is, however, a strong correlation of 0.9400 ($p < 0.05$) and the range of the limits of agreement in Figure 4 (405.6W, from -301.0W to 104.6W, 8.95% on either side of the bias) is smaller than the range of the limits of agreement using the force plate only method (Figure 2; range of 510.3W, -288.4W to 221.9W, 9.97% either side of the bias), suggesting a smaller difference between individuals. While the absolute difference seems quite big, the relative difference is only 1.02%, which is relatively small. The variation at a group level is also smaller in the SJ (as with the CMJ) using the combined method than the force plate method only, with confidence intervals from -184.7 to -11.7. Again, the mean difference using this method is greater than the mean difference using the force plate method only.

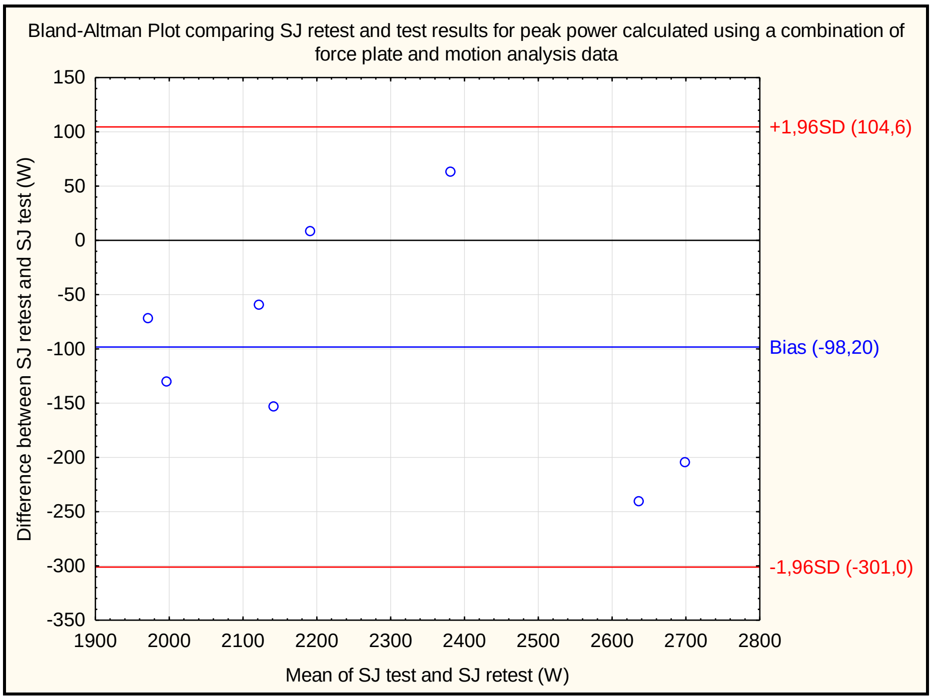

Figure 4: Bland-Altman Plot comparing SJ retest and test results for peak power calculated using a combination of force plate and motion analysis data, showing the bias (blue line) and the upper and lower limits of agreement (red lines).

3.3 Comparison of methods

In Table 3 it can be seen that there was a strong and significant correlation between the two methods used to calculate peak power in all tests. However, there was also a significant difference between the methods in all tests.

Table 3:Correlation of peak power calculated using force plate only and peak power calculated using a combination of force plate and motion analysis data.

Comparison of methods used to calculate peak power (force plate only vs combination of force plate and motion analysis data)	ICC results	T-Test
CMJ test	0.8333; $p < 0.05$	p = 0.0059
SJ test	0.8468; $p < 0.05$	p = 0.0003
CMJ retest	0.8054; $p < 0.05$	p = 0.0012
SJ retest	0.7153; $p < 0.05$	p = 0.0009

When comparing the results for peak power in the test CMJ when calculated using the force plate only with that calculated using a combination of force plate and motion analysis data, there was a strong correlation (0.8333, $p < 0.05$), as can be seen in Table 3. The mean difference in leg power is -196.1W (7.82%), with a confidence interval from -315.0 to -77.2. There was a significant difference between the two methods, as seen by the t-test result ($p < 0.05$). In this case a Scatterplot of the raw data, because a Bland-Altman plot showed a significant correlation (r=-0.8486, p<0.05), indicating heteroscedasticity, it now shows identity regression line with the 95% confidence intervals. The line of regression here, along with the confidence intervals, shows that there was greater variance between the methods for individuals with higher peak power. The standard error of estimate (SEE) was 66.19.

