Test-Retest reliability of a hip strength assessment system in varsity soccer players

Abstract

Objective: To investigate test-retest reliability of a hip strength assessment system (GroinBar).

Design: Test-retest reliability study.

Setting: Laboratory.

Participants: Twenty asymptomatic varsity soccer players.

Main Outcome Measures: Maximal isometric hip strength (adduction, abduction, internal and external rotation, flexion and extension) was assessed using the GroinBar. Intraclass correlation coefficient (ICC) and relative standard error of measurement (SEM) were calculated to evaluate reliability of peak ($ICC_{3,1}$) (highest peak within 3 trials) and average peak ($ICC_{3,3}$) (average of 3 trials) force and rate of force development (RFD). Hotelling's T^2, were also used to compare bilateral and reciprocal ratios between dominant and non-dominant leg.

Results: ICC for both peak force and RFD values revealed moderate to good reliability (0.53-0.88 and 0.61-0.84, respectively), whereas reliability was good to excellent regarding their average values (0.77-0.95 and 0.81-0.92, respectively). SEM of average peak force and RFD values (4.1-9.4% and 8.2-13.9%, respectively) were lower than that of peak force and RFD values (5.7-13.0% and 10.7-19.1%, respectively). No significant difference was found in bilateral and reciprocal force ratios between dominant and non-dominant leg.

Conclusions: The GroinBar is a reliable tool to assess hip muscle function in athletic populations and could be used for player screening and follow-up.

Keywords

Hip; Muscle function; Reliability; Soccer

1. Introduction

Soccer is a sport involving explosive-type muscle actions like kicking, side-to-side cutting, sprints, and sudden directional changes. As a result, hip and groin injuries are one of the most common injuries in soccer, accounting for up to 13% of all soccer injuries (Arnason, et al., 2004). According to Hölmich, Thorborg, Dehlendorff, Krogsgaard, and Gluud (2013), adductor-related injuries (51%) followed by iliopsoas-related injuries (30%) are the most frequent clinical presentations of groin injuries in male soccer players. These injuries represent a major problem as players with history of previous acute groin strain are significantly more likely to sustain a new injury (Arnason, et al., 2004; Engebretsen, Myklebust, Holme, Engebretsen, & Bahr, 2010). In addition to history of injury, strength deficits have been reported as a significant risk factor for groin injuries in field-based sports athletes (J. Ryan, DeBurca, & Mc Creesh, 2014). For instance, Engebretsen, et al. (2010) revealed that players with weak hip adductor muscles have a four times higher groin injury risk. Moreover, recent studies reported that soccer players who suffered from groin pain have lower maximal isometric hip adductor strength compared to control players, and that this deficit can be up to 12% or 15% for injuries of longer than six weeks (Esteve, et al., 2018; Malliaras, Hogan, Nawrocki, Crossley, & Schache, 2009; Moreno-Pérez, et al., 2019). In addition to hip adductor weakness, stronger hip flexors have also been found in soccer athletes with groin pain during isokinetic evaluation (Mohammad, Abdelraouf, Elhafez, Abdel-Aziem, & Nassif, 2014). Therefore, as highlighted by Moreno-Pérez, et al. (2019), hip muscle strength assessment represents one of the key features for groin injury prevention and appropriate rehabilitation, given that muscle strength, as a risk factor, can be altered. Besides clinical interest, athletes' muscle function assessment is essential for athletic performance. Indeed, peak force and rate of force development (RFD), that is how rapidly force can be developed, are important physiological characteristics and determinants of performance in modern soccer play (Gissis, et al., 2006). However, despite being essential for athletes engaged in sports involving explosive movements (sprints, changes of direction, kicks, etc.), where maximal force cannot be reached because of the short contraction times (Aagaard, Simonsen, Andersen, Magnusson, & Dyhre-Poulsen, 2002), few studies have focused on RFD around the hip joint during isometric tasks.

To assess muscle function, isokinetic dynamometers have been shown to be reliable and valid instruments, and are generally considered as the 'gold standard' (Stark, Walker, Phillips, Fejer, & Beck, 2011). Moreover, this kind of instrument can provide multiple parameters regarding muscle function (peak force, power, endurance, angle of maximal force, force curves), and under several contraction modes (isometric, concentric and eccentric). Despite the extensive data that isokinetic dynamometers can provide, they are expensive and non-portable devices, and therefore present some limitations for daily/weekly use on the field or in clinical environment. As a consequence, hand-held dynamometers (HHD) have been widely used, as they are cheap, portable, and easy to use. Although several studies reported the reliability of HHD for measuring hip muscle strength (Charlton, Mentiplay, Grimaldi, Pua, & Clark, 2017; Fulcher, Hanna, & Elley, 2010; Kelln, McKeon, Gontkof, & Hertel, 2008; Thorborg, Petersen, Magnusson, & Hölmich, 2010), uncertainties remain on the use of this technique, especially in athletic populations. Indeed, Krause, et al. (2014) showed that force values obtained through HHD are influenced by examiner strength and testing technique. Thus, evaluating hip strength of athletes, which are likely to be stronger, results in systematic bias between testers (Thorborg, Bandholm, Schick, Jensen, & Hölmich, 2013). Moreover, as highlighted by recent studies, there is a lack of standardization when assessing hip adductor strength as a risk factor for a new groin injury in soccer (Esteve, et al., 2018; Moreno-Pérez, et al., 2019). Yet, the adductor squeeze test has become popular ((Esteve, et al., 2018; Fulcher, et al., 2010; Mosler, et al., 2017), and appears to be reliable and valid for screening changes in adductor strength and groin pain in athletes (Verrall, Slavotinek, Barnes, & Fon, 2005).

Recently, a transportable hip strength assessment system, the GroinBar (Vald Performance, Queensland, Australia), has been designed to enable measurement of both limbs simultaneously, enhance measure standardization and reduce the variability related to examiner strength and testing technique. In addition, by the use of four independent load cells, this device allows to easily estimate reciprocal and bilateral force ratios around the hip joint. Therefore, the GroinBar is a relevant device as part of the adductor muscles strength assessment through the squeeze test, but also considering that force ratios between agonist and antagonist muscle groups, as well as asymmetries between dominant and non-dominant limb, are possible risks for groin injuries in soccer (Engebretsen, et al., 2010; Mohammad, et al., 2014). Further, using the GroinBar does not require an extensive training or expertise, making it directly accessible to coaches. Regarding the reliability of this device, S. Ryan, Kempton, Pacecca, and Coutts (2018) looked at adductor strength assessment and have shown that the system is reliable and provide a greater measurement precision than HHD. Yet, little is known about the reliability of this system for the assessment of other hip muscle groups' strength. Therefore, the purpose of this study was to evaluate the test-retest reliability of the GroinBar for measuring muscle strength and RFD and to estimate reciprocal and bilateral force ratios around the hip in varsity soccer players without groin pain.

