

**Feasibility and reliability of health-related physical fitness
tests among Colombian preschool children: the Fuprecol kids
study**

By

Julio Cesar Amado Pacheco

A thesis submitted to the School of Medicine and Health

Science of Universidad del Rosario

in partial fulfillment of the requirements for the
degree of Master of Physical Activity and Health

Daniel Humberto Prieto Benavides, MSc (Advisor)

Robinson Ramírez-Vélez, PhD (Co-advisor)

Center of Studies in Physical Activity Measurements

School of Medicine and Health Sciences

Universidad del Rosario, Bogotá D.C, Colombia

Center of Studies in Physical Activity Measurements

School of Medicine and Health Sciences

Universidad del Rosario

Bogotá D.C,

Colombia

2018

Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Universidad del Rosario, I hereby grant to Universidad del Rosario and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

ABSTRACT

Objective: to assess the reliability of health-related physical fitness field tests used in the “Fuprecol Kids” study among Colombian preschool children aged 3–5 years.

Methods: A total of 90 preschoolers aged 3-5 years participated in the study. Weight, height, waist circumference, cardiorespiratory fitness (CRF) (PREFIT 20 m shuttle run), musculoskeletal fitness (handgrip strength and standing broad jump), speed-agility (4 × 10 m shuttle run), and flexibility (sit and reach) components were tested twice (2 weeks apart). Each anthropometric component was used as an indicator of precision by technical error of measurement (TEM). The agreement between test–retest trials of all fitness tests was assessed following the Bland-Altman method. To determine the concordance correlation between test–retest measures, we used Lin's concordance correlation coefficient.

Results: There were no significant differences in test–retest mean differences, except for waist circumference ($P = 0.001$), CRF ($P = 0.018$), and flexibility ($P = 0.002$) in girls. The TEMs were 0.719 kg for the weight, 0.001 cm for the height, 0.053 kg/m² for the BMI, and 1.230 cm for the waist circumference. The reliability of anthropometric measurements was always higher than 99%, except in waist circumference (79%). Finally, when the fitness assessments were performed twice, the systematic error was nearly zero for all tests.

Conclusions: The results from this study indicate that the “Fuprecol kids” battery of tests, administered by physical education teachers, was reliable for measuring health-related components of fitness in preschoolers in a school setting in Colombia. However, the PREFIT 20 m shuttle run test showed poor reliability in our study in girls.

keywords: Reliability; health-related physical fitness; morphological; preschooler.

INTRODUCTION

Motor fitness (i.e. speed/agility), musculoskeletal fitness (MSF) and cardiorespiratory fitness (CRF) are the powerful health-related fitness components in youths ¹. In particular, low CRF and low MSF are independently associated with increased risk of cardio-metabolic disease ²⁻⁴ and mortality ^{5, 6} in both adolescent and adult populations. It has been consistently reported that a higher adiposity (i.e. central abdominal or fat mass) and metabolic risk factors are associated with lower CRF and MSF levels in young people ⁷. Based on this evidence, youth fitness assessment guidelines have called for a better understanding of the inter-relationship between physical fitness and body composition ⁸.

Owing to the importance of health-related fitness components for the current and future health in youths, it is important that intervention studies use feasible, reliable and valid measures to assess fitness. A review of previous literature relevant to establishing a fitness test battery in early ages, revealed the only reference used for field-based fitness-test battery in preschool children (<6 years old) was the PREFIT (Preschool children Fitness testing) battery ⁹.

The health-related fitness tests, included in the “Fuprecol study” (Asociación de la fuerza prensil con manifestaciones tempranas de riesgo cardiovascular en niños y adolescentes Colombianos "ESTUDIO FUPRECOL" in Spanish), have been previously validated in youths ¹⁰. However, the reliability of physical fitness tests has not been explored in preschool children. Thus, the present study addresses the reliability of the “Fuprecol kids” battery of tests and contributes to a better understanding for the researchers and practitioners about which health-related fitness tests should be used in preschool children (<6 years old). Identifying from an early age those who have a

healthy/unhealthy body composition or are fit/unfit is of crucial importance for public health.

