Greater understanding of normal hip physical function may guide clinicians in providing targeted rehabilitation programs

**ABSTRACT:**

**Objective:** This study investigated tests of hip muscle strength and functional performance. The specific objectives were to: i) establish intra- and inter-rater reliability; ii) compare differences between dominant and non-dominant limbs; iii) compare agonist and antagonist muscle strength ratios; iv) compare differences between genders; and v) examine relationships between hip muscle strength, baseline measures and functional performance.

**Design:** Reliability study and cross-sectional analysis of hip strength and functional performance.

**Methods:** In healthy adults aged 18-50 years, normalized hip muscle peak torque and functional performance were evaluated to: i) establish intra-rater and inter-rater reliability; ii) analyse differences between limbs, between antagonistic muscle groups and genders; and iii) associations between strength and functional performance.

**Results:** Excellent reliability (intra-rater ICC=0.77–0.96; inter-rater ICC=0.82–0.95) was observed. No difference existed between dominant and non-dominant limbs. Differences in strength existed between antagonistic pairs of muscles: hip abduction was greater than adduction (p<0.001) and hip ER was greater than IR (p<0.001). Men had greater ER strength (p=0.006) and hop for distance (p<0.001) than women. Strong associations were observed between measures of hip muscle strength (except hip flexion) and age, height, and functional performance.
Conclusions:

Deficits in hip muscle strength or functional performance may influence hip pain. In order to provide targeted rehabilitation programs to address patient-specific impairments, and determine when individuals are ready to return to physical activity, clinicians are increasingly utilising tests of hip strength and functional performance. This study provides a battery of reliable, clinically-applicable tests which can be used for these purposes.

Keywords: Hip joint, muscle strength dynamometer, rehabilitation
Introduction:

Reduced hip strength is a common feature of conditions involving the hip joint, such as hip osteoarthritis (OA)\textsuperscript{1,2} and femoroacetabular impingement (FAI)\textsuperscript{3,4}. In addition, reduced hip muscle strength has been shown to be associated with increased prevalence of lower limb injury in runners\textsuperscript{5}; and with subsequent acute groin injuries in professional ice hockey players\textsuperscript{6}. As a result, many prevention or intervention programs are targeted towards improving hip strength and physical function\textsuperscript{7,8}. In the clinical setting, the ability to measure and interpret hip strength and physical function is therefore necessary to guide and monitor appropriate treatments.

Muscle strength and functional performance of the hip can be difficult to assess reliably. Hand held dynamometry (HHD) is one option for measuring peak isometric hip muscle strength\textsuperscript{9-11}. Measurement of hip muscle strength is increasingly being used as a screening tool to potentially assist in the identification of individuals at risk of hip problems\textsuperscript{4}. Muscle strength assessment has been shown to be reliable with a single rater in active, healthy adults\textsuperscript{10}, in older adults with hip osteoarthritis\textsuperscript{12} and in footballers\textsuperscript{13}. However, not all studies report the standard error of measurement (SEM), and the inter-rater reliability of using HHD for measuring hip muscle strength has not been evaluated to date. Finally, tests of functional performance of the lower limb such as the single leg hop for distance\textsuperscript{14} and side bridge test\textsuperscript{15} may also be valuable to guide rehabilitation, but little is known about the reliability of these tests. Therefore, the reliability of hip strength and commonly used tests of hip functional performance needs further evaluation.

Hip muscle strength is also used to monitor progress and guide treatment following injury\textsuperscript{11,16}. Knowledge of the balance of hip muscle strength between dominant and non-dominant limbs and between agonist and antagonist muscle groups may assist clinical decision making. Studies have investigated dominant versus non-dominant and agonist versus antagonist strength differences for
the hip abductors and adductors\textsuperscript{11,16,17}. However these studies were undertaken in elite athletes. To our knowledge no studies have examined these factors in non-elite physically active adults. In addition, no studies have examined these relationships in other key muscle groups surrounding the hip. As a consequence, there is a need to develop a more complete understanding of both dominant versus non-dominant and agonist versus antagonist muscle strength profiles in all hip muscle groups in non-elite physically active adults.

