Journal of Science and Medicine in Sport 22 (2019) 1030–1037

Contents lists available at ScienceDirect

Journal of Science and Medicine in Sport

journal homepage: www.elsevier.com/locate/jsams

Review

Normative Yo-Yo Intermittent Recovery Level 1 and Yo-Yo Intermittent Endurance Level 1 test values of boys aged 9–16 years

Boris Schmitz[a,*,1], Carina Pfeifer[a,1], Kiana Kreitz[b], Matthias Borowski[b], Andreas Faldum[b], Stefan-Martin Brand[a]

[a] Institute of Sports Medicine, University Hospital Muenster, Germany
[b] Institute of Biostatistics and Clinical Research, University of Muenster, Germany

ARTICLE INFO

Article history:
Received 24 January 2019
Received in revised form 7 April 2019
Accepted 27 May 2019
Available online 31 May 2019

Keywords:
Adolescent
Child
Physical performance
Physical fitness
Physical activity

ABSTRACT

Objectives: To provide age- and sex-specific reference values of Yo-Yo tests in children and adolescents.
Design: Systematic review.
Methods: A literature search for articles on Yo-Yo Intermittent (YYI) tests was performed in MEDLINE, SPORTDiscus, Web of Science and Google Scholar. Original reports on healthy children/adolescents 6–16 years of age were eligible. For each test, age- and sex-related reference values were calculated using global means and percentiles.
Results: Ninety-two studies (7398 participants) fulfilled the eligibility criteria. The YYI tests most frequently used were the Yo-Yo Intermittent Recovery Level 1 test (YYIR1, 57.8%), Yo-Yo Intermittent Endurance Level 1 test (YYIE1, 14.7%), Yo-Yo Intermittent Recovery Level 1 Children's test (YYIR1C, 12.7%), Yo-Yo Intermittent Endurance Level 2 test (YYIE2, 8.8%) and the Yo-Yo Intermittent Recovery Level 2 test (YYIR2, 5.9%). Of these, 71.6% reported test results of boys, 17.6% reported mixed test results and 10.8% reported test results of girls. Smoothed centile curves for the YYIR1 and YYIE1 over the entire age range were generated for boys, revealing constantly increasing performance with increasing age.
Conclusions: YYI tests values differ with respect to age and sex. In boys, development of YYIR1 and YYIE1 test values (6–16 years of age) was different, suggesting better applicability of the YYIR1 test for boys >13 years of age. The results may be used to rate YYI test performance for continuous screening and to identify children with low physical fitness. Since limited data was available of females, further research on YYI tests is needed with respect to sex-specific results.

Practical implications

- We provide the first collection of normative data for Yo-Yo Intermittent (YYI) test results in children and adolescents.
- The results may be used to rate YYI test performance to monitor age-related development and identify children with low physical fitness levels.
- The provided data will also allow coaches to identify adolescents with outstanding physical fitness.

1. Introduction

Physical inactivity is not only a major risk factor for cardiovascular disease (CVD)[1] and all-cause mortality and morbidity[2–4] but has been described as a global pandemic.[5] Recent worldwide estimates indicate that 23% of men, 32% of women and over 80% of adolescents (11–17 years of age) do not meet the World Health Organisation (WHO) recommendations on physical activity.[6,7] Physical activity is an important determinant of physical fitness and physical fitness itself is one of the most important health markers and a predictor of morbidity and mortality of CVD and other diseases.[8,9] Moreover, physical fitness is a far more objective parameter considering that most epidemiologic evidence on physical activity and sedentary behaviour is based on self-reported data and thus affected by measurement/ rating error and recall bias.[10,11] In addition to the association between physical fitness and future health in adults, an association between low childhood fitness and increased risk for CVD, obesity, musculoskeletal and mental health problems in

* Corresponding author.
E-mail address: boris.schmitz@ukmuenster.de (B. Schmitz).
[1] Contributed equally to this work.

https://doi.org/10.1016/j.jsams.2019.05.016
1440-2440/

adulthood exists.[12–14] Therefore, timely diagnosis of substandard fitness levels or physical restraints in children and adolescents is important.

Physical fitness depends on different physiological systems including the respiratory, cardiovascular, musculoskeletal, nervous and endocrine-metabolic component and can thus be seen as an integrated measure of body function.[15] In addition to the association of physical fitness and mortality in general, a strong correlation between cardiorespiratory fitness (determined by maximal oxygen uptake [$V\dot{O}_{2max}$]) and mortality exists.[15] Hence, different tests have been applied to measure cardiorespiratory fitness also in children and adolescents[15–18] including the 20 m shuttle run test (20mSRT)[12,19,20] with high levels of test-retest reliability and validity.[21,22] The 20mSRT is a maximal multistage filed test,[23] which has been modified by including an active recovery period to form the Yo-Yo Intermittent Recovery (YYIR) test.[24] The main objective of the YYIR test is to determine the ability to repeatedly perform intense exercise and the potential to rapidly recover from such exercise.[25] It has thus been frequently used in intermittent sports such as soccer or basketball.[24,25] During the test, participants perform 2 × 20-m runs at increasing speed, which are interspersed by 10-s periods of active recovery (2 × 5 m).[24] The test is performed as maximal performance test (i. e. until total exhaustion is reached). Two different YYIR test levels exist, namely the Yo-Yo Intermittent Recovery Level 1 (YYIR1) test and the YYIR Level 2 (YYIR2) test. The YYIR1 test starts at a lower level[26] with 4 runs at 10 to 13 km h^{-1} (0–160 m) followed by 7 runs at 13.5–14 km h^{-1} (160–440 m) and proceeds with 0.5 km h^{-1} speed increments after every 8 running bouts.[25] The YYIR2 test starts with two initial runs of 13 and 15 km h^{-1}, respectively, followed by two runs at 16 km h^{-1}, three runs at 16.5 km h^{-1}, 4 runs at 17.0 km h^{-1} and proceeds with stepwise 0.5 km h^{-1} speed increments after every 8 running bouts.[24,27] Additional test modifications exist mainly for the determination of endurance capacity, the Yo-Yo Intermittent Endurance Level 1 (YYIE1) test (starting at 8 km h^{-1})[28] and the YYIE Level 2 (YYIE2) test (starting at 11.5 km h^{-1}).[29] During both YYIE tests, the recovery period has been shortened to 5 s (2 × 2.5 m) and the increase is reduced to 0.25 km h^{-1}.[30] Additionally, the Yo-Yo Intermittent Recovery Level 1 Children's test (YYIR1C) was developed based on the YYIR1 test, but with a shorter running distance of only 16 m and active recovery during 2 × 4.0 m.[31,32] All Yo-Yo Intermittent (YYI) tests have the advantages of a field test method. Testing does not require any special locations (indoor, outdoor) and can be carried out at low costs also in groups. However, some essential requirements for good testing practice exist. As discussed in detail recently[33] these include that participants need to understand the test settings and rules including criteria of termination. This might be of pivotal relevance predominantly in the lowest age groups and can best be addressed by practical familiarization with the test.

