Artificial Intelligence in Medicine (2008) xxx, xxx—xxx

ARTIFICIAL INTELLIGENCE IN MEDICINE

ELSEVIER

http://www.intl.elsevierhealth.com/journals/aiim

Artificial neural network-based equation for estimating VO_{2max} from the 20 m shuttle run test in adolescents

Jonatan R. Ruiz [a,b,1,*], Jorge Ramirez-Lechuga [c,1], Francisco B. Ortega [a,b], José Castro-Piñero [d], Jose M. Benitez [e], Antonio Arauzo-Azofra [e], Cristobal Sanchez [c], Michael Sjöström [b], Manuel J. Castillo [a], Angel Gutierrez [a], Mikel Zabala [c]

on behalf of the HELENA Study Group

[a] *Department of Physiology, School of Medicine, University of Granada, Granada 18012, Spain*
[b] *Unit for Preventive Nutrition, Department of Biosciences and Nutrition at NOVUM, Karolinska Institutet, Huddinge SE-14157, Sweden*
[c] *Department of Physical Education, School of Sport Sciences, University of Granada, Granada 18015, Spain*
[d] *Department of Physical Education, School of Education, University of Cadiz, Puerto Real 11519, Spain*
[e] *Department of Computer Science and Artificial Intelligence, Escuela Técnica Superior de Ingenierías de Informática y Telecomunicación, University of Granada, Granada 18071, Spain*

Received 23 November 2007; received in revised form 11 June 2008; accepted 16 June 2008

KEYWORDS
Cardiorespiratory fitness;
Maximal oxygen uptake;
Aerobic capacity test;
Exercise field test;
Artificial neural networks

Summary

Objective: To develop an artificial neural network (ANN)-equation to estimate maximal oxygen uptake (VO_{2max}) from 20 m shuttle run test (20mSRT) performance (stage), sex, age, weight, and height in young persons.
Methods: The 20mSRT was performed by 193 (122 boys and 71 girls) adolescents aged 13—19 years. All the adolescents wore a portable gas analyzer to measure VO_2 and heart rate during the test. The equation was developed and cross-validated following the ANN mathematical model. The neural net performance was assessed through several error measures. Agreement between the measured VO_{2max} and estimated VO_{2max} from Léger's and ANN equations were analysed following the Bland and Altman method.
Results: The percentage error was 17.13 and 7.38 for Léger and ANN-equation ($P < 0.001$), respectively, and the standard error of the estimate obtained with

* Corresponding author at: Department of Physiology, School of Medicine, University of Granada, Granada 18012, Spain. Tel.: +34 958 243 540; fax: +34 958 249 015.
E-mail address: ruizj@ugr.es (J.R. Ruiz).
[1] These authors contributed equally to this work.

doi:10.1016/j.artmed.2008.06.004

Please cite this article in press as: Ruiz JR, et al., Artificial neural network-based equation for estimating VO_{2max} from the 20 m shuttle run test in adolescents, Artif Intell Med (2008), doi:10.1016/j.artmed.2008.06.004

Léger's equation was 4.27 ml/(kg min), while for the ANN-equation was 2.84 ml/(kg min). A Bland–Altman plot for the measured VO_{2max} and Léger-VO_{2max} showed a mean difference of 4.9 ml/(kg min) ($P < 0.001$), while the Bland–Altman plot for the measured VO_{2max} and ANN-VO_{2max} showed a mean difference of 0.5 ml/(kg min) ($P = 0.654$). In the validation sample, the percentage error was 21.08 and 8.68 for Léger and ANN-equation ($P < 0.001$), respectively.

Conclusions: In this study, an ANN-based equation to estimate VO_{2max} from 20mSRT performance (stage), sex, age, weight, and height in adolescents was developed and cross-validated. The newly developed equation was shown to be more accurate than Léger's. The proposed model has been coded in a user-friendly spreadsheet.

1. Introduction

The maximal rate of oxygen uptake (VO_{2max}) is considered as a gold standard for measurement of cardiorespiratory fitness. Cardiorespiratory fitness is a direct marker of physiological status and reflects the overall capacity of the cardiovascular and respiratory systems and the ability to carry out prolonged exercise [1]. In addition, recent reports suggest that cardiorespiratory fitness is also an important health marker in young individuals [2,3]. High-cardiorespiratory fitness during childhood and adolescence has been associated with a favourable plasma lipid profile in both overweight and non-overweight adolescents [4], with total and central body fat [5,6], features of the metabolic syndrome [7,8], blood pressure [9], novel cardiovascular disease risk factors [10], and with arterial compliance in young people [11]. These findings support the concept that cardiorespiratory fitness may exert a protective effect on the cardiovascular system from an early age [2,3].

Cardiorespiratory fitness is one of the main health-related physical fitness components used in schools, sports centres and health institutions. One of the most widely used field tests for estimating cardiorespiratory fitness among adolescents is the 20 m shuttle run test (20mSRT), also called the "Course Navette" test [12,13]. The 20mSRT consists of 1-min stages of continuous incremental speed running. The initial speed is 8.5 km/h, and increases by 0.5 km/h per minute (1 min equal one stage), reaching 18.0 km/h at minute 20. The participants are required to run between two lines 20 m apart, while keeping the pace with audio signals emitted from a pre-recorded CD. The test ends when the participant fails to reach the end lines concurrent with the audio signals on two consecutive occasions. The 20mSRT, or a slight modification of it, has been included in several fitness batteries, such as the EUROFIT [14], and the FITNESSGRAM battery [15]. The 20mSRT is a feasible fitness test, since a large number of subjects can be tested at the same time, which enhances the motivation of the participants. It can be conducted indoors or outdoors in a relatively small area, and on different surfaces (slippery and rubber floors).

Several equations have been developed to estimate VO_{2max} from maximal speed attained during the 20mSRT (Table 1). Léger et al. [16] developed an equation based on a sample of 188 boys and girls aged 8–19 years to estimate the VO_{2max} from maximal speed attained during the 20mSRT, age and the speed and age interaction. However, Léger's equation has some limitations. Sex is not included in the model, yet it is well known that physical performance is different in boys and girls of all ages. Moreover, the estimates of VO_{2max} for low scores were based on extrapolated data from the study since the original study population did not have data for these points. The accuracy of the Léger's [16] prediction model has been examined by several researchers [17–26], but no attempts have been made to develop a more accurate model in a wide age range of young individuals.

It seems viable to develop a more accurate VO_{2max} equation for the adolescent period, while taking those variables which have been shown to have an impact on the level of cardiorespiratory fitness into account. Published equations for VO_{2max} have the shape of a linear or quasi-linear expression on different input variables (sex, age, body weight, and stage) (Table 1). These types of models have mainly been used because of their simplicity, ease of using, and familiarity. A way forward in obtaining an improved model could be done by exploring the feasibility of new methods. Recently, there has been a growing interest in artificial neural networks (ANN). ANNs have some theoretical advantages over more traditional regression methods [27], such as its capability of producing a nonlinear input–output mapping. Predictive models based on ANNs have been studied extensively in many areas of medicine (e.g. breast cancer diagnosis, mortality assessment in intensive care units, diagnostic scoring, renal function evaluation, etc.).

