



Original Research

Intrarater Test-Retest Reliability of Hip Abduction, Internal Rotation, and External Rotation Strength Measurements in a Healthy Cohort Using a Handheld Dynamometer and a Portable Stabilization Device: A Pilot Study



Aaron Byrne, BSc, Clare Lodge, MSc, DPT,
Jennifer Wallace, MSc

healthCORE, Department of Science & Health, Institute of Technology Carlow, Carlow, Ireland

KEYWORDS

Clinical decision-making;
Rehabilitation

Abstract Objective: To investigate the within-day and between-day test-retest reliability of hip abduction, internal rotation, and external rotation strength measurements taken using a portable device externally stabilizing a handheld dynamometer in healthy participants.

Design: Observational study.

Setting: Institute of Technology Carlow, Ireland, a third-level education institute.

Participants: Healthy participants (N=18; 11 male, 7 female) who participate in a field sport for more than 2 hours per week were recruited via convenience sampling.

Interventions: Not applicable.

Main Outcome Measures: Hip abduction, internal rotation, and external rotation peak force during a maximal voluntary isometric contraction (N). The 3 best values recorded for each movement for each day were used to analyze within-day and between-day test-retest reliability. Intraclass correlation coefficients (ICCs), coefficients of variance, standard error of measurement, and minimal detectable change statistics were also calculated.

Results: External fixation of a handheld dynamometer produced excellent test-retest reliability for within-day (ICC>0.934) and between-day (ICC>0.802) contexts.

Conclusions: Clinical measurements of hip strength can be performed reliably, efficiently, and cost effectively using the methods described. Furthermore, the use of external fixation eliminates the influence of tester strength on the handheld dynamometry measurements.

List of abbreviations: CV, coefficient of variance; ERot, external rotation; IRot, internal rotation; HHD, handheld dynamometry; HHDstab, HHD and stabilization device; ICC, intraclass correlation coefficient; MDC, minimal detectable change; PVC, polyvinyl chloride; SEM, standard error of measurement.

Disclosures: none.

Cite this article as: Arch Rehabil Res Clin Transl. 2020;2:100050.

<https://doi.org/10.1016/j.arrct.2020.100050>

2590-1095/© 2020 The Authors. Published by Elsevier Inc. on behalf of the American Congress of Rehabilitation Medicine. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

© 2020 The Authors. Published by Elsevier Inc. on behalf of the American Congress of Rehabilitation Medicine. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Hip strength is commonly measured in sports and musculoskeletal medicine as part of an objective assessment or as a marker for recovery. Hip strength has also been associated with injury incidence rates. Athletes who sustained a lower limb injury during a 2-season period reported significantly lower hip abduction strength ($P=.02$, 3% body weight) and hip external rotation (ERot) strength ($P=.001$, 2.7% body weight) when compared with their counterparts who did not sustain an injury.¹ Furthermore, when expressed as a percentage of body weight, hip abduction and ERot strength of less than 35.4% and 20.3%, respectively, classified an athlete as high risk for sustaining a noncontact anterior cruciate ligament injury.² Deficits in hip strength have also been associated with many other conditions, such as ankle ligament sprains, patella-femoral pain syndrome, iliotibial band syndromes, groin strains, hip pain, and low back pain.³⁻⁹ Furthermore, a recent consensus statement recommends future research to "investigate, report and improve the measurement properties of tests of...muscle strength and functional performance."^{9(p3)}

Lateral hip musculature, namely gluteus medius, is fundamental in hip abduction while also contributing to hip ERot and internal rotation (IRot) in varying proportions depending on hip position.^{10,11} Gluteus medius activity is notably high during single-leg tasks,^{12,13} illustrating its important contribution to lumbopelvic hip or core stability, which, along with hip strength, is a major target of many neuromuscular training programs used for injury prevention purposes.¹⁴⁻¹⁷ Therefore, reliable clinical strength measurements for all movements to which gluteus medius can contribute to are important for rehabilitation clinicians in assessment, tracking progress post injury, or monitoring the effects of interventions carried out, such as neuromuscular training programs.