In Colombia, a region which has undergone a well-documented epidemiologic transition fueling a non-communicable diseases epidemic, relatively little research on physical activity and physical fitness exists ^{11, 12}. This is important to assess, particularly in the context of a low-to-middle income country setting like Colombia ¹⁰. Therefore, the aim of the study was to assess the reliability of health-related physical fitness field tests used in the “Fuprecol kids” study among Colombian preschoolers.

METHODS

Participants and study design

The present cross-sectional study was conducted to evaluate test–retest reliability, with two complete sets of assessments included in the “Fuprecol kids” study. A total of 90 healthy preschool children (48 boys and 42 girls) aged 3–5 years (hereinafter called preschoolers) were recruited from different schools located in the Bogota capital district, in the Cundinamarca Department, Andean region. It is located at approximately 4° 35' 56" N 74° 04' 51" W, at an elevation of approximately 2,625 meters (min: 2,500, max: 3,250) above sea level.

A convenience sample of volunteers was included in groups by sex and age with 1-year increments (a total of three groups). Convenience sample size was estimated at 30 participants per age-sex group (~50% boys). The recruitment period lasted from June 2017 to January 2018. The Review Committee for Research Human Subjects at the University of Rosario (code N° CEI-ABN026-000262) approved the study. A comprehensive verbal description of the nature and purpose of the study and its experimental risks was given to the preschool children and their parents/guardians. All participants and their parents/legal guardians provided written informed consent before

entering the study. The protocol was in accordance with the latest revision of the Declaration of Helsinki.

Procedures

Consistent with a previous systematic review ¹³ and recommendations ⁸, we restricted our analysis to health-related field-based tests that have demonstrated adequate levels of criterion-related validity ^{9, 14–16}, and reliability ¹⁰ in the assessment of four components of the “Fuprecol kids” battery of tests. Thus, the following tests were included: weight, height and waist circumference to assess the morphological component; the PREFIT 20 m shuttle run test (PREFIT 20 m SRT) to assess the CRF component; handgrip strength and standing long jump tests to assess the musculoskeletal component (upper and lower limbs, respectively); 4 × 10 m shuttle run test (4 × 10 m SRT) to assess the speed-agility component; and finally, the sit and reach test to assess the flexibility component.

At each school, a team of trained CEMA center evaluators administered the tests in partnership with the physical education instructor. To determine the test–retest reliability of the “Fuprecol kids” battery of test, the assessments were administered twice (two weeks apart as previously done in similar reliability studies) ¹⁰ under the same physical conditions and by the same physical education instructor. Re-testing was performed at the same time of day to minimize circadian rhythm variability. The preschoolers wore sports clothing and footwear during testing.

Morphological component

Weight (Tanita, model BF689 Tokyo, Japan), height (Seca® 206, Hamburg, Germany) and waist circumference (WC) were measured without shoes and with light clothing. WC was measured using a metal tape measure (Lufkin W606PM®, Parsippany,

New Jersey, USA) at the level of the umbilicus zone in the horizontal plane. Body mass index (BMI) was calculated.

Cardiorespiratory component

CRF, PREFIT 20 m shuttle run ($ml \cdot kg^{-1} \cdot min^{-1}$). Preschoolers ran in a straight line between two lines 20 m apart while keeping pace with pre-recorded audio signals. The initial speed was 6.5 km/h and increased by 0.5 km/h per minute. The test was finished when the child failed to reach the end lines while keeping pace with the audio signals on two consecutive occasions or when the child stopped because of fatigue. The results were recorded to the nearest stage (minute) completed. The audio signal used will be freely available in Spanish for download at the website of Universidad de Granada research group: <http://profith.ugr.es/>. The feasibility, reliability and maximality of this test in preschoolers have been reported elsewhere ⁹.