The validity of the YYI tests can be assessed by comparison of test results with performance during actual competitive games. Moreover, YYI test variants have been reported valid to determine cardiorespiratory fitness and thus $V\dot{O}_{2max}$ has been used for test validation. For the YYIR1 test, good correlation with high-intensity activity and total distance covered during a match has been reported for young soccer players (age ~14 years, r = 0.77 and r = 0.65, respectively).[34] Also in young soccer players, a good correlation with time spent at sprinting speed (>18.0 km h^{-1}) during the match has been described (age ~15 years, r = 0.63).[35] For the YYIR2 test, performance correlated well with the time above 85% of maximal heart rate (HR_{max}) during a soccer match in young male soccer players (age ~14 years, r = 0.71).[36] In terms of cardiorespiratory fitness, YYIE1 test performance showed good correlation with $V\dot{O}_{2max}$ determined in a laboratory setting (soccer players, age ~14 years, r = 0.63)[37] and YYIR1C test performance also correlated well with $V\dot{O}_{2max}$ in children aged 6–9 years (r^2 = 0.47).[31] Overall, YYI tests show good validity for game-specific performance and $V\dot{O}_{2max}$, even if some studies exist that did not detect a correlation of YYI test performance and $V\dot{O}_{2max}$ for unknown reasons.[28,35,38]

Currently, no normative data for the different YYI tests is available for children and adolescents and individual test results are therefore hard to evaluate in terms of cardiorespiratory fitness. Thus, the primary aim of this study was to develop age- and sex-specific normative values for YYI tests in children and adolescents and to provide a basis for the identification of children with very low cardiorespiratory fitness by use of the YYI tests. Moreover, using this framework, children and adolescents with very high cardiorespiratory fitness could be identified, which may also be of interest for future athletic success.

2. Methods

2.1. Study design and participants eligibility criteria

A systematic review (PROSPERO, CRD42018084267) was performed to identify original articles reporting on performance testing of <16 years old using one of the five YYI test variants including the YYIR1, YYIR2, YYIE1, YYIE2 and YYIR1C test. Other test modifications, such as sub-maximal test versions or the original multistage 20mSRT without recovery periods, which has occasionally been described as 'Yo-Yo test' were not considered in this analysis. Only reports on healthy humans with no disability (n ≥ 5) were eligible. Data of participants, who were >16 years of age were not included as our previous normative analysis of YYI test performance in adults already considered this population.[33] Articles had to be original reports (not a conference abstract, review or book [chapter]) and be published in English. Grey literature, such as websites, reference lists, or theses were not included. Articles were excluded if they (1) reported test results in any other format than maximal distance (m), time (min), speed (m s^{-1}) or stages, (2) did not report original performance test results [i.e. percent changes, etc.], (3) were not available as full-text (after an attempt to contact the corresponding author), (4) presented performance data only in figures, or if (5) the test type was not clearly defined. Furthermore, articles were excluded if test performance mean or standard deviation (SD)/standard error of the mean (SEM) or the number of tested subjects were missing or not clearly reported in the full-text. These eligibility criteria were chosen also in accordance with the quality assessment (described below).

2.2. Search strategy and data sources

A systematic literature search was conducted (CP) using PubMed (MEDLINE database), Web of Science, SPORTDiscus with Full Text and Google Scholar for records published up until August 2018. The following key words were used: 'Yo-Yo intermittent test' or 'Yo-Yo intermittent' or 'Yo-Yo intermittent recovery test' or 'Yo-Yo intermittent recovery' or 'Yo-Yo intermittent endurance test' or 'Yo-Yo intermittent endurance' or 'Yo-Yo Intermittent Recovery Level 1 Children's' or 'YYIR1' or 'YYIR2' or 'YYIR' or 'YYIE1' or 'YYIE2' or 'YYIE' or 'YYIR1C' or 'Yo-Yo IR' or 'Yo-Yo IE' or 'Yo-Yo IR1' or 'Yo-Yo IE1' or 'Yo-Yo IR2' or 'Yo-Yo IE2' or 'Yo-Yo test' or 'YoYo test'. Manual searches based on references from identified articles were also performed. Two reviewers performed full-text screening on potential relevant reports. The individual steps of report identification, screening and processing are documented in the PRISMA flow-chart (Supplementary Fig. 1).[39] Search results and fulfilment of eligibility criteria were discussed if unclear (BS and CP) until consensus was achieved and a third person was consulted upon disagreement to determine inclusion.

2.3. Study selection and data extraction

Data were extracted using Excel (Microsoft Office 2016, Microsoft Corporation, Redmond, USA) into a standard extraction table by two reviewers (BS and CP) including information on first author, year of publication, description of participants (total number, anthropometric data, activity level), performance data and test type. In cases where articles reported on test results over different age groups, separate data on participants <16 years of age were extracted if possible. In case of redundant data reported in separate publications, only data from the earliest report were extracted. In studies where both girls and boys were tested in combination (mixed data) and data could not be extracted separately, test results were not included in the quantitative analysis (Supplementary Fig. 1). In cases where intra- and inter-seasonal performance changes were reported, the best test data were extracted. In studies where performance data were reported involving any type of experimental condition with hypothesized effect on test performance, only pre-intervention data were extracted if available or data of (untreated) control groups were used. Participants were classified by sex as well as active or inactive in accordance with author's description in the original article.