Gender may be an important factor in hip pathologies. The prevalence of symptomatic and asymptomatic cam-type FAI is greater in males than in females\textsuperscript{18}, and pincer type FAI is greater in females than males\textsuperscript{19}, while the prevalence of hip OA is reported to be higher in females\textsuperscript{20}. Such gender biases may be related, in part, to differences in hip muscle strength. Two studies have demonstrated gender-related differences in strength do exist for a cohort of healthy controls\textsuperscript{21,22}, however they examined hip abductor and external rotator muscle strength only. Another study demonstrated males with hip OA have reduced strength of the hip abductors, adductors and flexors compared to controls\textsuperscript{2}. Enhanced knowledge of gender-related differences in hip muscle strength and functional performance may assist in providing appropriate gender-specific rehabilitation programs.

The aims of this study were, in a cohort of healthy adults, to: (i) establish the intra- and inter-rater reliability of hip strength HHD assessments and functional performance measures; (ii) compare the hip strength and functional performance of the dominant versus non-dominant lower-limb; (iii) calculate the ratio between the strength of agonist and antagonist hip muscle groups; (iv) compare hip muscle strength and functional performance between genders; and (v) test whether any relationship exists between measures of hip muscle strength and participant baseline characteristics as well as functional performance.
METHODS:

Participants aged between 18 and 50 years were recruited from the community via advertisements in local print media and posters. A cohort of 15 people were recruited for the reliability study, (age 28 ±4 years; 1.73 ±0.06m; 68 ±13kg; BMI 22.4 ±3.3)kg/m²). A separate cohort of 57 people were recruited for Hip muscle strength and functional performance testing (age 34 ±9 years; 1.71 ±0.09m; 669 ±12kg; BMI 23.5 ±3.)kg/m² - The complete baseline characteristics are contained in Supplementary Material Table 1. Participants were excluded if they currently suffered low back pain or other lower limb injury, could not walk without assistance, had a history of hip pain in the previous 6 months, had undergone hip surgery, or had an inability to read or speak English. All participants provided written informed consent and the study was approved by the University of Melbourne Human Research Ethics Committee (HREC number 1033739 and 1033063).

Intra-rater reliability was evaluated by single tester performing two assessment sessions separated by one week. Inter-rater reliability was evaluated by two testers performing a single assessment session on the same day. Both examiners were female sports physiotherapists with at least 10 years of clinical experience. Measurements of hip strength and functional performance were collected by the same tester (JLK) using a standardised order of testing.

All strength tests were performed with a Commander Power track II (J-Tech medical) HHD. The tester matched the force generated by the participant performing an isometric muscle contraction (i.e. “make” test). The best of three tests was recorded. Visible marks were placed on the legs of participants to guide placement of the force transducer for the HHD strength assessments and to allow moment arm lengths to be calculated. Moment arm length was measured from the joint axis of rotation to the point of application of the force transducer for each test. For each test, torque was calculated by multiplying the force (measured in Newtons (N)) by the length of the moment arm.
(measured in metres (m)), and then data were normalized for body weight (measured in kilograms (kg)) (i.e. Nm/kg). All tests were conducted on both legs. Strength measures assessed included hip flexion, extension, external rotation in both neutral and 90 degrees of hip flexion, internal rotation in both neutral and 90 degrees of hip flexion, abduction and adduction. Strength testing set up procedures are contained in Supplementary Material Figures 1a-h.

Functional performance measures included the Side Bridge test\textsuperscript{15} and the single leg hop for distance (HFD) test\textsuperscript{14}. Testing procedures for Side Bridge, which was modified from its original description, are contained in Supplementary Material Figure 2, and HFD in Supplementary Material Figure 3.