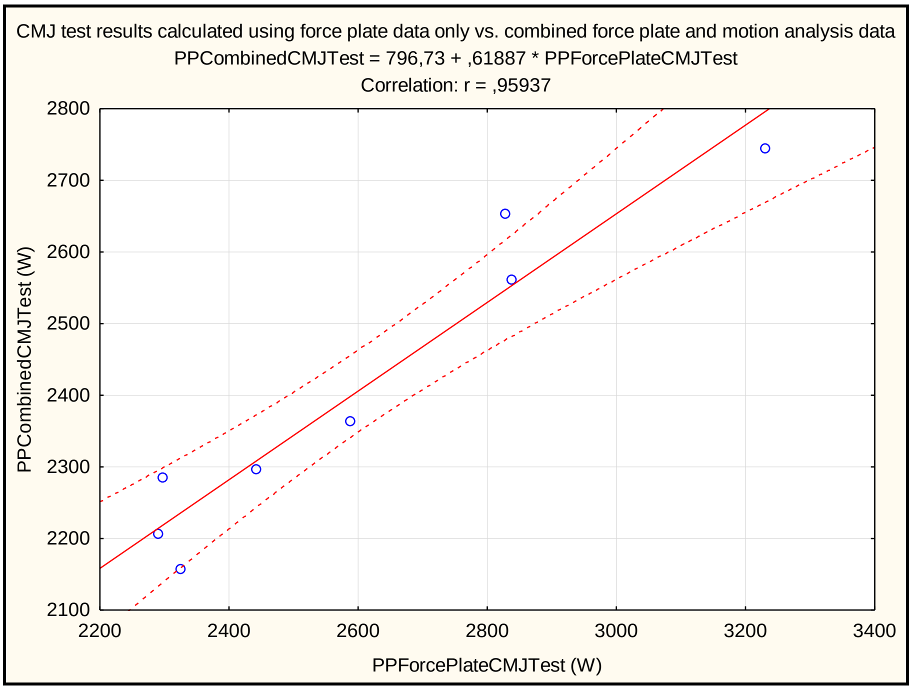

Figure 5: Scatterplot comparing raw data from CMJ test results for peak power when calculated using a combination of force plate and motion analysis data with those calculated using the force plate only, including 95% confidence intervals.

Table 3 also shows a high and significant correlation between peak power values in the test SJ when comparing the two methods of peak power calculation with a value of 0.8468 ($p < 0.05$). Figure 6 shows a bias of -260.4W (10.64%) and a range of 445.37W for the limits of agreement, from -483.1W to -37.73W (9.10% either side of the bias). The t-test result in Table 4 ($p < 0.05$) shows that the mean difference here is significant. The confidence interval in this case is from -355.4 to -165.4. The points here fall closer together at the lower means of peak power values, with half the points being at low values and between the bias line and upper limit of agreement. Furthermore, one point, that with the greatest mean of peak power values falling further out from the bias line, on the lower limit of agreement.