The current and most commonly used strength measurement technique is manual muscle testing,¹⁸ which consists of a clinician's subjective rating of force along the Oxford Muscle Grading Scale from 0-5, with 0 being no palpable muscle contraction and 5 being normal full muscle performance.¹⁹ Although widespread in clinical practice over a large array of professions, its subjective nature and inability to be used to truly express small strength differences are some of its limitations.²⁰

Previous research has led to the introduction and practice of handheld dynamometry (HHD) as an alternative to manual muscle testing, providing clinicians with an objective, numerical measurement of muscle generated force.^{18,21} HHD has also become more common in the scientific literature with normative HHD values reported for strength testing in the literature.²¹ HHD has previously been shown to be valid and comparable with the criterion

standard in strength testing and isokinetic dynamometry, without sacrificing ease of use, portability, or cost.^{22,23}

HHD is not without limitations. Research dating back to 1991 highlights the importance of tester strength in the accuracy of HHD measurements, particularly upper limb tester strength and its inverse relationship with strength values recorded by HHD.²⁴ These reliability discrepancies are most common in stronger movements of $>120\text{N}$,²⁵ as may be expected across lower limb movements or in highly trained individuals in particular.^{21,26}

These findings have led to the development of externally stabilized dynamometers. Examples include belt fixation to an adjacent wall²⁷ or fixation through the construction of cage-like structures around a treatment plinth.²⁸ Both the aforementioned studies resulted in satisfactory reliability for hip strength values (intraclass correlation coefficient [ICC], 0.76-0.95 and 0.73-0.89, respectively), but these procedures may not be as time efficient as traditional handheld measurement methods.

A much simpler solution was recently proposed by using a polyvinyl chloride (PVC) pipe-like structure that could be placed between the limb being tested and a wall.²⁹ One end was designed to accommodate the MicroFET2 dynamometer, and the other end, a flat plate, was designed to aid in its stability against the wall. Excellent reliability was established with ICCs for hip abduction and ERot strength measurements (ICC=0.96 and 0.98, respectively) across 30 limbs tested in 15 participants; however, researchers omitted IRot measurement and did not investigate the between-day reliability of these methods.

The aim of this current study was to establish intratester reliability when measuring the strength of hip abduction, IRot, and ERot by the use of a simple pipe-like stabilization device coupled with a MicroFET2 dynamometer and additionally to explore the between-day reliability of these strength values. This article was formulated in accordance to the Guidelines for Reporting Reliability and Agreement Studies.³⁰

Methods

Participants

Convenience sampling was used to recruit 18 (11 male, 7 female) participants from the Institute of Technology Carlow. Sample size requirements for ICCs were predetermined with $R_0=0.0$, $R_1=0.7$ (as established during pilot testing), and statistical power=0.9. The calculated sample size was 13; however, to allow for potential drop-outs, 18 was the target sample size.³¹ Subjects were deemed eligible if they participated in a field sport for more than 2 hours per week. Subjects were excluded if they

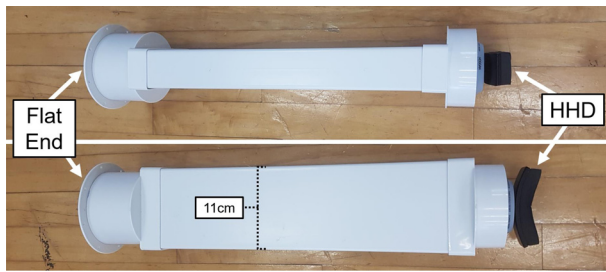


Fig 1 HHDstab construction. Abbreviation: HHDstab, HHD and stabilization device.

had any incidence of injury to the lower back, hip, knee, ankle, or foot of their self-selected preferred jumping leg (leg which they were most likely jump off) in the past 6 months.

Ethical considerations

Ethical approval for the study was granted by the Ethics in Research committee of the Institute of Technology Carlow, Ireland. Following a description of the study, individuals were recruited for participation. Written informed consent and medical screening questionnaires were also collected prior to the initiation of testing procedures. There was no financial inducement offered to participants, and no participants were in a dependent relationship to either the lead researcher or research supervisors at the time of testing. Participants were also free to withdraw from participation at any time. Personal information was protected in accordance to the Institute of Technology Carlow Data Protection Policy and General Data Protection Regulation guidelines. This study was conducted as part of a PhD research program funded by the President's Fellowship Scheme at the Institute of Technology Carlow, Ireland.

