Reliability and Validity of Three Clinical Methods to Measure Lower Extremity Muscle Power

Mattie E. Pontiff¹, Li Li², Noelle G. Moreau¹

¹Department of Physical Therapy, Louisiana State University Health Sciences Center, 1900 Gravier St., New Orleans, LA 70112
²Department of Health Sciences and Kinesiology, Georgia Southern University, 1332 Southern Dr, Statesboro, GA 30458

Corresponding Authors: Mattie E. Pontiff, PT, DPT, E-mail: mpont1@lsuhsc.edu

ARTICLE INFO
Article history
Received: October 02, 2020
Accepted: December 20, 2020
Published: January 31, 2021
Volume: 9 Issue: 1

Conflicts of interest: None.
Funding: The research is supported by the Department of Physical Therapy at LSU Health Sciences Center.
There was no external funding support for this study.

ABSTRACT
Background: Lower extremity muscle power is critical for daily activities and athletic performance in clinical populations. Objective: The purpose of this study was to determine the reliability and validity of 3 clinically feasible methods to measure lower extremity muscle power during a leg press. Methods: Ten of 26 subjects performed 2 sessions of 5 submaximal leg presses separated by 3-7 days in this repeated-measures cross-sectional design; the remaining performed 1 test session. Power was calculated independently for each method [simple video, linear position transducer, and accelerometer] and compared to the reference force plate. Test-retest reliability was evaluated using intraclass correlation coefficients (ICC). Pearson’s correlation coefficient (r), Bland-Altman plots with 95% limits of agreement (LOA), and mean bias percentages (%) were used to determine relative and absolute validity. Results: Power measures were reliable for all methods (ICC=.97-.99). All were highly correlated with the force plate (r=.96-.98). Mean bias was -0.8% (LOA: -16.57% to 14.98%) (video), -13.21% (LOA: -23.81% to -2.61%) (position transducer) compared to the force plate. Proportional bias was observed for accelerometer. Conclusion: All methods were reliable and highly correlated with the force plate. Only the video and position transducer demonstrated absolute validity. The position transducer was the most feasible method because of its simplicity and accuracy in measuring power.

Key words: Lower Extremity, Accelerometry, Muscles, Transducers, Weight Lifting

INTRODUCTION
Muscle power is defined as the amount of work performed per unit of time (Power = work/time) or the product of force and velocity (Bean et al., 2002). While it is related to muscle strength, muscle power is a measure of a muscle’s ability to generate force rapidly, whereas strength assesses a muscle’s ability to generate maximal force (Reid et al., 2014). Both are useful measures to assess skeletal muscle performance and are key components in functional activities (Tevald et al., 2016). Successful completion of activities such as brisk walking, transitioning from sit to stand, and maintaining balance require more rapid force generation (Moreau & Gannotti, 2015; Tevald et al., 2016). Thus, muscle power may be more important than muscle strength in performance of daily activities.

Deficits in lower extremity muscle power are linked to impairments in mobility, functional limitations, and increased disability across the lifespan (Bean et al., 2003; Kuo et al., 2006). Several studies have identified peak power as a predictor of poor physical performance and risk of falling in the elderly (Skelton, Kennedy, & Rutherford, 2002). Impaired lower extremity muscle power has also been associated with decreased knee confidence (Ageberg & Roos, 2016), decreased self-report function (Flosadottir, Roos, & Ageberg, 2016), and decreased participation in vigorous sporting activity in young adults with anterior cruciate ligament injuries (Flosadottir et al., 2016). Further, evidence suggests that power or high-velocity training is more effective in improving mobility and functional limitations than strength or functional training in several clinical populations (Bean et al., 2003; Corti, McGuirk, Wu, & Patten, 2012; Moreau, Holthaus, & Marlow, 2013; Tschopp, Sattelmayer, & Hilfiker, 2011).