Musculoskeletal component

Upper limbs, handgrip strength (kg). Handgrip strength (HGS) was measured using a standard adjustable handle analogue handgrip dynamometer T-18 TTK SMEDLY III[®] (Takei Scientific Instruments Co., Ltd, Niigata, Japan). The HGS was measured to the nearest kilogram (kg) twice on each hand (alternating). The test was done in the standing position, with the wrist in a neutral position and the elbow extended. Pupils were given verbal encouragement to “squeeze as hard as possible” and apply maximal effort for at least three seconds. Two trials were allowed for each limb and the average score recorded the peak HGS (kg). Thus, the HGS values presented here combine the results of left- and right-handed subjects, without consideration for hand dominance. We periodically calibrated the dynamometer against known weights and found no evidence of drift ¹⁰.

Lower limbs, standing broad jump (cm). The participant stood behind the starting line and was instructed to push off vigorously and jump as far as possible. In order to help and guide the preschoolers to jump, first we put a stick along the take-off line, so they could easily line up their feet close to it. The test was repeated twice, and the best score was retained to the nearest 0.1 cm, the distance was measured between toes at take-off and heels at landing or whichever body part landed nearest to the take-off spot ¹⁰.

Motor component

Speed/agility test (speed of movement, agility and coordination assessment). Two parallel lines were drawn on the floor 10 m apart. The preschooler ran as fast as possible from the starting line to the other line and returned to the starting line, crossing each line with both feet every time. This was performed twice, covering a total distance of 40 m (4 × 10 m). To make this test simpler, two physical education instructors were positioned at the two extremities and participants had to touch the hand of each instructor (placed behind the line) and go back at maximum speed. The best of two attempts was recorded (seconds). A slip-proof floor, four cones, a stopwatch and three sponges were used to perform the test ¹⁰.

Flexibility component

Hamstring and lumbar extensibility were measured using the sit and reach test. Participants were asked to sit on the floor with legs out straight ahead. Feet with shoes off were placed with the soles flat against the test device and shoulder-width apart. Both knees were held flat against the floor. With hands on top of each other and palms facing down, the preschooler reached forward along the measuring line as far as possible. The measuring stick on the device has the zero mark at 25 cm before the feet. The result was recorded directly from the meter on the device ¹⁰.

Statistical analyses

The data are presented as the means \pm SD, unless otherwise stated. The agreement between test–retest trials of all tests (CRF, musculoskeletal, speed-agility and flexibility component) was assessed following the Bland-Altman method ¹⁷. The analysis measures bias as estimated from mean differences, the 95% confidence interval for bias, the limits of agreement and ± 1.96 SD of the difference. Sex differences of the studied health-related physical fitness tests were analyzed by a t-test on inter-trial difference (test 2 – test 1, hereafter called T2 – T1). As no sex-specific effect on reliability of the studied physical fitness tests was found, except in WC, CRF, and flexibility, the analyses were performed for both boys and girls together.

The morphological component was used as an indicator of precision by technical error of measurement (TEM). It is based on at least two measurements taken of the same child by the same observer (intra-observer variability) or by at least two observers taking the same measurement of the same child (inter-observer variability). The calculations for intra- and inter-observer error are broadly the same. The coefficient of reliability (R) estimates the proportion of between-subject variance in a measured population that is free from measurement error. Measures of R can be used to match the relative reliability of different anthropometric measurements, as well as of the same measurements in different age groups, and to estimate sample size requirements in anthropometric studies. R as a percentage (R%) was calculated using the following equation: $R\% = 1 - (\text{total TEM}^2 / \text{SD}^2)$. To compare TEMs assessed for different measurements or different populations, absolute TEM was converted into relative TEM (%TEM) using the following equation: $\% \text{TEM} = (\text{TEM} / \text{mean}) \times 100$.

To determine the concordance-correlation between test–retest measures, we used Lin's concordance correlation coefficient (*pc*). All calculations were performed using

IBM SPSS Statistics 24 software for Windows (SPSS, Chicago, Illinois, USA). For all analyses, the significance level was 0.05.

RESULTS

The characteristics for the four components of the “Fuprecol kids” study (mean value \pm SD) assessed twice, as well as the mean inter-trial difference in the study, are shown in Table 1. Overall, there were no significant differences in test–retest mean differences, except in CRF ($P = 0.002$).

