Journal of Science and Medicine in Sport xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Journal of Science and Medicine in Sport

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

Review

A systematic review of methods to predict maximal oxygen uptake from submaximal, open circuit spirometry in healthy adults

Harrison J.L. Evans[a,*], Katia E. Ferrar[a], Ashleigh E. Smith[a], Gaynor Parfitt[a], Roger G. Eston[a,b]

[a] Exercise for Health and Human Performance Group, Sansom Institute for Health Research, School of Health Sciences, University of South Australia, GPO Box 2471, Adelaide, Australia
[b] Sport and Health Sciences, College of Life and Environmental Sciences, University of Exeter, Exeter, UK

ARTICLE INFO

Article history:
Received 20 December 2013
Received in revised form 27 February 2014
Accepted 7 March 2014
Available online xxx

Keywords:
Exercise tolerance
Exercise test
Oxygen consumption
Regression analysis
Predictive value of tests

ABSTRACT

Objectives: This systematic review aimed to (i) report the accuracy of submaximal exercise-based predictive equations that incorporate oxygen uptake (measured via open circuit spirometry) to predict maximal oxygen uptake ($\dot{V}O_{2\,max}$) and (ii) provide a critical reflection of the data to inform health professionals and researchers when selecting a prediction equation.
Design: Systematic review.
Methods: A systematic search of MEDLINE, EMBASE (via OvidSP), CINAHL, SPORTDiscus™ (via EBSCO Host) and Scopus databases was undertaken in February 2013. Studies were required to report data on healthy participants aged 18–65 y. Following tabulation of extracted data, a narrative synthesis was conducted.
Results: From a total of 7597 articles screened, 19 studies were included, from which a total of 43 prediction equations were extracted. No significant difference was reported between the measured and predicted $\dot{V}O_{2\,max}$ in 28 equations. Pearson's correlation coefficient between the predicted and measured $\dot{V}O_{2\,max}$ ranged from $r = 0.92$ to $r = 0.57$. The variables most commonly used in predictive equations were heart rate ($n = 19$) and rating of perceived exertion ($n = 24$).
Conclusions: Overall, submaximal exercise-based equations using open circuit spirometry to predict $\dot{V}O_{2\,max}$ are moderately to highly accurate. The heart rate and rating of perceived exertion methods of predicting $\dot{V}O_{2\,max}$ were of similar accuracy. Important factors to consider when selecting a predictive equation include: the level of exertion required; participant medical conditions or medications; the validation population; mode of ergometry; time and resources available for familiarisation trials; and the level of bias of the study from which equations are derived.

1. Introduction

Maximal oxygen uptake ($\dot{V}O_{2\,max}$), defined as the highest rate at which the body can transport and utilise oxygen during exercise,[1] is widely accepted as a criterion measure of cardiorespiratory fitness.[2] However, direct assessment of the $\dot{V}O_{2\,max}$ requires exercise to volitional exhaustion, which is not advisable for individuals who may be limited by pain, fatigue, abnormal gait or impaired balance.[3] Maximal effort tests also significantly increase the likelihood of adverse cardiac events in elderly and cardiac patients.[4] Furthermore, they require a very high level of motivation.[5,6] For these reasons a wide variety of submaximal testing protocols have been developed to reduce risk, testing time, costs and the reliance on participant motivation associated with more strenuous, exhaustive exercise.[3] It is evident that consideration of the exercise setting, modality and test protocol, in respect to the intended population, are important when selecting an appropriate prediction equation.[7]

The publication of equations that incorporate oxygen uptake ($\dot{V}O_2$) measured via open circuit spirometry during submaximal exercise to predict the $\dot{V}O_{2\,max}$, have increased in the last 10 y. This is largely a result of technological improvements, whereby computerised gas-analysis systems simplify the process of gathering and storing complex oxygen uptake data. However, there is little consensus within the literature regarding the accuracy of these equations. Additionally, there is a lack of guidance on the factors to be considered when selecting an equation. To our knowledge, submaximal, exercise-based prediction equations that incorporate direct measures of $\dot{V}O_2$ have yet to be systematically collated and

* Corresponding author. Tel.: +61 8 830 22910.
E-mail address: Harrison.Evans@mymail.unisa.edu.au (H.J.L. Evans).

http://dx.doi.org/10.1016/j.jsams.2014.03.006
1440-2440/

reviewed. Therefore, the primary aim of this systematic review is to collate and report on the accuracy (criterion validity) of these equations. A critical reflection of the data extracted in this review will also be used to help inform researchers and health professionals of the important factors, in addition to accuracy, that should be considered when selecting an equation to predict $\dot{V}O_{2\,max}$.

2. Methods

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses[8] (PRISMA) guidelines were used as a reporting structure for this systematic review. The inclusion criteria and analysis methods were designed a priori and documented in a PROSPERO protocol (CRD42013004182). With agreement from all authors, the protocol was amended to focus on adults (aged 18–65 y) and prediction equations that incorporated a direct measure of $\dot{V}O_2$ (measured via open circuit spirometry).