2. Methods

2.1 *Participants*

Twenty players from a university men's soccer team in first division Canadian league (age: 20.5 ± 2.2 years, height: 1.81 ± 0.06 m, body mass: 75.6 ± 8.2 kg, body mass index: 23.0 ± 2.2 kg/m^2) gave their written informed consent to participate in the study. All the players were free of any lower limb injury and had no history of surgery to the hip and groin region. Their dominant leg was determined as the preferred leg used to kick a ball. All testing procedures were approved by the local ethics committee.

2.2 *Protocol*

Isometric strength of hip muscles was assessed using the GroinBar at 400 Hz, a device previously described in O'Brien, Bourne, Heerey, Timmins, and Pizzari (2018). All tests involved three maximal voluntary isometric contractions (MVIC) lasting 5-s. Participants were strongly encouraged during MVIC and a minimum of 10-s rest was given between each contraction. The testing procedure included six isometric tests in a randomized order: adduction (ADD), abduction (ABD), external rotation (ER), internal rotation (IR), extension (EXT) and flexion (FLEX) of the hips as described in Fig 1. Hip flexors and extensors were evaluated one side at a time whereas the other muscle groups were assessed bilaterally. Test and retest were performed with a 15-min interval.

2.3 *Data processing*

Data analyses were achieved using Matlab software (R2016a, The Mathworks, Natick, MA). Force signal was low-pass filtered at 15 Hz using a 2nd order, zero lag, Butterworth filter. Peak force was defined as the single highest peak value of the three trials, for each muscle group, for each session of testing, namely test and retest. Average peak force was determined using the average of the three peaks, for each muscle group, during each session of testing. Following a similar method, maximal and average rate of force development (RFD) were calculated by scanning successive time intervals of 200-ms across the force signal to determine peak RFD as it has been shown to be the most reliable method (Mentiplay, et al., 2015). Force and RFD values were then normalized by body mass. Finally, reciprocal (agonist to antagonist muscles) and bilateral (dominant to non-dominant leg) force and RFD ratios were calculated to assess hip joint stability and asymmetries or imbalances. Bilateral ratios were determined using the formula $\frac{D-ND}{D} * 100$, where *D* and *ND* were the forces of the dominant and non-dominant leg respectively. For a secondary objective, reciprocal and bilateral ratios were computed for both test and retest and reported as the mean of the two.

2.4 *Statistical analysis*

All data are presented as mean ± SD, and were analysed using IBM SPSS Statistics version 25.0 (IBM Corp., Armonk, NY, USA). Test-retest relative reliability of force and RFD was assessed using intraclass correlation coefficients (ICC) and their corresponding 95% confidence intervals (95% CI), based on absolute agreement, two-way mixed-effects model. A single measure model ($ICC_{3,1}$) was used for peak values whereas an average measures model ($ICC_{3,3}$) was used for mean values. ICC were interpreted based on guidelines provided by Portney (2008): excellent (1.00 – 0.90), good (0.89 – 0.75), moderate (0.74 – 0.50), and poor (< 0.50). In addition, standard error of measurement (SEM), relative SEM (SEM%), and limits of agreement (LOA) were calculated to evaluate the absolute reliability. The SEM was calculated using the formula $SD_{av} * \sqrt{1 - ICC}$, where SD_{av} is the average of test and retest standard deviation and ICC is the calculated intraclass correlation coefficient (Atkinson & Nevill, 1998). The LOA was calculated using the equation: $MD \pm 1.96 * SD$, where MD is the mean difference between test and retest and SD the corresponding standard deviation (Atkinson & Nevill, 1998). Bland-Altman plots with 95% LOA were calculated for all the variables. Finally, Hotelling's T^2, corresponding to a multivariate t-test, were used to compare bilateral forces and RFD, as well as reciprocal force and RFD ratios between the dominant and non-dominant leg. The level of significance was set at $p < 0.05$.

3. Results

3.1 *Force and RFD reliability*

$ICC_{3,1}$ regarding peak force values revealed moderate to excellent reliability, ranging from 0.75 to 0.88 (95% CI: 0.58–0.86) (Table 1), except for FLEX where reliability was poor to moderate ($ICC_{3,1} = 0.53$; 95% CI: 0.26–0.72). Measurement variations (SEM) were between 5.7 and 13.0%. Test-retest reliability for average peak force values indicated good to excellent reliability, with $ICC_{3,3}$ varying from 0.86 to 0.95 (95% CI: 0.74–0.97) (Table 1), except for FLEX where reliability was moderate to good ($ICC_{3,3} = 0.77$; 95% CI: 0.55–0.88). The SEM were between 4.1 and 9.4%. For peak RFD, $ICC_{3,1}$

showed moderate to excellent reliability for most of the testing variables, ranging from 0.61 to 0.84 (95% CI: 0.53–0 .91) (Table 2). ABD and EXT were the two measurements where reliability was poor to good ($ICC_{3,1}$ = 0.61–0.68; 95% CI: 0.42–0.84). The SEM were between 11.6 and 19.1%. Test-retest reliability for average peak RFD revealed good to excellent reliability ($ICC_{3,3}$ = 0.81–0.92; 95% CI: 0.81–0.96), with ABD and EXT being the only tests where reliability was moderate to excellent ($ICC_{3,3}$ = 0.81–0.84; 95% CI: 0.63–0.92 (Table 2). The SEM were between 8.2 and 13.9%. Results from the Bland-Altman plots are provided in Supplementary file Fig. S1-4.

3.2 *Reciprocal ratios and bilateral forces*

Average peak force and RFD values, normalized to body mass, as well as reciprocal ratios are presented in Table 3 with division in leg dominance. As they were shown to be more reliable, and to avoid redundancy, data were only reported for average peak force and RFD values. Hotelling's T^2 tests revealed no significant difference ($p > 0.05$) between dominant and non-dominant leg as regards to reciprocal and bilateral ratios for both force and RFD values.

4. Discussion

The purpose of this study was to investigate the test-retest reliability of the GroinBar for measuring muscle strength and RFD around the hip. The results showed that the tested device had good to excellent test-retest reliability for average peak force and RFD values, and moderate to good reliability for maximal peak force and RFD values. Also, no significant difference was found when comparing muscle function between hip muscle groups of the dominant and non-dominant leg in varsity soccer players.