2.4. Quality assessment

The methodological quality of the studies was assessed using the critical appraisal tool by Brink and Louw.[40] The tool consists of 13 items to assess the quality of a study. Individual items can be scored as 'yes', 'no' or 'not applicable'. For our analysis, we determined (in accordance with the above mentioned inclusion criteria) that the following items (8 out of 13) had to be scored 'yes'. A detailed description of the subject sample was available, the reference standard was explained, the competence or qualification of rater(s) performing the test was clarified, the stability of the variable being measured was taken into account, the execution of the test was described in (sufficient) detail to permit replication, the execution of the reference standard was described in (sufficient) detail, the statistical methods were appropriate and withdrawals from the study were explained. Studies were rated by two reviewers (BS and CP). Disagreements were resolved by discussion if necessary. Researchers were not blinded to study authors, results or publication journal.

2.5. Data treatment and statistical data analyses

In most cases age was reported as age at last birthday given as mean and SD. Subjects were then assigned to respective age categories. The combined age category of 6–9 years included subjects older than 6 years of age who had not yet completed their 9th year of life. Each following age category then represents the next span of one year, respectively. In rare cases, age was already given as a span of a year (eg. 10–11 years) and was extracted without modifications. For overall analysis of tested participants by test type, a weighted mean age was generated and a corresponding SD of 0.25 was assumed if age was reported over a span of a year. Data for YYI test performance was analysed as maximal distance (m). Maximal test duration (min), speed ($m\,s^{-1}$) or stages were recalculated for comparison. SD was calculated from SEM using the equation $SD = \sqrt{n} \cdot SEM$ where n is the number of subjects. For each category (i. e. age, sex and test type) a global mean and global SD was calculated based on the reported means and SDs of the individual studies as weighted mean of the individual reported means, with weights built by the number of subjects per study as described.[33] Each global SD was calculated using the formula described in Supplementary File 1. Finally, global mean and global SD were translated into normal quantiles for each category. Individual forest plots were created for each category providing the individual mean and 95% confidence interval (CI) of each study as well as the calculated global mean and global SD. The I^2 statistic was calculated for each category to quantify heterogeneity within the respective individual study results.[41] All forest plots including I^2 statistics are available in the online repository.

3. Results

3.1. Study selection and characteristics

The procedure of study identification, selection and final inclusion is shown in Supplementary Fig. 1. During the screening stage, 591 records did not meet the inclusion criteria (see methods) and were removed. Of the remaining 537 reports, 445 studies were excluded (details in Supplementary Fig. 1) and 92 studies with a total of 7398 participants were included in the qualitative synthesis. Each individual included study (or sub-study) with information on author, mean age, sex, activity status and test results is presented in Supplementary Table 1, sorted by test and test results. Seventy-three studies were included in the quantitative synthesis since 19 studies reported mixed results of boys and girls.

3.2. Results by sex, test type and age

3.2.1. Results of girls

Of girls, only limited data (10.8% of all included studies) was available with respect to the different age categories and test types. For available data, test global means, SDs and quantiles for the YYIR1, YYIE1/2 and YYIR1C test are shown in Fig. 1, computed from 17 studies/subgroups with a total of 889 female participants.[42–49]

3.2.2. Results of boys

The majority of included studies reported test results of boys (71.6% of the studies) and results by test type and age were as follows. For the computation of YYIR1 test global means, SDs and quantiles (Fig. 2A), 92 studies or subgroups with a total of 2832 participants were available.[34,35,42,50–60–91] Computation of YYIR2 test global means, SDs and quantiles (Fig. 2B) involved 8 studies/subgroups reporting on 147 participants.[36,50,55,92–94] Twenty-four studies/subgroups reporting on 848 participants were available for generation of YYIE1 test global means, SDs and quantiles (Fig. 2C).[37,95–104] Computation of YYIE2 test global means, SDs and quantiles (Fig. 2D) included 13 studies/subgroups and 498 participants.[35,98,105–109] For the generation of YYIR1C test global means, SDs and quantiles (Fig. 2), 8 studies/subgroups and 665 participants were available.[47–49,98,110]

Since the YYIR1 and the YYIE1 test were the most frequently used test types, smoothed centile curves over the complete age range could be generated for these two tests and are presented in Fig. 3A/B. Boys' mean performance regarding YYIE1 and YYIR1 test improved steadily from age group 6–9 years to 15–16 years (Fig. 3C). For both test types, the average increase in mean test results seemed to decline when boys became older than 12–13 years. However, age-related changes, i.e. the boys' performance improvements were not identical between the YYIR1 and YYIE1 (Fig. 3).

3.3. Risk of bias

As detailed above, the quality of the included studies was assessed using the critical appraisal tool by Brink et al.[40] The tool does not incorporate a quality score and the impact of each item is considered individually. Since the current analysis includes test results obtained from single observations (i. e. does not include data from two consecutive tests), the following 5 items were found not