To be considered eligible, the original article must have reported primary data and be published in English, in a peer-reviewed journal. Studies were required to report data on participants between the ages of 18–65 y that were apparently healthy (asymptomatic of disease and free from acute or chronic injury). A submaximal, exercise-based method must have been used to predict $\dot{V}O_{2\,max}$. The prediction equation must have incorporated a direct measure of $\dot{V}O_2$ (measured via open circuit spirometry) and should be clearly reported or the method replicable. The actual $\dot{V}O_{2\,max}$ also had to be directly measured using open circuit spirometry. Studies were required to be non-interventional (e.g. observational, cross sectional). However, intervention studies were included if baseline data could be extracted. Studies were included if they reported one validity statistic (Pearson's correlation coefficient, intraclass correlation coefficient, Bland and Altman's[9] Limits of Agreement) and either the measured and predicted $\dot{V}O_{2\,max}$ values or a directional significant difference between the measured and predicted $\dot{V}O_{2\,max}$ values.

A systematic search of MEDLINE, EMBASE (via OvidSP), CINAHL, SPORTDiscus™ (via EBSCO Host) and Scopus databases was undertaken from the respective inception of each database through to February 2013 (HE and KF). Two experts in the field were contacted to identify any additional studies that were not captured from the electronic search and checking of reference lists of the included studies.

The search terms were developed by two authors (HE and KF) and reviewed by two experts on the research team (RE and GP). The electronic database searches were conducted by one author (HE). The search terms reflect the key word groupings of data collection methods (e.g. test, protocol, ergometry etc.), the outcome measure (e.g. $\dot{V}O_{2\,max}$, $\dot{V}O_{2\,peak}$ etc.) and the indirect calculation of the outcome measure (estimate and predict). Search strategies for all databases are available from the corresponding author (HE).

The title and abstract of each study were first screened for eligibility. Full-text manuscripts were then screened along with any additional studies identified from the checking of reference lists. Screening was conducted by two authors (HE and KF) and disagreements were discussed between three authors (HE, KF, AS) and resolved by consensus.

A data extraction sheet, pilot tested on ten studies, was developed and refined by the research team. Two authors (HE and KF) independently extracted data for half of the studies each. Two authors (KF and AS) then independently checked the extracted data for errors.

The data items extracted were: authors, year, title, sample size, gender split, age range, mean and standard deviation of the age, country of data collection, year of data collection, name of predictive method (if relevant), mode of ergometry, brief description of the exercise protocol, variables entered in to the prediction equation, criterion validity statistics, mean measured $\dot{V}O_{2\,max}$, mean predicted $\dot{V}O_{2\,max}$ and the direction of significant difference. Three authors (HE, KF, AS) resolved disagreements through discussion.

The Quality Assessment Tool for Quantitative Studies[10] was used to appraise the methodological rigour of the studies included in this review. This standardised critical appraisal tool was modified to suit the design of this systematic review. Specifically, three components of the tool that were relevant to this review are: (i) selection bias, (ii) data collection methods and (iii) withdrawals and dropouts. Two authors (KF and AS) independently appraised half of the included studies each and one author (HE) appraised all of the included studies against each of the three components. For each study, the components were individually rated as 'strong', 'moderate' or 'weak', based on the standard criteria.[11] A global rating for each study was then obtained based on the total number of weak ratings that were accumulated (two or more weak ratings = 'weak', one weak rating = 'moderate', zero weak ratings = 'strong'). No studies were excluded on the basis of risk of bias.

The primary outcome measures were the reported validity statistics between the measured and predicted $\dot{V}O_{2\,max}$ (Pearson's correlation coefficient (r), Intraclass Correlation Coefficient (ICC) and the bias of the Limits of Agreement [LoA]) and the differences between the two measures (either a statistical significance or directional trend).

Data were extracted for all unique prediction equations reported within the same study. If the same prediction equation was used across consecutive trials, data from the most accurate trial were extracted. If the same prediction equation was used with different ranges of data from the same trial, data from the most accurate range were extracted. For example, if low-intensity data and high-intensity data from the same exercise trial were separately entered in to the same equation, the data range with the strongest validity for predicting $\dot{V}O_{2\,max}$ were extracted. If the same prediction equation was used with different ranges of data and multiple trials, data from the most accurate data range by trial combination were extracted. Where possible, data were extracted from the entire sample of each study. When studies failed to report data on the entire sample, data for particular subgroups (i.e. males and females or active and sedentary) were extracted.

A narrative synthesis of results was conducted. Studies were ranked in order of accuracy. Studies were first ranked and grouped according to the significant difference between the measured and predicted $\dot{V}O_{2\,max}$ values, such that the highest ranked were those that reported no significant difference, followed by those which failed to conduct (or report) statistical analysis to determine a significant difference, followed by those that reported a significant difference. Within those three groups, the second level of ranking was based on the strength of the reported validity statistic, either r, ICC or LoA. Studies were ranked on the basis of only one validity statistic. Equations were ranked by r in the first instance. When no r was reported, equations were ranked in order of the highest ICC. When neither r nor ICC were reported, equations were ranked in order of the bias closest to zero of the LoA. If multiple equations were extracted for a single study, the study was ranked by the most accurate equation. The next stage of the narrative synthesis included a critical reflection of the data in an attempt to inform the reader of the important considerations that should be made when selecting an equation to predict $\dot{V}O_{2\,max}$.

3. Results

A total of 7597 studies were identified from preliminary searching. From the titles and abstracts 7397 obviously irrelevant studies were excluded. From the 200 full-text manuscripts reviewed, a total

Please cite this article in press as: Evans HJL, et al. A systematic review of methods to predict maximal oxygen uptake from submaximal, open circuit spirometry in healthy adults. *J Sci Med Sport* (2014), http://dx.doi.org/10.1016/j.jsams.2014.03.006