For the measurement of maximal absolute strength around the hip, analysis of the test-retest reliability of the GroinBar showed better results for both relative an absolute reliability using averaged measures than single measures. Indeed, ICC and SEM% with average measures ($ICC_{3,3}$ = 0.77–0.95 and SEM% = 4.7–9.4%) revealed higher and lower values respectively, compared to single measures ($ICC_{3,1}$ = 0.53–0.88 and SEM% = 5.7–13.0%), reflecting a reduced within-subject variability and measurement error with the use of averaged measures of maximal isometric force. However, despite of the use of an anchoring system, the GroinBar showed lower level of reliability than Scott, Bond, Sisto, and Nadler (2004) when assessing hip flexors strength. Difference in testing positions compared this previous study, and/or the modification of the device set-up to assess this muscle group could explain the lower values found in the present study. Overall, on the few studies carried out in athletic populations, measurements made with HHD have been reported to have similar or slightly lower level of reliability (ICC = 0.55–0.94) (Charlton, et al., 2017; Fulcher, et al., 2010; Malliaras, et al., 2009). Results were even worse when looking at inter-rater reliability (ICC = 0.40–0.83) (Fulcher, et al., 2010; Malliaras, et al., 2009). Regarding absolute reliability, values of standard error of measurement in the present study for averaged measures were acceptable (<10%) and comparable to those previously reported using HHD in a non-athletic population (Mentiplay, et al., 2015). Yet, the use of HHD, especially in athletic population, implies biased results due to tester's strength (Thorborg, et al., 2013). Moreover, using HHD, a question can arise about the complete commitment of athletes if they feel themselves overpowering the tester (Kelln, et al., 2008). Therefore, by allowing a strict isometric contraction and enhancing the standardization of tested positions, the GroinBar, while more expensive than HHDs, represents a reliable device to assess hip muscle strength in athletic population and could be then used for clinical management of athletes.

Despite its key role in sports involving explosive movements, the present study is, to our knowledge, the first one to investigate muscle RFD around the hip in elite soccer players. Results revealed that the GroinBar, when preferring the use of average measures on each degree of freedom, is a reliable device (ICC = 0.84-0.92; SEM = 8.2-13.9%) to assess RFD around the hip in soccer players. These results, can be compared to those from Mentiplay, et al. (2015) who were the first to investigate the reliability of HHD and validate its use to assess lower limb RFD in healthy participants. They found moderate to excellent reliability (ICC= 0.74-0.94), with SEM ranging from 9.5 to 16.9% for hip muscle groups. Yet, when using HHD to measure lower limb strength, the subject and especially athletes could overpower the tester. This would results in changes of joint angle as well as dissipation and thus attenuation of force during contractions, effects that are highly undesirable in the estimation of RFD (Maffiuletti, et al., 2016). In addition, according to Maffiuletti, et al. (2016), the force signal should be sampled at a high frequency to accurately measure the high RFD that human skeletal muscle is capable of producing. Thus by the use of a higher frequency rate than HHD (≤100 Hz), the GroinBar (400 Hz) could provide access to pertinent data regarding RFD in athletic population.

Looking at force values provided by the present study and contrary to Thorborg, et al. (2011) who reported stronger hip adductor and abductor muscles on the dominant leg for elite soccer players, no significant difference was found between the dominant and non-dominant leg for all the tested position. Normalized to body mass, ADD strength was 4.6 ± 1.0 N/kg on average, corresponding to a high force level following strength profiles established by Mosler, et al. (2017) in professional soccer players. This high level of force might be explained by the device used as a similar range of forces for ADD and ABD were reported recently with the use of the GroinBar in professional soccer players (O'Brien, et al., 2018). Regarding ADD/ABD ratio (1.05 ± 0.21), the present study revealed results in agreement with previous studies (O'Brien, et al., 2018; Thorborg, et al., 2014; Thorborg, et al., 2011). As regards to other degrees of freedom and reciprocal ratios, the present study is, to our knowledge, the first one to report values for the assessment of hip strength in FLEX, EXT, ER and IR in an athletic population during isometric tasks. This seemed essential to us since groin injuries are not only adductor-related but also ilio-psoas related (Hölmich, et al., 2013). On average, EXT/FLEX ratio was 1.33 ± 0.25 whereas ER/IR ratio was 0.92 ± 0.30, indicating stronger hip extensors compared to flexor muscles, and hip rotators having almost similar strength. More studies are needed to assess whether there is a relationship between isometric hip EXT/FLEX or ER/IR ratio and groin injuries in athletes.

Alike force data, results revealed no significant difference between the dominant and non-dominant leg regarding at RFD values. However, it can be observed that hip rotators produced generally smaller RFD (5.9 $N.s^{-1}.kg^{-1}$) than other hip muscle groups (11.6-16.5 $N.s^{-1}.kg^{-1}$). This difference might be explained by the possible relationship between peak force and RFD previously reported (Andersen & Aagaard, 2006). Muscle function or unusual testing positions for soccer players could also be factors explaining this difference. Regarding reciprocal ratios, results showed that hip extensors produced higher RFD than hip flexors (EXT/FLEX ratio = 1.46 ± 0.33) whereas other muscle groups exhibited substantially similar RFD (ADD/ABD ratio = 1.02 ± 0.26; ER/IR ratio = 1.15 ± 0.46). Considering its key role in sports involving explosive movements, further studies investigating the association between RFD, performance and groin injuries, or differences in RFD profiles based on the sports as well as the playing level could be interesting.

One of the limitations of the present study is that we only assessed the test-retest reliability during a single session and did not investigate the between-day reliability. Thus, more studies are needed to validate the use of the GroinBar for athlete follow-up. In addition, as pointed out by O'Brien, et al. (2018), several tests (ADD, ABD, ER, IR) were assessed bilaterally and could have impacted the reported force values. Indeed, it has been shown that the force produced during simultaneous maximal contraction of both limbs was lower than the sum of the forces produced by the left and right limbs separately (Škarabot, Cronin, Strojnik, & Avela, 2016). However, Simoneau-Buessinger, et al. (2015) revealed that this phenomenon, called the bilateral deficit, is not neural in origin and would be due to dynamometer mechanical configuration and body adjustments. Therefore, reliability results for these test positions are not directly applicable to unilateral assessment. Further studies are needed to compare force values obtained through unilateral and bilateral testing positions with the use of the GroinBar.

5. Conclusion

Assessment of strength and rate of force development around the hip exhibited good to excellent test-retest reliability in asymptomatic varsity soccer players using the GroinBar. Yet, averaged measures over three trials on each degree of freedom should be favoured over peak measures. More studies are needed regarding the assessment of hip flexors as well as the inter-session reliability of the device. Values for both strength and rate of force development were determined and exhibited no significant difference between the dominant and non-dominant leg. Findings of the present study may provide clinicians with the confidence for using the GroinBar in monitoring athletes and minimize their injury risks.

References

Aagaard, P., Simonsen, E. B., Andersen, J. L., Magnusson, P., & Dyhre-Poulsen, P. (2002). Increased rate of force development and neural drive of human skeletal muscle following resistance training. *Journal of applied physiology, 93*, 1318-1326.

Andersen, L. L., & Aagaard, P. (2006). Influence of maximal muscle strength and intrinsic muscle contractile properties on contractile rate of force development. *European journal of applied physiology, 96*, 46-52.

Arnason, A., Sigurdsson, S. B., Gudmundsson, A., Holme, I., Engebretsen, L., & Bahr, R. (2004). Risk factors for injuries in football. *The American journal of sports medicine, 32*, 5-16.

Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports medicine, 26*, 217-238.

Charlton, P. C., Mentiplay, B. F., Grimaldi, A., Pua, Y.-H., & Clark, R. A. (2017). The reliability of a maximal isometric hip strength and simultaneous surface EMG screening protocol in elite, junior rugby league athletes. *Journal of Science and Medicine in Sport, 20*, 139-145.

