Original Article

Evaluation of reliability and sensitivity of the Y-balance test for quantifying lower limb asymmetry in prepubescent soccer and basketball players

ZIED NEJI¹, AHMED ATTIA², YASSINE NEGRA³, SENDA SAMMOUD⁴, RAJA BOUGUEZZI⁵, YOUNES HACHANA⁶.

^{1,2,3,4,5,6}Research Unit (UR17JS01) "Sport Performance, Health & Society", Higher Institute of Sport and Physical Education of Ksar-Saïd, University of La Manouba, TUNISIA

Published online: June 30, 2020

(Accepted for publication: June 22, 2020)

DOI:10.7752/jpes.2020.04247

Abstract:

The Y-balance test (Y-BT) is a dynamic test that requires strength and proprioception and has been used to assess physical performance, identify chronic ankle instability, and identify athletes at greater risk for lower extremity injury. To date, research concerning the sensitivity and reliability of the Y-BT has been conducted only for adults. The aim of this study was to determine the reliability and sensitivity of the Y-BT maximal reach distances and their relationships with hamstring flexibility in young male team sport athletes. A total of 151 (78 basketball players and 73 soccer players) prepubescent soccer and basketball players, participated to this study. Hamstring flexibility was assessed using the sit and reach test. The intraclass correlation coefficient (ICC_{3,1}) of the Y-BT maximal reach distances and composite scores for dominant and non-dominant legs ranged from 0.97-0.98 and the standard error of measurement (SEM%) was <5% (0.55-1.47%). Moreover, the capacity to detect change was rated as "good" because the SEMs (0.007-0.94 cm) were smaller than the respective SWCs (0.01-1.52 cm). There were no significant or meaningful correlations between Y-BT with the sit and reach test performance ($p>0.05$). The unpaired Student t-test revealed that dominant leg dynamic limits and composite scores were greater than those of non-dominant leg ($p<0.05$). This result was rejected by a small d_z (<0.46) and an area under the ROC curve <0.7 . According to the results of the study, the Y-BT dynamic limits and composite scores have excellent reliability but poor discriminate ability when used in prepubescent male soccer and basketball players.

Key words: prepubescent athletes; dynamic balance; performance deficit

Introduction

Dynamic balance is the ability to stabilize the body and maintain center of mass stability during dynamic actions and is critical for the successful performance of fundamental movement skills (Faigenbaum, Bush, McLoone, Kreckel, Farrell, Ratamess, & Kang, 2015). Previous studies (Faigenbaum, Myer, Fernandez, Carrasco, Bates, Farrell, Ratamess, & Kang, 2014; Hianik, Šimonek, & Horicka, 2017; Palomares, Palomares, Uchôa, Andrade, Deana, & Alves, 2019) have shown that balance training should be incorporated in all stages of long-term athletic development, because balance training-induced gains may support young athletes during the acquisition phase of complex movements, and to withstand the demands of long-term athletic training and competition.

One tool that assesses dynamic postural control is the Low Quarter Y Balance Test (Y-BT) (Plisky, Gorman, Butler, Kiesel, Underwood, & Elkins, 2009). The Y-BT is scored by comparing reaching ability to normalize scores as well as identification of asymmetry between the right and left lower extremity i.e. dominant (DOML) and non-dominant leg (NDOML) (Faigenbaum, Bush, McLoone, Kreckel, Farrell, Ratamess, & Kang, 2015; Gonell, Romero, & Soler, 2015; Linek, Sikora, Wolny, & Saulicz, 2017).

The utility of the Y-BT depends in part on the ability of its dynamic limits of stability to reproduce error-free scores upon sequential administration (i.e. test-retest reliability) (Smith, Creps, Bean, Rodd, & Alsalaheen, 2018). The Y-BT dynamic reach directions were proven to be a reliable tool (Chimera, Smith, & Warren, 2015; Plisky, Gorman, Butler, Kiesel, Underwood, & Elkins, 2009). Smith et al. (2018) reported a good reliability (Intraclass correlation coefficient (ICC = 0.89) for the anterior (AT) direction, and moderate reliability (ICC= 0.76, 0.63, for posterolateral (PL), and posteromedial (PM) directions, respectively in high-school athletes. The dynamic limits of stability are often used to calculate Y-BT asymmetry index calculated as absolute difference between legs in AT, PL, and PM (Smith, Chimera, & Warren, 2015). Y-BT asymmetry index has emerged as a common screening tool to assess injury risk (Breen, Howell, Straccioli, Dawkins, & Meehan, 2016; Chimera, Knoeller, Cooper, Kothe, Smith, & Warren, 2017; Smith, Chimera, & Warren, 2015).

One another of the prerequisites of test/tool applicability in sport science, is its construct validity. The construct validity of the Y-BT asymmetry index has been previously established (Plisky, Rauh, Kaminski, &

Underwood, 2006). Plisky et al. (2006) showed that, individuals with anterior left/right asymmetries more than four cm on the Y-BT were 2.5 times more likely to sustain a lower extremity injury. More recently, Stiffler et al. (2017) revealed that asymmetry in the AT direction evaluated by the star excursion balance test (SEBT) may assist in identifying collegiate athletes who are at risk for sustaining non-contact injuries to the knee or ankle. To the authors' knowledge, there is no study available that operated control over construct (i.e., dominant leg) supporting variables testing DOML and NDOML in prepubertal athletes. Therefore, it is of a big interest to analyze if the Y-BT maximal reach directions are enough sensitive to detect small changes between maximal reach distances of DOML and NDOML legs used to calculate Y-BT asymmetry index in prepubertal athletes.