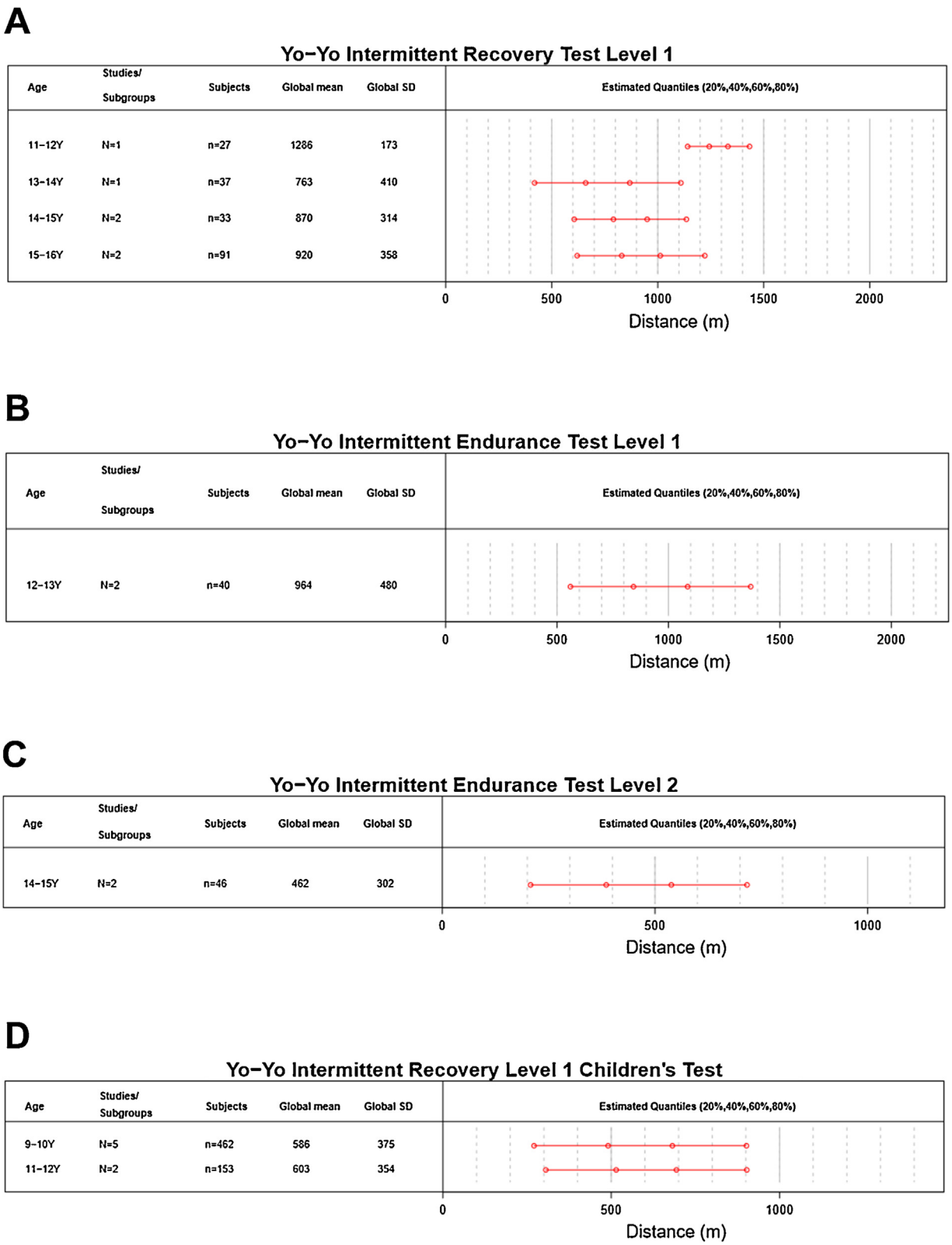

Fig. 1. Yo-Yo Intermittent test results of girls by test type and age. Qualitative ratings: values above the 80th centile, 'very high'; between the 60th and 80th centiles, 'high'; between the 40th and 60th centiles, 'moderate'; between the 20th and 40th centiles, 'low'; values below the 20th centile, 'very low'.

applicable: blinding of raters to their own prior findings, evaluation of time period between reference standard and index test, testing of interrater reliability and blinding of raters to the test results, variation of the order of examination, and independence of reference standard and index test. All other items were scored with 'yes' by definition of the inclusion criteria. Other criteria for the assessment of bias, such us blinding of test participants or raters to test results were unfeasible for this type of test.

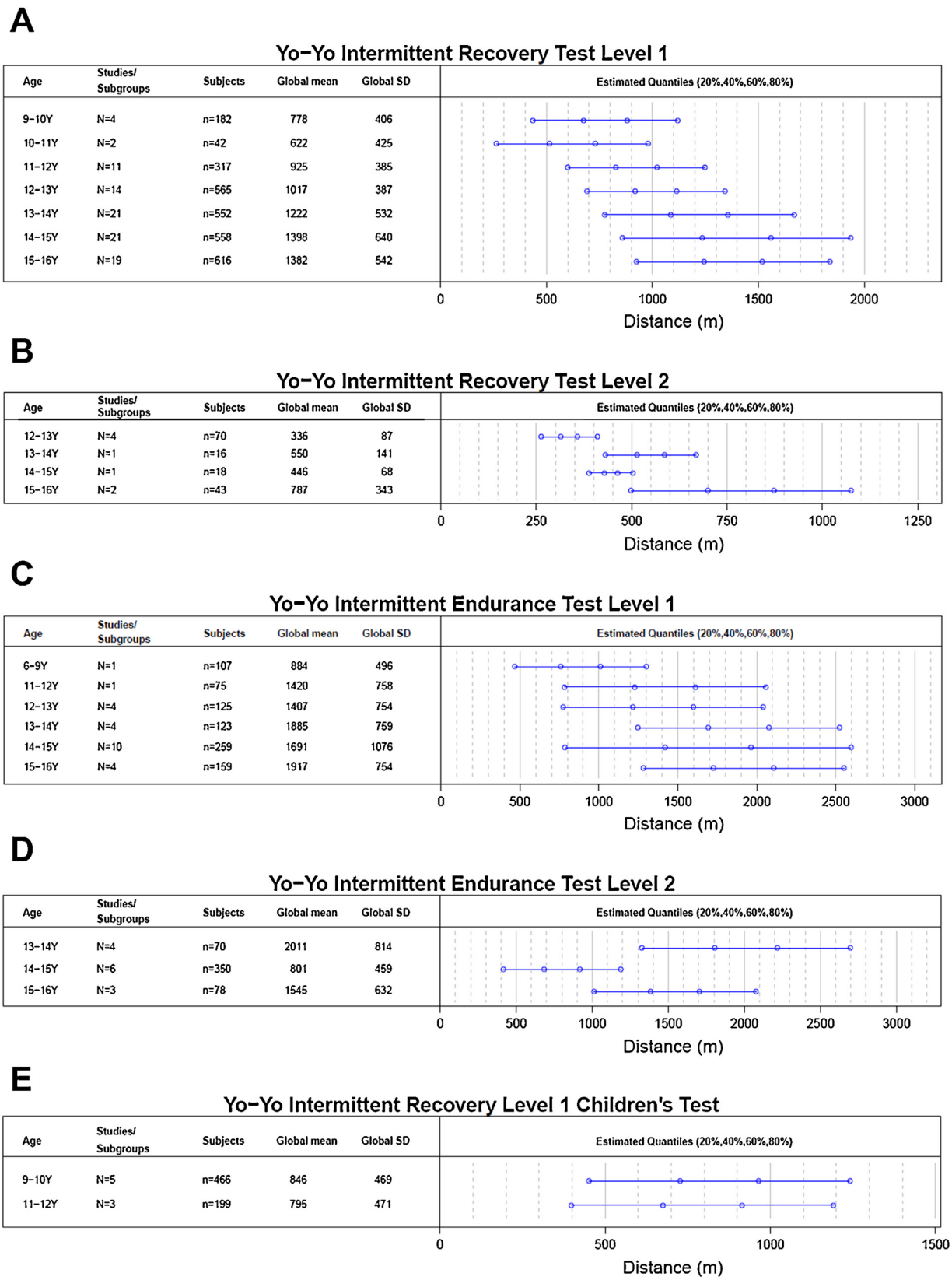

Fig. 2. Yo-Yo Intermittent test results of boys by test type and age. Qualitative ratings: values above the 80th centile, 'very high'; between the 60th and 80th centiles, 'high'; between the 40th and 60th centiles, 'moderate'; between the 20th and 40th centiles, 'low'; values below the 20th centile, 'very low'.