Instrumentation

A MicroFET2 dynamometer^a was used to obtain all strength measurements. The stabilization device was constructed using a PVC pipe, 11 cm wide, and adjoining duct pieces that were bonded together with adhesive so that 1 side contained a 100-mm diameter circular opening that accommodated the shape of the handheld dynamometer securely, and the opposing end consisted of a flat surface that would lay against the wall during testing procedures (fig 1).

An adjustable treatment plinth (Plinth 2000^b), which was sourced from a National Health Service approved supplier, was used for all participants.

Testing procedures

All measurements were performed on the participants' self-reported preferred jumping leg by a single tester, a certified athletic therapist. A predefined script was used to describe the tests so as not to bias efforts exerted by participants. Testing took place on 2 occasions, 3 days apart, in the physiology laboratory at Institute of Technology Carlow.

Procedures as outlined hereafter, plinth height and position in proximity to the wall, apparatus used, rest periods, and time of day were replicated between both testing days. Participants were also urged to abstain from high-intensity exercise for the 24 hours preceding both testing sessions.

Peak force in newtons over a 5-second maximal voluntary isometric contraction was recorded for each movement. For each measurement, the pad of the handheld dynamometer was positioned 5 cm proximal to the malleoli³² with the HHD and stabilization device (HHDstab) perpendicular to the wall and supported by the tester.²⁹ All trials were separated by a 30-second rest period. Four trials were recorded for each movement with the best (highest) 3 scores tabulated for analysis.^{27,32}

Hip abduction measure

Hip abduction strength was recorded with the participant lying supine on the treatment plinth, positioned parallel to the adjacent wall. A belt was secured around the participant and plinth, resting on both anterior superior iliac spines to limit lateral pelvic motion during testing. The HHDstab was positioned perpendicular to wall and the target leg, contacting the leg 5 cm proximal to the lateral malleolus (fig 2). The participant was then instructed to "cross your arms over your chest and push into the pad as hard as possible" for 5 seconds.

Hip internal rotation measure

Hip IRot strength was recorded with the participant seated on the end of the treatment plinth, thigh parallel to the adjacent wall and hip in neutral rotation. A belt was secured around participant and plinth, on the superior femur, with a standardized 11-cm wide piece of PVC positioned between the knees to maintain knee position. The HHDstab was positioned between the wall and the target leg, contacting the leg 5 cm proximal to the lateral malleolus (fig 2). The participant was then asked to "keep both hands on top of the pipe, squeeze both knees together and push into the pad as hard as possible" for 5 seconds.

Hip external rotation measure

Hip ERot strength was recorded with the participant seated on the opposite end of the treatment plinth to the IRot measurement position, with thigh parallel to the adjacent wall and hip in neutral rotation. For ERot, the target leg was the leg furthest away from the wall, and the longer length PVC device was used so that the plinth could remain in situ. A belt was secured around participant and plinth, on the superior femur, with a standardized 11-cm wide length of PVC positioned between the knees to maintain knee position. The HHDstab was positioned between the wall and the target leg, contacting the leg 5 cm proximal to the medial malleolus. The nontest leg was flexed so as to lie behind the HHDstab (see fig 2). The participant was then asked to "keep both hands on top of the pipe, squeeze both knees together and push into the pad as hard as possible" for 5 seconds.



Fig 2 Hip abduction, internal rotation, and external rotation strength testing positions.

Statistical analysis:

All data was tabulated and analyzed using SPSS version 23^c and Microsoft Excel 2013.^d Means, standard deviations, coefficients of variance percentage (CV%), ICCs, and respective 95% CIs were calculated within SPSS with $\alpha=0.05$ and $1-\beta=0.95$. ICC (3,1) was applied in within-day analyses, with ICC (3,k) applied in between-day analyses for intrarater reliability.³³⁻³⁵ ICC statistics were classified within the following ranges: poor (0-0.39), fair (0.4-0.59), good (0.6-0.74), or excellent (0.75-1).³³ The standard error of measurement (SEM) and minimal detectable change (MDC⁹⁵) were calculated for both within-day and between-day reliability analyses using the following formulas^{29,33,34}:

- $SEM = SD \times \sqrt{1-r}$, (with r being the ICC value calculated prior)
- $MDC^{95} = 1.96 \times \sqrt{2} \times SEM$

Results

Participant sex, age, preferred jumping leg, and body mass are presented in table 1.