According to Miller et al. (2017) the Y-BT assesses dynamic stability and balance, requiring a range of skills including neuromuscular control, strength, and flexibility. Recently, Hammami et al. (2016) confirmed the association between balance and muscle strength/power measures. The same authors concluded that this association increases with maturity and may imply transfer effects from balance to strength/power training and vice versa in young male soccer players. Flexibility is a neuromuscular capability that provides body motion with an absolute range of movement in a joint or series of joints that is attainable in a momentary effort, with the help of a partner or a piece of equipment (Sporis, Jukic, Milanovic, & Vucetic, 2010). Rehabilitation and fitness workers recognize flexibility to be a valuable part of an athletes' physical fitness program for the purpose of injury prevention, better and efficient performance, or a faster and safer rehabilitation (Sporis, Jukic, Milanovic, & Vucetic, 2010). Major morphological and neural changes occurring in youth play an important role in the ability to adapt to a specific training stimulus (Malina, Bouchard, & Bar-Or, 2004). Thus, individual physiological changes (e.g., hormonal, central nervous system myelination) may favor different types of adaptation depending on maturity status (Lloyd, Oliver, Faigenbaum, Howard, De Ste Croix, Williams, Best, Alvar, Micheli, Thomas, Hatfield, Cronin, & Myer, 2015), knowledge about the relationship between balance and flexibility, might provide greater insight into the way maturity modifies the effectiveness of different training regimes (e.g., balance and flexibility training). However, flexibility could be evaluated within the Y-BT test as there should be a great deal of shared variance between balance and flexibility test. Consequently, we would assume improving and/or evaluating dynamic balance ability through improving/evaluating flexibility.

The most common method of assessing hamstring flexibility in the field setting has been the floor sit-and-reach test (SRT) (Mayorga-Vega, Merino-Marban, & Viciania, 2014). This test is present in most health related fitness test batteries because it is believed that maintaining hamstring and low back flexibility may prevent acute and chronic musculoskeletal injuries and low back problems, postural deviations, gait limitations, and risk of falling (Baltaci, Un, Tunay, Besler, & Gerçek, 2003).

Based on the above-mentioned considerations, this study aimed to (1) investigate the reliability and construct validity (sensitivity) of the Y-BT maximal reach distance for each direction, (2) to explore the association between measures of dynamic balance (i.e., Y-BT) and flexibility (i.e., SRT). Finally, we hypothesized the good sensitivity of the Y-BT three reach distances, and the large association between Y-BT and SRT performances.

Materiel & Methods

Experimental approach to the problem

The evaluation of muscular asymmetry and the unilateral dynamic balance is particularly important relative to the growing interest in physical fitness testing and injury prevention strategies in youth (Faigenbaum, Lloyd, MacDonald, & Myer, 2016; Yiannis, Christos, Vassilis, Konstantinos, & Thomas, 2020). Since Y-BT is frequently used to properly assess and monitor childrens' dynamic balance (Faigenbaum, Lloyd, MacDonald, & Myer, 2016), additional research is needed to determine if the Y-BT maximal reach directions are enough sensitive to detect small changes between maximal reach distances of DOML and NDOML, used to calculate Y-BT asymmetry index in prepubescent athletes. Moreover, knowledge on performance in dynamic balance and low back flexibility and their interrelation in young athletes will be of particular interest particularly in sports in which the proficiency level is sufficiently determined by these neuromuscular capacities. To that end, variables of dynamic balance and muscle flexibility evaluated by the Y-BT and, the SRT, were assessed in a cross-sectional study. As a measure of dynamic balance, the Y-BT balance test was performed on the dominant and non-dominant legs with eyes opened. Measurement of flexibility was conducted with the SRT (Baltaci, Un, Tunay, Besler, & Gerçek, 2003). All Y-BT and SRT tests were administered by the same researcher and at the same time of day (i.e., 9-11 AM). The dominant leg was determined according to the lateral preference inventory (Coren, 1993). All the players undertook three familiarization sessions of all tests, within the 2 weeks preceding the experience. The subjects performed all Y-BT and SRT trials barefoot and wearing shorts and T-shirts for all test sessions (Coughlan, Fullam, Delahunt, Gissane, & Caulfield, 2012). Tests were performed in triplicate and the best trial on each side was retained for statistical analyses. Before each testing session, participants performed a 10 min warm-up (six min alternate step-up on 25 cm bench, static stretching and calisthenics) (Sarabon, Malaker, & Markovic, 2010).

Participants Data were collected from 151 healthy prepubescent boys (mean $\pm$ SD: age: 10.93 ± 0.64 years; body mass: 40.47 ± 9.98 kg; height: 149.2 ± 7.26 cm, maturity offset = -2.8 ± 0.6 ; age at peak height velocity = 13.7 ± 0.7 years). Inclusion criteria for the study required volunteers and active participants in sporting activity

three or more times per week with no known impairment of the musculoskeletal system affecting the spine or the lower extremities. The cohort study comprised 73 soccer players and 78 basket-ball players. All participants had been involved in their sports training regularly for more than two years before the study. Verbal and written informed consent was obtained from legal representatives and children after explanation of the experimental protocol and its potential benefits and harms.

Data collection

Anthropometric Measurements All measurements were taken by the same experienced researcher. Measured dimensions included stature, body mass and sitting height. Stretch stature was measured with a wall-mounted stadiometer (± 0.1 cm, Holtain Ltd, Crosswell, UK), sitting height with a stadiometer mounted on a purpose-built table (± 0.1 cm, Holtain Ltd, Crosswell, UK) body mass with a weighing device (± 0.1 kg). Maturation status was determined according to the predicting age at peak height velocity (Moore, McKay, Macdonald, Nettlefold, Baxter-Jones, Cameron, & Brasher, 2015).

Y-Balance-Test The Y-BT kit was used, which consists of three connected cylindrical tubular plastic bars marked in half centimeter increments. Each bar has a moveable indicator plate, which the subject moves by pushing with their foot/toes without bearing weight on the indicator (Gonell, Romero, & Soler, 2015). To perform the Y-BT, the athletes stood on the leg in the center of the platform with the most distal end of the longest toe just behind a marked line. While maintaining single-leg stance, the subject was instructed to reach with the free limb in the anterior direction for three trials, followed by three trials in postero-medial direction and then three trials in postero-lateral direction, all named in relation to the stance foot, per Y-BT protocol (Faigenbaum, Lloyd, MacDonald, & Myer, 2016). The athlete was instructed to push the distance indicator as far as possible towards the direction that was being evaluated. The subject was monitored by the researcher during testing, and was not allowed to move the indicator by kicking it or accelerating the indicator at the end of the push (Faigenbaum, Lloyd, MacDonald, & Myer, 2016). The maximal reach distance was recorded at the most distal point reached by the foot in the proximal edge of the indicator and was measured to the half centimeter. The specific testing order was right AT, left AT, right PM, left PM, right PL, and left PL. If a proper reach was not performed in three trials for any direction, an additional three trials were allowed. The furthest successful reach for each direction was used for analysis (Faigenbaum, Lloyd, MacDonald, & Myer, 2016).