4. Discussion

This study systematically analysed test data of the most common YYI tests from 7398 children and adolescents aged 6–16 years to generate normative age- and sex-specific reference values. Overall, the YYIR1 was the most frequently performed test variant in children and adolescents up to 16 years of age with 57.84%, followed by the YYIE1, which was performed in 14.71% of the

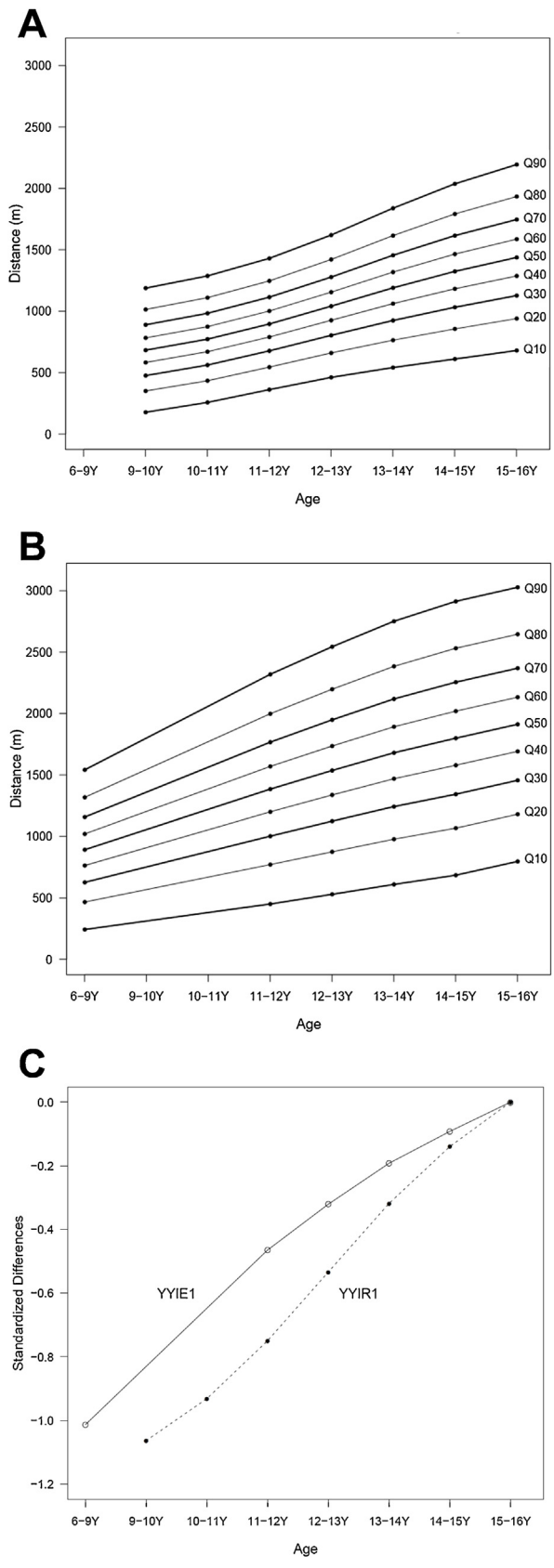

Fig. 3. Quantile curves for (A) Yo-Yo Intermittent Recovery Level 1 test and (B) Yo-Yo Intermittent Endurance Level 1 test results by age of boys. Qualitative ratings: values above the 80th centile, 'very high'; between the 60th and 80th centiles, 'high'; between the 40th and 60th centiles, 'moderate'; between the 20th and 40th centiles, 'low'; values below the 20th centile, 'very low'. (C) Standardised differences in mean Yo-Yo Intermittent Recovery Level 1 (YYIR1) test and Yo-Yo Intermittent Endurance Level 1 (YYIE1) test performance of boys by age.

studies. The YYIR1C was performed in 12.75% of the studies and the YYIE2 and YYIR2 were performed in 8.82% and 5.88% of the studies, respectively. This cross-sectional approach suggested that YYI tests performance increased with age and boys tended to perform better than girls, even though data on girls was scarce.

For the two most common tests, the YYIR1 and the YYIE1 test, a centile framework for boys aged 9–16 years was generated which allows qualitative interpretation of test results and the classification of age-specific physical fitness as follows. Values above the 80th centile are rated 'very high', values between the 60th and 80th centiles are rated 'high', values between the 40th and 60th centiles are rated 'moderate', values between the 20th and 40th centiles are rated 'low' and values below the 20th centile are rated 'very low'.[12] Moreover, normal maturation-related longitudinal changes can be monitored using these references and the effects of interventions (exercise training, general increase of regular physical activity) may be controlled. Of note, the majority of included studies described their participants as "physically active" and reported involvement in ball team sports most frequently represented by soccer, rugby/football and basketball. In detail, of all tested participants who performed the YYIR1, 98.76% were described as physically active (male: 98.51%, female: 100%). Among participants performing the YYIR2, 100% were physically active. For the YYIE1, 78.96% (male: 96.36%, female: 50%), for the YYIE2, 92.44% (male: 95.98%, female: 45.65%) and for the YYIR1C, 42.12% (male: 42.9%, female: 40.15%) were physically active. Therefore, individuals performing at values below the 20th centile of the YYIR1 or the YYIE1 may not necessarily be at a critical fitness level. However, physical activity in children and adolescents should be conceived as the 'norm' rather than physical inactivity.[6,7] This is also of relevance with respect to the reported association between physical fitness in childhood and future health in adults including the observation that low childhood fitness may lead to an increased risk for CVD, obesity, musculoskeletal and mental health problems in adulthood.[12–14] Of note, specific associations between 20mSRT performance and health indicators have also been described.[111] Even though the general description of "physical activity" of the included studies was not investigated in detail in the current analysis, an overall comparison of included study participants described as "inactive" to participants described as "active" suggested that regular physical activity may lead to at least 20% increased test results over all age categories.