Table 1. Test (T1), retest (T2), mean differences (T2-T1) and concordance correlation coefficient (pc) of “Fuprecol Kids” battery.

Component	1st Trial (T1)	2st Trial (T2)	Inter-trial difference (T2-T1)*	pc	P value
Morphologic					
Weight (kg)	18.00 (2.89)	17.99 (2.86)	-0.011 (0.015)	0.999	0.484
Height (m)	1.06 (0.07)	1.06 (0.07)	-0.002 (0.001)	0.998	0.004
Body mass index (kg/m ²)	16.05 (1.53)	16.01 (1.54)	-0.040 (0.022)	0.991	0.073
Waist circumference (cm)	51.76 (6.87)	52.50 (4.79)	0.740 (0.518)	0.698	0.157
Cardiorespiratory					
20-m shuttle run (stage)	5.89 (2.58)	6.18 (2.73)	0.289 (0.091)	0.949	0.002
Musculoskeletal					
Handgrip (kg) ^a	7.76 (2.21)	7.78 (2.28)	0.020 (0.069)	0.958	0.772
Standing broad jump (cm)	60.77 (18.73)	60.90 (18.72)	0.133 (0.409)	0.978	0.745
Motor					
Sit and reach (cm)	18.06 (3.11)	17.82 (2.70)	-0.133 (0.162)	0.933	0.413
4x10m shuttle run (s)	23.46 (4.26)	23.32 (3.93)	-0.234 (0.142)	0.902	0.103

*One-sample t-test. P value refers whether a mean difference is significantly different from 0 for all measures.

Table 2 shows the reliability statistics by sex. Overall, there were no significant differences in test–retest mean differences in the boys ($P > 0.05$), but there were differences in waist circumference ($P = 0.001$), CRF ($P = 0.018$), and flexibility ($P = 0.002$) variables, in the girls.

Table 3 shows the inter-observer TEM and R% for each morphological component variable. The TEMs were 0.719 kg for the weight, 0.001 cm for the height, 0.053 kg/m² for the BMI, and 1.230 cm for the WC. The reliability of anthropometric measurements was always higher than 99%, except in WC ($R\% = 79$).

Table 2. Test (T1), retest (T2), mean differences (T2-T1) and concordance correlation coefficient (*pc*) of “Fuprecol kids” battery, by sex.

Component	1st Trial (T1)		2nd Trial (T2)		Inter-trial difference (T2_T1)*			
	Boys (n=48)	Girls (n=42)	Boys	Girls	Boys	P value (<i>pc</i>) Boys	Girls	P value (<i>pc</i>) Girls
Morphologic								
Age (years)	4.04 (0.82)	3.95 (0.82)	-	-	-	-	-	-
Weight (kg)	18.54 (2.92)	17.36 (2.69)	-	-	-	0.422 (0.999)	-	0.910 (0.999)
Height (m)	1.07 (0.08)	1.04 (0.07)	-	-	-	0.059 (0.998)	-	0.057 (0.998)
Body mass index (kg/m ²)	16.17 (1.31)	15.92 (1.75)	-	-	-	0.307 (0.984)	-	0.105 (0.995)
Waist circumference (cm)	51.39 (8.38)	52.18 (4.66)	-	-	-	0.216 (0.596)	-	0.001 (0.997)
Cardiorespiratory fitness								
20-m shuttle run (stage)	6.56 (2.88)	5.12 (1.95)	6.81 (3.13)	5.45 (2.00)	0.250 (0.125)	0.051 (0.962)	0.333 (0.131)	0.018 (0.903)
Musculoskeletal								
Handgrip strength (kg) ^a	8.10 (2.40)	7.36 (1.93)	8.19 (2.47)	7.30 (1.96)	0.088 (0.084)	0.304 (0.972)	-0.06 (0.11)	0.611 (0.931)
Standing broad jump (cm)	63.67 (17.75)	57.45 (19.47)	64.23 (17.63)	57.10 (19.40)	0.563 (0.503)	0.269 (0.981)	-0.36 (0.66)	0.592 (0.976)
Motor								
4x10m shuttle run (s)	17.34 (2.45)	18.88 (3.59)	17.28 (2.59)	18.45 (2.72)	-0.063 (0.139)	0.651 (0.928)	-0.43 (0.26)	0.104 (0.896)
Flexibility								
Sit and reach (cm)	23.19 (4.27)	23.76 (4.28)	23.46 (3.84)	23.17 (4.08)	0.271 (0.245)	0.275 (0.918)	-0.60 (0.18)	0.002 (0.960)