Within-day test-retest reliability statistics for strength measurements were highly reliable with all ICC values >0.934 , CV% $<6.2\%$, and the largest MDC⁹⁵ value was 5.09 N, which was recorded in IRot strength.

Similar to within-day reliability, between-day reliability statistics for strength measurements were excellent, with all ICC values >0.802 , CV% $<14.7\%$, and the MDC⁹⁵ value was 13.41 N for ERot strength (table 2).

Discussion

Findings from this current study suggest that external stabilization of a handheld dynamometer provided excellent

reliability of measurements of hip abduction, IRot strength, and ERot strength in both within-day and between-day conditions. The methodologies conducted in this study took approximately 8 minutes to complete, including landmarking, positioning, and 4 repetitions of each specific movement with a minimum of 30 seconds rest allotted between repetitions, demonstrating its time efficient nature, which is ideal for clinical settings.

Within-day reliability for abduction and ERot strength was excellent³³ (ICC = 0.947 and 0.961, respectively) (fig 3). The abduction and ERot reliability observed in the current study was comparable with previous research using a similar stabilization device (ICC = 0.96 and 0.98, respectively).²⁹ In addition, IRot strength was measured with similarly excellent reliability (ICC = 0.934) as the aforementioned movements. MDC⁹⁵ values for within-day reliability were also low, the largest of which was in IRot at 5.09 N. Any change in hip strength seen immediately >5.09 N, or 3.85% of maximum muscle force, would suggest a change that cannot be attributed to measurement error alone.³⁴ The outlined procedures are therefore more sensitive to detect change than nonstabilized HHD measurements taken in comparable positions for

Table 1 Participant characteristics

Characteristics	Female (n=7)	Male (n=11)	Total (n=19)
Age (y), mean \pm SD	22.9 \pm 2.7	21.4 \pm 1.6	21.9 \pm 2.2
Weight (kg), mean \pm SD	73.2 \pm 17.1	75.4 \pm 12.2	74.6 \pm 13.9
Preferred jumping leg	L=4 R=3	L=8 R=3	L=12 R=6

Abbreviations: L, left; R, right.

Table 2 Within-day and between-day test-retest reliability statistics

Movement	Within-Day Reliability (n=18)						
	Trial 1 (N)	Trial 2 (N)	Trial 3 (N)	ICC (3,1) (95% CI)	CV (%)	SEM (N)	MDC ⁹⁵ (N)
Abduction	117.37±43.78	115.90±41.96	117.72±41.74	0.947 (0.887-0.978)	6.2	1.75	4.85
Internal rotation	132.24±36.77	134.69±37.01	129.50±32.95	0.934 (0.863-0.973)	5.2	1.84	5.09
External rotation	74.44±24.96	76.36±25.16	74.48±26.07	0.961 (0.917-0.984)	6.1	0.85	2.36
Movement	Between-Day Reliability (n=18)						
	Day 1 (N)	Day 2 (N)	ICC (3,k) (95% CI)	CV (%)	SEM (N)	MDC ⁹⁵ (N)	
Abduction	117.00±41.74	121.57±35.30	0.953 (0.875-0.982)	8.4	2.11	5.86	
Internal rotation	132.14±34.84	133.06±32.26	0.928 (0.806-0.973)	8.3	2.65	7.34	
External rotation	75.09±25.07	70.64±21.11	0.802 (0.470-0.926)	14.7	4.84	13.41	

Abbreviations: ICC (3,1), intraclass correlation coefficient, 2-way mixed-effects, single measures; ICC (3,k), intraclass correlation coefficient, 2-way mixed-effects, average measures.

abduction, ERot strength, and IRot strength (MDC⁹⁵ = 9.4N, 12.4N, and 26.6N, respectively),³² even when those non-stabilized measurements were taken by an experienced tester.

