Johnsen, M. B., Eitzen, I., Moksnes, H., Risberg, M. A. (2013). Inter- and intrarater reliability of four single-legged hop tests and isokinetic muscle torque measurements in children. *Knee Surgery, Sports Traumatology, Arthroscopy*. doi:10.1007/s00167-013-2771-x

---

Inter- and intrarater reliability of four single-legged hop tests and isokinetic muscle torque measurements in children.

Abstract

Purpose: Single-legged hop tests and isokinetic muscle torque are common outcome measures in the evaluation of knee function. The reliability of the single-legged hop tests in children has not been documented. The aim was to examine inter- and intrarater reliability of four single-legged hop tests and isokinetic muscle torque measurements in children.

Methods: Twenty-eight sports active children (12.4 ± 0.3 years old) were tested three times in two test sessions separated by one week. They performed four single-legged hop tests and concentric isokinetic torque measurements during knee extension and flexion. Inter- and intrarater reliability were calculated using the intraclass correlation coefficient, (ICC 2,1). Relative terms of the standard error of measurement (SEM%) and smallest real difference (SRD%) were emphasized to allow comparison between the different variables.

Results: Twenty-six children were included for statistical analysis. ICCs for inter- and intrarater reliability were moderate to high for the hop tests (0.62 to 0.91) and isokinetic measurements (0.76 to 0.87). SEMs% were low for the hop tests (3.9% to 7.4%) and the isokinetic measurements (5.2% to 8.9%). SRDs% were 20.5% or less for the hop tests, 15.7% or less for knee extension and 24.6% or less for knee flexion.

Conclusion: The single-legged hop tests and isokinetic muscle torque measurements demonstrated moderate to high reliability with low measurement error in sports active children. A change above 20.5% for the single-legged hop tests, 15.7% for knee extension, and 24.6% for knee flexion is necessary to represent a real change in knee function.

Level of Evidence III.

Keywords: Performance-based tests, Lower extremity, Hop tests, Isokinetic dynamometry, Muscle strength, Reproducibility of results

Introduction

Anterior cruciate ligament (ACL) injuries are among the most common sport injuries in adolescents and adults [2], and have been increasingly reported also in active children over the last decade [5, 17, 23]. The use of multifactorial outcome measures to capture the different aspects and consequences of an ACL injury has been established for the adult population [4]. Among these, some of the most commonly assessed performance-based tests are single-legged hop tests and isokinetic muscle torque measurements [9, 25]. These tests have also been implemented in the evaluation of knee function after ACL injury in children [13, 19]. High reliability has been documented for both of these outcome measures in adults [21, 25, 26, 30]. To our knowledge, the reliability and validity of the single-legged hop tests has not been investigated in children. The reliability of isokinetic peak torque has been documented in some studies in children [7, 14, 28, 32], however the results are generally based on small sample sizes and diversity in designs, instrumentation, and protocols. Previous studies of the single-legged hop tests and isokinetic muscle torque measurements have often been performed on an uninjured population to establish reliability [14, 21, 30]. As reliability cannot be conceived as an inherent property of any instrument or test, examination within different populations are essential [31]. Children may differ from adults in their ability to comprehend information, to focus on the task at hand, and recognize the movement skills required to perform the test. This emphasises the need to investigate and develop reliable and valid outcome measures to assess knee function in children [20]. Single-legged hop tests and isokinetic muscle torque measurements are used as part of a functional test battery determining lower extremity function at our institution. The importance and relevance of the present study is that results regarding measurement properties are valuable references to researchers and clinicians when interpreting results of hop performance and isokinetic torque measurements in children. It can also be considered as a first step in the research process to validate outcome measures used to assess knee function in children. Thus, the aims of this study were to examine inter- and intrarater reliability of four single-legged hop tests and selected isokinetic measurements, and to establish the standard error of measurement (SEM) and smallest real difference (SRD) for these outcome measures in a sample of uninjured children.