The trial was discarded and repeated if the athlete:

- Lost his balance during the exercise (reaching the maximal point and coming back at the initial position);
- Lifted the heel of the foot that was on the platform. The entire surface of the foot must have remained remain in contact with the platform throughout the entire duration of the movement;
- The foot did not maintain contact with the distance indicator while the indicator was in motion (e.g. the indicator was kicked);
- The distance indicator was used to maintain posture (e.g. the athlete supported their weight on the movement indicator);
- A loss of balance occurred during the return to the starting position once the distance had been marked.

With the test complete and all performances recorded, the test administrator can also calculate three other Y-BT performance scores:

- Absolute reach distance (ARD) =

$$ARD = \frac{\text{MaxAT} + \text{MaxPM} + \text{MaxPL}}{3}$$

- Relative (normalized) reach distance (RRD) =

$$RRD(\%) = \frac{ARD}{\text{Leg Length}} \times 100$$

- Composite reach distance (%) =

$$CS(\%) = \frac{(\text{MaxAT} + \text{MaxPM} + \text{MaxPL})}{\text{Leg Length} \times 3} \times 100$$

Sit and Reach Test The subject sat on the floor with his back and head against the wall. His legs were out straight ahead and his knees flat against the floor. The SRT box was placed against the athlete's feet whilst keeping his head and back against the wall and was instructed to stretch out his arms as far as he can towards the box without their head or back losing contact with the wall. Then, the ruler was placed on the box and the zero-end was moved towards the athlete fingertips to have the zero point. After that, the athlete was instructed to lean forward slowly as far as possible keeping his fingertips level with each other and his legs flat on the floor (his head and shoulders can come away from the wall now) and slowly reach as far as possible along the length of the ruler and hold for two seconds. The SRT score was the greatest distance contacted by the fingertips. Positive values were awarded if subjects could reach beyond their toes, and negative values were awarded if subjects could not reach beyond their toes (Castro-Piñero, Chillon, Ortega, Montesinos, Sjöström, & Ruiz, 2009; Hartman, & Looney, 2003). The SRT score was registered an accuracy of up to 0.1 cm.

Statistical Analysis Data analyses were performed using SPSS 19.0 program for Windows (SPSS, Inc., Chicago, IL, USA) with the level of statistical significance set a priori at 0.05. Descriptive statistics were generated for all

variables. The normality assumption was checked by the Kolmogorov-Smirnov test. The null hypothesis of test-retest differences was investigated by dependent *t* test. The effect size of the difference (*dz*) was determined using GPOWER software (Faul, Erdfelder, & GPOWER, 2004). The modified Hopkins scale (Hopkins, 2000) was used for the interpretation of *dz*: < 0.2 (*trivial*); 0.2-0.6 (*small*); 0.6-1.2, (*moderate*); 1.2-2 (*large*); > 2 (*very large*) (Hastad, & Lacy, 1994; Scanlan, Dascombe, & Reaburn, 2012). The relative reliability was determined by calculating the intraclass correlation coefficient (ICC_{3,1}) (Hopkins, 2000). We considered an ICC below 0.40 as poor, between 0.40 and, 0.70 as fair, between 0.70 and, 0.90 as good and ≥ 0.90 as excellent (Coppiters, Stappaerts, Janssens, & Jull, 2002). Absolute reliability was analyzed using the standard error of measurement (SEM) expressed as a coefficient of variation (CV). The SEM was calculated by dividing the standard deviation of the difference between scores by $\sqrt{2}$ (Atkinson, & Nevil, 1998). The smallest worth while change (SWC) was determined to establish the usefulness of the Y-BT. When $SEM \leq SWC$ the test's capacity to detect change is considered (*good*); when $SEM = SWC$ it is considered (*satisfactory*), and when $SEM \geq SWC$ the test is rated as (*marginal*) (Liow, & Hopkins, 2003). The MDC_{95} determines the minimum amount of change in the measurement that would be required to exceed the level of measurement error and was expressed in absolute and relative term for comparison between Y-BT. Pearson's correlations were used to determine the relationship between Y-BT and SRT performance scores. Cohen, (1992) suggested that an absolute correlation coefficient of 0.2-0.4 was small, 0.5-0.7 was moderate, and > 0.7 was large. The assumption of heteroscedasticity of the raw residuals was assessed via the Konker test. An unpaired Student t-test was used to compare dynamic limits and composite scores of DOML and NDOML. The external responsiveness of the dynamic limits and composite scores of DOML and NDOML of the Y-BT was analyzed using the receiver operator characteristics (ROC) curve (Deyo, & Centor, 1986; Mannion, Elfering, Staerkle, Junge, Grob, Semmer, Jacobshagen, Dvorak, & Boos, 2005). The ROC curve analysis determines the sensitivity and specificity of a tool to classify individuals according to a fixed criterion. It consists of a plot of true positive rate (sensitivity) vs. false positive rate (1-specificity) for each of several possible cut-off points in change score (Deyo, & Centor, 1986). The area under the ROC curve (AUC) was interpreted as the probability of correctly discriminating between young team sport athletes with a good and a poor outcome. The area of 0.5 is interpreted as no discriminatory accuracy and 1.0 as complete accuracy (De Vet, Bouter, Bezemer, & Beurskens, 2001). As general rules, the area under the ROC curve of 0.70 is commonly considered to indicate good discriminative ability (Mannion, Elfering, Staerkle, Junge, Grob, Semmer, Jacobshagen, Dvorak, & Boos, 2005).

Results

All participants completed the aforementioned study procedures as described. No unexpected events or injury were reported. All the 151 participants were categorized as pre-PHV with a maturity offset of -2.8 ± 0.6 and an age at peak height velocity of 13.7 ± 0.7 years.

Reliability and sensitivity analysis Results of the Y-BT and SRT scores obtained from the test and retest are presented in Table 1. The data suggest excellent reliability of all Y-BT maximal reach (AT, PM, and PL) directions while standing on the dominant and non-dominant foot and the SRT test (Table 2). Specifically, all ICC values were well above 0.90, whereas SEM values were about 0.9 cm (<5%).

Table 1. Test-Retest data, represented as means and standard deviations of the Y-BT and the SRT tests (n = 81)

			Test (SD)	Retest (SD)	p	dz
Y-BT	DOML	AT (cm)	69.51 (5.66)	69.53 (5.53)	0.86	0.003
		PM (cm)	63.22 (7.52)	63.49 (7.24)	0.05	0.03
		PL (cm)	73.69 (6.14)	73.88 (6.01)	0.14	0.031
		CS (%)	89.57 (6.89)	89.79 (6.81)	0.11	0.032
		ARD (cm)	68.81 (5.27)	68.97 (5.07)	0.09	0.031
		RRD (%)	0.89 (0.10)	0.90 (0.07)	0.17	0.11
	NDOML	AT (cm)	67.83 (6.14)	68.02 (5.82)	0.11	0.031
		PM (cm)	63.80 (7.76)	63.86 (7.52)	0.74	0.007
		PL (cm)	71.86 (6.33)	72.04 (6.13)	0.19	0.028
		CS (%)	88.28 (6.76)	88.48 (6.69)	0.11	0.029
		ARD (cm)	67.83 (5.37)	67.98 (5.15)	0.13	0.028
		RRD (%)	0.88 (0.07)	0.87 (0.10)	0.22	0.112
SRT			4.40 (2.60)	4.34 (2.57)	0.46	0.023
Notes: DOML: Dominant leg; NDOML: Non Dominant leg; AT: Anterior; PM: Posteromedial; PL: Posterolateral; CS: Composite reach distance; ARD: Absolute reach distance; RRD: Relative reach distance; p: p value; dz: Cohen d.						

Moreover, the ability of the Y-BT maximal reach directions, to detect small performance change was rated as good, and was not affected by limb symmetry ($SEM < SWC$) (Table 2). In addition, the ability to detect small performance change was also rated as good in the SRT test ($SEM < SWC$) (Table 2). In addition, the excellent test-retest reliability was not affected by the form of expression of the Y-BT outcome, beyond it was expressed as maximal reach distance or composite score performance. The CS, ARD and RRD of DOML and NDOML presented excellent reliability (Table 2).

Table 2. Test-retest reliability of the Y-BT and the SRT tests (n = 81)

	Leg	Score performance	ICC (95% CI)	SEM	SEM%	SWC	MDC ₉₅	MDC%
Y-BT	DOML	AT (cm)	0.97 (0.95-0.98)	0.69	0.74	1.11	1.91	2.74
		PM (cm)	0.98 (0.97-0.99)	0.61	0.96	1.47	1.69	2.66
		PL (cm)	0.98 (0.97-0.98)	0.55	0.74	1.21	1.52	2.06
		CS (%)	0.98 (0.97-0.99)	0.50	0.55	1.31	1.38	1.54
		ARD (cm)	0.98 (0.97- 0.99)	0.38	0.55	1.03	1.06	1.53
		RRD (%)	0.98 (0.97-0.98)	0.007	0.77	0.01	0.02	2.17
	NDOML	AT (cm)	0.98 (0.97-0.98)	0.51	0.75	1.19	1.41	2.08
		PM (cm)	0.97(0.96-0.98)	0.94	1.47	1.52	2.60	4.07
		PL (cm)	0.98 (0.97-0.98)	0.63	0.88	1.24	1.74	2.42
		CS (%)	0.98 (0.97-0.99)	0.52	0.58	1.34	1.44	1.36
		ARD (cm)	0.98 (0.97-0.99)	0.39	0.57	1.05	1.08	1.59
		RRD (%)	0.98 (0.97- 0.99)	0.007	0.79	0.01	0.02	2.27
SRT		0.98 (0.99-0.99)	0.07	1.61	0.19	0.19	4.44	
Notes: DOML: Dominant leg; NDOML: Non Dominant leg; AT: Anterior; PM: Postero medial; PL: Postero lateral; CS: Composite reach distance; ARD: Absolute reach distance; RRD: Relative reach distance. ICC: intraclass correlation coefficient; CI: confidence interval; SEM: standard error of measurement; SEM%: relative standard error of measurement; MDC ₉₅ =minimal detectable change at a 95% confidence level; MDC ₉₅ %: relative minimal detectable change at a 95% confidence level.								

Relationship between the Y-BT and SRT performance indices

The assumption of heteroscedasticity of the raw residuals was assessed via the Konker test which tests the null hypothesis that variance of the residuals is homogenous. The raw residuals for AT, PM, PL CS, ARD and RRD ($\chi^2=0.13, p=0.73$; $\chi^2=0.15, p=0.85$; $\chi^2=0.45, p=0.98$; $\chi^2=0.35, p=0.57$; $\chi^2=0.30, p=0.99$; $\chi^2=0.59, p=0.76$, respectively) were homoscedastic. Therefore, data from the present study have met the criteria of correlation diagnostics. Relationships between SRT and all Y-BT outcomes (for DOML) are presented in Table 3. Any statistical significance was observed. Correlations approached zero (Table 3).

Table 3. Correlations between Y-BT relative and absolute outcomes and the SRT performance (n=71)

		Y-BT score performance					
		CS (%)	ARD (cm)	RRD (%)	AT (cm)	PM(cm)	PL (cm)
SRT	R	0.07	-0.06	0.01	-0.04	-0.04	-0.08
	P	0.56	0.58	0.43	0.72	0.70	0.51

Notes: CS: Composite reach distance; ARD: Absolute reach distance; RRD: Relative reach distance; AT: Anterior; PM: Posteromedial; PL: Postero lateral; r: correlation coefficient; p: p value.

Construct Validity of Performance Scores on the Y-BT maximal reach distances

Mean ($\pm$ SD) Y-BT scores for DOML and NDOML are listed in Table 4. Normality of the distribution of residuals was verified using the Kolmogorov-Smirnov. The calculated D_{ratio} was less than D_{cri} value ($D_{cri}=0.11$ when $n=151$) represents normal distribution. The raw residual data for all the distributions were normally distributed. The D_{ratio} ranged from 0.052 to 0.102 with $p>0.05$. Independent sample t-test revealed that Y-BT DOML score performances were higher than Y-BT NDOML score performances. This significant bias ($p<0.05$) between DOML and NDOML performances was denied by a small estimated dz , which ranged from 0.33 to 0.46 making insinuation to a weak discriminate ability of the Y-BT score performances in young male team sport athletes (Table 4). Receiver operating characteristic analysis showed that the Y-BT score performances were insensitive in detecting differences between DOML and NDOML in this population of young soccer and basketball players. The AUC was under 0.7 for all the score performances. The AUC ranged from 0.6 to 0.63 (Table 4).