Moreover, from previous research that only examined within-day reliability for a similarly stabilized handheld dynamometer,²⁹ excellent between-day reliability was observed for abduction, IRot strength, and ERot strength (ICC = 0.953, 0.928, and 0.802, respectively) by comparing the averages of the 3 best scores recorded on each day. The largest MDC⁹⁵ value for between-day hip strength measurement was seen in ERot at 13.4 N, or 18.3% of maximum muscle force, indicating that if on measurement by a clinician, hip strength changed by greater than this MDC⁹⁵ value between days, one cannot attribute this change to measurement error alone.³⁴

Through the addition of IRot strength measurement, the protocol in this current study aims to build on previous research conducted on abduction and ERot strength measurement without sacrificing portability, cost, or time. The addition of a standard 11-cm wide pipe section keeps femoral position consistent across all tests unlike the nonuniform towel used previously.²⁹ The addition of IRot measurement to the already established abduction and ERot reliability provides clinicians with an accessible method to measure hip abduction and rotational strength, which may be of particular importance to rehabilitation clinicians.⁹

Study limitations

The findings from the current study, although encouraging, should be considered with caution. The current procedures

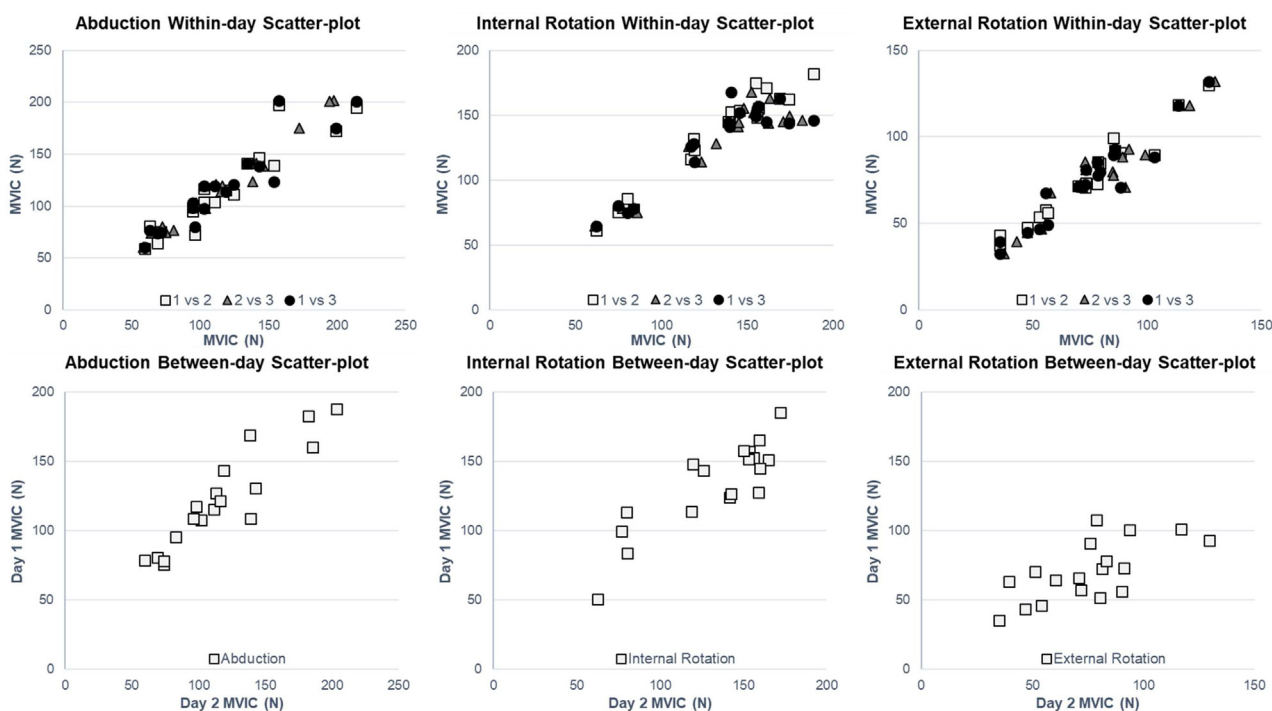


Fig 3 Within-day and between-day scatter plots for abduction, internal rotation, and external rotation.

were only carried out on a healthy, physically active cohort. These same methodologies should be investigated in populations with pathologies prior to its adaptation to clinical practice.