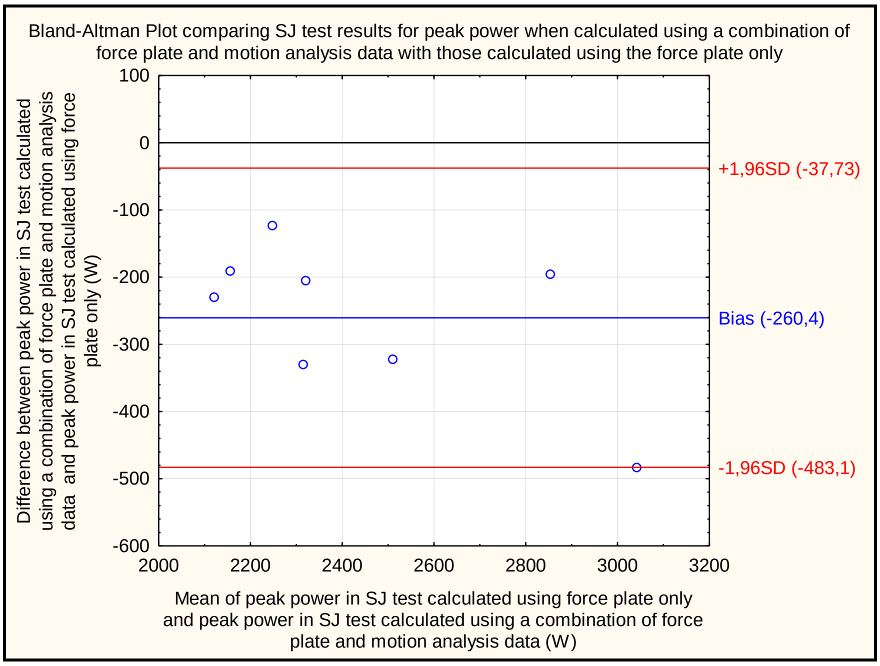

Figure 6: Bland-Altman Plot comparing SJ test results for peak power when calculated using a combination of force plate and motion analysis data with those calculated using the force plate only, showing the bias (blue line) and the upper and lower limits of agreement (red lines).

CMJ retest values for peak power when the two methods of calculation are compared also have a strong correlation, with a value of 0.8054 ($p < 0.05$; Table 3). The mean difference seen in the values of the retest CMJ (-234.3W, 9.50%) is greater than that in the test CMJ (-196.1W, 7.82%), and there is also a significant difference between the methods ($p<0.05$), with a confidence interval from -339.7 to -129.1. As with the CMJ test, the Bland-Altman plot in this case had a significant correlation ($r=-0.7441$, $p<0.05$), indicating heteroscedasticity. For this reason, again, a scatterplot of the raw data is shown in Figure 7, with the regression line and confidence intervals. The line with its confidence intervals shows, as in the CMJ test, that there is a greater difference between methods in those individuals that have greater leg power. The SEE was 76.90.

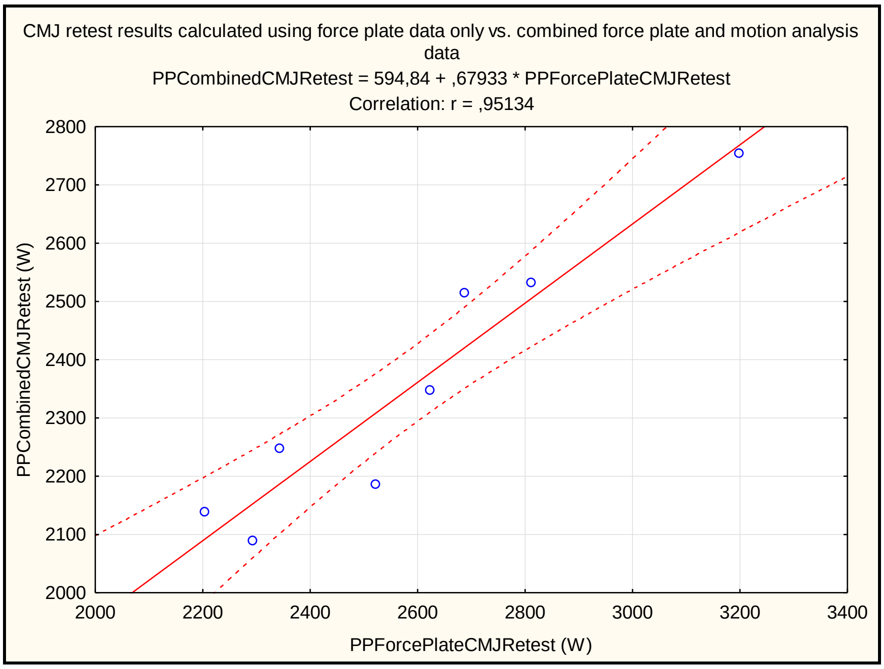

Figure 7: Scatterplot comparing raw data from CMJ retest results for peak power when calculated using a combination of force plate and motion analysis data with those calculated using the force plate only, including 95% confidence intervals.

Table 3 also shows a strong correlation in peak power values for SJ retest when comparing the two methods of calculation, with a value of 0.7153 ($p < 0.05$). Figure 8 shows a bias of -325.4W (13.64%), greater than that of the SJ test in Figure 6 (-260.4W, 10.64%), there is also a greater range in the limits of agreement in the SJ retest than in the SJ test, with a range in Figure 8 (SJ retest) of 657.3W, 14.07% either side of the bias, from -654.0W to 3.271W, compared with 445.4W, 9.10% either side of the bias, -483.1W to -37.73W, in Figure 6 (SJ test). The difference is not as big as that in the SJ test, however the confidence intervals are further apart (from -465.6 to -185.2), and the t-test results show there is a significant difference.