Engebretsen, A. H., Myklebust, G., Holme, I., Engebretsen, L., & Bahr, R. (2010). Intrinsic risk factors for groin injuries among male soccer players: a prospective cohort study. *The American journal of sports medicine, 38*, 2051-2057.

Esteve, E., Rathleff, M. S., Vicens-Bordas, J., Clausen, M. B., Hölmich, P., Sala, L., & Thorborg, K. (2018). Preseason adductor squeeze strength in 303 Spanish male soccer athletes: A cross-sectional study. *Orthopaedic journal of sports medicine, 6*, 2325967117747275.

Fulcher, M. L., Hanna, C. M., & Elley, C. R. (2010). Reliability of handheld dynamometry in assessment of hip strength in adult male football players. *Journal of Science and Medicine in Sport, 13*, 80-84.

Gissis, I., Papadopoulos, C., Kalapotharakos, V. I., Sotiropoulos, A., Komsis, G., & Manolopoulos, E. (2006). Strength and speed characteristics of elite, subelite, and recreational young soccer players. *Research in sports Medicine, 14*, 205-214.

Hölmich, P., Thorborg, K., Dehlendorff, C., Krogsgaard, K., & Gluud, C. (2013). Incidence and clinical presentation of groin injuries in sub-elite male soccer. *Br J Sports Med*, bjsports-2013-092627.

Kelln, B. M., McKeon, P. O., Gontkof, L. M., & Hertel, J. (2008). Hand-held dynamometry: reliability of lower extremity muscle testing in healthy, physically active, young adults. *Journal of sport rehabilitation, 17*, 160-170.

Krause, D. A., Neuger, M. D., Lambert, K. A., Johnson, A. E., DeVinny, H. A., & Hollman, J. H. (2014). Effects of examiner strength on reliability of hip-strength testing using a handheld dynamometer. *Journal of sport rehabilitation, 23*, 56-64.

Maffiuletti, N. A., Aagaard, P., Blazevich, A. J., Folland, J., Tillin, N., & Duchateau, J. (2016). Rate of force development: physiological and methodological considerations. *European journal of applied physiology, 116*, 1091-1116.

Malliaras, P., Hogan, A., Nawrocki, A., Crossley, K., & Schache, A. (2009). Hip flexibility and strength measures: reliability and association with athletic groin pain. *British journal of sports medicine*.

Mentiplay, B. F., Perraton, L. G., Bower, K. J., Adair, B., Pua, Y.-H., Williams, G. P., McGaw, R., & Clark, R. A. (2015). Assessment of lower limb muscle strength and power using hand-held and fixed dynamometry: a reliability and validity study. *PloS one, 10*, e0140822.

Mohammad, W. S., Abdelraouf, O. R., Elhafez, S. M., Abdel-Aziem, A. A., & Nassif, N. S. (2014). Isokinetic imbalance of hip muscles in soccer players with osteitis pubis. *Journal of sports sciences, 32*, 934-939.

Moreno-Pérez, V., Travassos, B., Calado, A., Gonzalo-Skok, O., Del Coso, J., & Mendez-Villanueva, A. (2019). Adductor Squeeze Test and Groin Injuries in elite Football players: A prospective study. *Physical Therapy in Sport*.

Mosler, A. B., Crossley, K. M., Thorborg, K., Whiteley, R. J., Weir, A., Serner, A., & Hölmich, P. (2017). Hip strength and range of motion: normal values from a professional football league. *Journal of Science and Medicine in Sport, 20*, 339-343.

O'Brien, M. M., Bourne, M. M., Heerey, M. J., Timmins, R., & Pizzari, T. (2018). A novel device to assess hip strength: Concurrent validity and normative values in male athletes. *Physical Therapy in Sport*.

Portney, L. G. (2008). Statistical measures of reliability. *Foundation of clinical research. Applications to practice*, 585-618.

Ryan, J., DeBurca, N., & Mc Creesh, K. (2014). Risk factors for groin/hip injuries in field-based sports: a systematic review. *Br J Sports Med, 48*, 1089-1096.

Ryan, S., Kempton, T., Pacecca, E., & Coutts, A. J. (2018). Measurement Properties of an Adductor Strength Assessment System in Professional Australian Footballers. *International journal of sports physiology and performance*, 1-13.

Scott, D. A., Bond, E. Q., Sisto, S. A., & Nadler, S. F. (2004). The intra-and interrater reliability of hip muscle strength assessments using a handheld versus a portable dynamometer anchoring station 1. *Archives of physical medicine and rehabilitation, 85*, 598-603.

Simoneau-Buessinger, E., Leteneur, S., Toumi, A., Dessurne, A., Gabrielli, F., Barbier, F., & Jakobi, J. M. (2015). Bilateral strength deficit is not neural in origin; rather due to dynamometer mechanical configuration. *PloS one, 10*, e0145077.

Škarabot, J., Cronin, N., Strojnik, V., & Avela, J. (2016). Bilateral deficit in maximal force production. *European journal of applied physiology, 116*, 2057-2084.

Stark, T., Walker, B., Phillips, J. K., Fejer, R., & Beck, R. (2011). Hand-held dynamometry correlation with the gold standard isokinetic dynamometry: a systematic review. *PM&R, 3*, 472-479.

Thorborg, K., Bandholm, T., Schick, M., Jensen, J., & Hölmich, P. (2013). Hip strength assessment using handheld dynamometry is subject to intertester bias when testers are of different sex and strength. *Scandinavian journal of medicine & science in sports, 23*, 487-493.

Thorborg, K., Branci, S., Nielsen, M. P., Tang, L., Nielsen, M. B., & Hölmich, P. (2014). Eccentric and isometric hip adduction strength in male soccer players with and without adductor-related groin pain: an assessor-blinded comparison. *Orthopaedic journal of sports medicine, 2*, 2325967114521778.

Thorborg, K., Petersen, J., Magnusson, S., & Hölmich, P. (2010). Clinical assessment of hip strength using a hand-held dynamometer is reliable. *Scandinavian journal of medicine & science in sports, 20*, 493-501.

Thorborg, K., Serner, A., Petersen, J., Madsen, T. M., Magnusson, P., & Hölmich, P. (2011). Hip adduction and abduction strength profiles in elite soccer players: implications for clinical evaluation of hip adductor muscle recovery after injury. *The American journal of sports medicine, 39*, 121-126.

Verrall, G. M., Slavotinek, J. P., Barnes, P. G., & Fon, G. T. (2005). Description of pain provocation tests used for the diagnosis of sports-related chronic groin pain: relationship of tests to defined clinical (pain and tenderness) and MRI (pubic bone marrow oedema) criteria. *Scandinavian journal of medicine & science in sports, 15*, 36-42.