Data analysis was performed using PASW for Graduate Students, SPSS for Windows, Version 18 (SPSS Inc., Chicago IL USA). The sample size for Part one was determined based on procedures described by Walter et al\textsuperscript{23}, whereby 15 participants was deemed to be acceptable for two raters, with $\alpha=0.05$ and $\beta=0.20$ (power = 0.8). To determine reliability, the intra-class correlation coefficient (ICC) and standard error of measurement (SEM) for each outcome measure were calculated. The SEM was calculated using the formula $SEM = SD \times \sqrt{1-ICC}$, where the SD is the SD of all of the scores for a particular test. Reliability was interpreted based on established criteria, where ‘good’ reliability was deemed to be ICC $\geq 75$%; ‘moderate’ reliability was within the range of ICC $\leq 75$% and $\geq 40$%; and ‘poor’ reliability was ICC $\leq 40$\textsuperscript{24}. The SEM was also expressed as a percentage, which was calculated by dividing the SEM with the average of the test and retest values.

Paired t-tests were performed to determine the mean difference between dominant versus non-dominant limbs and agonist versus antagonist muscle groups for all outcome measures. Independent t-tests were used to determine the mean difference between males versus females.
The dominant limb was deemed to be that used to kick a ball, as utilized previously by Crossley et al. Bivariate correlation (Pearson’s correlation coefficient (r)) was used to determine the association between measures of hip strength and measures of functional performance (for the dominant limb only). The critical value of r was determined at 0.325, greater than which the null hypothesis can be rejected. The significance criteria for all tests was determined a priori as p=0.01.

**Results**

The reliability of hip strength and functional measures are summarised in Supplementary Material Table 2. ICC values of greater than 0.77 for both intra-rater and inter-rater reliability were observed for all measures. All but two measures demonstrated ICC values greater than 0.85. The SEM ranged from 2.5% (HFD inter-rater reliability) to 14.4% (extension and adduction strength inter-rater reliability).

No significant between-limb differences were found in strength for any muscle groups (Figure 1), or functional performance measures (HFD mean difference (99% CI) 0 metres (-0.036 to 0.022, p=0.505); SB mean difference (99% CI) -5 seconds (-11.6 to 2.2), p=0.076). The mean difference [99% CI] in abduction to adduction strength (0.19 [0.10 to 0.28]; p<0.001) and ER to IR strength tested at a neutral hip position (0.20 [0.15 to 0.24]; p<0.001) were significantly different (Figure 1). The extension and flexion strength (0.03 [-0.08 to 0.14]; p=0.491) and ER to IR strength at 90° of hip flexion (0.03 [0.00 to 0.07]; p=0.024) were not significantly different. These findings are reflected in the ratios between agonist and antagonists; which were significantly greater than 1.0 for abduction to adduction (1.13) and ER to IR at neutral hip flexion (1.32), but not for flexion to extension (1.02) or ER to IR at 90° (1.04).

Gender-related differences in hip strength and functional performance for the dominant limb only, between males and females, are outlined in Table 1. Males exhibit significantly greater strength.
(peak torque normalised for body weight) than females only in ER (p=0.006), tested in a neutral hip position. Large differences were seen between males and females for functional performance in the single leg HFD (p<0.001) only.

Correlations between hip strength measures and participant characteristics as well as functional performance measures are outlined in Table 2. Age and height correlated significantly with all hip strength measures, with the exception of hip flexion strength. The HFD test and the side bridge test also correlated significantly with all measures of hip strength, with the exception of hip flexion.

**Discussion**

In a cohort of healthy individuals, aged 18-50 years, we identified that all measures of hip muscle strength and functional performance displayed good (ICC > 0.75) intra- and inter-rater reliability. While dominant and non-dominant limbs did not differ in hip strength or functional performance, agonist to antagonist ratios revealed abduction strength was greater than adduction strength and ER strength was higher than IR strength at neutral hip flexion. Men exhibited higher ER strength as well as greater HFD performance than women. Finally, strong associations were observed between all hip strength measures (except hip flexion) and the HFD and side bridge tests. This study offers new knowledge which may be useful for clinicians to facilitate the development of targeted rehabilitation programs and determine readiness for return to activity for people with physical impairments of the hip joint.