Table 4. The Y-BT score performances within dominant and non-dominant legs (n=151)

	Score performance	Dominant leg	Non-Dominant leg	<i>p</i>	<i>dz</i>	AUC (95% CI)	<i>P</i>
		mean ($\pm$ SD)	mean ($\pm$ SD)				
Y-BT	AT (cm)	70.30 ($\pm$ 6.19)	67.97 ($\pm$ 6.57)	0.001	0.38	0.60 (0.54-0.66)	0.001
	PM (cm)	64.77 ($\pm$ 8.15)	62.02 ($\pm$ 8.21)	0.003	0.33	0.59 (0.53-0.65)	0.002
	PL (cm)	73.86 ($\pm$ 7.98)	70.21 ($\pm$ 7.65)	0.000	0.46	0.62 (0.56-0.68)	0.000
	CS (%)	92.88 ($\pm$ 9.39)	89.07 ($\pm$ 8.98)	0.000	0.41	0.63 (0.57-0.69)	0.000
	ARD (cm)	69.64 ($\pm$ 6.46)	66.73 ($\pm$ 6.35)	0.000	0.45	0.63 (0.56-0.69)	0.000
	RRD (%)	0.91 ($\pm$ 0.095)	0.87 ($\pm$ 0.091)	0.000	0.43	0.62 (0.56-0.68)	0.000

Notes: AT: Anterior; PM: Postero-medial; PL: Postero lateral; CS: Composite reach distance; ARD: Absolute reach distance; RRD: Relative reach distance; *p*: p value; *dz*: Cohen *d*; AUC: area under the curve.

Discussion

The main finding of this study showed that the Y-BT score performances are reliable but not discriminate tools for monitoring asymmetry in prepubescent team sport athletes.

According to Impellizzeri & Marcora, (2009), reliability results are most often used to judge whether a test protocol should be widely used. The Y-BT maximal reach distances and composite scores were highly reliable for both DOML and NDOML in our team players (Table 2). In terms of relative reliability, the intraclass correlation coefficients between Y-BT test and retest results were “good”. ICC values exceeded 0.90. Our results are in line with the literature (Faigenbaum, Lloyd, MacDonald, & Myer, 2016; Filipa, Byrnes, Paterno, Myer, & Hewett, 2010; Geldhof, Cardon, De Bourdeaudhuij, Danneels, Vanderstraeten, & Clercq, 2006; Plisky, Rauh, Kaminski, & Underwood, 2006; Schwiertz, Brueckner, Schedler, Kiss, & Muehlbauer, 2019). In a previous study Faigenbaum et al. (2016) reported an $ICC_{(3,1)}$ between 0.75 to 0.82 in children of 6-12 years old. The $ICC_{(3,1)}$ were moderate-to-good for the AT (right=0.82; left=0.82), PM (right = 0.77; left = 0.75), PL (right 0.80; left = 0.77) reach directions. For instance, Filipa et al. (2010) reported an $ICC_{(3,1)}$ of 0.81 to 0.96 for the Y-BT reach directions in adolescent athletes. Similarly, excellent between session relative reliability for the Y-BT reach directions was reported in a sample of high school basketball players ($ICC = 0.82-0.87$) by Plisky et al. (2006). More recently Schwiertz et al. (2019) detected lower ICC values (0.40-0.90) in healthy adolescents from grade 6 to 11 years.

Relative reliability is a limited measure of reliability. It is affected by the magnitude of the between-individuals variation. Also, it does not indicate possible systematic bias and its analysis outcomes cannot be clinically interpreted. Consequently, absolute reliability is recommended to be used, in conjunction with relative reliability. If data are homoscedastic, as in the present study, an appropriate measure of absolute reliability is the SEM (Nevill, & Atkinson, 1997), which can be presented in absolute or relative expression. Thus, absolute reliability in the present study was evaluated by the SEM only. In the present study, absolute reliability was evaluated by the SEM. In fact, the absolute reliability of the Y-BT score performances was very good, with SEM% below the 5% and ranged from 0.55 to 0.98%. Compared to other studies our SEM% values are smaller to those reported by Faigenbaum et al. (2014) (3.03 to 12.32%) and Schwiertz et al. (2019) (1.77 to 5.81%). Of note, we detected the largest ICC /lowest SEM% values for the CS in both DOML and NDOML legs. Similar results have been reported by Schwiertz et al. (2019). According to these authors, this finding is most likely based on the higher number of data points (i.e., AT, PM, PL) compared to the measurements of each leg individually. Accordingly, practitioners are advised to use the CS to evaluate dynamic balance performance in healthy adolescents because this parameter showed the most reliable results. However, this finding did not correspond to the previous Y-BT reliability studies performed by Shaffer et al. (2013) on active military personnel and by Faigenbaum et al. (2016) on recreational children.

Our reliability results imply that the Y-BT is a reproducible test for the assessment of dynamic balance performance in prepubescent soccer and basketball players. Nevertheless, having good reliability does not mean that a test is useful, as a number of physical measures have high reliability but some are not sensitive measurements tools (Hopkins, Elske, Schabert, & Hawley, 2001; Schwenk, Gogulla, Englert, Czempik, & Hauer, 2012). In the current study, the sensitivity of the Y-BT was also analyzed. The SWC was smaller than the SEM for all Y-BT scores (Table 2). The usefulness of the Y-BT reach distances of DOML and NDOML legs and CSs was rated as “good”.

The MDC_{95} , refers to the minimal amount of change score outside of measurement error that may reflect true change with training (Schwenk, Gogulla, Englert, Czempik & Hauer, 2012; Haley & Fragala-Pinkham, 2003). In the current study, the MDC_{95} for individual and composite reach measures across the study population did not demonstrate side-to-side differences, except non-dominant PM maximal reach direction and ranged from 0.02 to 4.07%. Compared to other studies, our values are better than those (7.68 to 13.70%) reported by Linek et al. (2017), those stated (8.4 to 34.2%) by Calatayud et al. (2014) and those (4.90-16.10%) mentioned by Schwiertz et al. (2019). The differences between our values and those of the above mentioned studies might be based on the investigated cohort. From a practical perspective, our obtained $MDC_{95}\%$ values suggest that this is the minimum amount of change needed to exist between pre- and post-testing Y-BT performance scores for a change to be practically significant in pre-pubescent soccer and basketball players.