Figure

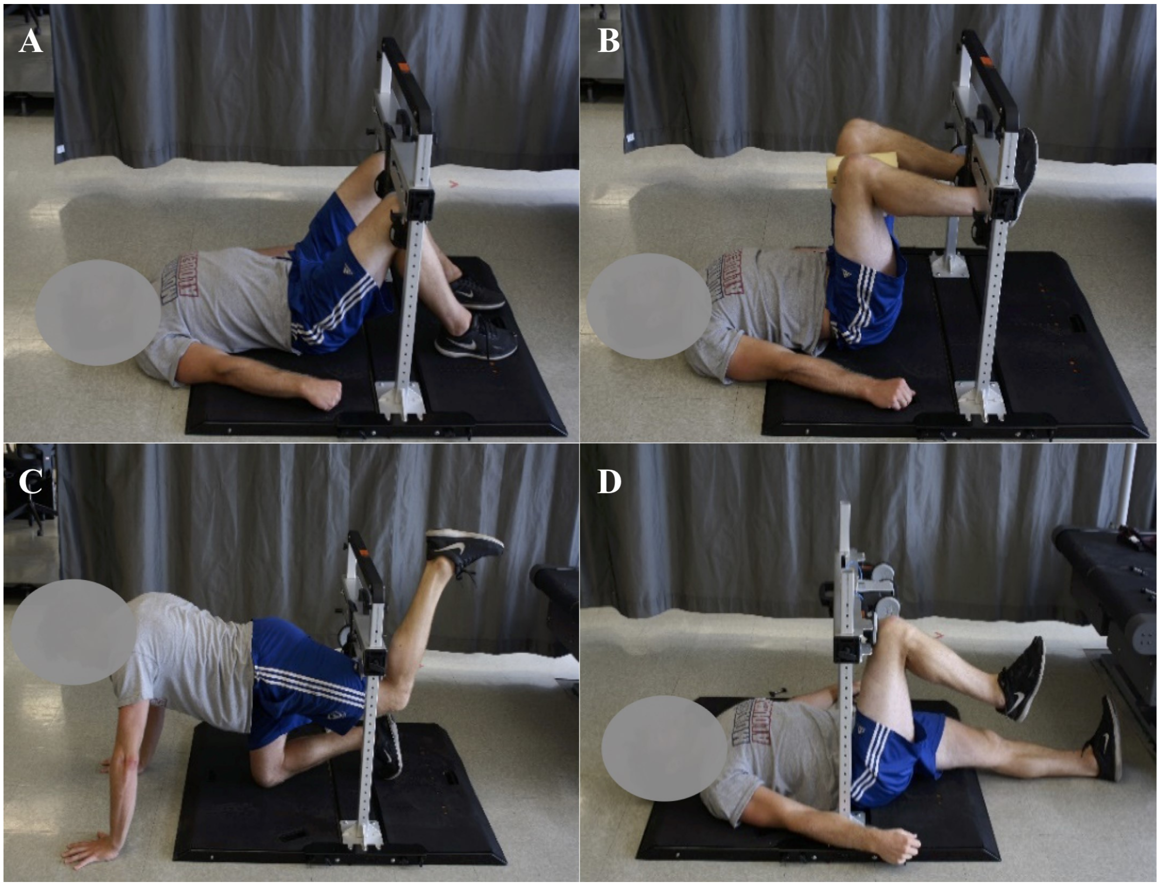

Fig 1. Testing positions for force and rate of force assessment in the GroinBar. (A) Hip adductors and abductors with the participant lying supine and hips and knees flexed at 45°. Sensors placed on the medial femoral condyles for adduction and lateral femoral condyles for abduction. (B) Hip internal and external rotators with the participant lying supine and hips and knees flexed at 90°. Sensors placed on the medial malleoli for internal rotation and lateral malleoli for external rotation. (B) Hip extensors with the participant in a quadruped position and hips extended and knees flexed at 90°. Sensor placed on the distal part of the hamstrings. (C) Hip flexors with the participant lying supine and hips and knees flexed at 90°. Sensor placed on the distal part of the quadriceps.

Table 1. Test-retest reliability of normalized hip force assessment using the GroinBar

Measures	Test* ($N.kg^{-1}$)	Retest* ($N.kg^{-1}$)	ICC (95% CI)	SEM ($N.kg^{-1}$)	SEM%	Bias	95% LOA	
							Lower	Upper
Peak force								
ADD	4.9 ± 1.1	4.8 ± 1.0	0.85 (0.74 – 0.92)	0.39	8.2	0.1	-0.9	1.2
ABD	4.7 ± 0.6	4.6 ± 0.7	0.82 (0.67 – 0.90)	0.29	6.2	0.1	-0.6	0.9
ER	1.8 ± 0.3	1.8 ± 0.3	0.88 (0.79 – 0.94)	0.10	5.7	0.0	-0.3	0.3
IR	2.1 ± 0.6	2.2 ± 0.6	0.77 (0.60 – 0.87)	0.28	13.0	-0.1	-0.9	0.6
EXT	4.2 ± 1.1	4.3 ± 1.0	0.75 (0.58 – 0.86)	0.52	12.2	-0.1	-1.5	1.4
FLEX	3.2 ± 0.6	3.1 ± 0.5	0.53 (0.26 – 0.72)	0.37	11.8	0.0	-1.0	1.1
Average peak force								
ADD	4.7 ± 1.1	4.5 ± 1.0	0.92 (0.85 – 0.93)	0.29	6.3	0.2	-0.9	1.2
ABD	4.5 ± 0.6	4.4 ± 0.7	0.90 (0.80 – 0.95)	0.21	4.7	0.1	-0.6	0.9
ER	1.7 ± 0.3	1.7 ± 0.3	0.95 (0.90 – 0.97)	0.07	4.1	0.0	-0.3	0.3
IR	1.9 ± 0.5	2.0 ± 0.5	0.90 (0.80 – 0.95)	0.17	8.8	-0.1	-0.7	0.5
EXT	3.9 ± 1.0	4.0 ± 1.0	0.86 (0.74 – 0.93)	0.37	9.4	0.0	-1.4	1.3
FLEX	3.0 ± 0.5	3.0 ± 0.5	0.77 (0.55 – 0.88)	0.25	8.4	0.0	-0.9	0.9

*Values are normalized forces of the left and right leg combined.

ABD: abduction, ADD: adduction, ER: external rotation, IR: internal rotation, EXT: extension, FLEX: flexion, ICC: intraclass correlation coefficient, 95% CI: 95% confidence interval, SEM: Standard error of measurement, Bias: difference between test and retest (test-retest), 95% LOA: 95% limits of agreement.