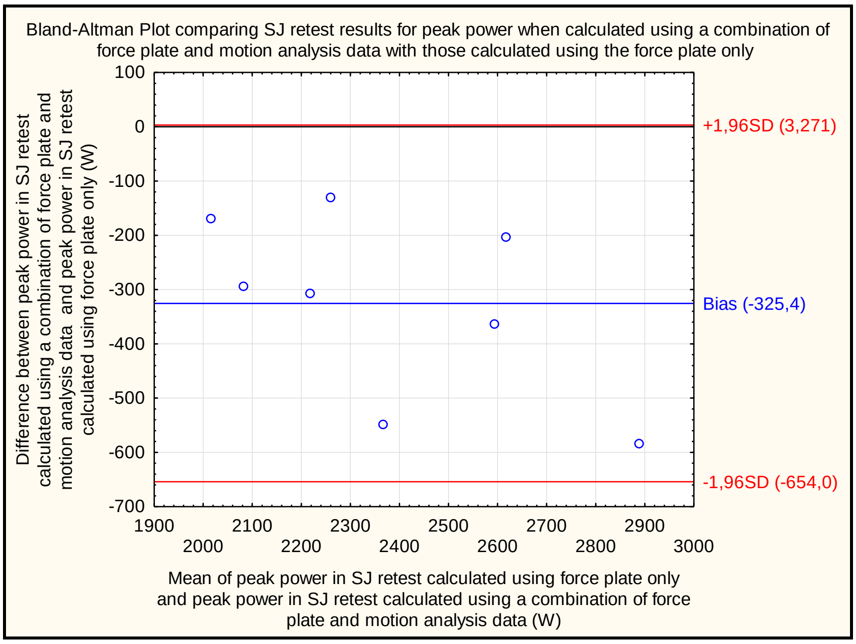

Figure 8: Bland-Altman Plot comparing SJ retest results for peak power when calculated using a combination of force plate and motion analysis data with those calculated using force plate only, showing the bias (blue line) and the upper and lower limits of agreement (red lines).

4. Discussion

The aim of this study was to investigate and describe the reliability and concurrent validity of two methods of calculating peak power during squat jumps and countermovement jumps in female second division volleyball players. The two methods used were calculation using force plate data only and calculation using a combination of force plate and motion analysis data. The main findings of the study were that in both SJ and CMJ, using both methods of calculation, peak power decreased from test to retest; and that the force plate data only method resulted in greater peak power values than the combined method for both jump types at both occasions (SJ, CMJ, test and retest). The results using the force plate only method were closer together, with smaller mean differences, but had a greater range in the limits of agreement, while the combined method had greater mean differences between measures, but a smaller range in the limits of agreement. The size of the range of the limits of agreement is an indication of the random error, with a greater range here indicating greater random error, while mean difference is an indication of systematic differences between tests. Both methods had high correlations between test and retest. The two calculation methods were also highly correlated to each other.

4.1. Reliability

The results for peak power when calculated using only the force plate could suggest that this method has high reliability, as seen by the lack of significant differences between test and retest data in both CMJ and SJ, which shows small systematic bias. However, the data is slightly more spread out using this method than the combined method, as seen in the large range between the limits of agreement, ±7.34% (Figure 1) and ±9.97% (Figure 2) respectively, which suggests a large individual difference between participants and a large random bias. As well as more variance at an individual level, the force plate only method also has larger confidence intervals, which suggest greater variation on a group level. There was no significant difference between test and retest in the CMJ using the combined method either, again suggesting that this method could also have good reliability, however, the results do tend towards difference, indicating a larger systematic bias. This could either mean that the reliability is not as good as the force plate only method, or it could suggest that there may have actually been a difference. This idea is further supported by the fact that there is a significant difference between test and retest using the combined method in the SJ. While there is a greater difference between test and retest using the combined method, the spread of the data, as seen by the width of the limits of agreement, is a bit less than when using the force plate method only, which suggests less random bias and less individual variation between participants. These results show a greater difference between the groups, and larger systematic bias, with a smaller difference within groups, both on a group and an individual level, and smaller random bias in both SJ and CMJ. This could indicate that there actually was a change from test to retest, however as there is no gold standard method, there is no way to know if the increased systematic difference was due to an actual change in leg power in the participants or due to an error in measuring. Due to the greater difference between the test and retest values for both CMJ and SJ in the combined method and as the limits of agreement range and confidence intervals were only slightly smaller when percentages were compared, it suggests that the force plate only method has slightly greater reliability than the combined method.