Lower extremity muscle power is typically measured using expensive research equipment, such as isokinetic testing devices (Kuo et al., 2006), force plate (Giroux, Rabita, Chollet, & Guilhem, 2015; Gorostiaga et al., 2012), and cycle ergometers (Astorino & Cottrell, 2012). While these methods are valid and reliable, the devices are costly and impractical for daily clinical use. To mitigate high equipment costs, researchers have used pedometric field tests to measure leg power in athletic populations using Newtonian laws (Giroux et al., 2015; Samozino, Morin, Hintzy, & Belli, 2008). Unfortunately, these field tests are often inappropriate to test clinical populations with mobility limitations or those in a post-operative or post-injury period. More recently, accelerometry has been used to measure peak power during sit to stand movements (Regterschot, Folkersma, et
al., 2014; Zijlstra, Bisseling, Schlumbohm, & Baldus, 2010). These methods demonstrate excellent test-re-test reliability (Regterschot, Zhang, Baldus, Stevens, & Zijlstra, 2014) and fair to excellent correlations with a force plate for calculating peak power (Zijlstra et al., 2010). However, it may be limited in its ability to detect other to detect change following a training program, especially in individuals with higher levels of pre-surgical or pre-injury levels function.

Due to the limitations in cost and complexity in processing biomechanical testing methods and the narrow scope of current clinical testing methods, additional tests which can measure lower extremity power over a wider range of physical abilities are needed. The purpose of this study is to examine the reliability and validity of three methods to measure lower extremity muscle power [a simple video method (SVM), linear position transducer method (LPT) and accelerometer method (ACM)] compared to a reference method using a force plate (FPM) during performance of a leg press.

**METHODS**

**Participants and Design of Study**

Twenty-six subjects were recruited for participation (11 males, 15 females; age: 27.9 ± 4.2 years; height: 170.7 ± 9.3 cm; body mass: 69.7 ± 13.0 kg). This repeated-measures cross-sectional design incorporated two identical test retest sessions separated by 3-7 days to evaluate the reliability of 4 methods to measure lower extremity power during a leg press activity. Validity of each test method to measure power was then compared to the gold standard measure, the force plate (Giroux et al., 2015). This study was approved by the Institutional Review Board (IRB) at Louisiana State University Health Sciences Center. Following consent, each subject was screened for eligibility. Healthy individuals between the ages of 18 to 45 were eligible for inclusion. Exclusion criteria included: body weight greater than 200 lbs., BMI greater than 28, lower extremity injury less than 3 months prior to participation, lower extremity fracture or surgery less than 12 months prior to participation, current systemic disease, and a score of less than 72/80 on the Lower Extremity Functional Scale (LEFS). Body weight was limited to 200 lbs. due to limitations in the weight capacity of the equipment (650 lbs. total: body weight plus external load).

**Procedures**

Weight, height, total leg length (anterior superior iliac spine to medial malleolus), and lower leg length (medial tibial plateau to the medial malleolus) were measured on each person in the supine position. Subjects were positioned an inclined leg press (Total Gym® GTS, Total Gym Global Corp., Carlsbad, CA) which allows individuals to perform a squat-like motion on an adjustable incline with a moving glideboard (Figure 1a). The incline of the glide board was elevated to the highest level (28.8° from the horizontal) for this study. The angle was calculated using simple trigonometry.

\[
\theta = \tan^{-1}(\text{height/length}) \quad \text{(Eq. 1)}
\]

Subjects performed 1 to 2 unloaded familiarization trials followed by 2 to 4 loaded practice trials with 100 to 150% of body weight to determine the load for test trials. Load was increased until the subject felt they could press quickly without jumping. Feedback on performance was given as needed to improve movement quality. Following a 3 to 5 minute rest break to provide adequate skeletal muscle recovery (de Salles et al., 2009), the subject was positioned on the sled with the knees flexed to 90° ± 5° and their feet on the force plate (Figure 1a). A strap was used to support the subject and external load in the starting position. The subject’s start and end distance was recorded using two devices: a tape measure fixed to the side of the Total Gym® and a linear position transducer fixed to the weight bar (Figure 1a). Subjects performed 5 power leg presses by extending their knees and pushing as hard and as fast as possible (Figure 1b). Data were simultaneously collected using a video camera, linear position transducer, an accelerometer, and a force plate (Figure 1b). Feedback on performance was given as needed to maintain proper form and adequate movement velocity during testing trials. To assess test-retest reliability, 10 subjects were re-tested 3 to 7 days following the initial testing session using identical experimental procedures.