Table 2. Test-retest reliability of hip RFD assessment using the GroinBar

Measures	Test ($N.s^{-1}.kg^{-1}$)	Retest ($N.s^{-1}.kg^{-1}$)	ICC (95% CI)	SEM ($N.s^{-1}.kg^{-1}$)	SEM%	Bias	95% LOA	
							Lower	Upper
Peak RFD								
ADD	13.9 ± 4.0	12.9 ± 4.5	0.81 (0.65 – 0.90)	1.8	13.8	1.0	-3.9	5.9
ABD	14.3 ± 2.8	13.2 ± 2.8	0.68 (0.42 – 0.83)	1.6	11.6	1.0	-3.1	5.2
ER	5.1 ± 1.3	5.4 ± 1.2	0.80 (0.65 – 0.89)	0.6	10.7	-0.3	-1.8	1.2
IR	5.3 ± 2.1	5.4 ± 1.9	0.84 (0.71 – 0.91)	0.8	15.4	-0.1	-2.4	2.2
EXT	13.9 ± 4.0	13.8 ± 4.4	0.61 (0.36 – 0.77)	2.6	19.1	0.1	-7.3	7.5
FLEX	9.8 ± 2.6	10.2 ± 2.8	0.72 (0.53 – 0.84)	1.4	14.4	-0.4	-4.3	3.6
Average peak RFD								
ADD	12.0 ± 3.7	11.2 ± 4.2	0.92 (0.85 – 0.96)	1.1	9.5	0.8	-3.1	4.7
ABD	12.6 ± 3.0	11.7 ± 2.6	0.84 (0.67 – 0.92)	1.1	9.2	0.9	-2.9	4.7
ER	4.6 ± 1.3	4.8 ± 1.2	0.91 (0.83 – 0.95)	0.4	8.2	-0.1	-1.6	1.3
IR	4.6 ± 1.8	4.7 ± 1.9	0.91 (0.84 – 0.95)	0.5	11.7	-0.1	-2.1	2.0
EXT	11.7 ± 3.3	12.0 ± 4.2	0.81 (0.63 – 0.90)	1.6	13.9	-0.3	-6.3	5.7
FLEX	8.6 ± 2.5	9.0 ± 2.5	0.90 (0.81 – 0.95)	0.8	8.8	-0.4	-3.2	2.4

*Values are normalized forces of the left and right leg combined.

ABD: abduction, ADD: adduction, ER: external rotation, IR: internal rotation, EXT: extension, FLEX: flexion, ICC: intraclass correlation coefficient, 95% CI: 95% confidence interval, SEM: Standard error of measurement, Bias: difference between test and retest (test-retest), 95% LOA: 95% limits of agreement.

Table 3. Hip muscle function (force and RFD) and reciprocal ratios between dominant and non-dominant legs

Measures	Dominant	Non-dominant	Asymmetry (%)
Average peak force ($N.kg^{-1}$)*			
ADD	4.6 ± 1.0	4.5 ± 1.0	2.0 ± 8.1
ABD	4.5 ± 0.7	4.4 ± 0.7	2.1 ± 4.5
ER	1.7 ± 0.3	1.7 ± 0.4	4.7 ± 5.4
IR	2.0 ± 0.5	1.9 ± 0.5	-0.5 ± 7.4
EXT	4.0 ± 0.9	3.8 ± 0.9	0.6 ± 14.1
FLEX	3.0 ± 0.4	3.0 ± 0.5	-2.3 ± 12.3
Reciprocal force ratios*			
ADD:ABD ratio	1.05 ± 0.20	1.05 ± 0.22	
ER:IR ratio	0.95 ± 0.31	0.88 ± 0.26	
EXT:FLEX ratio	1.36 ± 0.24	1.33 ± 0.27	
Average peak RFD ($N.s^{-1}.kg^{-1}$)*			
ADD	11.8 ± 4.0	11.4 ± 3.9	2.9 ± 8.6
ABD	12.3 ± 2.8	12.0 ± 2.5	2.1 ± 4.3
ER	4.8 ± 1.2	4.6 ± 1.2	4.9 ± 9.8
IR	4.6 ± 1.9	4.6 ± 1.7	-2.7 ± 10.0
EXT	11.9 ± 3.5	11.8 ± 3.4	-3.7 ± 21.7
FLEX	8.6 ± 2.4	8.9 ± 2.4	-5.0 ± 19.1
Reciprocal RFD ratios*			
ADD:ABD ratio	0.97 ± 0.23	0.96 ± 0.26	
ER:IR ratio	1.20 ± 0.51	1.12 ± 0.47	
EXT:FLEX ratio	1.39 ± 0.28	1.35 ± 0.31	

*Values are mean of test and retest.
ABD: abduction, ADD: adduction, ER: external rotation, IR: internal rotation, EXT: extension, FLEX: flexion