With respect to the yearly increase in test performance, it is of interest to compare the YYIR1 test data with normative data on 20mSRT performance recently presented by Tomkinson et al.[12] The group also provided centile curves of boys (age range 9–17 years) showing comparable characteristics to the YYIR1 test with the largest rate of increase at the age of 12 years. Since the YYIR1 was developed based on the 20mSRT, both tests are described to mainly estimate the anaerobic performance capacity of the participant. By contrast, the YYIE1 has been introduced to address aerobic endurance capacity and thus cardiorespiratory fitness. Interestingly, the development of age-related YYIE1 performance over all age groups showed different characteristics compared to the YYIR1 and 20mSRT in that a strong increase per age category in YYIE1 performance was observed already at younger age. Of note, the included studies for both tests, the YYIR1 and the YYIE1 test, involved a high number of "physically active" participants (both >95% for boys). Based on these observations and the different physiological requirements of the two tests, age-specific YYIR1 and YYIE1 test results might indicate different physiological stages during maturation. To this respect, it has been suggested that, irrespective of exercise training, aerobic capacity increases with age in childhood. In girls, a plateau for aerobic capacity and thus cardiorespiratory fitness is reached at about 15 years of age, while in boys

aerobic capacity continues to increase until the age of about 18 years.[112] Of note, the ratio of aerobic capacity to body mass in children is comparable to that of adults, while anaerobic capacity is significantly lower than that of adults both in comparison to aerobic capacity and in comparison to body mass.[113] So while children are very well adapted to aerobic exercise, anaerobic exercise quickly leads to fatigue.[112] This is also documented by the oxygen (O_2) uptake kinetics during exercise in children, which is much steeper early under load compared to adults.[114] Thus, the YYIR1C test has been developed with a shorter running distance of 16 m and shorter recovery distance of 2×4 m but similar speed increments as the YYIR1 test.[31] It is therefore suggested especially for the youngest and/ or less fit children.[31,46] However, our analysis did not identify a sufficient number of studies reporting on the YYIR1C test to generate centile curves over the entire age range of 9–16 years for comparison to the development of test results with age to the YYIR1 and YYIE1 test.

Since it has been suggested that response to different training strategies may be monitored by YYI tests[27] and YYI tests may be used to monitor age-specific development of physical fitness, effectors of test reproducibility are important. For the YYIR1 test, reported reproducibility was higher in competitive U17 soccer players (coefficient of variation [CV] = 7.9%) compared to U13 soccer players (CV = 17.3).[63] High reproducibility was also observed for different groups of elite youth soccer players performing the YYIE2 test (CV = 3.9%)[115] or the YYIE1 test (CV = 5.7%).[37] For the YYIR1C test, mainly performed in younger children, reported reproducibility was better in children aged 8–9 years (CV = 13.0%) compared to children between 6–7 years of age (CV = 26.0%).[31] Combined, these observations suggest that test reproducibility increases with age but may potentially be higher in more experienced participants already at younger age. This might be explained by the fact that, per definition, the YYI tests are maximal performance tests. It is thus an essential requirement that children perform YYI tests at a maximal effort, which may be evaluated through ratings of perceived exertion or the use of heart rate monitors. However, if the child is not used to all-out exercise efforts, perceived exertion might be problematic when performing the test for the first time. It could thus be necessary that full test familiarization is performed before the actual test (with sufficient offset, $\tilde{1}$ week). In general, it is mandatory that the test is supervised and documented by experienced personnel, at best by at least two raters. This will help to prevent documentation and procedure errors and will also limit the effect of observer errors including failure of early or late test termination. Moreover, YYI tests with shorter running distances, such as the YYIR1C test or lower initial running speeds, such as the YYIE1 should be used for younger or less fit children. Even if individual tests are performed considering these requirements, test performance values at the lowest level (at or below the 10% level) should be interpreted with great care. While it is possible that comparably low test performance reveals low physical fitness or might even indicate a certain delay in development or unknown health problems, individual circumstances leading to the test result should be identified and this should include detailed questioning of the participant. Any adverse events should be documented and physical examinations should be performed if indicated. When test results remain at a low level also after retest within a certain time period and/or show no improvement with age, it might be adequate to involve the child into regular or individually optimized physical activity programs. This could lead to increased test results and thus improved physical fitness to prevent future health risks. The here presented framework might also be used by coaches to identify children or adolescents with outstanding physical fitness (at or above the 90% level) and test results may be compared to the YYI test normative data of adults at different athletic levels presented in our previous study.[33]

5. Limitations

Some limitations for the presented analysis may exist. First, we have grouped tested subjects into age categories and inter-annual differences have thus not been taken into account, which leads to reduced accuracy especially in the youngest age group of 6–9 years. Available data on YYI test results of girls was scarce and more data will be needed to generate test-specific normative data of girls over all age groups. The generated centile curves for the YYIE1 did not include original data on age categories 9–11 years and may thus be affected by the conformation process. The necessary pooling and transformation of data may also have affected the presented results to some extent. Second, our analysis was based on the reported sex and the effect of gender has not been addressed. Third, besides methodological quality assessment of the included studies, selection bias within individual studies may have occurred in terms of test termination (i.e. the test is not stopped at the earliest time point violating test requirements) or data partitioning and reporting of data subsets (i.e. reporting on best test results). Reporting and publication bias may have affected the present analysis since some data/studies may have remained unreported or were not published because of unexpected/contradictory, negative or not significant test results. In addition, a number of studies reported mixed test results of boys and girls and were not included in our quantitative analysis. Furthermore, the record search was limited to studies published in English and inclusion of data reported in other languages may have altered the results of age groups with smaller sample sizes.

6. Conclusions

Of the available YYI test variants, the YYIR/2, the YYIE1/2 and the YYIR1C have been used in children and adolescents. Of these, the YYIR1 and the YYIE1 have been used most frequently and over the entire age range from 9 to 16 years in boys, while data on girls is underrepresented. Our analysis provides evidence that YYIR1 and YYIE1 test reference values differ depending on age and sex. The presented results may be used by practitioners and researchers to rate YYI test performance and monitor age-related development of physical fitness. With regard to varying reproducibility values, caution is warranted when using YYI tests to rate physical fitness especially in younger children.

Funding

None.

Acknowledgments

We thank Franziska Breulmann for reviewing the extracted data after entry into the databases.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jsams.2019.05.016.

References [2]

1. Hamer M, Chida Y. Walking and primary prevention: a meta-analysis of prospective cohort studies. *Br J Sports Med* 2008; 42(4):238–243.

[2] Further references can be found in Multimedia Component 3.

2. Biswas A, Oh PI, Faulkner GE et al. Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: a systematic review and meta-analysis. *Ann Intern Med* 2015; 162(2):123–132.
3. O'Donovan G, Hamer M, Stamatakis E. Relationships between exercise, smoking habit and mortality in more than 100,000 adults. *Int J Cancer* 2017; 140(8):1819–1827.
4. Lee IM, Shiroma EJ, Lobelo F et al. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet* 2012; 380(9838):219–229.
5. Kohl 3rd HW, Craig CL, Lambert EV et al. The pandemic of physical inactivity: global action for public health. *Lancet* 2012; 380(9838):294–305.
6. *Global Recommendations on Physical Activity for Health*, Geneva, World Health Organization, 2010.
7. *Global Status Report on Noncommunicable Diseases*, Geneva, World Health Organization, 2014.
8. Blair SN, Kohl 3rd HW, Paffenbarger Jr RS et al. Physical fitness and all-cause mortality. A prospective study of healthy men and women. *JAMA* 1989; 262(17):2395–2401.
9. Kriemler S, Manser-Wenger S, Zahner L et al. Reduced cardiorespiratory fitness, low physical activity and an urban environment are independently associated with increased cardiovascular risk in children. *Diabetologia* 2008; 51(8):1408–1415.
10. Jefferis BJ, Parsons TJ, Sartini C et al. Objectively measured physical activity, sedentary behaviour and all-cause mortality in older men: does volume of activity matter more than pattern of accumulation? *Br J Sports Med* 2018. http://dx.doi.org/10.1136/bjsports-2017-098733 [Epub ahead of print].
11. Arem H, Moore SC, Patel A et al. Leisure time physical activity and mortality: a detailed pooled analysis of the dose-response relationship. *JAMA Intern Med* 2015; 175(6):959–967.
12. Tomkinson GR, Lang JJ, Tremblay MS et al. International normative 20 m shuttle run values from 1 142 026 children and youth representing 50 countries. *Br J Sports Med* 2017; 51(21):1545–1554.
13. NCD Risk Factor Collaboration (NCD-RisC). Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19.2 million participants. *Lancet* 2016; 387(10026):1377–1396.
14. Ruiz JR, Sui X, Lobelo F et al. Association between muscular strength and mortality in men: prospective cohort study. *BMJ* 2008; 337:a439.
15. Ortega FB, Ruiz JR, Castillo MJ et al. Physical fitness in childhood and adolescence: a powerful marker of health. *Int J Obes (Lond)* 2008; 32(1):1–11.
16. Ruiz JR, Castro-Pinero J, Artero EG et al. Predictive validity of health-related fitness in youth: a systematic review. *Br J Sports Med* 2009; 43(12):909–923.
17. Catley MJ, Tomkinson GR. Normative health-related fitness values for children: analysis of 85347 test results on 9-17-year-old Australians since 1985. *Br J Sports Med* 2013; 47(2):98–108.
18. Artero EG, Lee DC, Ruiz JR et al. A prospective study of muscular strength and all-cause mortality in men with hypertension. *J Am Coll Cardiol* 2011; 57(18):1831–1837.
19. Leger LA, Mercier D, Gadoury C et al. The multistage 20 metre shuttle run test for aerobic fitness. *J Sports Sci* 1988; 6(2):93–101.
20. Olds T, Tomkinson G, Leger L et al. Worldwide variation in the performance of children and adolescents: an analysis of 109 studies of the 20-m shuttle run test in 37 countries. *J Sports Sci* 2006; 24(10):1025–1038.
21. Domone S, Mann S, Sandercock G et al. A method by which to assess the scalability of field-based fitness tests of cardiorespiratory fitness among schoolchildren. *Sports Med* 2016; 46(12):1819–1831.
22. Castro-Pinero J, Artero EG, Espana-Romero V et al. Criterion-related validity of field-based fitness tests in youth: a systematic review. *Br J Sports Med* 2010; 44(13):934–943.
23. Leger LA, Lambert J. A maximal multistage 20-m shuttle run test to predict VO_2 max. *Eur J Appl Physiol Occup Physiol* 1982; 49(1):1–12.
24. Bangsbo J, Iaia FM, Krustrup P. The Yo-Yo intermittent recovery test: a useful tool for evaluation of physical performance in intermittent sports. *Sports Med* 2008; 38(1):37–51.
25. Krustrup P, Mohr M, Amstrup T et al. The yo-yo intermittent recovery test: physiological response, reliability, and validity. *Med Sci Sports Exerc* 2003; 35(4):697–705.
26. Castagna C, Impellizzeri FM, Belardinelli R et al. Cardiorespiratory responses to Yo-yo Intermittent Endurance Test in nonelite youth soccer players. *J Strength Cond Res* 2006; 20(2):326–330.
27. Krustrup P, Mohr M, Nybo L et al. The Yo-Yo IR2 test: physiological response, reliability, and application to elite soccer. *Med Sci Sports Exerc* 2006; 38(9):1666–1673.
28. Castagna C, Impellizzeri FM, Chamari K et al. Aerobic fitness and yo-yo continuous and intermittent tests performances in soccer players: a correlation study. *J Strength Cond Res* 2006; 20(2):320–325.