Also, unlike the previous studies that validated HHD measurements by comparing its measurement to isokinetic dynamometry,²² this HHDstab method, to the authors' knowledge, is yet to be validated or directly compared with measurements taken with HHD without external stabilization.

Future research

Future research should focus on directly comparing HHDstab to strength measurements taken with the HHD stabilized manually by the tester. Moreover, validating HHDstab by comparing it with isokinetic dynamometry, and assessing HHDstab reliability in populations with pathologies should be performed prior to its wide-scale adaptation to clinical practice.

Conclusions

The addition of external fixation to HHD addresses a previously documented limitation of HHD. The removal of individual tester strength is possible and provides a high level of consistency in strength assessments about the hip. Hip abduction, iRot strength, and eRot strength can be reliably measured with minimal additional time or financial costs to either clinicians or patients, allowing such objective markers to guide clinical decision making in rehabilitation settings.

Suppliers

- a. MicroFET2 dynamometer; Hoggan Scientific.
- b. Plinth 2000 treatment plinth; Sports Physio Supplies.
- c. SPSS version 23; IBM.
- d. Excel 2013; Microsoft.

Corresponding author

Aaron Byrne, BSc, healthCORE, Department of Science and Health, K221, Dargan Centre, Institute of Technology Carlow, Kilkenny Road, Carlow, Ireland R93 V960. *E-mail address:* Aaron.Byrne@itcarlow.ie.

References

1. Leetun DT, Ireland ML, Willson JD, Ballantyne BT, Davis IM. Core stability measures as risk factors for lower extremity injury in athletes. *Med Sci Sports Exerc* 2004;36:926-34.
2. Khayambashi K, Ghoddosi N, Straub RK, Powers CM. Hip muscle strength predicts noncontact anterior cruciate ligament injury in male and female athletes. *Am J Sports Med* 2016;44:355-61.
3. Souza RB, Powers CM. Differences in hip kinematics, muscle strength, and muscle activation between subjects with and without patellofemoral pain. *J Orthop Sports Phys Ther* 2009;39:12-9.
4. Powers CM, Ghoddosi N, Straub RK, Khayambashi K. Hip strength as a predictor of ankle sprains in male soccer players: a prospective study. *J Athl Train* 2017;52:1048-55.
5. Fredericson M, Cookingham CL, Chaudhari AM, Dowdell BC, Oestreicher N, Sahrman SA. Hip abductor weakness in distance runners with iliotibial band syndrome. *Clin J Sport Med* 2000;10:169-75.
6. Tyler TF, Nicholas SJ, Campbell RJ, McHugh MP. The association of hip strength and flexibility with the incidence of adductor muscle strains in professional ice hockey players. *Am J Sports Med* 2001;29:124-8.
7. Nadler SF, Malanga GA, DePrince M, Stitik TP, Feinberg JH. The relationship between lower extremity injury, low back pain, and hip muscle strength in male and female collegiate athletes. *Clin J Sport Med* 2000;10:89-97.
8. Nadler SF, Ga Malanga, Solomon JL, Feinberg JH, Foye PM, Park YI. The relationship between lower extremity injury and the hip abductor to extensor strength ratio in collegiate athletes. *J Back Musculoskelet Rehabil* 2002;16:153-8.
9. Mosler AB, Kemp J, King M, et al. Standardised measurement of physical capacity in young and middle-aged active adults with hip-related pain: recommendations from the first International Hip-related Pain Research Network (IHIPRN) meeting, Zurich, 2018. *Br J Sports Med* 2019 Dec 19 [Epub ahead of print].
10. Delp SL, Hess WE, Hungerford DS, Jones LC. Variation of rotation moment arms with hip flexion. *J Biomech* 1999;32:493-501.
11. Drake R, Vogl W, Mitchell A. Gray's anatomy for students. 3rd ed. Philadelphia: Elsevier Churchill Livingstone; 2005.
12. Distefano LJ, Blackburn JT, Marshall SW, Padua DA. Gluteal muscle activation during common therapeutic exercises. *J Orthop Sports Phys Ther* 2009;39:532-40.
13. O'Sullivan K, Smith SM, Sainsbury D. Electromyographic analysis of the three subdivisions of gluteus medius during weight-bearing exercises. *Sports Med Arthrosc Rehabil Ther Technol* 2010;2:17.
14. Collard DCM, Verhagen EALM, Chinapaw MJM, Knol DL, van Mechelen W. Effectiveness of a school-based physical activity injury prevention program. *Arch Pediatr Adolesc Med* 2010;164:145-50.
15. Bizzini M, Junge A, Dvorak J. Implementation of the FIFA 11+ football warm up program: how to approach and convince the football associations to invest in prevention. *Br J Sports Med* 2013;47:803-6.
16. Sewry N, Verhagen E, Lambert M, van Mechelen W, Brown J. Evaluation of the effectiveness and implementation of the BokSmart Safe Six Injury Prevention Programme: a study protocol. *Inj Prev* 2017;23:428.
17. Schlingermann BE, Lodge CA, Gissane C, Rankin PM. Effects of the gaelic athletic association 15 on lower extremity injury incidence and neuromuscular functional outcomes in collegiate gaelic games. *J Strength Cond Res* 2018;32:1993-2001.
18. Bohannon RW. Manual muscle testing: does it meet the standards of an adequate screening test? *Clin Rehabil* 2005;19:662-7.
19. Hislop HJ, Montgomery J. Daniels and Worthingham's muscle testing techniques of manual examination. 6th ed. Philadelphia: Saunders; 1995.
20. Cuthbert SC, Goodheart GJ. On the reliability and validity of manual muscle testing: a literature review. *Chiropr Osteopat* 2007;15:4.
21. Bohannon RW. Reference values for extremity muscle strength obtained by hand-held dynamometry from adults aged 20 to 79 years. *Arch Phys Med Rehabil* 1997;78:26-32.
22. Stark T, Walker B, Phillips JK, Fejer R, Beck R. Hand-held dynamometry correlation with the gold standard isokinetic dynamometry: a systematic review. *PM R* 2011;3:472-9.