Supplementary file

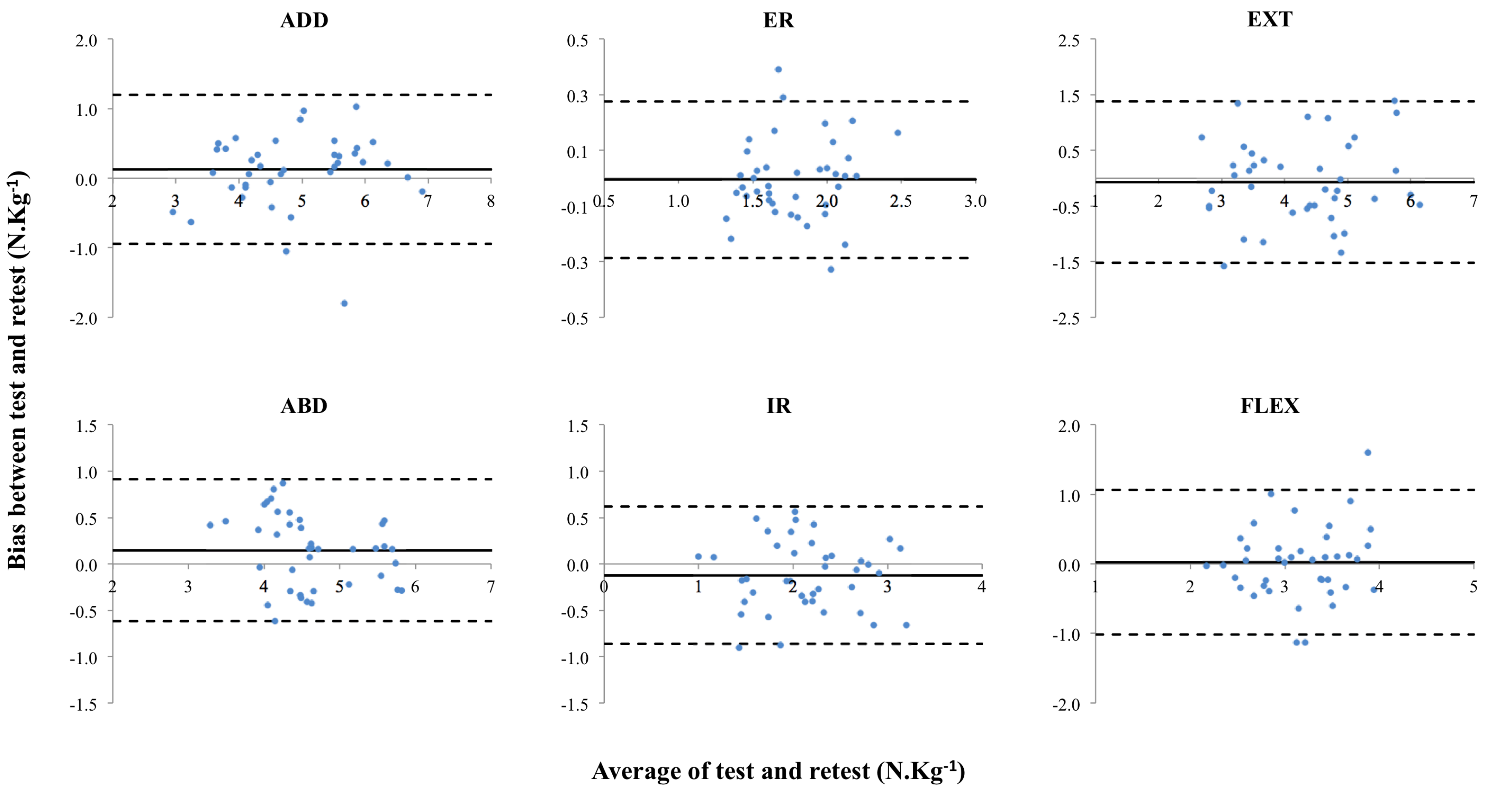

Fig. S1: Bland-Altman plots showing the difference between test and retest measurement of hip ADD, ABD, ER, IR, EXT and FLEX peak force. The solid line represents the mean difference (bias) and the dotted lines are the limits of agreement (±1.96 SD).

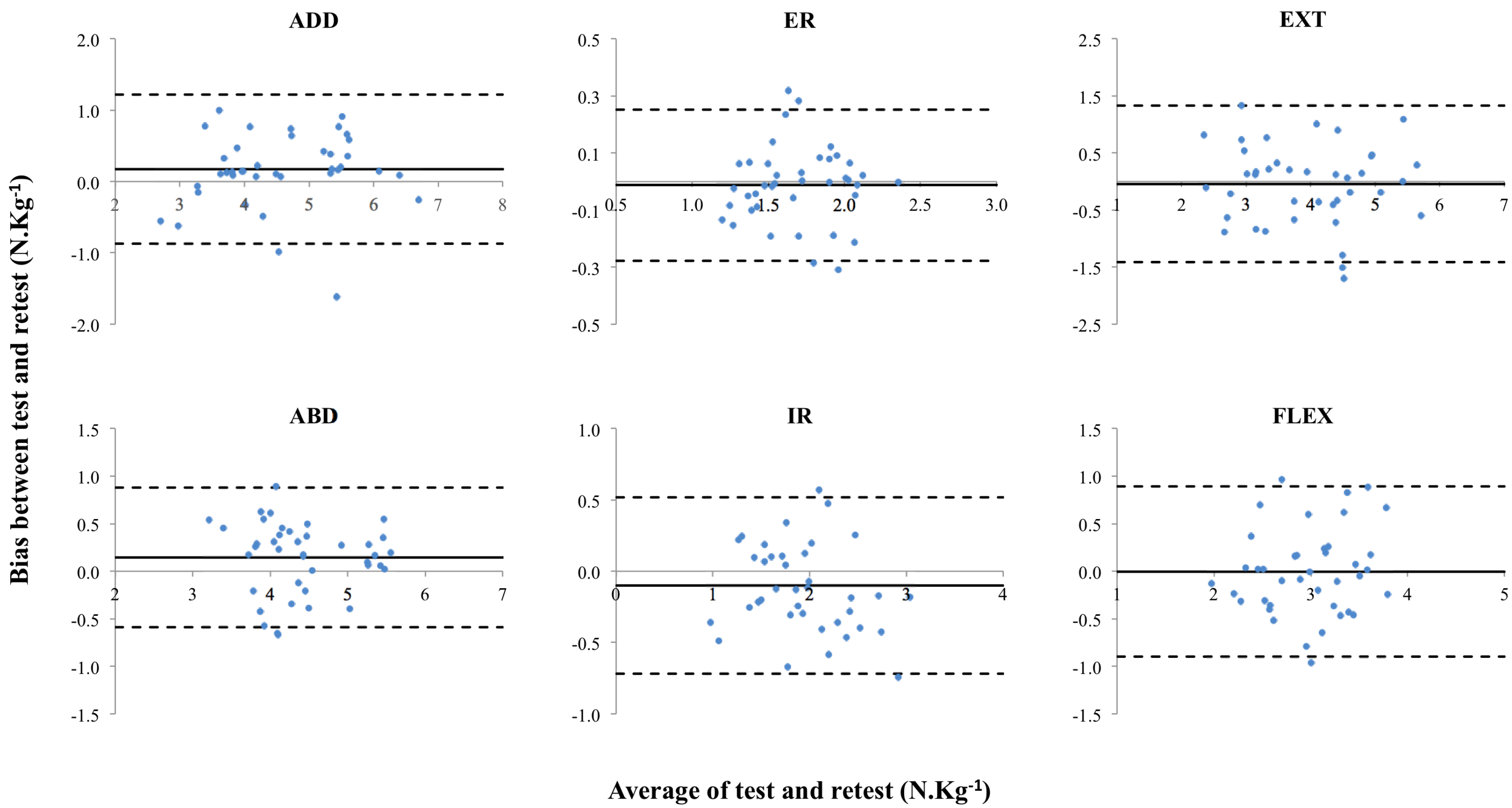

Fig. S2: Bland-Altman plots showing the difference between test and retest measurement of hip ADD, ABD, ER, IR, EXT and FLEX <u>average peak force</u>. The solid line represents the mean difference (bias) and the dotted lines are the limits of agreement (±1.96 SD).