29. Bradley PS, Bendiksen M, Dellal A et al. The application of the Yo-Yo intermittent endurance level 2 test to elite female soccer populations. *Scand J Med Sci Sports* 2014; 24(1):43–54.
30. Reilly T, Bangsbo J, Franks A. Anthropometric and physiological predispositions for elite soccer. *J Sports Sci* 2000; 18(9):669–683.
31. Ahler T, Bendiksen M, Krustrup P et al. Aerobic fitness testing in 6- to 9-year-old children: reliability and validity of a modified Yo-Yo IR1 test and the Andersen test. *Eur J Appl Physiol* 2012; 112(3):871–876.
32. Bendiksen M, Ahler T, Clausen H et al. The use of Yo-Yo intermittent recovery level 1 and Andersen testing for fitness and maximal heart rate assessments of 6- to 10-year-old school children. *J Strength Cond Res* 2013; 27(6):1583–1590.
33. Schmitz B, Pfeifer C, Kreitz K et al. The Yo-Yo intermittent tests: a systematic review and structured compendium of test results. *Front Physiol* 2018; 9:870.
34. Castagna C, Impellizzeri F, Cecchini E et al. Effects of intermittent-endurance fitness on match performance in young male soccer players. *J Strength Cond Res* 2009; 23(7):1954–1959.
35. Rebelo A, Brito J, Seabra A et al. Physical match performance of youth football players in relation to physical capacity. *Eur J Sport Sci* 2014; 14(Suppl 1):S148–S156.
36. Da Silva CD, Natali AJ, de Lima J Roberto Perrout et al. Yo-Yo IR2 test and margaria test: validity, reliability and maximum heart rate in young soccer players. *Revista Brasileira de Medicina do Esporte* 2011; 17(5):344–349.
37. Wong P, Chaouachi A, Castagna C et al. Validity of the Yo-Yo intermittent endurance test in young soccer players. *Eur J Sport Sci* 2011; 11(5):309–315.
38. Martinez-Lagunas V, Hartmann U. Validity of the Yo-Yo Intermittent Recovery Test Level 1 for direct measurement or indirect estimation of maximal oxygen uptake in female soccer players. *Int J Sports Physiol Perform* 2014; 9(5):825–831.
39. Moher D, Liberati A, Tetzlaff J et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009; 339:b2535.
40. Brink Y, Louw QA. Clinical instruments: reliability and validity critical appraisal. *J Eval Clin Pract* 2012; 18(6):1126–1132.
41. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002; 21(11):1539–1558.
42. Myburgh GK, Cumming SP, Silva MC et al. Maturity-associated variation in functional characteristics of elite youth tennis players. *Pediatr Exerc Sci* 2016; 28(4):542–552.
43. Moss SL, McWhannell N, Michalsik LB et al. Anthropometric and physical performance characteristics of top-elite, elite and non-elite youth female team handball players. *J Sports Sci* 2015; 33(17):1780–1789.
44. Wright MD, Hurst C, Taylor JM. Contrasting effects of a mixed-methods high-intensity interval training intervention in girl football players. *J Sports Sci* 2016; 34(19):1808–1815.
45. Attene G, Pizzolato F, Calcagno G et al. Sprint vs. intermittent training in young female basketball players. *J Sports Med Phys Fitness* 2014; 54(2):154–161.
46. Povoas SC, Castagna C, da Costa Soares JM et al. Reliability and construct validity of Yo-Yo tests in untrained and soccer-trained schoolgirls aged 9-16. *Pediatr Exerc Sci* 2016; 28(2):321–330.
47. Skoradal MB, Purkhus E, Steinholm H et al. "FIFA 11 for Health" for Europe in the Faroe Islands: effects on health markers and physical fitness in 10- to 12-year-old schoolchildren. *Scand J Med Sci Sports* 2018; 28(Suppl 1):8–17.
48. Larsen MN, Nielsen CM, Orntoft C et al. Fitness effects of 10-month frequent low-volume ball game training or interval running for 8–10-year-old school children. *Biomed Res Int* 2017; 2017, 2719752.
49. Geertsen SS, Thomas R, Larsen MN et al. Motor skills and exercise capacity are associated with objective measures of cognitive functions and academic performance in preadolescent children. *PLoS One* 2016; 11(8):e0161960.
50. Karakoc B, Akalan C, Alemdaroglu U et al. The relationship between the yo-yo tests, anaerobic performance and aerobic performance in young soccer players. *J Hum Kinet* 2012; 35:81–88.
51. Benounis O, Benabderrahman A, Chamari K et al. Association of short-passing ability with athletic performances in youth soccer players. *Asian J Sports Med* 2013; 4(1):41–48.
52. Spencer M, Pyne D, Santisteban J et al. Fitness determinants of repeated-sprint ability in highly trained youth football players. *Int J Sports Physiol Perform* 2011; 6(4):497–508.
53. Carvalho HM, Bidaurrazaga-Letona I, Lekue JA et al. Physical growth and changes in intermittent endurance run performance in young male Basque soccer players. *Res Sports Med* 2014; 22(4):408–424.
54. Deprez D, Vaeyens R, Coutts AJ et al. Relative age effect and Yo-Yo IR1 in youth soccer. *Int J Sports Med* 2012; 33(12):987–993.
55. Chuman K, Hoshikawa Y, Iida T et al. Relationships between Yo-Yo intermittent recovery tests and development of aerobic and anaerobic fitness in U-13 and U-17 soccer players. *Int J Sport Health Sci* 2011; 9:91–97.
56. Abian-Vicen J, Puente C, Salinero JJ et al. A caffeinated energy drink improves jump performance in adolescent basketball players. *Amino Acids* 2014; 46(5):1333–1341.
57. Deprez D, Fransen J, Lenoir M et al. The Yo-Yo intermittent recovery test level 1 is reliable in young high-level soccer players. *Biol Sport* 2015; 32(1):65–70.
58. Souhail H, Castagna C, Mohamed HY et al. Direct validity of the yo-yo intermittent recovery test in young team handball players. *J Strength Cond Res* 2010; 24(2):465–470.
59. Casamichana D, Castellano J. Time-motion, heart rate, perceptual and motor behaviour demands in small-sides soccer games: effects of pitch size. *J Sports Sci* 2010; 28(14):1615–1623.
60. Doncaster G, Iga J, Unnithan V. Assessing differences in cardiorespiratory fitness with respect to maturity status in highly trained youth soccer players. *Pediatr Exerc Sci* 2018; 30(2):216–228.