Material and methods

Twenty-eight sports active children (22 boys and 6 girls) volunteered to participate in the study. We included boys and girls together, as no gender differences were expected regarding performance on the tests. The children were recruited from two sport clubs; a soccer academy and an alpine skiing team in Oslo, Norway. Invitation to participate in the study was distributed by email and on the website of the sport clubs. All children and their parent/guardian signed an informed consent prior to inclusion. Exclusion criteria were injury to the lower extremity in the last six months. Two children were excluded during testing because of pain that affected their performance. The remaining 26 children (20 boys and 6 girls; 13 soccer players and 13 alpine skiers) were included in the statistical analysis. The majority of the children (77%, 20 of 26) reported participating in sports training 4-5 days a week. Characteristics of the included children are shown in Table 1.

Procedures

Data was collected in November 2010 and January 2011 at the Norwegian Sports Medicine Clinic (NIMI). Each child underwent testing at three occasions distributed over two test days with one week interval. Both test days included isokinetic testing in a Biodex dynamometer (Biodex 6000, Shirley, New York, USA), followed by four single-legged hop tests, previously described by Noyes et al. [22] Two sports physiotherapists (IE=Rater A and HM=Rater B) with extensive experience with the test procedures conducted all tests on day 1 to assess interrater reliability. The testing was conducted in a crossover design with a random test rater order. The order of raters was randomized by a simple randomization, determining who would start test 1 on day 1. The order of which rater started on test 2 was determined by the order on test 1. A standardised rest period of 30-60 minutes was given between test 1 and test 2 on day 1. Rater B performed all testing on day 2. The intrarater reliability was calculated from the measures of Rater B from day 1 and day 2. Anthropometric data was measured on day 1. Both legs were tested in a randomised order, which remained the same throughout the two test days. All children performed a standardised warm-up of 5 minutes on a stationary bike. The tests were performed in a successive manner, starting with the isokinetic muscle torque measurement and followed by the single-legged hop tests.

The isokinetic muscle torque measurements

Concentric isokinetic muscle torque during knee flexion and extension were performed with 5 repetitions at 60°/second. Peak torque (PT), average peak torque (AVG PT) and total work (TW) were calculated. The range of motion was set from 90° knee flexion to full knee extension (0°). Seating adjustments were made with each child positioned in 80° of hip flexion and 90° of knee flexion in the test chair. The test position with regard to the depth and height of the test chair, side-to-side placement of the dynamometer, and the length of the attached arm was adjusted for each child prior to test 1. Individual adjustments were registered and saved by the first rater testing, and the exact same settings used for the retest measures. Effects of gravity were corrected by static weighing of the lower leg with the knee joint at 10° flexion. Straps across the chest, waist, and thigh were re-adjusted between testing of each leg to ensure that the seating position did not change during testing. Standardised verbal instructions before and during the test were given. The children performed four practice repetitions followed by one minute of rest before five measured repetitions. During the test the rater said: “Start” and then “up” and “down” for every repetition, and “stop” at the end of the test. No other verbal or visual feedback was given. Real strength values in Newton meter and Joules, given as PT (Nm), AVG PT (Nm), and TW (J) was measured with one decimal.

The single-legged hop tests

The single hop (SH), crossover hop (CH), triple hop (TH), and 6-meter timed hop (6MTH) were performed in the given order (Fig. 1). A six meter long and 15 cm wide tape on the floor marked the test-course. In addition, a tape measurement was placed in the longitudinal direction to measure the distance hopped. We used a manual stopwatch (Rucanor) to record the 6MTH. Each hop began with the big toe of the leg to be tested placed on the start line. After landing a line was drawn from the back of the heel onto the tape measure to record the distance hopped. The children were first given a demonstration, followed by one practice trial. The instruction for the SH was to hop as far as possible forward and land on the same leg. For the CH the children hopped forward three times, alternately crossing the 15 cm wide line. For the TH they hopped three consecutive hops along the tape measure. A hop was considered valid if the final landing was maintained for two seconds without any excessive balance movements. Hops were continued until two valid trials were recorded, with a maximum of five trials possible. For the 6MTH the children hopped forward as quickly as possible along the six meter test-course. The rest period between trials was the time needed for the child to walk back and resume the start position. The rest between the different hops was in addition the time it took to demonstrate a new hop. There were no restrictions of arm movements and the children wore the same shoes on both test days. The distance hopped was measured to the nearest 0.5 cm, and time recorded to the nearest 0.01 second for the 6MTH.