23. Martins J, Da Silva JR, Da Silva MRB, Bevilaqua-Grossi D. Reliability and validity of the belt-stabilized handheld dynamometer in hip-and knee-strength tests. *J Athl Train* 2017;52:809-19.
24. Thorborg K, Bandholm T, Schick M, Jensen J, Hölmich P. Hip strength assessment using handheld dynamometry is subject to intertester bias when testers are of different sex and strength. *Scand J Med Sci Sports* 2013;23:487-93.
25. Wikholm JB, Bohannon RW. Hand-held dynamometer measurements: tester strength makes a difference. *J Orthop Sports Phys Ther* 1991;13:191-8.
26. Charlton PC, Mentiplay BF, Grimaldi A, Pua YH, Clark RA. The reliability of a maximal isometric hip strength and simultaneous surface EMG screening protocol in elite, junior rugby league athletes. *J Sci Med Sport* 2017;20:139-45.
27. Thorborg K, Bandholm T, Hölmich P. Hip- and knee-strength assessments using a hand-held dynamometer with external belt-fixation are inter-tester reliable. *Knee Surg Sports Traumatol Arthroscop* 2013;21:550-5.
28. Tourville TW, Smith HC, Shultz SJ, et al. Reliability of a new stabilized dynamometer system for the evaluation of hip strength. *sports health: a multidisciplinary approach* 2013;5:129-36.
29. Jackson SM, Cheng MS, Smith AR, Kolber MJ. Intrarater reliability of hand held dynamometry in measuring lower extremity isometric strength using a portable stabilization device. *Musculoskelet Sci Pract* 2017;27:137-41.
30. Kottner J, Audige L, Brorson S, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *Int J Nurs Stud* 2011;48:661-71.
31. Bujang MA. A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. *Arch Orofac Sci* 2017;12:1-11.
32. Thorborg K, Petersen J, Magnusson SP, Hölmich P. Clinical assessment of hip strength using a hand-held dynamometer is reliable. *Scand J Med Sci Sports* 2010;20:493-501.
33. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res* 2005; 19:231-40.
34. Portney LG, Watkins MP, Portney GWP. *Foundations of clinical research: applications to practice*. Philadelphia, PA: Pearson/Prentice Hall; 2009.
35. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15:155-63.