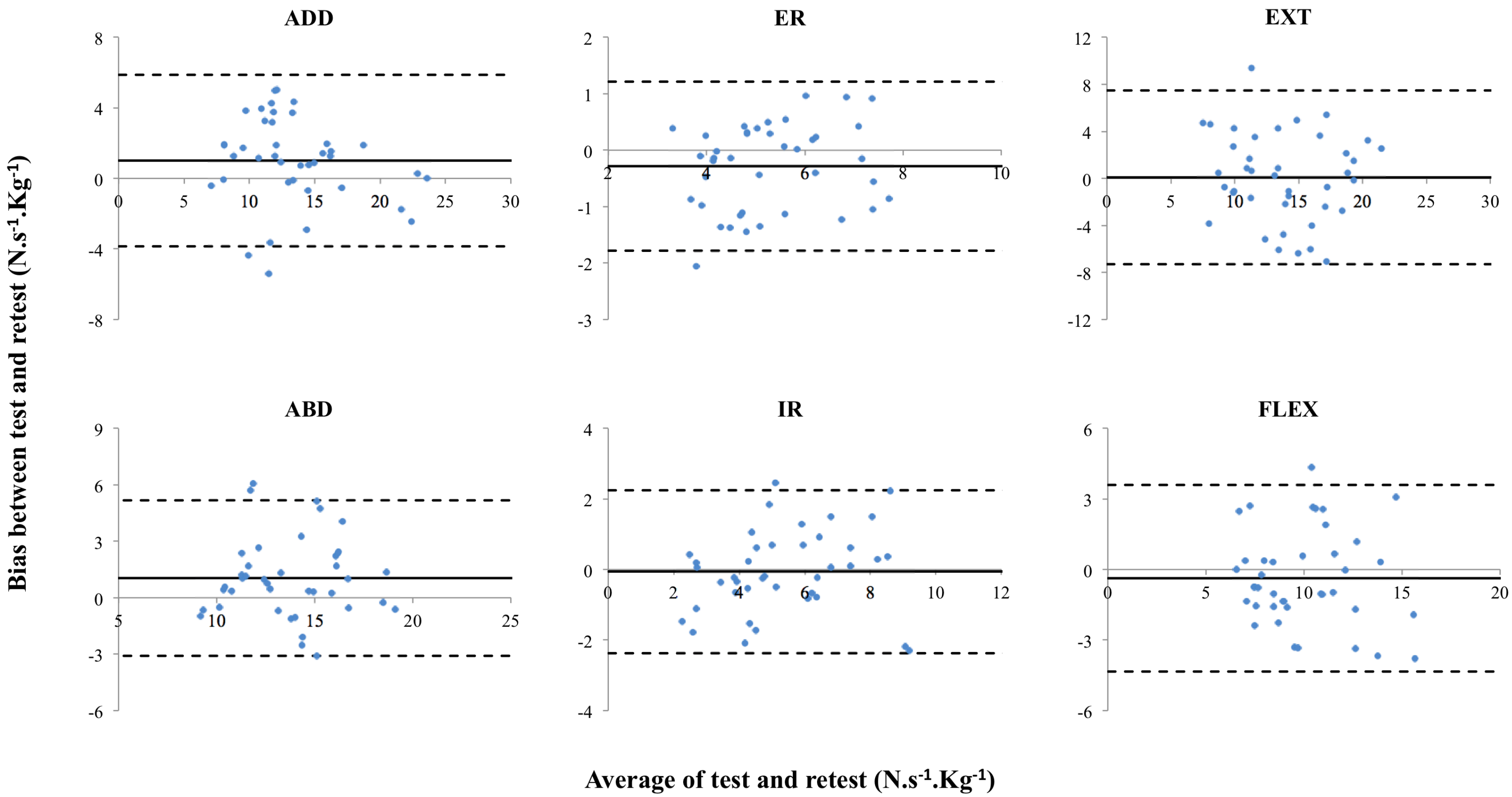

Fig. S3: Bland-Altman plots showing the difference between test and retest measurement of hip ADD, ABD, ER, IR, EXT and FLEX peak rate of force development. The solid line represents the mean difference (bias) and the dotted lines are the limits of agreement (±1.96 SD).

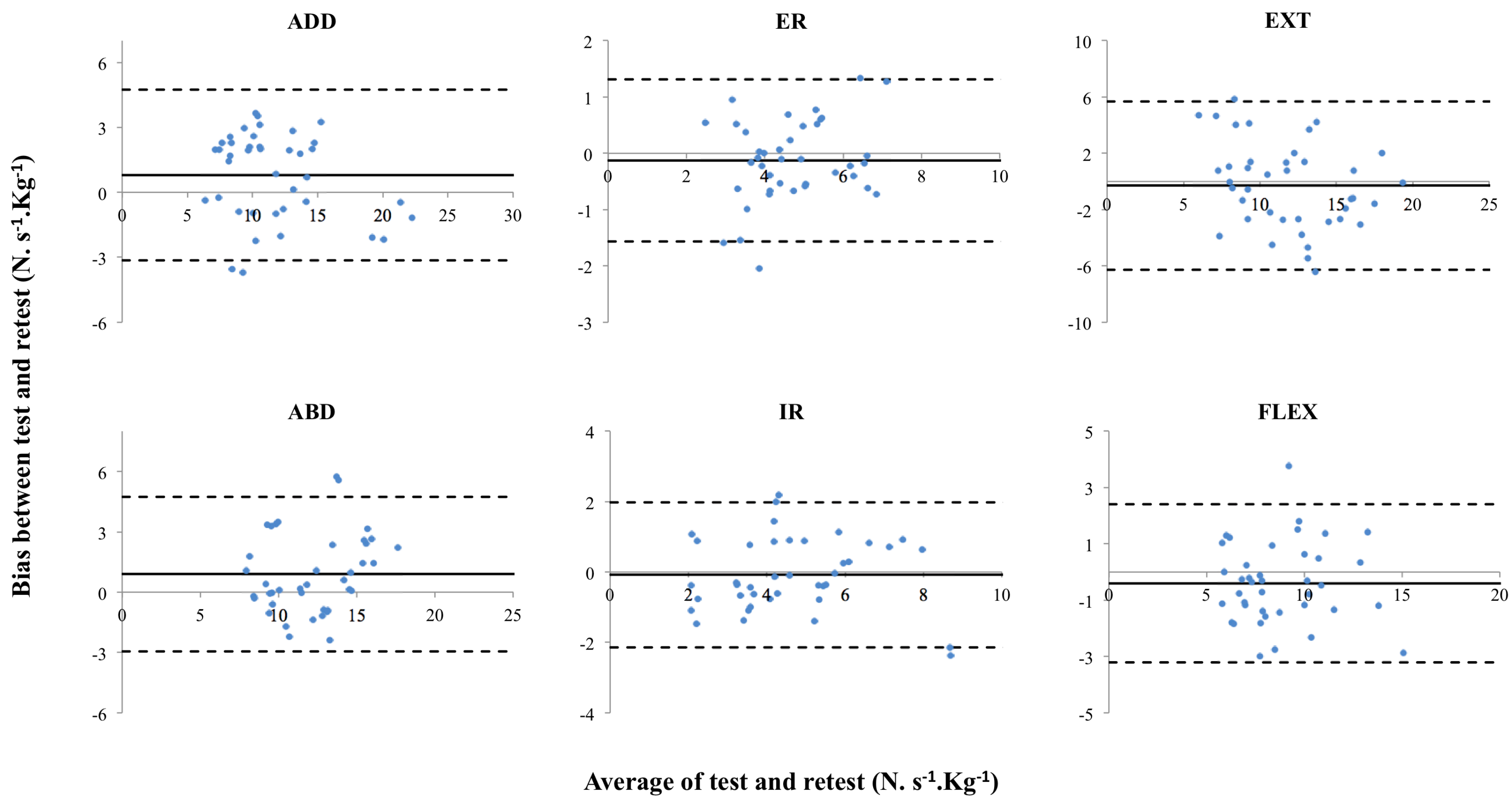

Fig. S4: Bland-Altman plots showing the difference between test and retest measurement of hip ADD, ABD, ER, IR, EXT and FLEX <u>average peak rate of force development</u>. The solid line represents the mean difference (bias) and the dotted lines are the limits of agreement (±1.96 SD).