*One-sample t-test. P value refers whether a mean difference is significantly different from 0 for all measures.

Table 3. Inter-observer TEM, relative TEM and intra-observer morphologic component assessments of Fuprecol kids health-related physical fitness

	Mean (SD)	Inter-observer		Intra-observer
		TEM	%TEM	R%
Weight (kg)	17.99 (2.87)	0.719	3.994	0.999
Height (m)	1.06 (0.07)	0.001	0.117	0.999
Body mass index (kg/m ²)	16.03 (1.53)	0.053	0.330	0.995
Waist circumference (cm)	52.13 (5.39)	1.230	2.359	0.792

TEM: technical error of measurement

The Bland-Altman plots (Figure 1) graphically show the reliability patterns, in terms of systematic errors (bias or mean inter-trial differences) and random error (95% limits of agreement), of the “Fuprecol kids” battery of test. The systematic error when fitness assessment was performed twice was nearly zero for all the tests.

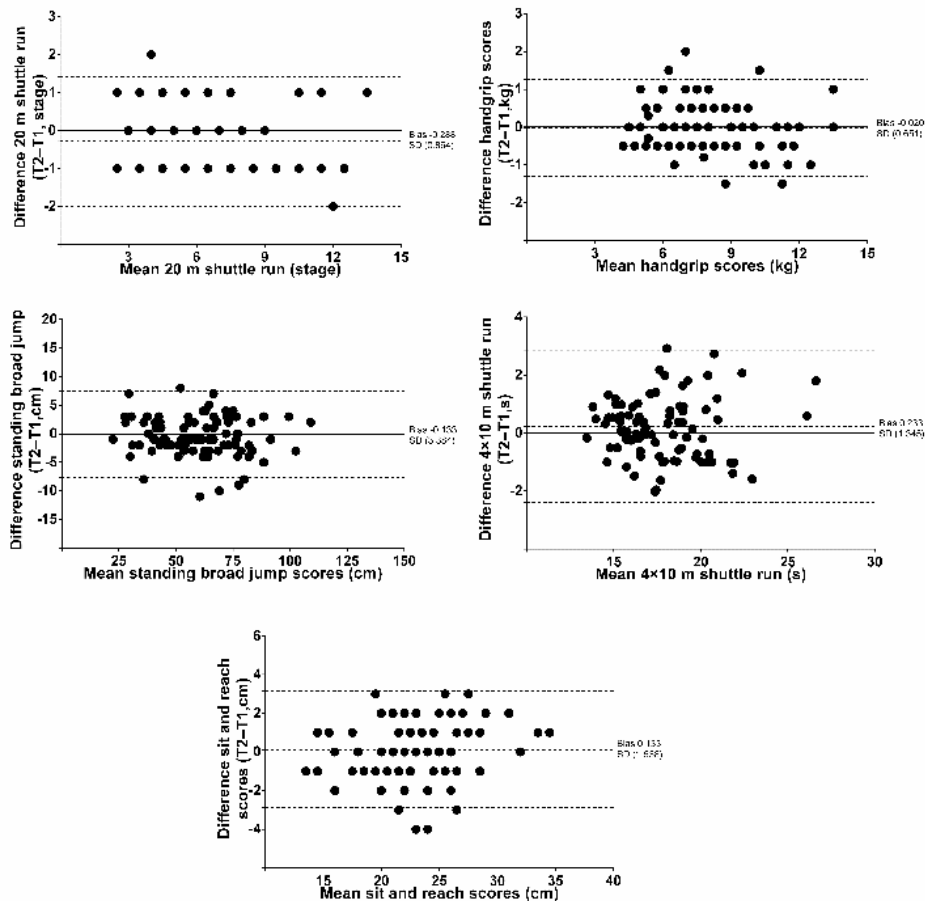


Figure 1. Bland–Altman plot of the PROFIT 20-m shuttle run, handgrip strength, standing broad jump, 4x10m shuttle and sit and reach among Colombian Preschool Children: The FUPRECOL Kids study