We demonstrated that commonly used clinical tests of physical function of the hip can be reliably performed between sessions and between testers. In addition, the SEM calculated in the present study provides clinicians with a guide to the measurement error for these tests, enabling them to determine whether a true change has been seen in the physical outcome following an intervention.
Whilst the SEM has been established previously for intra-rater reliability\(^\text{10}\), our study is the first to our knowledge to also establish this value for inter-rater reliability. The inter-rater SEM is of importance when testing is undertaken by different clinicians. Our results are similar to Thorborg et al.\(^\text{10}\) who tested the reliability of HHD, despite differences in test set up, and activity levels of cohorts. Our results indicate that changes of up to 14% for most measures would be required for observed differences to be greater than error. With an increasing need to justify the effects of treatments to third parties funding physical rehabilitation programs, it is essential for clinicians to have a battery of reliable measures of physical function capable of being used to demonstrate a meaningful change.

When assessing hip muscle strength and functional performance in those with or at risk of hip pathology, knowledge of the relative strength between dominant and non-dominant limbs or antagonists and agonist muscle groups can assist in the interpretation of test results. We observed no differences between the dominant versus non-dominant lower-limb for any measure of isometric hip muscle strength or functional performance. Our findings are similar to those by Thorborg et al.\(^\text{16}\), who examined isometric adductor and abductor strength in elite soccer players. This consistency was true despite differences in the activity levels of the cohort in our study and the cohort studied by Thorborg et al. Our results, combined with the findings of Thorborg et al., indicate that the use of the unaffected limb as a comparator when examining patients with unilateral lower-limb injury would be considered reasonable, in both athletic populations as demonstrated by Thorborg et al.\(^\text{16}\), and non-elite physically active adults as the current study demonstrates.

Knowledge of the agonist to antagonist hip strength ratios in healthy individuals may provide comparator baseline measures for the clinician. Our observed ratio of 1.13 for the abductor to adductor strength indicates that in our cohort the abductors were stronger than the adductors (\(p<0.001\)). Despite differences in the demographic variables and activity levels, this result is
consistent with the findings of Tyler et al.\textsuperscript{27}, who found that hip adductor strength was 95% of hip abductor strength in painfree ice hockey players, and Thorborg et al.\textsuperscript{16}, who reported a ratio of 1.05 for abductor strength compared to adductor strength in soccer players without pain. Both previous studies observed lower abductor to adductor strength ratios athletes with groin pain\textsuperscript{27}\textsuperscript{16}, indicating that targeted rehabilitation may be required until the abductors are stronger than the adductors in this patient population. In addition to establishing the abductor to adductor strength ratios in non-elite physically active adults, we also observed that the hip ER relative to IR strength ratio was 1.32, which has not been previously documented. Thus, rehabilitation and preventative programs should aim to restore these strength ratios. This information may be important when determining whether individuals with hip pathology or who have undergone hip surgery have recovered, or are ready to return to physical activity\textsuperscript{16}\textsuperscript{27}. In addition, these ratios may be helpful when comparing strength test results between therapists of differing stature, as these ratios should remain independent of tester strength.

We expected men to have greater hip muscle strength than women\textsuperscript{28}, even after accounting for size variation between genders by normalising torque measures for body weight. Our study confirmed that ER strength measures were higher in men, which was similar to the findings of a previous study\textsuperscript{21}. However, the current study did not demonstrate significant differences between genders for hip abduction, adduction, flexion, extension or internal rotation strength. Our results differ from those of Leetun et al.\textsuperscript{21} and Jacobs et al.\textsuperscript{22} in that we did not find a difference between gender in hip abduction strength. However, Leetun et al.\textsuperscript{21} tested trained athletes with a mean age of 19 years, which differs vastly from our population. In addition these studies used different testing protocols and strength testing devices when measuring hip abduction strength. This knowledge is of importance, as performance on hip ER and IR strength by patients with hip pathology can be expected to be different for each gender. The incidence and type of hip pathology varies between males and females\textsuperscript{20}, and a greater understanding of differences in physical performance between
genders may assist clinicians in the provision of individualised rehabilitation programs. In addition, it is possible that the differences seen in the incidence of hip pathology between males and females may be associated with the normal variations in strength and physical function reported in the present study.

When examining the relationship between hip strength, participant characteristics and functional performance, the HFD correlated strongest with ER strength in 0˚ of hip flexion \((r=0.624)\), whereas the side bridge test correlated strongest with ER \((r=0.463)\) and IR \((r=0.496)\) strength. In conditions where reduced strength of the hip ER is common (for example: patello-femoral joint pain\(^{29}\)), the HFD and side bridge tests may be employed by clinicians as a gross surrogate indicator of hip ER strength in the absence of HHD assessment.

The present study has some limitations. Due to the large age range of participants in this study, the results can be generalised to a range of people, but are not specific to an age group that may suffer a particular problem, for example athletic groin pain. In addition, as the two testers were female therapists of similar stature, the inter-rater reliability results may not be generalised to testers of different gender or body sizes. Further studies are needed to assess hip strength and functional performance in specific athletic populations, as the unique demands of elite sport will invariably mean that the results of strength and functional performance in uninjured athletes will be different from the results of the present study.

**Conclusion**

This study is important as it provides clinicians with a number of reliable, clinically-applicable tests to examine hip strength and functional performance. This knowledge is useful for clinicians as it may facilitate the development of targeted rehabilitation programs for people with physical impairments.
of the hip. In addition, gender-related differences were found, which may assist future research in establishing causative factors for differences seen in incidence and types of hip pathology between males and females. Finally, it provides values that can be used as a reference for assessment and re-assessment purposes, which may ultimately enable clinicians to provide targeted rehabilitation programs for their patients, and to set realistic objective goals in physical function.

Practical Implications

- Reliable and meaningful measures of hip physical function are essential for clinical practice, as they may guide optimal management and prevention of hip-related problems.
- This study comprehensively examined both hip muscle strength and functional performance in healthy adults.
- It provides values for reliable, clinical tests that can be used by clinicians as a reference for assessment and re-assessment purposes.
- The results may ultimately guide clinicians in providing targeted rehabilitation programs for their patients.

Acknowledgements:

We gratefully acknowledge the assistance of in conducting reliability testing. is the recipient of an Grant.

References:


### Table 1: Between-group differences in strength and functional performance between genders

<table>
<thead>
<tr>
<th>TEST</th>
<th>MALES*</th>
<th>FEMALES*</th>
<th>MEAN DIFFERENCE (99% CI)</th>
<th>p-VALUE**</th>
<th>%DIFFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extension†</td>
<td>1.62(0.46)</td>
<td>1.49(0.44)</td>
<td>0.13 (-0.19 to 0.44)</td>
<td>0.305</td>
<td>8%</td>
</tr>
<tr>
<td>Flexion‡</td>
<td>1.50(0.47)</td>
<td>1.39(0.50)</td>
<td>0.11 (-0.21 to 0.43)</td>
<td>0.358</td>
<td>7%</td>
</tr>
<tr>
<td>ER (90° flex)†</td>
<td>0.86(0.24)</td>
<td>0.71(0.21)</td>
<td>0.15 (-0.04 to 0.27)</td>
<td>0.057</td>
<td>17%</td>
</tr>
<tr>
<td>IR (90° flex)†</td>
<td>0.79(0.26)</td>
<td>0.69(0.21)</td>
<td>0.10 (-0.07 to 0.27)</td>
<td>0.119</td>
<td>13%</td>
</tr>
<tr>
<td>ER (0° flex)†</td>
<td>0.94(0.32)</td>
<td>0.74(0.20)</td>
<td>0.20 (0.01 to 0.39)</td>
<td>0.006**</td>
<td>21%</td>
</tr>
<tr>
<td>IR (0° flex)†</td>
<td>0.68(0.22)</td>
<td>0.56(0.15)</td>
<td>0.12 (-0.02 to 0.25)</td>
<td>0.024</td>
<td>18%</td>
</tr>
<tr>
<td>ABD‡</td>
<td>1.84(0.44)</td>
<td>1.65(0.35)</td>
<td>0.19 (-0.09 to 0.47)</td>
<td>0.077</td>
<td>10%</td>
</tr>
<tr>
<td>ADD‡</td>
<td>1.68(0.48)</td>
<td>1.46(0.36)</td>
<td>0.22 (-0.08 to 0.51)</td>
<td>0.055</td>
<td>13%</td>
</tr>
<tr>
<td>SB (seconds)</td>
<td>101(38)</td>
<td>77(43)</td>
<td>24 (-8 to 56)</td>
<td>0.053</td>
<td>24%</td>
</tr>
<tr>
<td>HFD (cm)</td>
<td>153(26)</td>
<td>111(31)</td>
<td>42 (20 to 63)</td>
<td>&lt;0.001**</td>
<td>27%</td>
</tr>
</tbody>
</table>

* = mean (SD); ** statistically significant difference (p<0.01); † = peak torque normalized for body weight (Nm/kg); EXT = extension; FLEX = flexion; ER = external rotation; IR = internal rotation; ABD = abduction; ADD = adduction; SB = side bridge; HFD = single leg hop for distance.
### Table 2: Associations between strength measures, participant characteristics and physical function (dominant limb)

<table>
<thead>
<tr>
<th>Strength measure (^\d)</th>
<th>Age ((r))</th>
<th>Height ((r))</th>
<th>HFD ((r))</th>
<th>SB ((r))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extension</td>
<td>-0.338(^*)</td>
<td>0.377(^*)</td>
<td>0.442(^*)</td>
<td>0.321(^*)</td>
</tr>
<tr>
<td>Flexion</td>
<td>-0.197</td>
<td>0.272</td>
<td>0.301</td>
<td>0.231</td>
</tr>
<tr>
<td>ER@90</td>
<td>-0.392(^*)</td>
<td>0.340(^*)</td>
<td>0.436(^*)</td>
<td>0.347(^*)</td>
</tr>
<tr>
<td>IR@90</td>
<td>-0.331(^*)</td>
<td>0.383(^*)</td>
<td>0.386(^*)</td>
<td>0.404(^*)</td>
</tr>
<tr>
<td>ER@0</td>
<td>-0.331(^*)</td>
<td>0.490(^*)</td>
<td>0.624(^*)</td>
<td>0.463(^*)</td>
</tr>
<tr>
<td>IR@0</td>
<td>-0.360(^*)</td>
<td>0.440(^*)</td>
<td>0.565(^*)</td>
<td>0.496(^*)</td>
</tr>
<tr>
<td>Abduction</td>
<td>-0.437(^*)</td>
<td>0.403(^*)</td>
<td>0.404(^*)</td>
<td>0.377(^*)</td>
</tr>
<tr>
<td>Adduction</td>
<td>-0.398(^*)</td>
<td>0.382(^*)</td>
<td>0.440(^*)</td>
<td>0.397(^*)</td>
</tr>
</tbody>
</table>

\(\d\)=peak torque(Nm) adjusted for body mass

\(^*\)=p<0.01; \(\^\d\)=r greater than critical value of 0.325 to reject the null hypothesis

External rotation measured at 90 degrees hip flexion; Internal rotation measured at 90 degrees hip flexion; External rotation measured at zero degrees hip flexion; Internal rotation measured at zero degrees hip flexion.
Figure Legends

Figure 1a

Strength values dominant vs. non-dominant leg

ER@90 = external rotation measured at 90 degrees hip flexion; IR@90 = internal rotation measured at 90 degrees hip flexion; ER@0 = external rotation measured at zero degrees hip flexion; IR@0 = internal rotation measured at zero degrees hip flexion.

Figure 1b

Strength values agonist vs. antagonist muscle group ratio

Ext:Flex = hip extension to hip flexion strength ratio; Abd:Add = hip abduction to hip adduction strength ratio; ER:IR@0 = hip external rotation to hip internal rotation strength ratio measured in 0˚ hip flexion; ER:IR@90 = hip external rotation to hip internal rotation strength measured at 90˚ hip flexion.