The International Fitness Scale (IFIS): usefulness of self-reported fitness in youth

Francisco B Ortega,1,2* Jonatan R Ruiz,1,3 Vanesa España-Romero,1,4 Germán Vicente-Rodríguez,5,6 David Martínez-Gómez,7 Yannis Manios,8 Laurent Béghin,9 Dénes Molnar,10 Kurt Widhalm,11 Luis A Moreno,5,12 Michael Sjöström,1 Manuel J Castillo2 and on behalf of the HELENA study group†

1Department of Biosciences and Nutrition, Unit for Preventive Nutrition, Karolinska Institutet, Huddinge, Sweden, 2Department of Physiology, School of Medicine, University of Granada, Granada, Spain, 3Department of Physical Education and Sport, School of Physical Activity and Sport Sciences, University of Granada, Granada, Spain, 4Department of Exercise Science, University of South Carolina, Columbia, SC, USA, 5Growth, Exercise, Nutrition and Development (GENUD) Research Group, Universidad de Zaragoza, Spain, 6Faculty of Health and Sport Science (FCSD), Department of Physiotherapy and Nursing, Universidad de Zaragoza, Huesca, Spain, 7Department of Metabolism and Nutrition, Immunonutrition Research Group, Instituto del Frio, Institute of Food Science, Technology and Nutrition (ICTAN), Spanish National Research Council (CSIC), Madrid, Spain, 8Department of Nutrition and Dietetics, Harokopio University, Athens, Greece, 9Unité U995 Institut National de la Santé et de La Recherche Médicale and Institut de Médecine Prédicte et Thérapeutique, Faculté de Médecine, Université Lille Nord de France, France, Centre d’Investigation Clinique CIC-9301, Centre Hospitalier and Universitaire de Lille, France, 10Department of Paediatrics, Medical Faculty, University of Pécs. József A, Pécs, Hungary, 11Division of Clinical Nutrition and Prevention, Department of Paediatrics, Medical University of Vienna, Vienna, Austria and 12Department of Physiotherapy and Nursing, School of Health Sciences (EUCS), Universidad de Zaragoza, Zaragoza, Spain

*Corresponding author. Department of Biosciences and Nutrition at NOVUM, Karolinska Institutet, 14157 Huddinge, Sweden.
E-mail: ortegaf@ugr.es

†The members of the HELENA Study Group are listed in the Appendix 1 available as Supplementary Data at IJE online.

Accepted 14 February 2011

Background We examined: (i) the usefulness of the International Fitness Scale (IFIS) to correctly rank adolescents into physical fitness levels; (ii) the capacity of the IFIS for predicting cardiovascular disease (CVD) risk; and (iii) the reliability of the IFIS in adolescents.

Methods The study comprised 3059 adolescents (aged 12.5–17.5 years) from nine European countries (HELENA study). Blood samples were collected in one-third of the participants (randomly selected, n = 981). Test–retest reliability of the IFIS was studied in a separate sample of 277 adolescents. Physical fitness—cardiorespiratory fitness (CRF), muscular fitness (MF), speed–agility (SP–AG), flexibility and overall fitness—was self-reported using 5-point Likert-scale questions (1 = very poor, 5 = very good) and measured using standard field-based tests. The CVD risk factors measured included total/central adiposity indices and mean arterial pressure, total and high density lipoprotein cholesterol, triglycerides, insulin resistance (HOMA) and C-reactive protein.

Results Analysis of covariance showed that adolescents reporting better fitness had higher measured fitness levels for all the variables studied (all P < 0.001), regardless of gender, age and weight status. Adolescents reporting very good overall fitness, CRF and SP–AG had a healthier cardiovascular profile in eight out of nine CVD risk factors studied. Perfect agreement (same test–retest answer) was observed in 65% of the adolescents and perfect-acceptable agreement (same answer or ±1) in 97% of the adolescents.
Conclusions (i) The IFIS is able to correctly rank adolescents according to their measured physical fitness levels; (ii) adolescents reporting a good/very good overall fitness, CRF or SP–AG have a more favorable cardiovascular profile; and (iii) The IFIS is reliable in adolescents.

Keywords Adolescent, physical fitness, questionnaires, risk factors, self-report

Introduction

Physical fitness is a powerful predictor of detrimental health outcomes that include all-cause mortality, cardiovascular disease (CVD) events and cancer events.1–3 Interestingly, there is evidence that physical fitness is a marker of health in youth4 and that levels of physical fitness at these ages are associated with health in adulthood.5 Therefore, the assessment of physical fitness in youth throughout the world is of public health interest, and timely research on different methods for physical fitness assessment is needed.

Physical fitness can be objectively and accurately measured through laboratory tests. However, their use is limited in population-based studies due to the high cost, the time required to complete laboratory tests and the impracticality of measuring a large number of youths. Field-based tests provide a reasonable alternative to laboratory tests. Another alternative method that could be used in large epidemiological surveys is self-reported physical fitness. An accurate self-report measure of fitness specifically for youths could be used to identify low-fit (and perhaps low-active) adolescents for targeted physical activity and obesity prevention interventions.

A questionnaire translated into different languages, to assess overall physical fitness and the main components—cardiorespiratory fitness (CRF), muscular fitness (MF), speed–agility (SP–AG) and flexibility—in youth is lacking. Such a fitness scale, which is valid and reliable and shown to relate to established CVD risk factors, would be a useful epidemiological measurement tool.4,5

In the present study, we investigate the International Fitness Scale (IFIS). This novel scale is adapted for youths, measures the main physical fitness components and can be completed within 5 min. We specifically aimed to examine: (i) the ability of the IFIS to correctly rank adolescents into physical fitness levels; (ii) if the IFIS is related to CVD risk in adolescents; and (iii) the test–retest reliability of the IFIS in adolescents.

Methods

Study design and sample

The Healthy Lifestyle in Europe by Nutrition in Adolescence cross-sectional study (Aims 1 and 2)
The Healthy Lifestyle in Europe by Nutrition in Adolescence cross-sectional study (HELENA-CSS) obtained standardized data from a random sample of European adolescents on a broad battery of nutrition and health-related parameters.6,7 Data collection took place during 2006 and 2007 in 10 European cities: Athens (inland city) and Heraklion (Mediterranean island city) in Greece, Dortmund in Germany, Ghent in Belgium, Lille in France, Pecs in Hungary, Rome in Italy, Stockholm in Sweden, Vienna in Austria and Zaragoza in Spain. The adolescents were sampled from classes in the school setting. Full details on all aspects of the HELENA-CSS have been published elsewhere.6,8

All adolescents meeting the general HELENA inclusion criteria (not participating simultaneously in another clinical trial, free of any acute infection or disease lasting <1 week before the inclusion)6 and with valid data for age, gender and body mass index (BMI), were considered the final HELENA-CSS sample: 3528 adolescents aged 12.5–17.5 years. For the purpose of the current study, the participants also had to have valid data on all IFIS items, and have completed at least one of the fitness tests. Based on these criteria, a total of 3059 adolescents (52% girls) participated in this study. Body weight, height and BMI in the current study sample were similar to the original HELENA-CSS sample (all P>0.05).

In the HELENA study, one-third of the classes were randomly selected for blood collection, resulting in a total of 1089 participants providing fasting blood.9 Sampling of these participants was balanced by centre, age and gender.8 From this sub-sample, 981 adolescents (53% girls) had valid data on self-reported fitness variables and were included in the current study. This sub-sample also had similar body weight, height, BMI and physical fitness (all the fitness variables) compared with the original HELENA-CSS sample (all P>0.05).

Reliability study (Aim 3)
The HELENA adolescents were not included for test–retest reliability purposes, since the physical fitness tests took place immediately after the IFIS was administered and this could influence the perception on their fitness levels at the retest. Therefore, we asked adolescents from five different cities in southern Spain, not involved in the HELENA study, to complete the IFIS twice, 2 weeks apart. A total of 277 adolescents (51% girls) aged 12.5–17.5 years successfully completed the IFIS on the two occasions and were
included in the current study. None of these adolescents was participating in another study.

**Ethics**

After receiving complete information about the aims and methods of the study, all parents/guardians signed a consent form, and all adolescents gave assent to participate in the study. The study was performed following the ethical guidelines of the Declaration of Helsinki 1961 (revision of Edinburgh 2000), the Good Clinical Practice and the legislation about clinical research in humans in each of the participating countries. The protocol was approved by Human Research Review Committees at the centres involved.

**Description of the IFIS**

We aimed to develop a simple and short self-administered scale to assess physical fitness, which could be completed in a few minutes (the IFIS). To be of use to researchers, the IFIS should be able to accurately rank the adolescents according to their overall physical fitness and by the main physical fitness components. The IFIS is composed of five Likert-scale questions asking about the perceived adolescents’ overall fitness, CRF, MF, SP–AG and flexibility in comparison with their friends’ physical fitness (very poor, poor, average, good and very good). The IFIS was originally written in English and then culturally adapted and translated (and reverse-translated) to all the languages involved in the HELENA study: German, Austrian German, Greek, Flemish, French, Hungarian, Italian, Spanish and Swedish. Versions of the IFIS in the nine different languages are available in the Supplementary Data 1 (available at *IJE* online), and also at the HELENA study website (www.helenastudy.com/IFIS).

**Physical fitness**

A detailed manual of operations on the field-based fitness tests was read by every researcher involved in the fieldwork before data collection started. In addition, workshop training lasting 1 week was carried out in Zaragoza (Spain) in January 2006, in order to standardize and harmonize the measurement of the physical fitness tests. The instructions given to the participants in every test were standardized for all the cities and were translated into the local language, to ensure that the same verbal information was given to all participants in the HELENA study.

Physical fitness characteristics of the study sample, as well as all the procedures used for assessing physical fitness in the HELENA study, have been published elsewhere. In brief, ‘CRF’ was assessed by the 20 m shuttle run test. The equation reported by Léger et al. was used to estimate the maximum oxygen consumption (VO$_{2\text{max}}$ ml/kg/min) from the 20-m shuttle run test scores. Adolescents were classified according to their CRF levels based on the FITNESSGRAM Standards for Healthy Cardiorespiratory Fitness Zone. Males with a VO$_{2\text{max}}$ of $\geq 42$ ml/kg/min were classified as having a healthy CRF level. Females aged 12 and 13 years with a VO$_{2\text{max}}$ of 37 and $\geq 36$ ml/kg/min, respectively, were classified as having a healthy CRF level. Females aged $\geq 14$ years with a VO$_{2\text{max}}$ of $\geq 35$ ml/kg/min were classified as having a healthy CRF level. For exploratory purposes, we additionally estimated VO$_{2\text{max}}$ using the equation developed by Ruiz et al., in order to test if the results could be affected by the equation chosen. ‘MF’ was assessed by means of the handgrip strength and standing long jump tests. ‘SP–AG’ was assessed with the 4 x 10-m shuttle run test, and ‘flexibility’ with the back-saver sit and reach test. The scientific rationale for the selection of all these tests has been previously published. The physical fitness tests mentioned above have shown to be valid and reliable in young people. Further information about the validity and reliability of the fitness tests used in this study are available in the Supplementary Data 2 (available at *IJE* online).

Sex- and age-specific ‘z-scores’ were computed for all the fitness tests in order to allow comparisons among the tests. We computed an ‘MF’ variable as the average of the two MF tests (z-score variables), and an ‘overall physical fitness’ variable as the average of all four physical fitness components studied (z-score variables).

**CVD risk factors**

**Physical examination**

The anthropometric methods followed in the HELENA study has been described in detail by Nagy et al. Briefly, height was measured to the nearest 0.1 cm in barefeet using a stadiometer (SECA 225). Body weight was measured to $\sim 0.05$ kg using an electronic scale (SECA 861), with the participants in their underwear. BMI was calculated as body weight (kg) divided by height (m) squared. Waist circumference was measured in triplicate at the midpoint between the lowest rib and the iliac crest, using an anthropometric tape (SECA 200). Waist-to-height ratio was computed as dividing waist circumference (cm) by height (cm), and this provides a surrogate measure of central body fat. Skinfold thicknesses were measured to $\sim 0.2$ mm, in triplicate, at the triceps, biceps, subscapular, suprailiac, thigh and calf on the left side of the body, using a Holtain Caliper (Crymych, UK). Body fat percentage was calculated by the equation described by Slaughter et al. and fat mass index (fat mass in kg/height in m$^2$) was calculated. A sum of the six skinfolds measured was also computed and used in the analyses. Pubertal status was measured according to the methods described by Tanner and Whitehouse. Briefly, pubertal status was assessed during a medical examination by a physician/paediatrician. Degree of pubertal development ranged from 1 to 5 (stages), and classification was based on the development of the testes, scrotum and penis size in the boys, and on breast development and pubic hair in the girls.
Blood pressure
Systolic and diastolic blood pressure was measured with an automatic oscillometric device (OMRON M6). The adolescents sat on a chair quietly for 5 min before the measurements were conducted on the right arm in an extended position. Two measures were taken 5 min apart, and the lowest value was recorded in mmHg. The mean arterial pressure, defined as the average arterial pressure during a single cardiac cycle, was calculated using the following equation: diastolic blood pressure + [0.333 × (systolic blood pressure–diastolic blood pressure)].

Biochemical analyses
A detailed description of the blood sampling procedures has been published elsewhere. In short, serum concentrations of cardiovascular risk factors were measured in centralized laboratories, from the fasting blood samples. Serum total cholesterol, high-density lipoprotein (HDL) cholesterol, triglycerides and glucose were measured on the Dimension RxL clinical chemistry system (Dade Behring, Schwalbach, Germany) with enzymatic methods that used the manufacturer's reagents and instructions. Insulin was measured by a solid-phase two-site chemiluminescent immunometric assay with an Immulite 2000 analyzer (DPC Biermann GmbH, Bad Nauheim, Germany). The homeostasis model assessment (HOMA) was also calculated. High sensitive C-reactive protein was measured by immunoturbidimetry (Olympus AU2700 Analyzer, Olympus UK Ltd, Watford, UK).

Socio-economic status
Maternal education was used as an indicator of socio-economic status, consistent with previous studies. The highest level of education was reported by the mothers and coded as 1 = lower education, 2 = lower secondary education, 3 = higher secondary education and 4 = higher education or university.

Statistical analysis
The IFIS and measured fitness (Aim 1)
The capacity of the IFIS to correctly rank adolescents into appropriate physical fitness levels was determined by means of analysis of variance without any adjustment (ANOVA) and after adjustment for the following confounders [analysis of covariance (ANCOVA)]: sex, age, sexual maturation status and centre. Measured fitness variables were entered as dependent variables and self-reported fitness variables as fixed factors. Differences in measured fitness z-scores among self-reported fitness categories were analysed by ANOVA (unadjusted models) and ANCOVA (models adjusted for the set of confounders). We performed additional analyses that were stratified by gender, age (< 14.5 and ≥ 14.5 years, based on the median) and weight status (non-overweight and overweight/obesity). We used binary logistic regression to determine the odds ratios (ORs) and 95% confidence intervals (CIs) of having a healthy CRF according to self-reported CRF.

The IFIS and CVD risk factors (Aim 2)
We studied the association between self-reported fitness and CVD risk factors by means of ANCOVA after adjustment for the set of confounders. Since blood sampling was carried out in one-third of the study sample, the adolescents in the ‘very poor’ and ‘poor’ categories were merged to allow for a sufficient sample size for analytical purposes.

Reliability (Aim 3)
The test–retest reliability of the IFIS was examined by means of per cent agreement and weighted Kappa coefficients. For per cent agreement, we calculated the difference between the initial test (T1) and the retest (T2). A difference (T2–T1) equal to 0 was called ‘perfect’ agreement (same test–retest answer); and a difference of 0 ± 1 was called ‘perfectly acceptable’ agreement. We also calculated weighted κ-coefficients, which is more appropriate when dealing with ordered categorical data. Cohen’s weighted κ not only accounts for strict agreement (as does the ‘unweighted’ κ), but also provides weighting to adjacent categories. For example, exact agreements might be given full weight, one-category difference given weight 1/2 and so on. Of the two ways of weighting, linear and quadratic, we chose linear weights as this is recommended when the difference between the first and second category has the same importance as a difference between the second and third category. For all the analyses, we used SPSS 17.0 software for Windows (SPSS Inc., Chicago, IL, USA) and the significance of the tests was interpreted as suggested by Sterne and Davey-Smith. Cohen’s weighted κ is not available in the standard SPSS package, but command syntax is available from the ‘Knowledgebase’ at SPSS.com. Data for imputation into the syntax were generated from cross-tabulation.

Results
The distributions of the answers to each of the five items assessed by the IFIS are shown in Supplementary Data 3 (Figure S1; available at IJE online). The distributions for all the items were negatively skewed, with a small number of adolescents reporting very poor fitness (percentage ranging from 1.0 to 2.7 of the study sample). Higher scores were reported for the boys compared with the girls with regard to CRF, MF, SP–AG and overall physical fitness (Chi-squared tests, P < 0.001). No differences between girls and boys were observed for self-reported flexibility (P = 0.114) (data not shown). Overweight/obese adolescents reported higher muscular strength than non-overweight adolescents (P < 0.001).
The IFIS and measured fitness (Aim 1)

Comparisons between self-reported fitness and measured fitness, without and with adjustment for a set of confounders (sex, age, sexual maturation and centre), are shown in Table 1. For all the fitness variables studied, the adolescents reporting a good or very good fitness had better measured fitness compared with those reporting poor or very poor fitness level (all \( P < 0.001 \)). No differences were observed between the categories very poor and poor with regard to measured fitness. Overall, the associations between self-reported and measured fitness were slightly strengthened after adjusting for the confounders.

The associations between self-reported and measured physical fitness, without and with adjustment for the confounders, are graphically shown in Figure 1. In general, a linear dose–response relationship between self-reported and measured fitness was observed. The ORs for having a healthy CRF (based on the FITNESSGRAM cut-offs) in adolescents reporting a good or very good CRF were 4.6 (95% CI 2.6–8.3) and 7.3 (95% CI 4.0–13.5) respectively, compared with those reporting a very poor CRF (Figure 2). After adjusting for the confounders, the corresponding ORs increased to 3.6 (95% CI 3.0–10.5) and 12.0 (95% CI 6.2–23.1), respectively (Figure 2). When Ruiz’s equation instead of Leger’s equation was used in the analyses, the result did not substantially change (data not shown).

Stratified analyses showed consistent results for boys and girls, younger and older adolescents (<14.5 and \( \geq 14.5 \) years), and non-overweight and overweight/obese adolescents (data not shown). Likewise, further adjustment for maternal education did not influence the results.

The IFIS and CVD risk factors (Aim 2)

The association of self-reported fitness with adiposity indices and other CVD risk factors are shown in Figures 3 and 4, respectively. Adolescents reporting a very good overall physical fitness had healthier levels in most of the CVD risk factors studied, compared with their peers reporting very poor/poor overall physical fitness level. Similar findings were observed for self-reported CRF and self-reported SP–AG. Self-reported MF was positively related to BMI. The use of the sum of six skinfolds instead of fat mass index provided identical results (data not shown).

Adolescents reporting a very good CRF, SP–AG and overall fitness level had an 80, 84 and 87%, respectively, lower risk of being overweight/obese compared with adolescents reporting a very poor/poor fitness (OR 0.20, 95% CI 0.14–0.30; OR 0.16, 95% CI 0.11–0.24 and OR 0.13, 95% CI 0.08–0.19, respectively; Figure S2, Supplementary Data 4; available at IJE online). Self-reported MF was not associated with the risk of being overweight, in line with the results presented in Figure 3. Further adjustment for maternal education did not influence any of the results described above.

### Table 1

<table>
<thead>
<tr>
<th></th>
<th>Unadjusted models</th>
<th>Adjusted modelsa</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Pair-wise comparisons</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Very poor (1)</td>
<td>Poor (2)</td>
<td>Average (3)</td>
<td>Good (4)</td>
<td>Very good (5)</td>
<td></td>
<td>1–2</td>
<td>2–3</td>
<td>3–4</td>
<td>4–5</td>
</tr>
<tr>
<td>CRF: 20-m shuttle run (stage)</td>
<td>2635</td>
<td>3.0 (0.2)</td>
<td>3.4 (0.1)</td>
<td>4.2 (0.1)</td>
<td>5.5 (0.1)</td>
<td>6.9 (0.1)</td>
<td>0.852</td>
<td>&lt;0.001</td>
<td>0.852</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>MF: handgrip (kg)</td>
<td>3043</td>
<td>24.0 (0.4)</td>
<td>26.3 (0.4)</td>
<td>28.6 (0.1)</td>
<td>32.2 (0.1)</td>
<td>36.3 (0.1)</td>
<td>0.491</td>
<td>&lt;0.001</td>
<td>0.491</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Standing long jump (cm)</td>
<td>3018</td>
<td>142.2 (4.5)</td>
<td>149.6 (1.8)</td>
<td>155.5 (1.0)</td>
<td>169.5 (1.1)</td>
<td>181.9 (2.0)</td>
<td>0.744</td>
<td>0.042</td>
<td>0.744</td>
<td>0.042</td>
</tr>
<tr>
<td>SP–AG: shuttle run 4 × 10 m (s)</td>
<td>2919</td>
<td>13.4 (0.2)</td>
<td>12.9 (0.1)</td>
<td>12.6 (0.1)</td>
<td>12.0 (0.1)</td>
<td>11.6 (0.1)</td>
<td>0.056</td>
<td>0.020</td>
<td>0.056</td>
<td>0.020</td>
</tr>
<tr>
<td>Flexibility: back-saver sit and reach (cm)</td>
<td>2917</td>
<td>15.8 (0.8)</td>
<td>18.3 (0.4)</td>
<td>21.8 (0.2)</td>
<td>24.5 (0.2)</td>
<td>26.4 (0.4)</td>
<td>0.046</td>
<td>&lt;0.001</td>
<td>0.046</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

SE: standard error.

aAnalysis of covariance adjusted for sex, age, sexual maturation status and centre.

bThe lower the score (time in seconds), the better the performance.
Reliability (Aim 3)

On average, perfect agreement was observed in 65% (range 61–70) of the adolescents and perfect-acceptable agreement in 97% (range 96–99) for the self-reported overall physical fitness and the main components (Figure S3, Supplementary Data 5; available at IJE online). Test–retest weighted $k$-coefficients for CRF, MF, SP–AG, and flexibility and overall fitness were 0.58, 0.54, 0.60, 0.59 and 0.65 (all $P<0.001$), respectively. These figures were similar in males and females, and in younger and older adolescents (data not shown).

Discussion

The IFIS and measured fitness

The findings in the current study suggest that the IFIS is a valid tool for ranking adolescents according to their objectively measured physical fitness levels. These findings indicate that adolescents properly understood the terminology used in the IFIS, and had a correct self-concept about their levels of CRF, MF, SP–AG, and flexibility. Interestingly, the single question about their self-perceived overall fitness level was strongly associated with their actual overall fitness. We observed a dose–response relationship between reported and measured fitness in boys and girls, younger and older adolescents and overweight and non-overweight participants.

Currently available self-report fitness methods greatly differ in content and structure. Some questionnaires, such as the Huet questionnaire, are composed of 10 questions to estimate CRF (expressed as $\text{VO}_{2\text{max}}$ in ml/kg/min). Similarly, Mikkelsson et al. used several questions to estimate specific physical fitness...
fitness components, such as self-reported endurance (also called CRF), which was assessed by computing an endurance index from several questions regarding self-perceived distance that the participants could complete running, cycling, skiing and walking in a continuous way. The approach used to design the IFIS was totally different. The adolescents were directly asked how they perceived their own CRF compared with their friends, which estimates their fitness level at a given age. A similar approach has been used in another study in adults. We postulate that the human brain might be more accurate at classifying than at quantifying information, yet this needs to be scientifically proved. For instance, it seems easier to classify people into weight status categories (normal-weight vs overweight) than to estimate their weight. The IFIS was designed based on this principle.

The questionnaires/scales mentioned above were designed and tested in adults. There are few available studies examining the validity of self-reported fitness in young people. Jürimäe and Saar compared self-reported and measured physical fitness components and found a good association for all of them, except for MF (handgrip strength). In contrast, we found an association between self-reported and measured MF, including handgrip strength test. In agreement with our results, the perceived fitness of Australian school children was associated with measured CRF and MF. Earlier physical self-perception instruments include Fox and Corbin’s Physical Self-Perception Profile, a 30-item measure of physical self-perception, which includes five subscales: physical self-worth, physical strength, sports competence, physical condition and body attractiveness. The Physical Self-Description Questionnaire is another physical self-concept instrument designed to measure 11 scales, including strength, endurance/fitness and flexibility, with the questionnaire showing good internal consistency, reliability and construct validity in high-school students. Collectively, these results contribute to the growing body of support for the validity of physical self-concept responses.

The IFIS and CVD risk factors

We studied nine established CVD risk factors in relation with self-reported fitness, including total and central adiposity indices, total and HDL cholesterol, triglycerides, blood pressure, insulin resistance (HOMA) and C-reactive protein. Measured CRF, MF and SP–AG are associated with most of these CVD risk factors in adolescent populations. We hypothesized that valid self-reported fitness measures should be associated with these risk factors in a similar fashion. Self-reported flexibility was not analysed, since there is no scientific evidence or physiological rationale supporting the association between measured flexibility and the CVD risk factors studied.

We found that high levels of self-reported CRF, SP–AG and overall physical fitness were associated with...
healthier levels in most of the CVD risk factors—i.e., total cholesterol $n=855$, high density lipoprotein cholesterol (HDLc) $n=855$, triglycerides $n=855$, mean arterial pressure $n=2727$, insulin resistance $n=839$ and C-reactive protein $n=594$—according to categories of self-reported physical fitness in adolescents. *$P<0.05$, †$P<0.01$, ‡$P<0.001$, **$P<0.1$ between ‘Very good’ and ‘Very poor/poor’. Analysis of covariance adjusting for sex, age, sexual maturation and centre. Overall, overall physical fitness.

Figure 4 Differences in several CVD risk factors—i.e., total cholesterol $n=855$, high density lipoprotein cholesterol (HDLc) $n=855$, triglycerides $n=855$, mean arterial pressure $n=2727$, insulin resistance $n=839$ and C-reactive protein $n=594$—according to categories of self-reported physical fitness in adolescents. *$P<0.05$, †$P<0.01$, ‡$P<0.001$, **$P<0.1$ between ‘Very good’ and ‘Very poor/poor’. Analysis of covariance adjusting for sex, age, sexual maturation and centre. Overall, overall physical fitness.

SP–AG or overall fitness who were 80–87% less likely to be overweight/obese than their peers reporting a very poor/poor fitness level. This finding suggests that asking adolescents about their perceived...
fitness levels provides additional information about their weight status.

Self-reported MF was not associated with cardiovascular profile. It is possible that the adolescents interpreted the question about MF as absolute strength, which has shown to be higher in overweight/obese adolescents. In this study, overweight/obese adolescents reported a higher strength than non-overweight adolescents (chi-squared test, \( P < 0.001 \)). It is well known that overweight/obesity has an adverse effect on health; therefore, the fact that the high self-reported MF group was over-represented by overweight/obese adolescents, may partially explain the lack of association between self-reported muscular strength and CVD risk.

Borodulin et al., using an identical scale for self-reported overall physical fitness that was used in the current study (from 1 = very poor to 5 = very good), also observed an inverse association with C-reactive protein in a sample of Finnish adults. Recently, Phillips et al. found that self-reported fitness is a predictor of mortality in middle-age adults, suggesting that where the objective assessment of aerobic fitness is not feasible, a simple measure of subjective fitness could prove a useful alternative. To the best of our knowledge, this is the first study examining the associations between a complete set of self-reported fitness components and CVD risk in young population.

Reliability
When the test–retest reliability of IFIS was examined on 277 adolescents, on average, perfect agreement was observed in 65%, and perfect-acceptable agreement on 97% of them. Test–retest weighted Kappa coefficients ranged from 0.54 to 0.65, which can be considered ‘moderate’ to ‘good’ agreement.

Limitations
A limitation of the IFIS is that the output can be affected by the average fitness level of region/country differences. For this reason, the IFIS should be mainly used for categorizing a study population into different fitness levels and to relate this with different health outcomes. The percentage of people reporting a ‘very poor’ fitness was very small (from 1.0 to 2.7%). Our interpretation of this result is that adolescents (probably as anybody else) do not like to recognize/report that they are ‘very bad’ at something. For this reason, when interpreting the results from the IFIS, the categories ‘very poor’ and ‘poor’ can be merged in a single ‘very poor/poor’ fitness group. No accurate information about ethnicity is available in the HELENA study; future studies should test the validity and reliability of the IFIS in different ethnic groups. Also, the reliability study was performed in Spain; further reliability studies in other countries are needed. It is important to highlight the need for additional cross-validation testing in different samples and ethnicities before the IFIS can be considered acceptable for global use.

Strengths and practical implications
The findings in this study suggest that the IFIS is a useful and reliable tool for estimating fitness in adolescent populations. The structure and content of the IFIS is simple and it can be completed within 1–5 min. A major strength of the current study was that IFIS was tested in nearly 3000 adolescents from northern, southern, western and eastern Europe, and is available in nine different languages (English, German, Austrian German, Greek, Flemish, French, Hungarian, Italian, Spanish and Swedish). In addition, we conducted a health-related validity study and found a strong association between self-reported physical fitness and the CVD risk factors studied.

The field-fitness tests used in this study have been widely used and tested for validity and reliability. Our group has recently carried out a number of systematic reviews and methodological studies in order to test the validity and reliability of fitness testing in young people. As a result of these investigations we believe that the tests used in this study provide a good measure of physical fitness in adolescents, and they can therefore be used to test the research question addressed in this study. The Supplementary Data 2 (available at *IJE* online) includes further discussion and rationale about the validity and reliability of the fitness test selected for this study.

The IFIS can be considered a useful and meaningful tool for the subjective assessment of physical fitness. Large surveys that focus on youth, in which field-based fitness measures cannot be included for time or practical reasons, will have the opportunity to use the IFIS as an alternative. Further, the IFIS could be used to target populations with low physical fitness level and implement physical fitness enhancing programmes in such populations where interventions are most needed.

Conclusion
The results suggest that: (i) the IFIS is a simple and useful assessment tool able to correctly rank adolescents according to their measured physical fitness levels; (ii) adolescents reporting a good/very good overall fitness, CRF or SP–AG have a healthier cardiovascular profile; and (iii) the IFIS is a reliable method to be used in adolescents.

Supplementary data
Supplementary data are available at *IJE* online.
Funding
European Community Sixth RTD Framework Programme (Contract FOOD-CT-2005-007034). Grants from the Spanish Ministry of Education (EX-2009-0899, EX-2008-0641, AP2006-02464), the Spanish Ministry of Science and Innovation (RYC-2010-05957), the Swedish Heart-Lung Foundation (20090635), the Swedish Council for Working Life and Social Research, and the assessing levels of physical activity in a population level study, a European Union-funded study, in the framework of the Public Health Programme (Ref: 2006120), and the Spanish Ministry of Health: Maternal, Child Health and Development Network (number RD08/0072).

Acknowledgements
The authors gratefully acknowledge all participating children and adolescents, and their parents and teachers, for their collaboration. They also acknowledge all the members involved in field work for their efforts and great enthusiasm. They thank JM Fernández, O Requena, J Peña, DN Ardoy, MA Fernández, V Molina and P Oller for their participation in the reliability study. They also wish to thank Jonathan Mitchell for his valuable comments and English revision.

Conflict of interest: None declared.

KEY MESSAGES
- This study provides a valid and reliable self-report tool, the IFIS, for correctly ranking adolescents according to their measured physical fitness levels.
- The IFIS is available in nine different languages (English, German, Austrian German, Greek, Flemish, French, Hungarian, Italian, Spanish and Swedish), requires 1–5 min to complete and can be a useful assessment tool in large surveys in which fitness cannot be measured for time or practical reasons.
- Adolescents reporting good/very good CRF, SP–AG or overall fitness, as assessed by the IFIS, have a more favourable cardiovascular profile.

References
34 Borodulin K, Laatikainen T, Salomaa V, Jousilahti P. Associations of leisure time physical activity, self-rated physical fitness, and estimated aerobic fitness with serum C-reactive protein among 3803 adults. Atherosclerosis 2006;185:381–87.