A possible reason for the greater reliability in the force plate only method is that there may have been small differences in the placement of the markers in the combined method. This would have led to slightly different values on the different testing occasions. While efforts were made to ensure the markers were placed in the same place, (using the same person, who was experienced at placing markers) there was no way to ensure these were in exactly the same place. This could possibly have been changed if photos had been take to see exactly where the markers were for the first testing session so these could have been replicated, however even then there would have been some margin of error in the exact positioning.

The relative consistency is high in both methods, suggesting they both have reasonable reliability, but this is again higher in the force plate only method (ICC=0.9797 and 0.9688) than in the combined method (ICC=0.9590 and 0.9400). This high relative consistency in the force plate method (0.94) is also supported by the results of Hansen, et al. (2011), who obtained slightly higher values for the combination of force plate and vertical displacement method (0.95) than the force plate only. All of these values (anything between 0.9-1.0) were classed as having a “practically perfect” correlation by Hansen, et al. (2011).

While there were no significant differences between test and retest peak power values, in the CMJ for either the force plate only method or the combined method, or in the SJ using the force plate only method, the values were still lower in the retest than in the test. This was also the case for the SJ using the combined method, but the difference here was significant. The fact that this was seen using both methods and in both types of jump suggests that leg power may have actually decreased. A decrease in leg power could be due to the fact that the original tests were carried out at the end of the competitive season, with retests carried out 1-6 weeks after this, which could have allowed for some deconditioning between the tests and therefore lower peak power outputs in the retests. If this is the case, it could support that the combined method is more valid and that, due to the smaller limits of agreement range and lower random bias, the small change due to deconditioning would be more likely to be significant. This means that if the difference was indeed due to deconditioning, the combined method would be more likely to pick up on it, whereas the force plate only method would not show it as significant, suggesting the combined method could be more reliable. However, there is no way to determine if this really was the reason for the differences. With the variation in times between tests (1-6 weeks), and as participants stated that they had not changed their training between tests, there may not actually have been a change in leg power either. The decreases seen could be due to another factor. Participants may not have been as motivated at during the retesting session, due to the season being over or because it was the last testing session and they had already done it before. This could have resulted in decreased effort and thus decreased values for leg power.

The range of the limits of agreement in the Bland-Altman plot, not just the bias, plays a large role in the reliability of the results, as suggested by Owen, et al. (2014). Connelly (2008) also suggests this, stating that a narrower interval is better when comparing values. This is because if there is a large range in the limits of agreement, it means the data is very spread out, and changes are more difficult to detect. For example, if the limits of agreement range is 200W either side of the mean of means/bias line, and peak power is approximately 2500W, an improvement of 10% (250W, for example with training) would fall outside of the limits of agreement, so one could be reasonably sure that there really was a difference/improvement. However, if the improvement is only 5%, this would still fall within the limits of agreement, so it would be uncertain whether there really was a difference/improvement or whether the change was just due to variation/error in measurement. If a range is smaller, for example 100W on either side of the bias, this would mean that an even smaller difference could be recognised. With previous training studies reporting significant improvements in leg power of anywhere between 5-18% (Luebbers, et al., 2003, Sheppard, et al., 2007 & McBride, et al., 2002), this suggests a smaller range would be needed in order to be able to recognise some of the differences seen in other studies.

Previous studies (Dugan, et al., 2004 & Hansen, et al., 2011) have found that a combination of force plate and vertical displacement data has shown greater reliability. However, all methods seem to have high reliability (McMaster, et al., 2014). The results of this study also suggest that both methods are reliable, however neither would have high reliability. They also suggest that the combined method, that is, the method using both a force plate and vertical displacement (motion analysis) has greater reliability than the force plate only.