The exact same procedures were performed for the single-legged hop tests and isokinetic muscle torque measurements at the retests. The study was approved by the Regional Committee for Medical Research Ethics for South-Eastern Norway (2010/1855) and carried out in accordance to the Declaration of Helsinki.

Statistical analysis

The hop score of each leg was reported as the better of two measured trials recorded in absolute distance (cm) for the SH, CH and TH, and in time (seconds) for the 6MTH. Isokinetic muscle torque was expressed as PT and AVG PT in Nm and TW in J. Means and standard deviations (SD) were presented for each of the variables.
The sample size in the present study was based on previous reliability studies using the same methods [7, 14, 21, 30], and on relevant literature regarding sample size in reliability studies [12, 34]. Based on discussion with a statistician, 30 children were a priori decided to be included. We used a paired *t* test to examine the differences between the absolute scores on day 1, test-retest, and between the right and left leg. The order of rater testing, and differences between boys and girls, were assessed with an independent *t* test. Significance level was set at $p<0.05$. Inter- and intrarater reliability were expressed by the intracorrelation coefficient (ICC 2,1) with 95% confidence intervals (CI) [29]. The SEM was calculated from the square root of the mean square of error (MSE) term derived from the analysis of variance (ANOVA) [35]. The SRD with 95% confidence was calculated as $1.96 \times \sqrt{2} \times \text{SEM}$. The SEM% and SRD% were calculated (SEM or SRD/mean of all scores) to represent measurement error in relative terms [15], and allow comparison across different variables. All analyses were performed in the Statistical Package for Social Science (SPSS), version 18.

Results

Twenty-six children completed testing. There were no significant differences in anthropometrics ($p=0.11$ to 0.95) between boys and girls (Table 1). We found no significant differences between boys and girls in absolute scores of the single-legged hop tests or isokinetic muscle torque measurements. Thus, data were combined for further analysis. There were no significant differences between the right and left leg. No effect was detected regarding the test rater order ($p=0.09$ to 0.94) for interrater testing, except for the 6MTH ($p=0.03$ and 0.04, left and right leg, respectively).

Reliability of the single-legged hop tests

There was a significant difference for the 6MTH ($p=0.01$) at day 1 (Table 2). One child did not succeed a valid TH. We found moderate to high ICCs for inter- (0.72 to 0.91) and intrarater (0.62 to 0.88) testing, with large corresponding 95% CIs (Table 3). SEM% values ranged from 3.9% to 7.4%, and SRD% from 10.8% to 20.5% for both test days. The CH produced the largest measurement errors.

Reliability of isokinetic muscle torque measurements

Results of PT are presented, as there were no differences between PT, AVG PT, and TW regarding reliability assessment, and since PT is the most commonly used isokinetic measurement. The reliability of AVG PT and TW are available in Appendix 1. There were significant differences for PT knee extension ($p=0.001$ to 0.02) on both test days (Table 4). The scores of Rater B were generally higher than of Rater A. We found high ICCs for inter- (0.76 to 0.82) and intrarater (0.77 to 0.87) testing, with large 95% CI to the ICCs (Table 5). Knee flexion produced the largest measurement errors with SEM% from 8.1% to 8.9%, and SRD% from 22.3% to 24.6%. Knee extension showed SEMs% from 5.2% to 5.6% and SRDs% from 14.4% to 15.7%.

Discussion

The most important finding of the present study was that single-legged hop test and isokinetic muscle torque measurements demonstrated to be reliable methods for assessing knee function in sports active children. This is the first study to examine the reliability of the single-legged hop tests in children. Furthermore, it is one of few studies performed on children that have given a comprehensive analysis of reliability by including complementary statistics to the ICC, like the SEM and SRD. Moderate to high inter- and intrarater reliability was established for both the single-legged hop tests and isokinetic muscle torque measurements. The 95% CI to the ICCs were in general wide, which indicated large variation among the estimates of the reliability coefficient. No absolute consensus exists regarding the standards for the magnitude of ICC, although to consider ICCs below 0.50 as poor, 0.50 to 0.75 as moderate, and values above 0.75 as high reliability have been proposed [24]. Physical performance-based tests with well-defined scoring procedures tend to give ICCs in the upper part of 0.70, to the lower parts of 0.90, in young individuals [3]. The ICC measures the strength of the relationship between subjects in repeated measures, but not agreement between individual scores. An absolute measure of reliability, like the SEM, is therefore recommended to estimate the precision of individual scores. Since the different hop tests and isokinetic measurements have inherent different magnitude of scores, we emphasised reporting SEM and SRD in relative terms (%) to allow comparison of results across the different variables. In total, our SEM% values can be characterised as low, and comparable to previous results in adults [8, 25, 30].