**Data Processing**

Mean power, mean velocity, and mean force were calculated for each method for each trial. Peak power could be calculated for ACM and FPM only. A minimum of 3 valid trials were averaged for each method from each testing session and these values were used in statistical analyses. Those subjects with less than 3 valid trials were excluded from the study.

**Simple video method (SVM)**

Movement performance of all trials were reviewed using a digital camera (Sony Handycam® HDR-HX250; New York, NY) recording at 29.97 frames per second. Trials were eliminated if the subject’s heels came off of the force plate during the press. Displacement (cm) for the concentric portion of each press was calculated by subtracting the starting and ending position of the sled measured with the tape measure. Total time for each press was calculated using Pinnacle Video Analysis Software (Corel Inc., Menlo Park, CA, USA). The point of first visual muscle activation plus one frame was the starting time and the point of first knee extension marked the end time of each press. Mean power (P) was calculated as the change in potential energy (PE) divided by time (t) using the following steps:

\[
E = \Delta PE + \Delta KE \quad \text{(Eq. 2)}
\]

The change in kinetic energy (\(\Delta KE\)) = 0 because the subject is not moving at the start and end of the press. The change in potential energy (\(\Delta PE\)) is defined as the product of the mass of the system (m), gravity (g) and height (h).

\[
\Delta PE = m \cdot g \cdot h \quad \text{(Eq. 3)}
\]

Where m is the mass of the system (sum of the mass of the subject, the external load and the sled); g is acceleration due to gravity.
due to gravity (-9.81 m/s\(^2\)) and \(h\) is defined as the change in height of the system at an angle (\(\theta\)) of 28.8°.

\[
h = d \sin(\theta) \quad \text{(Eq. 4)}
\]

where \(d\) is the displacement of the sled during a press. After all variables are derived, mean power (\(P\)) can be calculated as the change in potential energy over change in time (\(\Delta t\)).

\[
P = \frac{\Delta PE}{\Delta t} \quad \text{(Eq. 5)}
\]

This can be re-written as:

\[
P = \frac{mg\Delta h}{\Delta t} \quad \text{(Eq. 6)}
\]

**Linear position transducer method (LPT)**

The linear position transducer (TE Connectivity\(^\text{®}\), SGD 120-in Cable Actuated Sensor; Chatsworth, CA) was positioned on the floor with the cable attached to the weight bar (Figure 1a). Procedures were the same as the SVM method except displacement (\(d\)) for each press was calculated using the LPT. Mean power was calculated as the product of force (\(F\)) and velocity (\(v\)).

\[
P = F \cdot v \quad \text{(Eq. 7)}
\]

Where \(v = \frac{d}{\Delta t}\), and force is defined as:

\[
F = mg \sin(\theta) \quad \text{(Eq. 8)}
\]

**Accelerometer method (ACM)**

A triaxial accelerometer, GeneActiv Wireless (Activinsights Ltd., Kimbolton, UK) was fixed to the glide board (Figure 1a) and sampling frequency was set at 500Hz. Based on axis orientation, acceleration in the x and y directions were used to calculate the resultant acceleration (\(a_r\)) of the subject.

\[
a_r = \sqrt{a_x^2 + a_y^2} \quad \text{(Eq. 9)}
\]

The \(a_r\) was then integrated to calculate instantaneous velocity (\(v\)) of the system at time \(t\):

\[
v = \int a_r(t) + v_0 \quad \text{(Eq. 10)}
\]

Where \(v_0\) is initial velocity at \(t=0\). Power was then calculated as the product of net force (\(F_{net}\)) and velocity (\(v\))

\[
P = F_{net} \cdot v \quad \text{(Eq. 11)}
\]

Where \(F_{net}\) was defined as the sum of the forces created by the subject’s press (\(F_p\)) and the force of gravity (\(F_g\)). \(F_p\) was defined as the product of \(m\) and \(a\). \(F_g\) was defined in Eq. 8 (Figure 2). Data was filtered using a low-pass Butterworth filter with a cutoff frequency of 4 Hz. Refer to Figure 3c for the cut points for calculating mean and peak power.

**Force plate method**

An Accugait (AMTI, Watertown, MA, USA) force plate was used as the reference method in this study(Giroux et al., 2015) with NetForce Software for data acquisition by AMTI (Watertown, MA, USA). The plate was leveled and fixed to the foot plate of the Total Gym\(^\text{®}\) and positioned perpendicular to the glide board (Figure 1a). Sampling frequency was set at 500 Hz for all data collection. \(F_x\), \(F_y\), and \(F_z\) components were used to calculate the instantaneous acceleration of the system (subject and external load) at 28.8° from the horizontal:

\[
a = \frac{(F_{net} - F_e)}{m} \quad \text{(Eq. 12)}
\]

Where \(F_{net}\) is the sum of force vectors (\(\sqrt{(F_x^2 + F_y^2 + F_z^2)}\)), \(F_e\) is the effective weight of the system (in Newtons) at rest supported by the strap attached to the top of the glide board, and \(m\) is the total mass (in kg) at 28.8° from the horizontal. Acceleration and power were calculated as in ACM (Eq. 10 and 11). Refer to Figure 3a and 3b for the cut points for calculating mean power and Figure 3b for peak power.

**Statistical Analyses**

Reliability between test sessions was evaluated using intra-class correlation coefficient (ICC) for each method. Pearson’s correlation coefficient (\(r\)) was used to determine relative validity between the FPM (reference method) and each
method (SVM, LPT and ACM) for mean power and peak power, where applicable. Absolute validity of each method with the FPM was assessed using Bland-Altman plots with 95% limits of agreement (LOA) (Bland & Altman, 1986).

All statistical analyses were performed using IBM® SPSS® Statistics, Version 25.

RESULTS

One subject was removed from the study for having only two valid trials resulting in a total of 10 subjects for reliability testing and 25 for validity testing. Mean external load pressed was 86.5 ± 27.0 kg and mean percentage body weight pressed was 121.9% ± 18.6%. Averages for mean power, mean force, mean velocity and peak power for the SVM, LPT, ACM and FPM are reported in Table 1.

Reliability

The ICC values ranged from .967 to .995 for mean power and the measures of peak power had ICC values of .983, demonstrating excellent test-retest reliability for each method (Table 1).

Means and SD for average power, force, and velocity and peak power for each method (n=25). Reliability (ICC) of 4 methods (SVM, LPT, ACM and FPM) for calculating mean power and peak power (n=10). W= watts, N= Newtons.

Validity

Relative validity was excellent between the reference method and the SVM (r=.974; p < .001), the LPT (r=.989; p < .001) and the ACM (mean power: r=.984, p < .001; peak power: r=.993; p < .001). Mean bias and 95% LOA for each method compared to the FPM are shown in Figure 4, a-d. Greater than 95% of the differences fell within 95% LOA for the SVM (-58.2 W to 65.8 W), the LPT (-93.27 W to 1.72 W), and the ACM (mean power: -261.98 W to 32.02 W; peak power: -0.56 W to -287.0 W) (Figure 4, a-d). Because proportional bias was noted for the ACM for measurement of mean and peak power (Figure 4, c-d), transformation was needed. While log transformation is commonly used to address proportional bias, it can be difficult to interpret (Bland & Altman, 1986). Therefore, differences were expressed as percentages of the power values on the y-axis versus the mean of the two measurements on the x-axis (Figure 5, a-d). This method has been recommended for data that demonstrates proportional variability in the differences (Giavarina, 2015). Mean bias % and 95% LOA for each method compared to the FPM for mean power are shown in Figure 5, a-c. and for peak power for the ACM in Figure 5d. Proportional bias remained for mean power values for the ACM (Figure 5c).

DISCUSSION

Feasible, valid, and reliable measures of lower extremity muscle power, which utilize non-plyometric testing methods, are critical for assessment in clinical populations. All 4 methods used in this study demonstrated excellent test-retest reliability in measuring power during performance of a power leg press. When relative validity was examined, all 3 test methods (SVM, ACM, LPT) were highly correlated with a gold standard method using a force plate. However, only the position transducer and video methods demonstrated absolute validity.

Previous studies have reported excellent reliability when measuring mean and peak power with ICCs ranging from 0.84—0.99 (Bean et al., 2003; Giroux et al., 2015; Gomez-Piraz, 2013; Thompson & Bemben, 1999). Our values fall within the upper end of this range for measurement of mean and peak power. The novelty of the SVM makes it difficult to compare to previously published work. However, our findings are consistent with the reliability values reported by Samozino et. al for a vertical jumping task (Samozino et al., 2008) which also used a simple displacement measure and video method to calculate mean power.
Table 1. Descriptive and reliability data

<table>
<thead>
<tr>
<th>Variable</th>
<th>SVM</th>
<th>LPT</th>
<th>ACM</th>
<th>FPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Power±SD (W)</td>
<td>328.2±119.9</td>
<td>371.1±133.3</td>
<td>440.3±187.8</td>
<td>325.3±117.8</td>
</tr>
<tr>
<td>Mean Force±SD (N)</td>
<td>826.1±190.3</td>
<td>828.8±192.0</td>
<td>919.4±267.7</td>
<td>807.3±175.0</td>
</tr>
<tr>
<td>Mean Velocity±SD (m/s)</td>
<td>0.388±0.07</td>
<td>0.436±0.07</td>
<td>0.485±0.09</td>
<td>0.393±0.07</td>
</tr>
<tr>
<td>Peak Power±SD (W)</td>
<td>-</td>
<td>-</td>
<td>1035.8±410.7</td>
<td>892.0±353.7</td>
</tr>
<tr>
<td>ICC Mean Power</td>
<td>0.987</td>
<td>0.982</td>
<td>0.967</td>
<td>0.995</td>
</tr>
<tr>
<td>ICC Peak Power</td>
<td>-</td>
<td>-</td>
<td>0.983</td>
<td>0.983</td>
</tr>
</tbody>
</table>

Figure 4. Bland-Altman Plots. Each plot displays the difference between the two methods (reference – new method) versus the average of each method (reference – new method/2). The solid line depicts the mean bias and the dashed lines display the 95% LOA. (FPM=Force plate method, SVM = simple video method, ACM= accelerometer method, P = power, LPT = Linear position transducer method, W = watts)

Figure 5. Bland-Altman Plots as % differences between methods. Bland-Altman method plotting differences between the two methods (reference – new method/average %) as a percent versus the average of each method (reference – new method/2). The solid line depicts the mean bias % and the dashed lines display the 95% LOA %. (FPM=Force plate method, SVM = simple video method, ACM= accelerometer method, P = power, LPT = Linear position transducer method, W = watts)
All test methods demonstrated excellent relative validity compared to the FPM (ACM: $r = 0.984$, LPT: $r = 0.989$, SVM: $r = 0.974$) for measurement of mean and peak power. This is consistent with previous literature which also reports high correlations when comparing force plate methods to position transducers ($r = 0.87-0.89$) and accelerometers ($r = 0.87-0.95$) (Crewther et al., 2011; Giroux et al., 2015). Similarly, simple video methods using a vertical jumping task are also highly correlated to force plate methods for measurement of mean power ($r = 0.98$) (Samozino et al., 2008) which is consistent with our findings (SVM: $r = 0.974$).

For absolute validity, each method overestimated mean power compared to the reference method with an average overestimation 0.8% (SVM), 13.2% (LPT), and 27.6% (ACM). The ACM also overestimated peak power by 14.9% compared to the FPM. Despite attempts to transform bias, proportional bias in mean power was observed for the ACM with greater bias at higher power values compared to lower power values. The SVM and LPT demonstrated a constant bias in that differences between measures were consistent across a range of power values. Overestimation of mean and peak power by each test method is consistent with findings from previous validity studies, which used the force plate as a reference method (Bean et al., 2003; Choukou, Laffaye, & Taia, 2014; Giroux et al., 2015; Regterschot, Zhang, Baldus, Stevens, & Zijlstra, 2016; Zijlstra et al., 2010). Giroux et al. reported overestimation of mean power by 9.3% for the Samozino method, 10.1% for the position transducer, and 14.2% for the accelerometer compared to a force plate (Giroux et al., 2015). The average magnitude of overestimation in our data is larger than those previously reported for accelerometers (Giroux et al., 2015) and position transducers (Garcia-Ramos et al., 2016) which may be partly explained by device position. The placement of the accelerometer and position transducer evaluated the center of gravity of the sled, whereas the force plate assessed the center of gravity of the system, which can lead to velocity discrepancies as reported in previous studies investigating a squat-jump activity using similar methods (Garcia-Ramos et al., 2016; Hori et al., 2007). Fixation of the LPT and accelerometer near the center of mass of the system may improve these discrepancies in future studies.