4.2. Concurrent validity

While both calculation methods seem to be reliable to a certain degree, there is no gold standard method to measure jump height and therefore the validity aspect studied here is the concurrent validity of the two methods. While the correlations were high and significant between the two methods in test and retest in both SJ and CMJ, they were not as strong as they were when looking at the methods individually. The biases seen in the Bland-Altman plots also suggest the concurrent validity is not perfect, with greater biases seen between methods than within methods. This is also suggested by the greater ranges in the limits of agreement (±11.12% in the CMJ test, ±9.10% in the SJ test, ±9.99% in the CMJ retest, ±14.07% in the SJ retest) when comparing the two methods than when looking at the methods individually, and with significant differences between methods in CMJ and SJ in test and retest, with all values being significantly greater when calculated using the force plate only method than when using the combined method. This significant difference between peak power measured using force plate only compared to peak power measured using a combination of force plate and vertical displacement measure was also seen by Hansen, et al. (2011).

As there is a greater difference between methods, this suggests you should not compare results from one method with those from another, for example when comparing results from before and after a training intervention, the same method should be used, but if reliability is high within each method, you could use either method, as long as you use the same method before and after, and only compare the results with those that have been obtained using the same method (Hansen, et al., 2011 & McMaster, et al., 2014).

4.3. Limitations

This study has a number of limitations, which could have affected the results found. The biggest limitation is the amount of time between test and retest sessions, with the participants having between one and six weeks between testing sessions. This could affect the results in two ways, one being that they had different amounts of time between tests, the second being that for some participants there was quite a long time between testing sessions. Having different amounts of time between tests means that some people may have had changes and others would have had less time for any changes to take place. This follows on to the second point that for some participants this was quite a long period between testing sessions, thus giving a greater opportunity for changes in fitness. This supposition also leads to another factor that could have affected the results, which is that the first testing session took place at the end of their competitive season, so, while the participants were asked not to change anything in their training between testing sessions and the assumption was made that they followed this instruction, they may have decreased their training a little bit anyway, or the fact that they were no longer playing games may have led to some deconditioning. This could explain the fact that all values measured and calculated in both SJ and CMJ and using both the force plate only method and the combined method of calculation decreased from test to retest.

Another limitation was the small sample size. With only 8 participants it is hard to say whether these results are representative of a larger group or whether these are particular to this specific sample, however, this could be an area for further research in the future. A larger sample size would also make it easier to identify trends, for example, if the combined method really does give lower values than the force plate only method. This would be because a larger sample size could decrease the random error.

The fact that the order of SJ and CMJ in the testing was not randomized could also have affected the results. The order was kept the same, CMJ then SJ, for all participants in both test and retest, in an attempt to standardise it, which may have helped the reliability, however, it may have affected the validity of the results. If this study was to be carried out again the order could be randomized, with half of the participants carrying out the SJ first and the other half carrying out the CMJ first, but keeping the orders for each participant the same for test and retest, to ensure reliability.

Another aspect of the CMJ and the SJ that could have improved the reliability and validity of the results is in the standardisation of the torso angle. Instructions could have included to keep the torso upright, which could have been applied to both types of jump. This study included very little standardisation of the CMJ, which could have allowed other factors, such as depth of jump to affect the leg power, in future it might be better to standardise the depth in the CMJ as well as the SJ.

4.4. Suggestions for future research

As this was such a small group and due to the participants being from a very specific group, the results from this study can only be applied to female volleyball players at the sub elite level. More elite players may be more familiar with the tests and therefore may also have greater validity and reliability, whereas complete novices may show less reliability and validity due to their inexperience in performing these tests. Using players at different levels or from different sports in future studies could help to further determine which method is better, in terms of validity and reliability, to use when measuring peak leg power.

The results presented here could be used to perform power calculations to determine how many participants would be required for an intervention study using these testing methods.

5. Conclusion

The narrow limits of agreement together with high ICC suggests that the reliability may be slightly better for the force plate only method than the method combining force plate and motion analysis. While there was a strong correlation between methods, the concurrent validity was poor since the peak power values were always significantly higher as measured with the force plate only method. This suggests that the methods are not interchangeable, and if a study uses one method for one testing session it needs to use the same method for the next lot of testing as values cannot be reliably compared when they have been obtained using different methods.

6. Acknowledgments

I would like to sincerely thank my supervisor, Dr. Maria Ekblom for all the time, energy and knowledge she has invested in this project, as well as the encouragement and advice she has given me along the way. I would also like to express my gratitude to Olga Tarassova for taking the time to help me with the design of the project, data collection, and analysis, as well as advice throughout the project. I would, of course, also like to thank the participants of this project, without whom none of this would have been possible.