Please cite this article in press as: Ruiz JR, et al., Artificial neural network-based equation for estimating VO_{2max} from the 20 m shuttle run test in adolescents, Artif Intell Med (2008), doi:10.1016/j.artmed.2008.06.004

Table 1 Equations to estimate $VO_{2max\ or\ peak}$ from the 20 m shuttle run test

Study	Sample	Age (year)	Input variables	Equation to estimate $VO_{2max\ or\ peak}$ (ml/(kg min))
Léger et al. [16]	188 boys and girls	8–19	Speed and age	Boys and girls: $VO_{2max} = 31.025 + 3.238 \times S - 3.248 \times A + 0.1536 \times S \times A$ (*A* the age; *S* the final speed ($S = 8 + 0.5 \times$ last stage completed)
Barnett et al. [17]	27 boys, 28 girls	12–17	Gender and skinfold and speed	Boys and girls: $VO_{2peak} = 28.3 - 2.1 \times G - 0.7 \times Z + 2.6 \times S$ (*G* is gender (male = 0, female = 1); *Z* is triceps skinfold; *S* the final speed)
			Gender and body weight and speed	Boys and girls: $VO_{2peak} = 25.8 - 6.6 \times G - 0.2 \times BM + 3.2 \times S$ (*G* is gender (male = 0, female = 1); BM the body mass (kg); *S* the final speed)
			Gender and age and speed	Boys and girls: $VO_{2peak} = 24.2 - 5.0 \times G - 0.8 \times A + 3.4 \times S$ (*G* is gender (male = 0, female = 1); *A* the age; *S* the final speed)
Stickland et al. [24]	63 boys, 62 girls	18–38	Last half-stage completed and gender	Males: $VO_{2max} = 2.75 \times X + 28.8$. Females: $VO_{2max} = 2.85 \times X + 25.1$ (*X* the last half-stage completed)
Fluoris et al. [25]	110 boys	21 ± 2.5	Speed	Boys: $VO_{2max} = (S \times 6.65 - 35.8) \times 0.95 + 0.182$ (*S* the maximal attained speed)
Matsuzaka et al. [26]	62 boys, 70 girls	8–17	Gender and age and body mass index and speed	Boys and girls: $VO_{2peak} = 25.9 - 2.21 \times G - 0.449 \times A - 0.831 \times BMI + 4.12 \times S$ (*G* is gender (male = 0, female = 1); *A* the age; BMI is body mass index; *S* the maximal running speed)
	56 boys, 99 girls	18–23	Gender and age and body mass index and speed	Boys and girls: $VO_{2peak} = 61.1 - 2.20 \times G - 0.462 \times A - 0.862 \times BMI + 0.192 \times S$ (*G* is gender (male = 0, female = 1); *A* the age; BMI is body mass index; *S* is number of laps completed)
			Gender and body mass index and speed	Males and females: $VO_{2peak} = -2.19 - 3.46 \times G - 0.416 \times BMI + 5.22 \times S$ (*G* is gender (male = 0, female = 1); BMI is body mass index; *S* the maximal running speed)
			Gender and body mass index and speed	Males and females: $VO_{2peak} = 42.4 - 2.85 \times G - 0.488 \times BMI + 0.247 \times S$ (*G* is gender (males = 0, female = 1); BMI is body mass index; *S* is number of laps completed)
Mahar et al. [23]	61 boys, 74 girls	12–14	Laps completed and gender and body weight	Boys and girls: $VO_{2peak} = 47.438 + (S \times 0.242) + (G \times 5.134) - BM \times 0.197$ (*S* is number of laps completed; *G* is gender (male = 1, female = 0); BM is body mass (kg))

Please cite this article in press as: Ruiz JR, et al., Artificial neural network-based equation for estimating VO_{2max} from the 20 m shuttle run test in adolescents, Artif Intell Med (2008), doi:10.1016/j.artmed.2008.06.004

The aim of this study was to develop an ANN-equation to better estimate VO_{2max} from 20mSRT performance (stage), sex, age, weight, and height in adolescents.

2. Methods

2.1. Subjects

A total of 203 adolescents (127 boys and 76 girls) aged 13—19 years volunteered to participate in the study after receiving a detailed explanation about the aim and the clinical implications of the investigation. A comprehensive verbal description of the nature and purpose of the study was also given to the teachers. Written informed consent was obtained from parents, and verbal assent was obtained from participants. The criteria for inclusion were: smoking, no personal history of cardiovascular or metabolic disease, free of disease, any muscular or skeletal injuries, medication at the time of the study and pregnancy. The experimental protocol was approved by the Review Committee for Research Involving Human Subjects at the University of Granada, Spain.

A total of five adolescents discontinued the test because of discomfort or distress. Several ($n = 5$) technical problems also occurred during the test or when downloading the data, which probably yielded inaccurate VO_{2max} results. Therefore, the final sample was confined to 193 (122 boys and 71 girls) adolescents with reliable measures of VO_{2max}.

2.2. Procedure

All participants performed the 20mSRT as previously described by Léger et al. [12]. Participants were required to run between two lines 20 m apart, while keeping the pace with audio signals emitted from a pre-recorded CD. The initial speed was 8.5 km/h, which was increased by 0.5 km/h per minute (1 min equal one stage). The CD used was calibrated over 1 min of duration. Participants were instructed to run in a straight line, to pivot on completing a shuttle, and to pace themselves in accordance with the audio signals. The test was finished when the participant failed to reach the end lines concurrent with the audio signals on two consecutive occasions. Otherwise, the test ended when the subject stopped because of fatigue. All measurements were carried out under standardized conditions on an indoor rubber floored gymnasium. The participants were encouraged to keep running as long as possible throughout the course of the test.

All participants were familiar with the test. The 20mSRT is one of the fitness tests included in the curriculum of Physical Education in Spain. However, 1 week prior the test, participants received a comprehensive instruction after which they also practiced the test. Subjects were instructed to abstain from strenuous exercises 48 h prior to the test. All the tests were conducted by the same investigators and at the same time for each subject (between 10:00 to 13:00 h).

2.3. Physiological measurements

Heart rate was recorded every 5 s throughout the 20mSRT using a Polar telemetry system (Polar 610i). Moreover, participants wore a portable gas analyzer (K4b^2, Cosmed, Rome, Italy), the purpose of which was to measure the VO_2 during the 20mSRT. Respiratory parameters were recorded breath-by-breath, which in turn were averaged over a 10-s period. VO_{2max} was the main parameter determined using the open circuit method. Exhaustion was confirmed when: (1) maximal heart rate was greater than 185 beats/min, (2) respiratory exchange ratio was greater than 1.1, and/or (3) a detection of a plateau in the VO_2 curve, defined as an increase of VO_2 less than 2 ml/(kg min) with a concomitant increase in stage.

The weight of the Cosmed K4b^2 equipment is 1.5 kg including the battery and a specially designed harness. It has been proven to be a valid device when compared with the Douglas bag method [28]. Wearing the portable gas analyzer during the 20mSRT does not significantly alter the subjects' energy demands, as it has been reported [25].

Before each test was conducted, the oxygen and carbon dioxide analyzers were calibrated according to the manufacturer's instructions. This consisted of performing a room air calibration and a reference gas calibration using 15.93% oxygen and 4.92% carbon dioxide. The flow turbine was then calibrated using a 3-l syringe (Hans-Rudolph). Finally, a delay calibration was performed to adjust for the lag time that occurs between the expiratory flow measurement and the gas analyzers. During each test, a gel seal was used to help prevent air leaks from the face mask.

The total time (in s) and the last half-stage completed (here called "stage") were recorded. Measured VO_{2max} was obtained directly from the K4b^2 data. Estimated VO_{2max} was calculated by the Léger's equation [16] (Table 1). VO_{2max} was estimated by the most widely used equation (i.e. Léger's equation) in order to assess the error of the VO_{2max} measurements obtained from Léger's and from the new equation to be developed.

Please cite this article in press as: Ruiz JR, et al., Artificial neural network-based equation for estimating VO_{2max} from the 20 m shuttle run test in adolescents, Artif Intell Med (2008), doi:10.1016/j.artmed.2008.06.004

Body weight was measured to 0.1 kg using a standard beam balance, and body height was measured to the 0.1 cm using a transportable stadiometer, with the participants clad only in their underwear. These measures were taken prior the test.

2.4. Validation sample

To confirm the usefulness of the ANN-equation, an additional group of 29 adolescents (16 males, 13 females) of the same ages volunteered to perform the 20mSRT and a maximal graded treadmill test in a random order within a 2-week period. The 20mSRT was performed following the same procedures as explained above but without wearing the portable gas analyzer. The maximal graded treadmill test (Eric Jaeger, GmbH & Co., Wurzburg, Germany) started with a 3-min warm-up at 6 km/h at 3% grade. The speed was increased 1 km/h every 1 min, and the grade was maintained at 3% throughout the test. The test was terminated when the subject was unable to continue despite verbal encouragement. Heart rate was measured by JECg 12 Channels electromyography (Eric Jaeger, GmbH & Co., Wurzburg, Germany). Oxygen uptake was measured via open circuit spirometry using an automated gas analyzer (Oxycon Delta Version 4.3, Eric Jaeger, GmbH & Co., Wurzburg, Germany) previously calibrated with standard gases. Respiratory parameters were recorded breath-by-breath, which in turn were averaged over a 10-s period. Exhaustion was confirmed using the same criteria as described before.

2.5. Statistical analyses

The mathematical model used to build the new equation to estimate VO_{2max} from 20mSRT performance (stage), sex, age, weight, and height in adolescents was an ANN. An ANN is a mathematical model that emulates some of the observed properties of biological nervous system and draw on the analogies of adaptive biological learning. The ANN modelling procedure has been described in detail elsewhere [29]. In brief, to solve a problem using ANN, a number of steps must be taken:

1. Select the type of neural net for the type of regression problem to be solved, i.e. identification of a VO_{2max} estimator. One of the best options for that purpose is to use a multilayered perceptron.
2. *Data preprocessing*. The data gathered for this study consists of a set of 193 instances, each instance being composed of six variables. All variables were originally expressed in their original units, i.e. sex (boys/girls), age (year), weight (kg), height (cm), stage (last half-stage completed), and VO_{2max} (ml/(kg min)). The sample data was afterwards normalized to the [0.1, 0.9] interval, which simplified the learning of the ANN regression model.
3. *Network design*. The ANN architecture, i.e. the number of input and output variables is set by the problem. There are plenty of different models of neural networks to choose from, each one having its specific properties and advantages for its particular application. One of the most successful and most popular is the feed-forward multilayered perceptron. In this network, the computing units are arranged into three layers, which are conveniently ordered. The information flows forward from the five neurons of the input layer to the three connecting neurons of the hidden layer and finally, to the single neuron of the output layer using no backward connection. The first layer (the input layer) corresponds to the independent variables (sex, age, weight, height, and stage), while the third layer (the output layer) corresponds to the dependent variable score (VO_{2max}). The intermediate layer, which is a hidden layer, consists on all possible connections between the input and the output layer, and allows for a combined impact of a multiple set of independent variables on the output layer. This would be the same as testing all possible interactions in a regression model, but without adding any extra degrees of freedom. To perform the final model selection, i.e. setting the size of the hidden layer, we conducted a cross-validation process. The architecture of the network used in this study is a multilayered perceptron (5-3-1) (Fig. 1).

   The activation function for the hidden and output nodes in the logistic function is: $f(x) = 1/(1+\exp(-x))$. Hence the function computed by the network is ANN $(x_1, \ldots, x_5) = f(v_1h_1 + v_2h_2 + v_3h_3)$, where $x_1, \ldots, x_5$ are the input variables (sex, age, weight, height, and stage); v_1, v_2, and v_3 are the weights of the links from hidden units to the output unit; and h_i is the function computed by the hidden unit i, $h_i(x_1, \ldots, x_5) = f(w_{1i}x_1 + w_{2i}x_{12} + w_{3i}x_3 + w_{4i}x_4 + w_{5i}x_5)$, where w_{ji} is the weight of the link from input j to hidden unit i.
4. *Learning algorithm parameters*. In order to obtain the synaptic weights of the ANN, we used the backpropagation algorithm [30]. The value for the algorithm parameters are 0.2 for the learning rate, and 0.5 for the momentum term. The training of the network is stopped when the sum of squared errors (SSE) falls below 0.00001

Please cite this article in press as: Ruiz JR, et al., Artificial neural network-based equation for estimating VO_{2max} from the 20 m shuttle run test in adolescents, Artif Intell Med (2008), doi:10.1016/j.artmed.2008.06.004

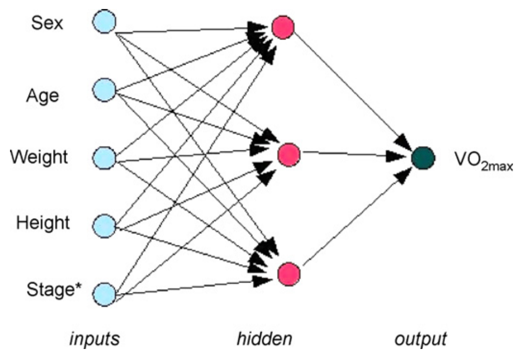

Figure 1 Neural network architecture. *Last half-stage completed.

or when 1500 training epochs have been performed.

5. *Training of the network*. The ANN-model is identified by means of a data-driven process, where a fraction of the available data set is used for designing the model and it is referred to as the *training set*. The remaining set of data is not used in the design of the model as such but rather for evaluating its validity once it is ready. This particular data set is called the *test set*.
6. *Validation of the model*. In order to validate the feasibility of the ANN-model for this problem, a cross-validation technique was applied [31]. It means that the total dataset (composed of 193 samples) was randomly split into k parts with the same number of samples, except one of them ($C = c_1, \ldots, c_k$). The process consists of building k different neural networks. For the model i, with $i = 1, \ldots, k$ the part c_i is used as the test set, and the remainder (all but c_i) are used as the training set. In our experiments, the value we have used for k is the total number of samples in the data set ($n = 193$). Thus each of the nets are built with different training sets, and evaluated on different and independent test sets.

The overall evaluation of the methodology is measured as the average of the performance on the test sets. Then we conducted another cross-validation series with a k value of 10. The results did not materially changed to those obtained for $k = 193$.

The neural net performance was assessed through an error measure. Suppose that N cases are available to evaluate the model, where y is the actual output (the measured VO_{2max}) and $\hat{y}$ is the output computed by the net (estimated VO_{2max} from the ANN-equation). Then, a common measure is the SSE (sum of squared errors) defined as

$$\mathrm{SSE} = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2$$

An easier way of understanding the expression for the error is to use the percentage error, which can be computed as follows: first, the SSE is averaged over the number of cases, rendering the mean sum of squared errors (MSE):

$$\mathrm{MSE} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2$$

MSE is then converted into domain units by taking the root square and yielding the root mean sum of squared errors (RMSE):

$$\mathrm{RMSE} = \sqrt{\mathrm{MSE}}$$

The percentage error should intuitively serve as a good indicator of the performance of a given model:

$$\%\mathrm{Error} = \frac{\mathrm{RMSE}}{\mathrm{domain\ width}} \times 100$$

The standard error of estimate (SEE), is another way to illustrate the performance of the ANN-model, which also serves for comparative purpose:

$$\mathrm{SEE} = \mathrm{SD}_y \sqrt{(1 - R^2_{y\hat{y}})}$$

where SD is the standard deviation of the estimated VO_{2max} from the ANN-model, and R^2 is the correlation between the measured VO_{2max} and the estimated VO_{2max} from the ANN-model.

The SSE difference between the Léger's equation and the ANN-model was examined by paired t-test. A second ANN-model was built with the same procedure and variables as the previous one, but instead of the last half-stage completed, the last stage completed was used.

Sex differences were analysed by one-way analysis of variance (ANOVA), and adjusted for mass significance as described by Holm [32]. Bivariate correlation analysis was done in order to examine the relationship between the measured VO_{2max} and the input variables (age, weight, height, and stage) in boys and girls. The relationship between measured VO_{2max} and estimated VO_{2max} from Léger's equation and the ANN-model was also examined. The overall differences between measured VO_{2max} and estimated VO_{2max} from Léger's equation and ANN-model were calculated by means of ANOVA for repeated measures. The agreement between measured VO_{2max} and estimated VO_{2max} from Léger's and the ANN equation was assessed according to the Bland and Altman method [33,34]. The association between the difference and the magnitude of the measurement (i.e. heteroscedasticity) was examined by regression analysis. $P \leq 0.05$ was considered significant. "P" is the probability that a variate would assume a value greater than or equal to the observed value strictly by chance: $P(z \geq z_{\mathrm{observed}})$ [35].

Please cite this article in press as: Ruiz JR, et al., Artificial neural network-based equation for estimating VO_{2max} from the 20 m shuttle run test in adolescents, Artif Intell Med (2008), doi:10.1016/j.artmed.2008.06.004

Table 2 Physical characteristics and 20 m shuttle run performance of the study participants by gender

	All (n = 193)	Males (n = 122)	Females (n = 71)
Age (year)	16.1 ± 1.2	16.2 ± 1.3	15.9 ± 1.1
Height (cm)	168.3 ± 9.1	172.5 ± 6.7	161.0 ± 8.2*
Weight (kg)	64.6 ± 13.3	68.5 ± 13.5	58.0 ± 9.8*
Stage	6.5 ± 2.4	8.0 ± 1.7	4.0 ± 1.1*
Speed (km/h)	11.3 ± 1.2	12.0 ± 0.9	10.0 ± 0.5*
Time (min)	6.6 ± 2.4	8.0 ± 1.7	4.1 ± 1.1*
Heart rate (beats/min)	197.7 ± 7.9	198.6 ± 7.9	196.2 ± 7.7
Léger-VO_{2max} (ml/(kg min))	43.0 ± 6.8	47.0 ± 5.0	36.2 ± 2.9*
Measured VO_{2max} (ml/(kg min))	47.7 ± 10.0	53.9 ± 6.2	37.1 ± 5.0*

Data are mean ± S.D.
* $P < 0.001$ from comparisons between sexes.

3. Results

Physical characteristics and the 20mSRT performance of the participants are presented in Table 2. Boys and girls were similar in age, but boys were significantly taller and heavier than girls. Moreover, boys had significantly higher values in all the 20mSRT performance-related variables. A bivariate correlation analysis between the measured VO_{2max}, age, weight, height, and stage in boys and girls is presented in Table 3. VO_{2max} was significantly associated with age, weight, and stage in both sexes. A borderline significant association was found between VO_{2max} and height in both boys and girls. Fig. 2(A) shows the relationship between the measured VO_{2max} and the estimated VO_{2max} from the Léger's equation, and Fig. 2(B) shows the relationship between the measured VO_{2max} and the estimated VO_{2max} from the ANN-equation. Estimated VO_{2max} from both the Léger's and the ANN-equation were significantly correlated with the measured VO_{2max} (r = 0.90 and 0.96, respectively, both $P < 0.001$).

Table 3 Bivariate correlation analysis between measured VO_{2max} (ml/(kg min)), age, weigh, height, and speed in males and females

	Age	Weight	Height	Stage[a]
Males (n = 122)				
VO_{2max}, r	−0.238*	−0.517***	−0.160	0.736***
Age, r		0.414***	0.252**	0.057
Weight, r			0.550***	−0.195*
Height, r				−0.070
Females (n = 71)				
VO_{2max}, r	0.501***	−0.241*	0.219	0.813***
Age, r		−0.081	0.147	0.418***
Weight, r			0.183	−0.118
Height, r				0.249*

[a] Refers to the last half-stage completed.
* $P < 0.05$.
** $P < 0.01$.
*** $P < 0.001$.

The evaluation of the error of the VO_{2max} measurements obtained from Léger's and the ANN-equation is presented in Table 4. The SSE was significantly higher in Léger's equation than in the ANN-equation ($P < 0.001$). The SSE obtained from the ANN-model built with the last stage completed was significantly higher than the SSE obtain from the ANN-model built with the last half-stage completed (1699.48 vs. 1600.91, respectively, P = 0.002). The ANN-equation to estimate VO_{2max} (ml/(kg min)) from 20mSRT performance (stage), sex, age, weight, and height in adolescents aged 13—19 years is depicted in Table 5. The proposed model has been coded in a user-friendly spreadsheet, and can be found at http://

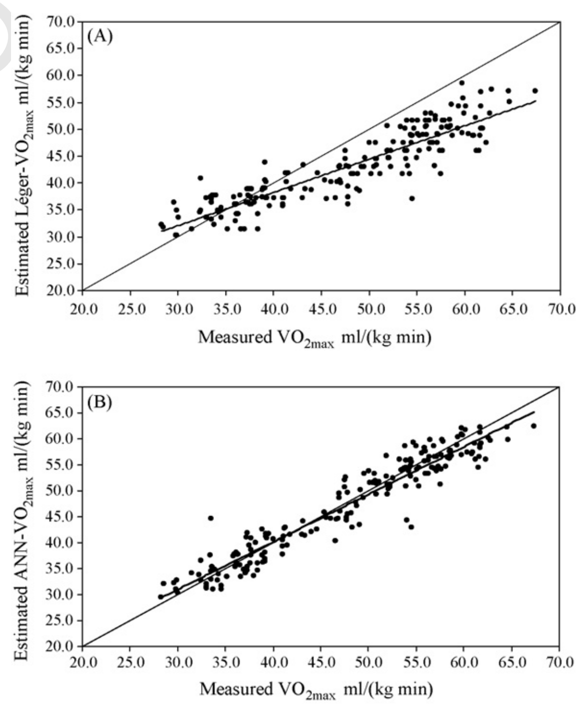

Figure 2 (A) Relationship between estimated VO_{2max} from Léger's equation and measured VO_{2max}. (B) Relationship between estimated VO_{2max} from artificial neural network (ANN)-equation and measured VO_{2max}.

Table 4 Evaluation of the error of the VO_{2max} measurements obtained from Léger's equation and the artificial neural network (ANN)-equation ($n = 193$)

Error measure	Equation	
	Léger	ANN
$SSE = \sum_{i=1}^{N}(y_i - \hat{y}_i)^2$ (ml/(kg min))^2	8663.14	1600.91
$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$ (ml/(kg min))^2	44.89	8.29
$RMSE = \sqrt{MSE}$ (ml/(kg min))	6.70	2.88
$\%Error = \frac{RMSE}{domain\ \ width} \times 100$ (%)	17.13	7.37
$SEE = SD_y\sqrt{(1 - R^2_{y\hat{y}})}$ (ml/(kg min))	4.27	2.84

SSE, sum of squared errors; MSE, mean of squared errors; RMSE, root mean squared errors; SEE, standard error of estimate. N are the cases available to evaluate the model where y is the actual output (measure VO_{2max}) and $\hat{y}$ is the output computed by the either Léger's equation or the net (ANN-VO_{2max}).

www.helenastudy.com/scientific.php (accessed 11 June 2008).

The Bland–Altman plot for the measured VO_{2max} and the estimated VO_{2max} from Léger's equation showed a mean difference of 4.9 ml/(kg min) (Fig. 3(A)). The 95% limits of agreement ranged from −4.3 to 14.1 ml/(kg min). There was a statistically significant difference between the measured VO_{2max} and the estimated VO_{2max} from the Léger's equation (47.7 ml/(kg min) vs. 43.0 ml/(kg min), $P < 0.001$). The Bland–Altman plot for the measured VO_{2max} and the estimated VO_{2max} from the ANN-equation showed a mean difference of 0.5 ml/(kg min) (Fig. 3(B)). The 95% limits of agreement ranged from −5.1 to 6.1 ml/(kg min). There was no statistical significance difference between the measured VO_{2max} and the estimated VO_{2max} from the ANN-equation (47.7 ml/(kg min) vs. 47.2 ml/(kg min), $P = 0.654$). There was an association between the difference and the magnitude of the measurement (i.e. heteroscedasticity) for the Léger's equation ($P < 0.001$), but not for the ANN-equation ($P > 0.5$).

The error assessment of the VO_{2max} measurements obtained from Léger's and the ANN-equation in an independent validation sample is presented in Table 6. The SSE was significantly higher in Léger's equation than in the ANN-equation ($P < 0.001$). There was a statistically significant difference between the measured VO_{2max} and the estimated VO_{2max} from the Léger equation (54.9 ml/(kg min) vs. 47.1 ml/(kg min), $P < 0.001$). The mean difference was 7.8 ml/(kg min), and the 95% limits of agreement ranged from 6.1 to 8.5 ml/(kg min). There was a difference between the measured VO_{2max} and the estimated VO_{2max} from the ANN-equation (54.9 ml/(kg min) vs. 57.4 ml/(kg min), $P < 0.05$). The mean difference was −2.4 ml/(kg min), and the 95% limits of agreement ranged from −3.5 to −1.3 ml/(kg min).

4. Discussion

In this study, an ANN-based equation to estimate VO_{2max} from 20mSRT performance (stage), sex, age,

Table 5 Syntax (Excel spreadsheet) of the artificial neural network-based equation to estimate VO_{2max} (ml/(kg min)) from 20mSRT performance (stage), sex, age, weight, and height in adolescents aged 13–19 years

VO_{2max} (ml/(kg min)) = (1/(1 + exp(−(1/(1 + exp(−((A1 × 0.8+ (−0.7)) × (−1.03329) + (B1 × 0.114285714286 + (−1.38571428571)) × 0.54719 + (C1 × 0.012213740458 + (−0.406870229008)) × 0.61542 + (D1 × 0.0195598978221 + (−2.76356892177)) × −0.51381 + (E1 × 0.0842105263158 + (−0.0684210526316)) × (−0.92239) + (−0.34242)))) × (−0.95905) + 1/(1 + exp(−((A1 × 0.8 + (−0.7)) × (−1.19367) + (B1 × 0.114285714286 + (−1.38571428571)) × (−1.54924) + (C1 × 0.012213740458 + (−0.406870229008)) × (−3.18931) + (D1 × 0.0195598978221 + (−2.76356892177)) × 0.77773 + (E1 × 0.0842105263158 + (−0.0684210526316)) × 3.31887 + (−0.55696)))) × 2.19501 + 1/(1 + exp(−((A1 × 0.8 + (−0.7)) × 1.38191 + (B1 × 0.114285714286 + (−1.38571428571)) × (−2.14449) + (C1 × 0.012213740458 + (−0.406870229008)) × 0.0485 + (D1 × 0.0195598978221 + (−2.76356892177)) × 0.10879 + (E1 × 0.0842105263158 + (−0.0684210526316)) × (−4.90052) + 0.53905))) × (−2.567) + (−0.05105)))) − (−0.478945173945))/0.0204587840012

A1 = sex (boys = 1; girls = 2); B1 = age (year, age range 13–19 years); C1 = weight (kg); D1 = height (cm); E1 = stage (0.5). A user-friendly spreadsheet can be found in http://www.helenastudy.com/scientific.php (accessed 11 June 2008).

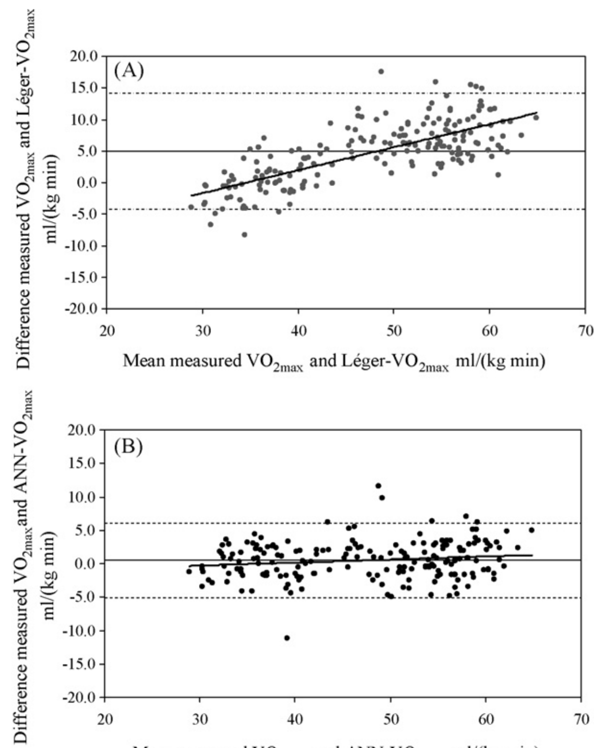

Figure 3 (A) Bland—Altman plot for the measured VO_{2max} and estimated VO_{2max} from Léger's equation. Central line represent the mean difference between equations (4.9 ml/(kg min)) and broken lines represent upper and lower 95% limits of agreement (−4.3 to 14.1 ml/(kg min)). (B) Bland—Altman plot for the measured VO_{2max} and estimated VO_{2max} from artificial neural network (ANN)-equation. Central line represent the mean difference between equations (0.5 ml/(kg min)) and broken lines represent upper and lower 95% limits of agreement (−5.1 to 6.1 ml/(kg min)).

weight, and height in a sample of 193 adolescents aged 13—19 years was developed and cross-validated. The equation is based on: (1) direct VO_2 data collected while the adolescents performed the 20mSRT. (2) The use of a numerical procedure to build the ANN-equation. (3) A fairly large amount of adolescents participating in the test. (4) The inclusion of variables that have previously shown to influence the VO_{2max} for the particular age group being tested. All variables included in the equation are easy to be measured in population-based studies and no specific equipment is required to collect the data. All the technical and environmental variables that may have an influence on the results were carefully controlled in order to obtain highly reliable VO_2 measures. (5) The use of a precise method for assessing agreement between two methods. The most frequently used summary statistics to assess overall agreement between the measurements of different methods has been correlation coefficient. However, correlation is a measure of the strength of association between two variables but not necessarily a measure of agreement [36].

The ANN-based equation proved to be more accurate for a prediction of the VO_{2max} value than Léger's equation for the samples of adolescents studied. Léger's equation had an error of 17.13%, while the ANN-equation had an error of 7.38%. The SEE calculated from Léger's equation was almost twice as high as that obtained with the ANN-equation (4.27 ml/(kg min) vs. 2.84 ml/(kg min), respectively). Moreover, Léger's equation significantly underestimated VO_{2max} by 4.9 ml/(kg min) when compared with the measured VO_{2max} ($P < 0.001$), while the ANN-equation slightly underestimated VO_{2max} by 0.5 ml/(kg min) ($P = 0.654$). The results of this study are in alignment with previous research, which has shown a systematic underestimation of the VO_{2max} value calculated from Léger's equation [24,37]. Similar results were obtained when the equation was applied to an independent group of adolescents.

Differences between the results obtained from the ANN-equation and those obtained from Léger's

Table 6 Evaluation of the error of the VO_{2max} measurements obtained from Léger's equation and the artificial neural network (ANN)-equation in the validation sample ($n = 29$)

Error measure	Equation: Léger	Equation: ANN
$SSE = \sum_{i=1}^{N}(y_i - \hat{y}_i)^2$ (ml/(kg min))^2	2310.85	391.55
$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$ (ml/(kg min))^2	67.97	11.52
$RMSE = \sqrt{MSE}$ (ml/(kg min))	8.24	3.39
$\%Error = \frac{RMSE}{domain\ width} \times 100$ (%)	21.08	8.68
$SEE = SD_y\sqrt{(1 - R^2_{y\hat{y}})}$ (ml/(kg min))	2.86	2.78

SSE, sum of squared errors; MSE, mean of squared errors; RMSE, root mean squared errors; SEE, standard error of estimate. *N* are the cases available to evaluate the model where *y* is the actual output (measure VO_{2max}) and $\hat{y}$ is the output computed by the either Léger's equation or the net (ANN-VO_{2max}).

Please cite this article in press as: Ruiz JR, et al., Artificial neural network-based equation for estimating VO_{2max} from the 20 m shuttle run test in adolescents, Artif Intell Med (2008), doi:10.1016/j.artmed.2008.06.004

equation may be partly explained by the test protocols and the gas analysis procedures used for the tests. Léger et al. recorded VO_{2max} by using the backward extrapolation technique [16]. This technique has been extensively validated, but it can only be considered as an estimate of the actual VO_{2max}. The present method seems to be a more sensitive method, since data were averaged every 10 s, which allowed for the detection of a plateau in the VO_2 over the final workloads.

The ANN-equation has other advantages over Léger's equation, and also on more recently published regression equations (Table 1), such as the inclusion of several variables that influence the level of VO_{2max}. The reason why sex, age, weight, height, and stage were used as predictive input variables for estimating VO_{2max} in the ANN-equation is argued below.

Sex. As it could be expected, there was a significant difference between boys and girls in the measured VO_{2max} value. This finding is also consistent with normative data showing lower levels of VO_{2max} for girls than for boys [38]. However, Léger's equation does not account for sex. Factors explaining the lower VO_{2max} values observed in girls may be related to the fact that girls usually have a lower development of muscular mass and higher fraction of body fat [39]. Moreover, it has been suggested that women may be more prone to pulmonary limitations during heavy exercise (and perhaps submaximal intensities) than men, which is supposedly due to the influence of the reproductive hormones (estrogen and progesterone) in combination with a reduced pulmonary capacity [40]. A greater ventilatory work associated with an increased expiratory flow limitation during the exercise and gas exchange impairments seems to be of primary importance. The influence of sex on VO_{2max} has also been taken into account in others published equations [17,23,24,26]. Stickland et al. [24] developed two sex-specific equations with similar slopes for both men and women aged 18—38 years. They found a slightly lower *Y*-intercept value for women, which is in agreement with our findings. Mahar et al. [23] developed an equation based on a sample consisting of 61 boys and 74 girls aged 12—14 years in which sex, number of laps completed, and body weight were included as independent variables (Table 1).

Age. Léger et al. [12] included age as one of the independent variables in their model, which was not the case in other published equations [23—25]. The age range of the adolescents involved in the present study was similar to the study made by Léger et al. [12]. However, the youngest adolescent in our study was 13 years old, while the youngest person in Léger's study was 8 years old. Findings from cross-sectional and longitudinal studies have shown that age is associated with VO_{2max} in both adolescents and adults [38,41,42].

Adolescence represents a period of life where many changes occur. Therefore, age might be an important factor to control for in order to understand the contribution of those age-dependent factors. It has been suggested that rather than using the chronological age as a measure for this variable, sexual maturation (i.e. biological age) may be a more accurate marker of the physiological status of the person in this particular period of life [43]. However, findings from cross-sectional studies examining the influence of sexual maturation on VO_{2max} have shown that sexual maturation may account for only a small proportion of the variance in the measured VO_{2max} value [44], and that weight and height are primarily responsible for variation in VO_{2max} throughout maturation [45]. Another reason why sexual maturation was not included in the equation was due to the suspected inaccuracy in self-reporting tanner stage in some circumstances, and the need for a paediatrician or trained physician to make an objective measurement, which is in most setting not feasible.

Body size: *weight and height*. The increases in VO_{2max} are influenced by changes in body weight and height. Controlling the effect of changes in body size in growing adolescents is critical in order to understand the relative contributions of other factors influencing changes in VO_{2max}, such as sex, maturation, habitual physical activity, and functional cardiorespiratory improvements. The conventional (ratio) approach for controlling or "normalizing" VO_{2max} for body size has been to divide the VO_{2max} value by kilogram of body weight. However, in walking/running activities, height also has been shown to have an impact on the performance, and specially in those activities incorporating shuttle running such as the 20mSRT [25,46]. Body weight has usually been used as a measure of body size. Yet, it has also been suggested that height could be used when scaling body size to account for possible disproportionate changes in muscle mass with increasing body size [47]. This study shows that both body weight and height are significantly correlated with the 20mSRT performance (Table 3). Body weight was negatively correlated with VO_{2max} ($r = -0.517$, $P < 0.001$; $r = -0.241$, $P < 0.043$ for boys and girls, respectively), while the correlation between height and VO_{2max} was less evident ($r = -0.170$, $P = 0.079$; $r = 0.219$, $P = 0.066$ for boys and girls, respectively). It is worth noting that height is negatively correlated with VO_{2max} for boys, while the opposite is true for girls. Girls had significantly lower height than boys (161.0 cm vs. 172.5 cm, respectively, $P < 0.001$), which may indicate that height has a positive contribution on the 20mSRT performance up to a certain level after which it has a negative impact. It is tenable that various biomechanical complexities of shuttle running may account for this. Other approaches have recommended the use of allometric scaling exponents [48] or accounting for fat-free mass [49] in order to allow for a more appropriate study on the impact of body size differences on VO_{2max}. However, the allometric scaling exponents have not been universally reported [47], and the use of fat-free mass needs either expensive instrumentation or trained evaluators (when derived from anthropometric measurements) which is not often a feasible choice, especially in schools settings.

There are several equations including body size measurements in the model [17,23,26]. The equation developed by Mahar et al. [23] includes both weight and height as single variables and as a ratio [body weight in kg divided by height in m^2 (BMI)] (Table 1). They also developed another equation where only weight is used as a predictive variable in the model. A SEE of 6.38 and 6.35 ml/(kg min) was reported for the first and second equation, respectively. These results are slightly higher than the SEE obtained in the present study by means of both the Léger's and the ANN-equation (4.27 and 2.84 ml/(kg min), respectively). Some aspects of the methods used may explain the observed differences in the SEE values. Mahar et al. [23] used a multiple regression model to predict the measured VO_{2max} from the number of laps completed on the 20mSRT. The following variables were included: sex and body mass or BMI. The dependent variable in the regression model was measured VO_{2max}, which was collected while running until exhaustion on the treadmill. Energy demands during shuttle running have been reported to be higher when compared with treadmill running [25], which can be attributed to the mode of exercise, technique, and musculature employed in the two conditions. This may be another source or error of the equations built with VO_{2max} values collected from treadmill-based protocols [17,23–26]. This may also explain the differences (−2.4 ml/(kg min)) found in the validation sample between the measured VO_{2max} in the treadmill and the estimated VO_{2max} from the ANN-equation.

Stage. In the ANN-equation, the maximal 20mSRT performance attained is calculated from the last half-stage completed, so it allows credit for 30 s when participants fall short of completing a full stage. This increased precision should help in detecting changes in fitness in interventional studies, follow-up studies, in athletes before and after a period of training, etc. The Léger's equation used maximal speed calculated from the last stage completed. Therefore, subjects falling just short of completing a full 1-min stage would be ascribed to the previously completed stage. Consequently, the ANN-equation may allow for a greater sensitivity in the estimation of VO_{2max} when compared with Léger's equation. Stickland et al. [24] also used the last half-stage completed as the measure of the 20mSRT performance to build a prediction model for adults, and it allowed for a higher degree of accuracy when compared with Léger's equation.

Constraints. It is important to acknowledge that the 20mSRT is a test requiring maximal effort. Special attention has to be paid during the course of the test as such, since today there are at least three major variants of the test available. Special attention should also be on the cassette or CDs to be used. Methodological variations in these cassettes (e.g. calling the stage number at the start vs. the finish of each stage; using only full minutes vs. both full minutes and half minutes to indicate completed stages) mean that identical performances are reported in different ways.

The main limitation of the ANN is its complexity and its ''black box'' nature. The complexity of the ANN-equation may become rather inconvenient when applied in the field. However, even when using Léger's equation in a relatively big sample of subjects, a programmable device (spreadsheet) is required. Similarly, the estimation of VO_{2max} using the proposed ANN-equation can be done by means of a spreadsheet.

Some of the advantages of using an ANN-model need special attention: (1) its capability of producing a nonlinear input—output mapping. A neural network computes a function, which maps its inputs variables with its output. A nonlinear relationship could exist between the input and the output variable. However, ANNs are especially suitable for modelling highly nonlinear maps. (2) Its learning ability (adaptivity). A neural network can be trained to perform a specific task, for example, reproducing an unknown input—output mapping. There is always a neural network which will match your input variables as closely as possible with your output for a given set of data. In other words, you can approximate a given input—output map with a network as precise as you need. (3) The ability to generalize. An ANN-model can be set up to be trained to produce a correct output for a given set of input data. The applications from the present investigation would be further increased by performing validation studies in specific populations and in different countries.

5. Conclusions

In this study an ANN-equation to estimate VO_{2max} from 20mSRT performance (stage), sex, age, weight, and height in adolescents aged 13—19 years was developed and cross-validated. The newly developed equation was shown to be more accurate than Léger's equation in the sample of adolescents studied. All variables included in the equation are usually measured in population-based studies, no specific equipment is required to collect the data, and is not time-consuming. The proposed model has been coded in a user-friendly spreadsheet, and can be found at http://www.helenastudy.com/scientific.php (accessed 11 June 2008).

Conflict of interest

None of the authors had any conflict of interest.

Acknowledgements

The present paper is published on behalf of the HELENA (Healthy Lifestyle in Europe by Nutrition in Adolescence) Study Group (http://www.helenastudy.com/list.php, accessed 11 June 2008). The HELENA Study takes place with the financial support of the European Community Sixth RTD Framework Programme (contract FOOD-CT-2005-007034). The content of this article reflects only the authors' views and the European Community is not liable for any use that may be made of the information contained therein. This study was also partially supported by the European Union in the framework of the Public Health Executive Agency (PHEA) DG Sango, Health Information Strand (ALPHA project, ref: 2006120); the Ministerio de Ciencia y Teconología (TIC2003-04650 and TIN2004-07236), FEDER funds (ERGOLAB-project CIT300100-2005-23); Centro Andaluz de Medicina del Deporte, Junta de Andalucía (ref. JA-CTD2005-01); and the Spanish Ministry of Education (EX-2007-1124, AP2004-2745). No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

References

[1] Taylor HL, Buskirk E, Henschel A. Maximal oxygen intake as an objective measure of cardio-respiratory performance. J Appl Physiol 1955;8:73—80.
[2] Ortega FB, Ruiz JR, Castillo MJ, Sjöström M. Physical fitness in childhood and adolescence: a powerful marker of health. Int J Obes (Lond) 2008;32:1—11.
[3] Castillo-Garzon M, Ruiz JR, Ortega FB, Gutierrez-Sainz A. A Mediterranean diet is not enough for health: physical fitness is an important additional contributor to health for the adults of tomorrow. World Rev Nutr Diet 2007;97:114—38.
[4] Mesa JL, Ruiz JR, Ortega FB, Warnberg J, Gonzalez-Lamuno D, Moreno LA, et al. Aerobic physical fitness in relation to blood lipids and fasting glycaemia in adolescents: influence of weight status. Nutr Metab Cardiovasc Dis 2006;16:285—93.
[5] Ruiz JR, Rizzo NS, Hurtig-Wennlöf A, Ortega FB, Warnberg J, Sjöström M. Relations of total physical activity and intensity to fitness and fatness in children: the European Youth Heart Study. Am J Clin Nutr 2006;84:299—303.
[6] Ortega FB, Tresaco B, Ruiz JR, Moreno LA, Martin-Matillas M, Mesa JL, et al. Cardiorespiratory fitness and sedentary activities are associated with adiposity in adolescents. Obesity (Silver Spring) 2007;15:1589—99.
[7] Brage S, Wedderkopp N, Ekelund U, Franks PW, Wareham NJ, Andersen LB, et al. Features of the metabolic syndrome are associated with objectively measured physical activity and fitness in Danish children: the European Youth Heart Study (EYHS). Diab Care 2004;27:2141—8.
[8] Ruiz JR, Ortega FB, Rizzo NS, Villa I, Hurtig-Wennlöf A, Oja L, et al. High cardiovascular fitness is associated with low metabolic risk score in children: the European Youth Heart Study. Pediatr Res 2007;61:350—5.
[9] Ruiz JR, Ortega FB, Loit HM, Veidebaum T, Sjöström M. Body fat is associated with blood pressure in school-aged girls with low cardiorespiratory fitness: the European Youth Heart Study. J Hypertens 2007;25:2027—34.
[10] Ruiz JR, Sola R, Gonzalez-Gross M, Ortega FB, Vicente-Rodriguez G, Garcia-Fuentes M, et al. Cardiovascular fitness is negatively associated with homocysteine levels in female adolescents. Arch Pediatr Adolesc Med 2007;161:166—71.
[11] Reed KE, Warburton DE, Lewanczuk RZ, Haykowsky MJ, Scott JM, Whitney CL, et al. Arterial compliance in young children: the role of aerobic fitness. Eur J Cardiovasc Prev Rehabil 2005;12:492—7.
[12] Léger LA, Lambert A, Goulet A, Rowan C, Dinelle Y. Capacity aerobic des Quebecois de 6 a 17 ans - test Navette de 20 metres avec paliers de 1 min. Can J Appl Sport Sci 1984;9:64—9.
[13] Tomkinson GR, Léger LA, Olds TS, Cazorla G. Secular trends in the performance of children and adolescents (1980—2000): an analysis of 55 studies of the 20 m shuttle run test in 11 countries. Sports Med 2003;33:285—300.
[14] Council of Europe Committee for the Development of Sport. Eurofit. Handbook for the EUROFIT tests of physical fitness. Rome, Italy: Edigraf editoriale grafica Rome; 1988.
[15] The Cooper Institute for Aerobics Research. FITNESSGRAM test administration manual. Champaign, IL: Human Kinetics; 1999.
[16] Léger LA, Mercier D, Gadoury C, Lambert J. The multistage 20 metre shuttle run test for aerobic fitness. J Sports Sci 1988;6:93—101.
[17] Barnett A, Chan LYS, Bruce IC. A preliminary study of the 20-m multistage shuttle run as a predictor of peak VO_2 in Hong Kong Chinese students. Pediatr Exerc Sci 1993;5:42—50.
[18] Van Mechelen W, Hlobil H, Kemper HCG. Validation of two running tests as estimates of maximal aerobic power in children. Eur J Appl Physiol Occup Physiol 1986;503—6.
[19] Anderson GS. The 1600-m run and multistage 20-m shuttle run as predictive tests of aerobic capacity in children. Pediat Exerc Sci 1992;4:312—8.

[20] Liu NY, Plowman SA, Looney MA. The reliability and validity of the 20-meter shuttle test in American students 12 to 15 years old. Res Q Exerc Sport 1992;63:360—5.
[21] Pitetti KH, Fernhall B, Figoni S. Comparing two regression formulas that predict VO_2 peak using the 20-m shuttle run for children and adolescents. Pediat Exerc Sci 2002;125—34.
[22] Suminski RR, Ryan ND, Poston CS, Jackson AS. Measuring aerobic fitness of Hispanic youth 10 to 12 years of age. Int J Sports Med 2004;25:61—7.
[23] Mahar MT, Welk GJ, Rowe DA, Crotts DJ, McIver KL. Development and validation of a regression model to estimate VO_{2peak} from PACER 20-m shuttle run performance. J Phys Act Health 2006;S34—46.
[24] Stickland MK, Petersen SR, Bouffard M. Prediction of maximal aerobic power from the 20-m multi-stage shuttle run test. Can J Appl Physiol 2003;28:272—82.
[25] Flouris AD, Metsios GS, Koutedakis Y. Enhancing the efficacy of the 20 m multistage shuttle run test. Br J Sports Med 2005;39:166—70.
[26] Matsuzaka A, Takahashi Y, Yamazoe M, Kumakura N, Ikeda A, Wilk B, et al. Validity of the multistage 20-m shuttle-run test for Japanese children, adolescents, and adults. Pediat Exerc Sci 2004;16:113—25.
[27] Goodman CS, Ahn R. Methodological approaches of health technology assessment. Int J Med Inform 1999;56:97—105.
[28] McLaughlin JE, King GA, Howley ET, Bassett Jr DR, Ainsworth BE. Validation of the Cosmed K4b^2 portable metabolic system. Int J Sports Med 2001;22:280—4.
[29] Haykin S. Neural networks, a comprehensive foundation. Prentice-Hall; 1999.
[30] Rumelhart DE, Hinton GE, Williams RJ. Learning representations of back-propagation errors. Nature 1985;323:533—6.
[31] Mitchell T. Machine learning. WCB/McGraw-Hill; 1997.
[32] Holm S. A simple sequentially rejective multiple test procedure. Scand J Statist 1979;6:65—70.
[33] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;8:307—10.
[34] Bland JM, Altman DG. Comparing methods of measurement: why plotting difference against standard method is misleading. Lancet 1995;346:1085—7.
[35] Weisstein EW. "P-Value." From MathWorld—a Wolfram web resource. http://mathworld.wolfram.com/P-Value.html (accessed June 11, 2008).
[36] Rothwell PM. Analysis of agreement between measurements of continuous variables: general principles and lessons from studies of imaging of carotid stenosis. J Neurol 2000;247: 825—34.
[37] St Clair Gibson A, Broomhead S, Lambert MI, Hawley JA. Prediction of maximal oxygen uptake from a 20-m shuttle run as measured directly in runners and squash players. J Sports Sci 1998;16:331—5.
[38] Ortega FB, Ruiz JR, Castillo MJ, Moreno LA, Gonzalez-Gross M, Warnberg J, et al. Low level of physical fitness in Spanish adolescents. Relevance for future cardiovascular health (AVENA Study). Rev Esp Cardiol 2005;58:898—909.
[39] Moreno LA, Mesana MI, Gonzalez-Gross M, Gil CM, Fleta J, Warnberg J, et al. Anthropometric body fat composition reference values in Spanish adolescents. The AVENA Study. Eur J Clin Nutr 2006;60:191—6.
[40] Harms CA. Does gender affect pulmonary function and exercise capacity? Respir Physiol Neurobiol 2006;151:124—31.
[41] Hasselstrom H, Hansen SE, Froberg K, Andersen LB. Physical fitness and physical activity during adolescence as predictors of cardiovascular disease risk in young adulthood. Danish Youth and Sports Study. An eight-year follow-up study. Int J Sports Med 2002;23:S27—31.
[42] Bradshaw DI, George JD, Hyde A, LaMonte MJ, Vehrs PR, Hager RL, et al. An accurate VO_{2max} nonexercise regression model for 18—65-year-old adults. Res Q Exerc Sport 2005;76:426—32.
[43] Preece MA. Evaluation of growth and development. In: Avner ED, Barratt TM, Holliday MA, editors. Pediatric nephrology. 3rd ed., Philadelphia: Williams & Wilkins; 1994
[44] Mota J, Guerra S, Leandro C, Pinto A, Ribeiro JC, Duarte JA. Association of maturation, sex, and body fat in cardiorespiratory fitness. Am J Hum Biol 2002;14:707—12.
[45] Jones MA, Hitchen PJ, Stratton G. The importance of considering biological maturity when assessing physical fitness measures in girls and boys aged 10 to 16 years. Ann Hum Biol 2000;27:57—65.
[46] McCann DJ, Adams WC. The size-independent oxygen cost of running. Med Sci Sports Exerc 2003;35:1049—56.
[47] Welsman JR, Armstrong N, Nevill AM, Winter EM, Kirby BJ. Scaling peak VO_2 for differences in body size. Med Sci Sports Exerc 1996;28:259—65.
[48] Vanderburgh PM, Mahar MT, Chou CH. Allometric scaling of grip strength by body mass in college-age men and women. Res Q Exerc Sport 1995;66:80—4.
[49] Neder JA, Lerario MC, Castro ML, Sachs A, Nery LE. Peak VO_2 correction for fat-free mass estimated by anthropometry and DEXA. Med Sci Sports Exerc 2001;33:1968—75.

Please cite this article in press as: Ruiz JR, et al., Artificial neural network-based equation for estimating VO_{2max} from the 20 m shuttle run test in adolescents, Artif Intell Med (2008), doi:10.1016/j.artmed.2008.06.004