The ICCs for the single-legged hop tests were similar to values found in young recreational athletes [21], although higher ICCs (0.82 to 0.97) have been reported after ACL reconstruction [25], and in healthy adults.[27] The 6MTH produced the lowest ICCs, as shown previously [21]. However, measurement error was small, indicating acceptable within-subject variability. The 6MTH was also the only hop with significant difference across trials at day 1 (p=0.01). There was a tendency for all hop scores to increase from day 1 to day 2, although the effect was not significant. Previous studies on adults have also shown a tendency of scores to increase across trials [21, 27]. However, there is no agreement on the optimal number of trials to minimise the learning effect. Munro et al. [21] reported that three to four trials were required to stabilise hop performance in adults, with the CH needing the most trials. The children in our study used on average two to four trials to achieve a valid hop, with the CH needing the most trials. The CH also showed the largest values of SEM%, suggesting that this hop demonstrated lowest absolute reliability. The SRDs% in the present study indicated that values below 7.4% most likely represented only "measurement noise". Improvement should exceed approximately 11% to 21% to reflect a real change in knee function beyond measurement error. In comparison, Reid et al. [25] reported a minimal detectable change (90% MDC) of 8% to 13% for the single-legged hop tests in adults after ACL reconstruction. Although the SEM and SRD are important benchmarks towards establishing the minimal important change (MIC), it should further be defined from an anchor-based approach using an external criterion to interpret whether a particular change is significant [6, 33].

The magnitude of the ICCs for isokinetic knee extension and flexion torque in our study are comparable to earlier studies on children [7, 14, 28], although somewhat lower than reported in adults (0.85 to 0.94) [10, 30]. We found higher ICC values for PT knee extension than knee flexion, which has also previously been documented in children [7, 14, 16], and adults [8, 10]. The difference was also reflected by the size of the SEMs%, with smaller measurement errors for knee extension. As reported in older adults [8, 10], children may intuitively understand the concept of maximal effort in knee extension better than flexion, since the latter movement is assisted by gravity. However, the magnitude of the absolute scores in our study are comparable to normative data from Norwegian school children at the same chronological age (12 years old) [11], and are in that way a valid expression of strength for the age group. The SEM for isokinetic measurements of knee extension or flexion has generally not been documented in children. Instead, the coefficient of variance (CV) has been used to express variability within subjects [7, 16, 18]. The CV for PT in these studies has ranged from 4.0% to 12.0%. The SEM% and CV will generally produce similar values [15], as shown in the study by Santos et al. [28] where SEM% ranged from 10.4% to 16.8%, and CV from 9.4% to 15.0%. This indicates that our SEM% values are comparable to previous results in children [7, 16, 18, 28]. Our SRD% values suggest that a change above 14% to 16% for knee extension and 22% to 25% for flexion is necessary to reflect a real change in knee function. It demonstrates that knee extension has better ability to detect small, but potentially meaningful changes. In comparison, Sole et al. [30] suggested a change of 18% to 19% (95% SRD) for knee extension and flexion in young adults. As we found no clear benefit of one isokinetic torque variable over another regarding reliability (Appendix 1), it seems like any of the outcome variables included in this study can be used as an expression of strength in children. Isokinetic dynamometers may also be used to measure isotonic, isometric and eccentric isokinetic contractions. Only a few studies have examined the reliability of eccentric isokinetic knee extension and flexion torque in children, with conflicting results [7, 14]. To explore whether different isokinetic test conditions would give different results regarding reliability is important. However, since the aim was to investigate reliability of our current protocol, only concentric actions were tested in this present study.