References

Ackland, T. R., Elliot, B. C. & Bloomfield, J., (2009). *Applied Anatomy and Biomechanics in Sport.* 2nd ed. Massachusetts: Blackwell Publishing.

Connelly, L. M., (2008). Bland-Altman plots. *Medsurg Nursing,* 17(3), pp. 175-176.

Cormie, P., Deane, R. & McBride, J. M., (2007a). Methodological concerns for determining power output in the jump squat. *Journal of Strength and Conditioning,* 21(1), pp. 424-430.

Cormie, P., McBride, J. M. & McCaulley, G. O., (2007b). Validation of power measurement techniques in dynamic lower body resistance exercises. *Journal of Applied Biomechanics,* Volume 23, pp. 103-118.

Dugan, E. L. et al., (2004). Determining the optimal load for jump squats: A review of methods and calculations. *Journal of Strength and Conditioning Research,* 18(3), pp. 668-674.

Hansen, K. T., Cronin, J. B. & Newton, M. J., (2011). The reliability of linear position transducer, force plate and combined measurement of explosive power-time variables during a loaded jump squat in elite athletes. *Sports Biomechanics,* 10(1), pp. 46-58.

Harris, N. K., Cronin, J. B. & Hopkins, W. G., (2007). Power outputs of a machine squat jump across a spectrum of loads. *Journal of Strength and Conditioning Research*, 21(4), pp. 1260-1264.

Kawamori, N. et al., (2005). Influence of different relative intensities on power output during the hang power clean: Identification of the optimal load. *Journal of Strength and Conditioning,* 19(3), pp. 698-708.

Luebbers, P. E. et al., (2003). Effects of plyometric training and recovery on vertical jump performance and anaerobic power. *Journal of Strength and Conditioning Research,* 17(4), pp. 704-709.

Markovic, G., Dizdar, D., Jukic, I. & Cardinale, M., (2004). Reliability and factorial validity of squat and countermovement jumps tests. *Journal of Strength and Conditioning Research,* 18(3), pp. 551-555.

Marques, M. C., Van Den Tillaar, R., Vescovi, J. D. & Gonzalez-Badillo, J. J., (2008). Changes in strength and power performance in elite senior female professional volleyball players during the in-season: A case study. *Journal of Strength and Conditioning Research* , 22(4), pp. 1147-1155.

McBride, J. M., Triplett-McBride, T., Davie, A. & Newton, R. U., (2002). The effect of heavy- vs. light-load jump squats on the development of strength, power, and speed. *Journal of Strength and Conditioning Research,* 16(1), pp. 75-82.

McMaster, D. T., Gill, N., Cronin, J. & McGuigan, M., (2014). A brief review of strength and ballistic assessment methodologies in sport. *Sports Medicine,* Volume 44, pp. 603-623.

Newton, R. U., Kraemer, W. J. & Häkkinen, K., (1999). Effects of ballistic training on preseason preparation of elite volleyball players. *Medicine & Science in Sports & Exercise,* 31(2), pp. 323-330.

Newton, R. U. et al., (2006). Four weeks of optimal load ballistic resistance training at the end of season attenuates declining jump performance of women volleyball players. *Journal of Strength and Conditioning Research,* 20(4), pp. 955-961.

Nibali, M. L., Chapman, D. W., Robergs, R. A. & Drinkwater, E. J., (2013). A rationale for assessing the lower body power profile in team sport athletes. *Journal of Strength and Conditioning Research,* 27(2), pp. 388-397.

Owen, N. J. et al., (2014). Development of a criterion method to determine peak mechanical power output in a countermovement jump. *Journal of Strength and Conditioning Research,* 28(6), pp. 1552-1558.

Samozino, P., Morin, J., Hintzy, F. & Belli, A., (2008). A simple method for measuring force, velocity and power output during squat jump. *Journal of Biomechanics,* Volume 41, pp. 2940-2945.

Sheppard, J., Newton, R. & McGuigan, M., (2007). The effect of accentuated eccentric load on jump kinetics in high-performance volleyball players. *International Journal of Sports Science & Coaching,* 2(3), pp. 267-273.

Vaverka, F. et al., (2013). The influence of additional load on time and force changes in the ground reaction force during the countermovement vertical jump. *Journal of Human Kinetics,* Volume 38, pp. 191-201.