Legend. The central dotted line represents the mean differences between the second trial (T2) and the first trial (T1); the upper and lower dotted lines represent the upper and lower 95% limits of agreement (mean differences \pm 1.96 SD of the differences), respectively

DISCUSSION

To our knowledge, this is the first study assessing the reliability of health-related fitness tests, including its five components (morphological, musculoskeletal, motor, flexibility and CRF) in Latin American preschoolers. The main finding of our study shows that the “Fuprecol kids” battery of tests administered by physical education teachers is reliable for assessing the levels of physical fitness in preschoolers in a school environment in the Colombian setting. However, both the sit and reach test and the PREFIT 20 m shuttle run test showed poor reliability in our study in girls.

The present study showed that the intra-rater and inter-rater TEM and R% values were above the required levels. An allowance for measurement error might be up to 10% of the observed variance, which is equivalent to an R value of 90% or greater¹⁸. However, only when R is 99% is such an error unlikely¹⁹. Specifically, the TEMs for weight, height and BMI were frequently lower than 1 cm and the R% greater than 99% (inter-rater), except WC. Our results are similar to those found in other studies carried out with Colombian children and adolescents¹⁰ and European adolescents^{20–23}. In Colombian youths, older than those in the present study, Ramirez-Vélez et al.¹⁰ reported that TEMs were small and reliability was greater than 95% in all cases for height and the waist and hip circumferences, in contrast to our results for WC in girls. Another study in preschoolers also showed high reliability for anthropometric measures in both sexes⁹.

In contrast, for the musculoskeletal and speed-agility components, we found adequate reliability patterns, in terms of systematic errors (bias) and random error (95% limits of agreement), but not for CRF and sit and reach tests. For the CRF, we observed a systematic error of 0.333 in girls ($P = 0.018$), but no pattern of heteroscedasticity was

observed. Despite strong evidence indicating that the 20 m shuttle run produces results with good test–retest reliability in children and adolescents aged 8–18 years ^{13, 24}, in preschoolers the reliability of this test is questionable. However, it is important to highlight that in Spanish preschoolers, Cadenas-Sánchez et al. ⁹ observed a mean difference of +2 laps in the PREFIT 20 m shuttle run test, considering the different age groups of the participants and that assessments were 2 weeks apart, this might not be meaningful.

On the other hand, scientific evidence indicates that strength tests have produced moderate test–retest reliability. The handgrip strength test reported a mean difference of 0.02 kg and a high correlation ($r = 0.972$ and 0.931 for boys and girls, respectively). The mean difference and concordance correlation coefficient were lower than those showed by Spanish preschoolers of 0.24 kg ⁹ (Table 4) and 0.86 for both sexes, respectively, but were similar to those which were reported among Colombian children and adolescents ¹⁰.

Table 4. Comparison of the reliability of fitness tests between Fuprecol Kids study and PREFIT study.

Tests	Mean differences			
	Boys		Girls	
	Fuprecol	PREFIT	Fuprecol	PREFIT
CRF (stage)#	0.25 ± 0.12	1.02 ± 7.98	$0.33 \pm 0.13^*$	3.41 ± 6.87
Handgrip strength (kg)	0.08 ± 0.08	-0.38 ± 1.32	-0.06 ± 0.11	-0.05 ± 1.21
Standing jump (cm)	0.56 ± 0.50	$-7.51 \pm 14.34^*$	-0.36 ± 0.66	$-7.03 \pm 14.59^*$
4×10 m shuttle run (s)	-0.06 ± 0.14	0.07 ± 0.92	-0.43 ± 0.26	0.21 ± 1.15

*In these cases we consider that the reliability of fitness tests markedly differs.