The relative nature of the ICC is reflected by the fact that the size of the ICC is highly dependent on the heterogeneity of the test sample. Low between-subject variability will decrease the value of the ICC even if the test-retest variability is low [35]. The somewhat moderate ICCs in our study were not associated with large measurement error. This may reflect our homogenous sample of sports active children with both low between- and within subject variability. The opposite is demonstrated in a study of a heterogeneous sample of children with cerebral palsy, were the reliability of isokinetic testing during knee extension and flexion were high (ICC 0.90 to 0.99), but with corresponding large CV (19% to 24%), reflecting large variability between the individual scores [1].

Some studies have reported to intentionally use only one rater to exclude the rater-linked variation. For us, it was important to establish interrater reliability for future validity of results in prospective studies. There were some significant differences ($p<0.05$) in scores between the two raters at day 1. The crossover design was a strength with regard to distributing potential differences within subjects randomly between the two raters. Thus, it is possible that these differences reflect disparities between the raters themselves, rather than the order of testing. We used standardised procedures to minimise protocol-related measurement error. All children were tested with a test interval of one week and at the same time of day to minimise biological variations. Previous reliability studies in children have often included a separate familiarisation session and visual feedback during isokinetic testing [7, 14]. We did not include a separate familiarisation session because this is not included in our current test protocol. The results revealed, however, that our protocol may need adjustments to allow more familiarisation with a maximal effort performance during isokinetic knee flexion. There were no substantial differences between the estimates of inter- and intrarater reliability. Overall, both the hop tests and isokinetic torque measurements produced reliable measures that can assist researchers and clinicians when interpreting results of hop performance and isokinetic knee extension and knee flexion torque in children.

Thus, from a clinical perspective this study suggests that single-legged hop tests and isokinetic muscle torque measurements can be used as part of a test battery when assessing knee function in children. Both methods showed high precision of individual scores, which is a necessity when determining outcome on an individual level. It is important for clinicians to be familiar with the measurements properties of the outcome measure that is used, to quantify the functional ability of a child, and assist in the decision-making regarding exercise progression or return to more advanced levels of activity.

There are some limitations to this study. Compared to previous studies [7, 14, 18, 32], we tested a large sample of children (n=26). Still, a sample size of 30 to 50 has been suggested to form practically useful 95% SRDs [12]. The limited between-subject variability in this study may have been a result of our convenience sample of children. Sports active children may be more familiar with physical tasks than less active children. Further, children with injuries may be anxious to perform testing because of pain or fear of re-injury. Consequently, our results can primarily be generalized to similar groups of sports active children.

Conclusions

Single-legged hop test and isokinetic muscle torque measurements are reliable methods for assessing knee function in sports active children. Moderate to high inter- and intrarater reliability and low measurement errors were established for both outcome measures. A change above 20.5% for the hop tests, 15.7% for isokinetic knee extension, and 24.6% for knee flexion (95% SRDs) is necessary to represent a real change in knee function.

Acknowledgments

We would like to thank Kristin Bølstad for help with the preparation of the study. Maren Hjelle Guddal, Arnhild Bakken, Nina Markussen, and Carina Wold are acknowledged for their assistance on the test days. Thanks to Professor Ingar Holme, at the Norwegian School of Sport Sciences, for helpful advice on the statistical analysis. Additionally, the Norwegian Sports Medicine Clinic (NIMI) is acknowledged for the contribution of infrastructure.

Funding

None.

Conflict of interests

The authors declare that there are no conflicts of interest.

Figure Captions

Fig. 1 Presentation of the four single-legged hop tests: single hop, crossover hop, triple hop, and 6-m timed hop

References

1. Ayalon M, Ben-Sira D, Hutzler Y, Gilad T (2000) Reliability of isokinetic strength measurements of the knee in children with cerebral palsy. Dev Med Child Neurol 42(6):398-402

2. Bahr R, Holme I (2003) Risk factors for sports injuries--a methodological approach. Br J Sports Med 37(5):384-392

3. Baumgartner TA (2003) Measurement for evaluation in physical education and exercise science.McGraw-Hill, Boston

4. Bent NP (2009) Selecting outcome measures in sports medicine: a guide for practitioners using the example of anterior cruciate ligament rehabilitation. British Journal of Sports Medicine 43(13):1006-1012

5. Caine D, Maffulli N, Caine C (2008) Epidemiology of injury in child and adolescent sports: injury rates, risk factors, and prevention. Clin Sports Med 27(1):19-50, vii

6. de Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM (2006) Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. Health Qual Life Outcomes 4:54

7. Deighan MA, De Ste Croix MBA, Armstrong N (2003) Reliability of isokinetic concentric and eccentric knee and elbow extension and flexion in 9/10 year old boys. Isokinetics & Exercise Science 11(2):109-115

8. Eitzen I, Hakestad KA, Risberg MA (2012) Inter- and intrarater reliability of isokinetic thigh muscle strength tests in postmenopausal women with osteopenia. Arch Phys Med Rehabil 93(3):420-427

9. Eitzen I, Moksnes H, Snyder-Mackler L, Engebretsen L, Risberg MA (2010) Functional tests should be accentuated more in the decision for ACL reconstruction. Knee Surg Sports Traumatol Arthrosc 18(11):1517-1525

10. Hartmann A, Knols R, Murer K, De Bruin ED (2009) Reproducibility of an isokinetic strength-testing protocol of the knee and ankle in older adults. Gerontology 55(3):259-268

11. Holm I, Fredriksen P, Fosdahl M, Vollestad N (2008) A normative sample of isotonic and isokinetic muscle strength measurements in children 7 to 12 years of age. Acta Paediatrica 97(5):602-607

12. Hopkins WG (2000) Measures of reliability in sports medicine and science. Sports Med 30(1):1-15

13. Janarv PM, Nystrom A, Werner S, Hirsch G (1996) Anterior cruciate ligament injuries in skeletally immature patients. J Pediatr Orthop 16(5):673-677

14. Kellis E, Kellis S, Gerodimos V, Manou V (1999) Reliability of isokinetic concentric and eccentric strength in circumpubertal soccer players. / Fiabilite des mesures de la force isocinetique concentrique et excentrique chez des joueurs de football autour de la puberte. Pediatric Exercise Science 11(3):218-228

15. Lexell JE, Downham DY (2005) How to assess the reliability of measurements in rehabilitation. Am J Phys Med Rehabil 84(9):719-723

16. Merlini L, Dell'Accio D, Granata C (1995) Reliability of dynamic strength knee muscle testing in children. J Orthop Sports Phys Ther 22(2):73-76

17. Mohtadi N, Grant J (2006) Managing anterior cruciate ligament deficiency in the skeletally immature individual: a systematic review of the literature. Clin J Sport Med 16(6):457-464

18. Mohtadi NG, Kiefer GN, Tedford K, Watters S (1990) Concentric and eccentric quadriceps torque in pre-adolescent males. Can J Sport Sci 15(4):240-243

19. Moksnes H, Engebretsen L, Eitzen I, Risberg MA (2013) Functional outcomes following a non-operative treatment algorithm for anterior cruciate ligament injuries in skeletally immature children 12 years and younger. A prospective cohort with 2 years follow-up. Br J Sports Med 47(8):488-494

20. Moksnes H, Engebretsen L, Risberg MA (2012) Management of anterior cruciate ligament injuries in skeletally immature individuals. J Orthop Sports Phys Ther 42(3):172-183

21. Munro AG, Herrington LC (2011) Between-Session Reliability of Four Hop Tests and the Agility T-Test. J Strength Cond Res 25(5):1470-1477

22. Noyes FR, Barber SD, Mangine RE (1991) Abnormal lower limb symmetry determined by function hop tests after anterior cruciate ligament rupture. Am J Sports Med 19(5):513-518

23. Parkkari J, Pasanen K, Mattila VM, Kannus P, Rimpela A (2008) The risk for a cruciate ligament injury of the knee in adolescents and young adults: a population-based cohort study of 46 500 people with a 9 year follow-up. Br J Sports Med 42(6):422-426

24. Portney LG, Watkins MP (2000) Foundations of clinical research: applications to practice.Prentice Hall, Upper Saddle River, N.J.