One of the findings of this study was that the SRT is a reliable tool for the evaluation of hamstring flexibility in prepubescent soccer and basketball players. Previous studies have analyzed test-retest reliability of SRT (Ayala, Sainz de Baranda, De Ste Croix, & Santonja, 2012; Davis, Quinn, Whiteman, Williams, & Young, 2008; Hui, Morrow, & Jackson, 1999; Hui, & Yuen, 2000). These studies have subsequently reported an excellent reliability, with ICCs ranging from 0.91 to 0.99. Similar results have been found in our study, with ICC scores of 0.98. From these mentioned above studies, only Ayala et al. (2012) have reported absolute reliability of SRT, expressed as coefficient of variation in young male and female adults ($CV = 8.5\%$). In fact, the reported absolute reliability coefficient of the SRT in the current study was very good, with SEM far below 5% (Table 2). The SRT has demonstrated in our athletes “good” sensitivity. The calculated SWC (0.19 s) was smaller than its respective SEM (0.07 s). Thus, the usefulness of the SRT for the evaluation of hamstring flexibility in prepubescent soccer and basketball players can be rated as “good” (Hopkins, Elske, Schabert, & Hawley, 2001) in our team players.

The MDC_{95} value in this investigation informs that 95% of the SRT outcome will demonstrate a random variation as a result of measurement error of less than 4.44%. To our knowledge, no study has examined the MDC_{95} of the SRT in young team soccer and basketball players, so it is difficult to compare this result to those of the specific literature. This limitation illustrates the necessity of having an MDC value that is specific to particular characteristics of athletes, i.e. specialty gender, team sport, age and level of expertise. Thus, future studies are needed to attempt to quantify a specific MDC_{95} differentiated by, team sport, age, gender and expertise.

An important prerequisite of a test or experiment measure applicability in sport and medicine sciences is its construct validity. Construct validity is commonly analyzed by testing the construct that it claims to be measuring. In the research design of our study we operated control over leg dominance (i.e. the construct).

Results showed significant differences in the Y-BT scores between the two legs (i.e., dominant vs. non-dominant) (Table 4). The ROC analysis is currently used to study the discriminate validity of performance testing (Castagna, Impellizzeri, Bizzi, Weston, & Manzi, 2011; Hachana, Chaabène, Ben Rajeb, Khelifa, Aouadi, Chamari, & Gabbett, 2014; Menespa, Aldo, & Impellizzeri, 2010; Stiffler, Bell, Sanfilippo, Hetzel, Pickett, & Heiderscheit, 2017). A test was considered as discriminate if its area under the ROC curve was equal to or greater than 0.70 (Castagna, Impellizzeri, Bizzi, Weston, & Manzi, 2011). The observed difference between DOML and NDOML in our study was not supported by ROC analysis. The area under the ROC curve was below 0.70 for all Y-BT scores (Table 4) suggesting that the Y-BT maximal reach distances and composite scores have a weak discriminate ability and cannot effectively guide coaches, clinicians and pediatric researchers to monitor legs asymmetry in prepubescent soccer and basketball male players. Due to the absence of any earlier study dealing with construct validity of Y-BT maximal reach distances and composite scores between dominant and non-dominant leg, the present research is unable to compare our results with previous findings.

To the authors' knowledge, our study is the first investigation to examine relationships between measures of dynamic balance and hamstring flexibility in prepubescent male soccer and basketball players. Our findings revealed a non-significant association between dynamic measure of balance and hamstring flexibility in these team sport athletes (Table 3). Likewise, Y-BT scores (DOML) independently of their form of expression were not associated to SRT performance. Due to the absence of any earlier study dealing with the relationship between Y-BT outcomes and SRT, the present research is unable to compare results with previous findings.

One of the potential limitations of this study is the population used. Although this investigation used a relative large sample size ($n = 151$), the participants were homogenous based on age, which limits the external validity of the results. Furthermore, only soccer and basketball players took part in this study, thus practitioners should not extrapolate our findings to other populations. Only the Y-BT was used in the current study, thus, the results should not be generalized to other dynamic balance tests.

Conclusion

In conclusion, this study provides estimates of reliability and measurement error of the Y-BT maximal reach directions and composite scores in a cohort of 151 prepubescent soccer and basketball male players. All Y-BT maximal reach directions and composite scores for dominant and non dominant legs demonstrated good levels of absolute and relative reliability. Moreover, the $MDC_{95\%}$ of the Y-BT score performance were rated as "satisfactory", and their capacities to detect change are considered "good". However, construct validity of Y-BT maximal reach distances and composite scores was weak in detecting differences between DOML and NDOML in our population of young soccer and basketball players. These findings strengthen our understanding of Y-BT as a tool, for the evaluation of leg asymmetry in prepubescent soccer players. However, additional studies are needed to confirm or deny the results of this study.

Conflicts of interest

There is no conflict of interest including any financial, personal or other relationships with other people or organizations that have influenced the performance of this work.

Acknowledgements

Authors express their gratitudes to coaches and participants for their adherence to this study.

References

- Atkinson, G., & Nevil, A.M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med.* 26(4), 217-38.
- Ayala, F., Sainz de Baranda, P., DeSte Croix, M., & Santonja, F. (2012). Reproducibility and criterion-related validity of the sit and reach test and toe touch test for estimating hamstring flexibility in recreationally active young adults. *Phy Ther in Sport.* 13(4), 219-26.
- Baltaci, G., Un, N., Tunay, V., Besler, A., & Gerçeker, S. (2003). Comparison of three different sit and reach in tests for measurement of hamstring flexibility female university students. *Br J Sports Med.* 37(1), 59-61.
- Breen, E.O., Howell, D.R., Straccolini, A., Dawkins, C., & Meehan, W.P. (2016). Examination of age-related differences on clinical tests of postural stability. *Sports Health.* 8(3), 244-49.

- Calatayud, J., Borreani, S., Colado, J.C., Martin, F., & Flandez, J. (2014). Test-retest reliability of the Star Excursion Balance Test in primary school children. *Phys Sportsmed.* 42(4), 120-124.
- Castagna, C., Impellizzeri, F.M., Bizzi, M., Weston, M., & Manzi, V. (2011). Applicability of a change of direction ability field test in soccer assistant referees. *J Strength Cond Res.* 25(3), 860-66.
- Castro-Piñero, J., Chillon, P., Ortega, F.B., Montesinos, J.L., Sjöström, M., & Ruiz, J. R. (2009). Criterion-related validity of sit-and-reach and modified sit-and-reach test for estimating hamstring flexibility in children and adolescents aged 6-17 years. *Int J Sports Med.*30 (9), 658-62.
- Chimera, N.J., Knoeller, S., Cooper, R., Kothe, N., Smith, C., & Warren, M. (2017). Prediction of fonctionel movement screen performance from lower extremity range of nation and core tests. *Int J Sports Phys Ther.* 12(2), 173-181.
- Chimera, N.J., Smith, C.A., & Warren, M. (2015). Injury history, sex, and performance on the functional movement screen and Y balance test. *J Athl Train.* 50(5), 475-85.
- Cohen, J. (1992). A power primer. *Psychological Bulletin.* 112(1), 155-159.
- Coppieters, M., Stappaerts, K., Janssens, K., & Jull, G. (2002). Reliability of detecting onset of pain and sub-maximal pain during neural provocation testing of the upper quadrant. *Physiother Res Int.* 7(3), 146-56.
- Coren, S. (1993). The lateral preference inventory for measurement of handedness, footedness, eyedness, and earedness: Norms for young adults. *Psycho Social Bulletin.*31 (1), 1-3.
- Coughlan, G.F., Fullam, K., Delahunt, E., Gissane, C., & Caulfield, B.M. (2012). A comparison between performance on selected directions of the star excursion balance test and the y balance test. *J Athl Train.* 47(4), 366-371.
- Davis, D.S., Quinn, R.O., Whiteman, C.T., Williams, J.D., & Young, C.R. (2008). Concurrent validity of four clinical tests used to measure hamstring flexibility. *J Strength Cond Res.* 22(2), 583-88.
- De Vet, H.C., Bouter, L.M., Bezemer, P.D., & Beurskens, A.J. (2001). Reproducibility and responsiveness of evaluative outcome measures. Theoretical considerations illustrated by an empirical example. *Int J Technol Assess Health Care.* 17(4), 479-487.
- Deyo, R.A., & Centor, R.M. (1986). Assessing the responsiveness of functional scales to clinical change: An analogy to diagnostic test performance. *J Chronic Dis.* 39(11), 897-906.
- Faigenbaum, A.D., Bush, J., McLoone, R., Kreckel, M., Farrell, A., Ratamess, N., & Kang, J. (2015). Benefits of strength and skill based training during primary school physical education. *J Strength Cond Res.* 29(5), 1255-1262.
- Faigenbaum, A.D., Lloyd, R.S., Mac Donald, J., & Myer, G.D. (2016). Citius, Altius, Fortius. Beneficial effects of resistance training for young athletes: Narrative review. *Br J Sports.* 50(1), 3-7.
- Faigenbaum, A.D., Myer, G.D., Fernandez, I.P., Carrasco, E.G., Bates, N., Farrell, A., Ratamess, N.A., & Kang, J. (2014). Feasibility and reliability of dynamic postural control measures in children in first through fifth grades. *Int J Sports Phys Ther.* 9(2), 140-8.
- Faul, F., Erdfelder, E., & GPOWER. (2004). A Priori, Post-Hoc, and Compromise Power Analyses for MS-DOS (Computer Program). *Bonn, FRG: Bonn University, Department of Psychology.*
- Filipa, A., Byrnes, R., Paterno, M., Myer, G., & Hewett, T. (2010). Neuromuscular training improves performance on the star excursion balance test in young female athletes. *J Orth Sports PhyTher.* 40(9), 551-58.
- Geldhof, E., Cardon, G., De Bourdeaudhuij, I., Danneels, L., Vanderstraeten, G., & De Clercq, D. (2006). Static and dynamic standing balance: test-retest reliability and reference values in 9 to 10 year old children. *Eur J Pediatrics.* 165(11), 779-86.
- Gonell, A.C., Romero, J.A.P., & Soler, L.M. (2015). Relationship between the Y balance test scores and soft tissue injury incidence in a soccer team. *Int J Sports PhyTher.* 10(7), 955-966.
- Hachana, Y., Chaabène, H., Ben Rajeb, G., Khlifa, R., Aouadi, R., Chamari, K., & Gabbett, T.J. (2014). Validity and Reliability of New Agility Test among Elite and Subelite under 14-Soccer Players. *PLoS One.* 9(4), e95773.
- Haley, S.M., & Fragala-Pinkham, M.A. (2003). Interpreting change scores of tests and measures used in physical therapy. *Phy Ther.* 86(5), 735-43.
- Hammami, R., Chaouachi, A., Makhoulouf, I., Grancher, U., & Behm, D.G. (2016). Associations Between Balance and Muscle Strength, power Performance in Male Youth Athletes of Different Maturity Status. *Pediatr Exerc Sci.* 28(4), 521-534.
- Hartman, J.G., & Looney, M. (2003). Norm-reference and criterion-referenced reliability and validity of the back-saver sit-and-reach. *Meas Phys Educ Exerc Sci.*7(2), 71-87.
- Hastad, D.N., & Lacy, A.C. (1994). Measurement and evaluation in physical education and exercise science (2ndedn.). *Scottsdale, AZ: Gorsuch Scarisbric.*
- Hianik, J., Šimonek, J., & Horicka, P. (2017). Impact of a specialized conditioning program on the development of postural stability of soccer referees. *J Phys Educ Sport.* 17(1), 135 - 141.
- Hopkins, W.G. (2000). Measures of reliability in sports medicine and science. *Sports Med.* 30(1), 1-15.
- Hopkins, W.G., Elske, J., Schabert, E.J., & Hawley, J.A. (2001). Reliability of power in physical performance tests. *Sports Med.* 31(3), 211-34.

- Hui, S.C., Morrow, J.R., & Jackson, A.W. (1999). Comparison of the criterion-related validity of sit-and-reach tests with and without limb length adjustment in Asian adults. *Res Q Exerc Sport*. 70(4), 401-6.
- Hui, S.C., & Yuen, P.Y. (2000). Validity of modified back-server sit and reach test: a comparison with other protocols. *Med Sci Sports Exerc*. 32(9), 1655-59.
- Impellizzeri F.M. & Marcora S.M. (2009). Test validation in sport physiology: Lessons learned from clinimetrics. *Int J Sports Physiol Perform*. 4(2), 269-277.
- Linek, P., Sikora, D., Wolny, T., & Saulicz, E. (2017). Reliability and number of trials of Y Balance Test in adolescent athletes. *Musculoskelet Sci Pract*. 31, 72-75.
- Liow, D.K., & Hopkins, W.G. (2003). Velocity specificity of weight training for kayak sprint performance. *Med Sci Sports Exerc*. 35(7), 1232-37.
- Lloyd, R.S., Oliver, J.L., Faigenbaum, A.D., Howard, R., De Ste Croix, M.B., Williams, C.A., Best, T.M., Alvar, B.A., Micheli, L.J., Thomas, D.P., Hatfield, D.L., Cronin, J.B., & Myer, G.D. (2015). Long-term athletic development- part 1: a pathway for all youth. *J Strength Cond Res*. 29(5), 1439-50.
- Malina, R.M., Bouchard, C., & Bar-Or, O. (2004). Growth. Maturation and physical activity. *Champaign, IL: Human Kinetics*.
- Mannion, A.F., Elfering, A., Staerkle, R., Junge, A., Grob, D., Semmer, N.K., Jacobshagen, N., Dvorak, J., & Boos, N. (2005). Outcome assessment in low back pain: how low can you go? *Eur Spine J*. 14(10), 1014-26.
- Mayorga-Vega, D., Merino-Marban, R., & Viciania, J. (2014). Criterion-Related Validity of Sit-And-Reach Tests for Estimating Hamstring and Lumbar Extensibility: A Meta-Analysis. *J Sports Sci Med*. 13(1), 01-14.
- Menespa, P., Aldo, S., & Impellizzeri, F.M. (2010). Aerobic fitness variables do not predict the professional career of young cyclists. *Med Sci Sports Exerc*. 42(4), 805-12.
- Miller, M.M., Trapp, J.L., Post, E.G., Trigsted, S.M., McGuine, T.A., Brooks, M.A., & Bell, D.R. (2017). The Effects of Specialization and Sex on Anterior Y-Balance Performance in High School Athletes. *Sport Health*. 9(4), 375-382.
- Moore, S.A., McKay, H.A., Macdonald, H., Nettlefold, L., Baxter-Jones, A.D., Cameron, N., & Brasher, P.M. (2015). Enhancing a somatic maturity prediction model. *Med Sci Sports Exerc*. 47(8), 1755-64.
- Nevill, A., & Atkinson, G. (1997). Assessing agreement between measurements recorded on a ratio scale in sports medicine and sports science. *Br J Sports Med*. 3(4), 314-18.
- Palomares, B.R.A., Palomares, E.M.D.G., Uchôa, F.N.M., Andrade, R.D.A., Deana, N.F., & Alves, N. (2019). Effectiveness of the conjugate influence method in improving static and dynamic balance in rhythmic gymnastics gymnasts. *J Phys Educ Sport*, 19(4), 1407 – 1417.
- Plisky, P.J., Gorman, P.P., Butler, R.J., Kiesel, K.B., Underwood, F.B., & Elkins, B. (2009). The reliability of an instrumented device of measuring components of the Star Excursion Balance Test. *Am J Sports Phys Ther*. 4(2), 92-99.
- Plisky, P.J., Rauh, M.J., Kaminski, T.W., & Underwood, F.B. (2006): Star Excursion Balance Test as a predictor of lower extremity injury in high school basketball players. *J Orthop Sports Phys Ther*. 36(12), 911-9.
- Sarabon, N., Mlaker, B., & Markovic, G. (2010). A novel tool for the assessment of dynamic balance in healthy individuals. *Gait & Posture*. 31(2), 261-264.
- Scanlan, A.T., Dascombe, B.J., & Reaburn, P.R.J. (2012). The construct and longitudinal validity of the basketball exercise simulation test. *J Strength Cond Res*. 26(2), 523-30.
- Schwenk, M., Gogulla, S., Englert, S., Czempik, A., & Hauer, K. (2012). Test-retest reliability and minimal detectable change of repeated sit-to-stand analysis using one body fixed sensor in geriatric patients. *Physiol Meas*. 33(11), 1931-46.
- Schwartz, G., Brueckner D, Schedler S, Kiss R, & Muehlbauer T. (2019). Performance and reliability of the Lower Quarter Y Balance Test in healthy adolescents from grade 6 to 11. *Gait Posture*. 67,142-146.
- Shaffer, S.W., Teyhen, D.S., Lorensen, C.L., Warren, R.L., Koreerat, C.M., Straseske, C.A., & Childs, J.D. (2013). Y-balance test: a reliability study involving multiple raters. *Military Med*. 178(11), 1264-70.
- Smith, C.A., Chimera, N.J., & Warren, M. (2015). Association of y balance test reach asymmetry and injury in division I athletes. *Med Sci Sports Exerc*. 47(1), 136-141.
- Smith, L.J., Creps, J.R., Bean, R., Rodda, B., & Alsalaheen, B. (2018). Performance and reliability of the y-balance test in high school athletes. *J Sports Med Phys Fitness*. 58(11), 1671-1675.
- Sporis, G., Jukic, I., Milanovic, L., & Vucetic, V. (2010). Reliability and factorial validity of agility tests for soccer players. *J Strength Cond Res*. 24(3), 679-86.
- Stiffler, M.R., Bell, D.R., Sanfilippo, J.L., Hetzel, S.J., Pickett, K.A., & Heiderscheit, B.C. (2017). Star Excursion Balance Test Anterior Asymmetry Is Associated With Injury Status in Division I Collegiate Athletes. *J Orthop Sports Phys Ther*. 47(5), 339-346.
- Yiannis, M., Christos, S., Vassilis, P., Dimitrios, M., Konstantinos, M., & Thomas, M. (2020). Association between jump asymmetry and reduced performance in the change of direction tests of youth soccer players. *J Phys Educ Sport*. 20(3), 1362 – 1368.