The PREFIT study used laps instead stage, which was used in the present study

Therefore, these results confirm that when performing the handgrip strength test with the TKK dynamometer adapted to the hand size, the agreement between test and retest is the same throughout the range of measured values (homoscedasticity) ^{13, 24}. In addition, evidence has shown no significant differences in test–retest for the standing broad jump for European ^{13, 24} and Colombian ¹⁰ youths, this is in agreement with our results in preschool children. However, in this population other studies show

contradictory results. For example, Oja and Jürimäe ¹⁶ showed that the standing broad jump was highly reliable in four- and five-year-old preschool children, but the coefficient of variation was higher in girls than in boys. In contrast, for Spanish preschoolers the authors reported a systematic error of 7.31 cm (Table 4), suggesting that the reliability of this test is questionable, due to the higher coordination patterns needed for the standing long jump test and the difficulty observed in the preschool stage to perform it correctly ¹⁰.

Likewise, for the 4 × 10 m shuttle run test, the mean difference between measures in the 4 × 10 m SRT was -0.234 s with a high concordance correlation coefficient ($r = 0.90$). Other authors reported the reliability of this test in preschool years and concluded that the test showed an acceptable reliability ^{9, 16}, therefore, this test seems to be considered easy to measure.

Lastly, the reliability of the sit and reach test was analyzed in previous studies ^{10, 16}. Contrary to our findings, Oja and Jürimäe ¹⁶ demonstrated good test-retest reliability in 61 boys and girls aged 4–5 years ($r = 0.75$ to 0.93). The present study shows a systematic error of -0.60 cm in girls ($P = 0.002$), reaffirming the results in Colombian older schoolchildren ¹⁰.

In conclusion, the “Fuprecol kids” battery of tests is reliable for assessing the levels of physical fitness in preschoolers in a school environment in the Colombian setting. However, the PREFIT 20 m shuttle run and the sit and reach tests showed poor reliability in our study in girls. Therefore, the results from these two tests (PREFIT 20 m shuttle run and the sit and reach) should be interpreted cautiously when used in preschool girls.

REFERENCES

1. Ruiz JR, Castro-Piñero J, Artero EG, Ortega FB, Sjöström M, Suni J, et al. Predictive validity of health-related fitness in youth: a systematic review. *British journal of sports medicine*. 2009.
2. Ramírez-Vélez R, Peña-Ibagón JC, Martínez-Torres J, Tordecilla-Sanders A, Correa-Bautista JE, Lobelo F, et al. Handgrip strength cutoff for cardiometabolic risk index among Colombian children and adolescents: The FUPRECOL Study. *Scientific reports*. 2017;7:42622.
3. Agostinis-Sobrinho CA, Ramírez-Vélez R, García-Hermoso A, Moreira C, Lopes L, Oliveira-Santos J, et al. Low-grade inflammation and muscular fitness on insulin resistance in adolescents: Results from LabMed Physical Activity Study. *Pediatric diabetes*. 2017.
4. Lobelo F, Pate RR, Dowda M, Liese AD, Daniels SR. Cardiorespiratory fitness and clustered cardiovascular disease risk in US adolescents. *Journal of Adolescent Health*. 2010;47(4):352-359.
5. Kodama S, Saito K, Tanaka S, Maki M, Yachi Y, Asumi M, et al. Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women: a meta-analysis. *Jama*. 2009;301(19):2024-2035.
6. García-Hermoso A, Caverro-Redondo I, Ramírez-Vélez R, Ruiz J, Ortega FB, Lee D-C, et al. Muscular strength as a predictor of all-cause mortality in apparently healthy population: a systematic review and meta-analysis of data from approximately 2 million men and women. *Archives of Physical Medicine and Rehabilitation*. 2018.
7. Arsénault JE, Mora-Plazas M, Forero Y, Lopez-Arana S, Jáuregui G, Baylin A, et al. Micronutrient and anthropometric status indicators are associated with physical fitness in Colombian schoolchildren. *British journal of nutrition*. 2011;105(12):1832-1842.
8. Pillsbury L, Oria M, Pate R. *Fitness measures and health outcomes in youth*: National Academies Press; 2013.
9. Cadenas-Sánchez C, Martínez-Tellez B, Sánchez-Delgado G, Mora-González J, Castro-Piñero J, Löf M, et al. Assessing physical fitness in preschool children: Feasibility, reliability and practical recommendations for the PREFIT battery. *Journal of science and medicine in sport*. 2016;19(11):910-915.
10. Ramírez-Vélez R, Rodrigues-Bezerra D, Correa-Bautista JE, Izquierdo M, Lobelo F. Reliability of health-related physical fitness tests among Colombian children and adolescents: the FUPRECOL study. *PLoS One*. 2015;10(10):e0140875.
11. Rivera JÁ, de Cossío TG, Pedraza LS, Aburto TC, Sánchez TG, Martorell R. Childhood and adolescent overweight and obesity in Latin America: a systematic review. *The lancet Diabetes & endocrinology*. 2014;2(4):321-332.
12. Hallal PC, Parra DC, Azevedo MR, Pratt M, Brownson RC. Research on physical activity and health: where is Latin America. *J Phys Act Health*. 2010;7(Suppl 2):S129-S130.
13. Ortega FB, Cadenas-Sánchez C, Sánchez-Delgado G, Mora-González J, Martínez-Téllez B, Artero EG, et al. Systematic review and proposal of a field-based physical fitness-test battery in preschool children: the PREFIT battery. *Sports medicine*. 2015;45(4):533-555.