25. Reid A, Birmingham TB, Stratford PW, Alcock GK, Giffin JR (2007) Hop testing provides a reliable and valid outcome measure during rehabilitation after anterior cruciate ligament reconstruction. Phys Ther 87(3):337-349

26. Ross MD, Irrgang JJ, Denegar CR, McCloy CM, Unangst ET (2002) The relationship between participation restrictions and selected clinical measures following anterior cruciate ligament reconstruction. Knee Surg Sports Traumatol Arthrosc 10(1):10-19

27. Ross MD, Langford B, Whelan PJ (2002) Test-retest reliability of 4 single-leg horizontal hop tests. J Strength Cond Res 16(4):617-622

28. Santos AN, Pavao SL, Avila MA, Salvini TF, Rocha NA (2013) Reliability of isokinetic evaluation in passive mode for knee flexors and extensors in healthy children. Braz J Phys Ther 17(2):112-120

29. Shrout PE, Fleiss JL (1979) Intraclass correlations: uses in assessing rater reliability. Psychol Bull 86(2):420-428

30. Sole G, Hamren J, Milosavljevic S, Nicholson H, Sullivan SJ (2007) Test-retest reliability of isokinetic knee extension and flexion. Arch Phys Med Rehabil 88(5):626-631

31. Streiner DL, Norman GR (2008) Reliability. In: Streiner DL, Norman GR, (eds) Health measurement scales: a practical guide to their development and use, 4th edn. Oxford University Press, Oxford, pp 167-210

32. Tsiros MD, Grimshaw PN, Schield AJ, Buckley JD (2011) Test-retest reliability of the Biodex System 4 Isokinetic Dynamometer for knee strength assessment in paediatric populations. J Allied Health 40(3):115-119

33. Turner D, Schunemann HJ, Griffith LE, Beaton DE, Griffiths AM, Critch JN, Guyatt GH (2010) The minimal detectable change cannot reliably replace the minimal important difference. J Clin Epidemiol 63(1):28-36

34. Walter SD, Eliasziw M, Donner A (1998) Sample size and optimal designs for reliability studies. Stat Med 17(1):101-110

35. Weir JP (2005) Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. Journal of Strength & Conditioning Research (Allen Press Publishing Services Inc) 19(1):231-240

Table 1 Characteristics of the included children, n=26

	Boys (n=20)	Girls (n=6)	p
Age (years)	12.4 ± 0.3 (11.8-12.9)	12.4 ± 0.2 (12.2-12.8)	n.s.
Height (cm)	154.6 ± 0.1 (144.0-166.0)	157.8 ± 0.1 (147.0-164.0)	n.s.
Weight (kg)	43.9 ± 6.4 (33.5-55.0)	41.8 ± 2.5 (38.0-45.0)	n.s.
BMI (kg/m^2)	18.3 ± 2.0 (14.3-22.4)	16.8 ± 1.2 (15.2-18.5)	n.s.

Values are mean ± SD and (minimum – maximum), BMI=body mass index

Table 1 Means, SDs and p-values with regard to inter- and intrarater reliability of the single-legged hop tests, n=26

	Day 1			Rater B		
	Rater A	Rater B	p	Day 1	Day 2	p
Single hop (cm)						
Right leg	126.7 ± 16.9	129.5 ± 17.1	n.s.	129.5 ± 17.1	127.9 ± 15.2	n.s.
Left leg	126.2 ± 14.7	126.9 ± 15.9	n.s.	126.9 ± 15.9	128.0 ±14.5	n.s.
Crossover hop (cm)						
Right leg	383.7 ± 50.8	374.2 ± 57.7	n.s.	374.2 ± 57.7	388.5 ± 53.9	n.s.
Left leg	381.3 ± 50.5	378.5 ± 51.3	n.s.	378.5 ± 51.3	387.4 ± 49.9	n.s.
Triple hop (cm)[a]						
Right leg	415.0 ± 51.9	414.6 ± 50.7	n.s.	414.4 ± 49.7	415.6 ± 53.8	n.s.
Left leg	414.8 ± 46.2	410.1 ± 50.2	n.s.	410.1 ± 50.2	416.6 ± 45.2	n.s.
6-m timed hop (s)						
Right leg	1.91 ± 0.20	1.85 ± 0.16	0.01*	1.85 ± 0.16	1.87 ± 0.21	n.s.
Left leg	1.90 ± 0.21	1.86 ± 0.17	n.s.	1.86 ± 0.17	1.85 ± 0.19	n.s.

*Significant difference ($p<0.05$), values are mean ± SD, a = for the triple hop n=25