Proportional bias of mean and peak power using accelerometry has been previously reported with greater overestimation of power occurring at higher velocities of movement (Crewther et al., 2011; Giroux et al., 2015). Regterschot et al. also reported greater overestimation with increasing peak power values, which is consistent with our results (Regterschot et al., 2016). Because it is difficult to correct proportional bias, we concluded the ACM did not demonstrate absolute validity for measurement of mean power.

Limitations

Several limitations exist in our current study. First, we assumed similar velocities of the center of gravity of the sled and the system. While this may have led to velocity discrepancies in power calculations between methods, we felt that fixation of the accelerometer and position transducer to the sled would improve trial to trial consistency by eliminating any unwanted movement of the body or trunk which occurs during rapid presses. Secondly, we only evaluated our testing methods on healthy individuals which limits the generalizability of these findings to clinical groups of people. Despite this limitation, our methods used a weight training machine and submaximal testing procedures which improve the safety and feasibility of performing this test in both clinical and non-clinical settings. Weight machines have been recommended for novice weight lifters which may include both patient and non-patient groups of people.

Only the SVM and LPT demonstrated absolute validity when measuring mean power. While the SVM demonstrated the smallest mean bias, a disadvantage to the simplicity of this method was that it was the most labor intensive in both collection and post-processing of the video data. In addition, the use of a tape measure to measure distance may increase error due to the difficulty of reading the ruler attached to the sled and due to the movement of the subject after the press is complete. In comparison to the SVM, the LPT provided a more accurate measure of distance with minimal post-processing of data.

Practical Implications of the Study

Prior studies utilizing biomechanical devices to measure power have focused on vertical plane analysis during plyometric tasks such as squat jumps (Garcia-Ramos et al., 2016; Giroux et al., 2015; Gomez-Piraz, 2013; Hori et al., 2007; Samozino et al., 2008). The strengths of our lower extremity power testing method can be summarized into two major points. First, our methods utilize a non-plyometric functional task that can be performed on an adjustable incline. These features have several practical implications. In the clinical setting, many patient groups are too weak to initiate explosive resistance training. The adjustable incline may allow rehabilitation specialists to measure lower extremity power and initiate explosive resistance training earlier in treatment as the external load can be varied with incline changes. In addition, our testing methods may be used to evaluate muscle power in active older adults who may not be able to participate in jumping assessments. Muscle power is a critical component of muscle performance in the elderly and in several patient groups. Muscle power deficits have been linked to increased fall risk in the elderly (Skelton et al., 2002). Moreover, power training has been shown to lead to greater functional improvements compared to strength training (Bean et al., 2003; Dorsch, Ada, & Alloggia, 2018; Scianini, Butler, Ada, & Teixeira-Salmela, 2009). Secondly, weight can be added to the system’s weight bar to test and address a range of muscle performance deficits and changes in muscle performance that occur with training. Power production is a key component of athletic performance, therefore, evaluating this element of muscle performance may improve training prescription in those participating in many different sports. The combination of an adjustable incline with the option to add external load allows training specialists to use this test to develop an exercise prescription for clients with a wide range of abilities. Lastly, our testing methods are cost effective and time efficient. We utilized readily available fit-
ness equipment combined with inexpensive biomechanical equipment to reduce cost and produce a precise and reliable power measure. In regards to time, a single test session takes approximately 15 minutes (including processing of LPT data) to complete making this test feasible to perform in both training and rehabilitation settings.

CONCLUSION

All 3 methods demonstrate excellent reliability and relative validity when measuring lower extremity power during the performance of a power leg press; however, only the SVM and LPT methods demonstrated absolute validity. The ACM demonstrated greater error with higher mean power output during the leg press, and thus is not recommended when absolute power values are needed. The LPT demonstrated excellent reliability, relative validity, and absolute validity while requiring the least amount of post-processing making it the most feasible for clinical use. Simple cost-effective measures of lower extremity power that are valid and reliable can enhance exercise prescription and outcome assessments across a variety of clinical populations.

REFERENCES