14. Cadenas-Sanchez C, Alcántara-Moral F, Sanchez-Delgado G, Mora-Gonzalez J, Martinez-Tellez B, Herrador-Colmenero M, et al. Assessment of cardiorespiratory fitness in preschool children: adaptation of the 20 metres shuttle run test. *Nutricion hospitalaria*. 2014;30(6):1333-1343.
15. Sanchez-Delgado G, Cadenas-Sanchez C, Mora-Gonzalez J, Martinez-Tellez B, Chillón P, Löf M, et al. Assessment of handgrip strength in preschool children aged 3 to 5 years. *Journal of Hand Surgery (European Volume)*. 2015;40(9):966-972.
16. Oja L, Jürimäe T. Assessment of motor ability of 4-and 5-year-old children. *American Journal of Human Biology*. 1997;9(5):659-664.
17. Bland JM, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *The lancet*. 1986;327(8476):307-310.
18. Ulijaszek SJ, Kerr DA. Anthropometric measurement error and the assessment of nutritional status. *British Journal of Nutrition*. 1999;82(3):165-177.
19. Ulijaszek SJ, Lourie JA. anthropometric measurement. *Anthropometry: the Individual and the Population*. 2005;14:30.
20. Nagy E, Vicente-Rodriguez G, Manios Y, Béghin L, Iliescu C, Censi L, et al. Harmonization process and reliability assessment of anthropometric measurements in a multicenter study in adolescents. *International Journal of Obesity*. 2008;32(S5):S58.
21. Moreno LA, Joyanes M, Mesana MI, González-Gross M, Gil CM, Sarriá A, et al. Harmonization of anthropometric measurements for a multicenter nutrition survey in Spanish adolescents. *Nutrition*. 2003;19(6):481-486.
22. Stomfai S, Ahrens W, Bammann K, Kovacs E, Mårild S, Michels N, et al. Intra- and inter-observer reliability in anthropometric measurements in children. *International journal of obesity*. 2011;35(S1):S45.
23. Miguel-Etayo D, Mesana M, Cardon G, De Bourdeaudhuij I, Gózdź M, Socha P, et al. Reliability of anthropometric measurements in European preschool children: the ToyBox-study. *Obesity reviews*. 2014;15(S3):67-73.
24. Artero E, Espana-Romero V, Castro-Pinero J, Ortega F, Suni J, Castillo-Garzon M, et al. Reliability of field-based fitness tests in youth. *International journal of sports medicine*. 2011;32(3):159.