Wilson, G. J., Newton, R. U., Murphy, A. J. & Humphries, B. J., (1993). The optimal training load for the development of dynamic athletic performance. *Medicine and Science in Sports and Exercise*, Volume 25, pp. 1279-1286.

Appendix 1 Literature Search

Syfte och frågeställningar: *How does the validity and reliability of a force platform alone compare to that of the combined force plate and motion analysis method of calculating leg power?*

Vilka sökord har du använt?

Leg power, measuring leg power, force plate validity, force plate reliability, motion analysis validity, motion analysis reliability, methods of measuring leg power, Bland-Altman plots, leg power in volleyball players, measuring leg power in volleyball players, plyometric training volleyball

Var har du sökt?

GIH:s bibliotekskatalog, PubMed, Ebsco, Google Scholar

Sökningar som gav relevant resultat

All key words in the above section gave results

Kommentarer

A lot articles were found via Google Scholar and GIH:s bibliotekskatalog, but most of the relevant articles were found by looking through the reference lists of other relevant articles.

Appendix 2 Informed consent form

This study was originally intended to be part of a training study, which is why participants filled out an informed consent form for a training study. However, with the withdrawal of one of the teams and the subsequent decrease in participant numbers, the training study was no longer feasible. Participants were verbally informed of the changes to the study.

Loaded Plyometric Training for Leg Power and Jump Height in Volleyball Players

Information

Background and Purpose

Jumping is an important part of many sports, and as such this is an area where training can be useful in order to help improve performance. Traditional training methods often include strength training for legs in order to help increase jump height, however, as jump height is very much dependent on leg power, which involves not only force production but also the speed of the movement, simply increasing leg strength may not be the best method to help increase jump height.

Plyometric training is a method of training which has been suggested to be an effective way of training leg power and jump height, with a number of studies suggesting there may be increased effects with the use of added weight during the training. There have been a number of studies that have looked at the effects of loaded plyometric training on leg power and jump height, however there have been no studies measuring its effectiveness in volleyball players, so that is what this study is focussing on.

How will the study be conducted?

The study will involve a six week training period, during which players will either be in the control group or in the training group, which will take part in loaded plyometric training, using 10% of body mass as the load. Training will take place two times per week and will consist of a warm up followed by three sets of six repetitions of both squat jumps and countermovement jumps. There will be a three minute rest in between each set. Pre- and post-tests will be carried out at GIH and will be used to determine maximal jump height and leg power. Tests used during the pre- and post-tests are the squat jump, and the countermovement jump.

We ask that during the period of this study players refrain from taking part in any additional leg training outside of the team training sessions regardless of whether they are in the training or the control group.

What are the risks?

The risks involved with this study are quite low. The main risk to players would be the risk of injury, however we will carry out familiarization in order to ensure they are performing the exercises using correct technique before the study starts, we will also be happy to answer any questions you or the players may have throughout the study.

Players may also notice some discomfort during the training sessions due to fatigue, as is normal during resistance training, however if they experience pain they will be asked to stop the exercise. As well as this, players may experience stiffness in the day or two following training, however we will be ensuring they warm up and cool down correctly in order to minimize this.

What are the advantages?

If the results show that loaded plyometric training does increase leg power and jump height this will help you, as it may give you a new training method to use during your training.

The results of this study will be useful both in adding to the information already available on loaded plyometric training and in helping to further the sport by giving possible training methods.

Dealing with data and confidentiality

All personal information will be computerized and stored electronically. GIH is responsible for all personal data, and no unauthorized person will have access to any personal information or results. Only the researcher will have access to personal information, which will be kept confidential and no identifying information will be published in the report, only average values will be used.

How do I obtain information about the result of the study?

Players will be able to have access to their own results, on request and will have access to the processed results if they wish.

Insurance, compensation

In the case of injury, players will be covered under the insurance of GIH.

Voluntariness

Participation in this study is voluntary, and if they do decide to take part, they can terminate their involvement in the study at any point without explanation.

Responsibility

The responsibility of this study is with Anja Zoellner (anja.zoellner@student.gih.se) as the primary researcher and Maria Ekblom (maria.ekblom@gih.se) as supervisor, as well as with Gymnastic- och Idrottshögskolan as the responsible research body.

If you have any questions please don't hesitate to contact either Anja or Maria.

Consent

I, ____,

hereby confirm that I have been given information, and the opportunity to ask questions, have received answers to my questions and I give consent to participate in this study.

Signed: Date:

Phone: Email: