ORIGINAL ARTICLE
EXERCISE

Peak oxygen uptake reference values for cycle ergometry for the healthy Dutch population: data from the LowLands Fitness Registry

Caspar Frederik Mylius[1,2], Wilhelmus Petrus Krijnen[1,3], Cornelis Peter van der Schans[1,4,5] and Tim Takken[6]

Affiliations: [1]Hanze University of Applied Sciences, Research Group Healthy Ageing, Allied Health Care and Nursing, Groningen, The Netherlands. [2]Centre of Expertise Primary Care Groningen (ECEZG), Groningen, The Netherlands. [3]Johann Bernoulli Institute for Mathematics and Computer Science, University of Groningen, Groningen, The Netherlands. [4]University of Groningen, University Medical Center Groningen, Dept of Rehabilitation Medicine, Groningen, The Netherlands. [5]University of Groningen, University Medical Center Groningen, Health Psychology Research, Groningen, The Netherlands. [6]University Medical Center Utrecht, Utrecht, The Netherlands.

Correspondence: Caspar Frederik Mylius, Petrus Driessenstraat 3, 9714 CA Groningen, The Netherlands. E-mail: c.f.mijlius@pl.hanze.nl

ABSTRACT Peak oxygen uptake ($V'O_2peak$) is recognised as the best expression of aerobic fitness. Therefore, it is essential that $V'O_2peak$ reference values are accurate for interpreting a cardiopulmonary exercise test (CPET). These values are country specific and influenced by underlying biological ageing processes. They are normally stratified per paediatric and adult population, resulting in a discontinuity at the transition point between prediction equations. There are currently no age-related reference values available for the lifespan of individuals in the Dutch population. The aim of this study is to determine the best-fitting regression model for $V'O_2peak$ in the healthy Dutch paediatric and adult populations in relation to age.

In this retrospective study, CPET cycle ergometry results of 4477 subjects without reported somatic diseases were included (907 females, age 7.9–65.0 years). Generalised additive models were employed to determine the best-fitting regression model. Cross-validation was performed against an independent dataset consisting of 3518 subjects (170 females, age 6.8–59.0 years).

An additive model was the best fitting with the largest predictive accuracy in both the primary (adjusted R^2=0.57, standard error of the estimate (SEE)=556.50 mL·min^{-1}) and cross-validation (adjusted R^2=0.57, SEE=473.15 mL·min^{-1}) dataset.

This study provides a robust additive regression model for $V'O_2peak$ in the Dutch population.

@ERSpublications
Peak oxygen uptake has a nonlinear dependence on years of age in the paediatric and adult Dutch populations http://ow.ly/H3fH30nIjRy

Cite this article as: Mylius CF, Krijnen WP, van der Schans CP, *et al.* Peak oxygen uptake reference values for cycle ergometry for the healthy Dutch population: data from the LowLands Fitness Registry. *ERJ Open Res* 2019; 5: 00056-2018 [https://doi.org/10.1183/23120541.00056-2018].

This article has supplementary material available from openres.ersjournals.com

Received: April 16 2018 | Accepted after revision: Feb 06 2019

Introduction

Peak oxygen uptake ($V'O_{2peak}$) represents the functional limit of the body's ability to deliver and extract oxygen in muscles in order to satisfy the metabolic demands of vigorous exercise; it is recognised as the best expression of aerobic fitness [1]. $V'O_{2peak}$ is increasingly utilised to optimise risk stratification and to facilitate clinical decision making because it reflects therapeutic response and predicts adverse events such as post-operative complications and mortality after abdominal and thoracic surgery [2–5].

For the interpretation of a cardiopulmonary exercise test (CPET), it is essential to have accurate $V'O_{2peak}$ reference values corresponding with the ergometer used [7, 8]. These values are region or country specific, and change over time due to cultural differences and evolving population characteristics [8, 9]. Therefore, each country must have specific updated $V'O_{2peak}$ reference values that optimally reflect the characteristics of the current population tested, the equipment and the methodology utilised [8–10]. Although multiple countries provided up-to-date $V'O_{2peak}$ reference values derived from large cohorts exceeding 4000 participants [11–13], $V'O_{2peak}$ reference values from 1985 are the most commonly used in clinical settings in the Netherlands as there are none available derived for the Dutch adult population [14]. These commonly used $V'O_{2peak}$ reference value prediction equations were obtained from a relatively small sample of 100 participants from the North American population [14].

$V'O_{2peak}$ is highly influenced by underlying biological ageing processes such as physical development, pubertal status, age-induced neuromuscular deterioration, sarcopenia and cardiopulmonary decline [7, 15, 16]. It has been hypothesised in both the paediatric and adult populations that $V'O_{2peak}$ develops in a nonlinear and interrelated manner with the progression of age [7, 17–20]. Linear regression models are predominantly used to determine $V'O_{2peak}$ reference value prediction equations depending upon sex, age, height, weight and physical activity levels [9, 14, 16, 21].

The frequently used age stratification between the paediatric and adult populations is somewhat arbitrary, and it introduces a discontinuity at the transition point between the two equations, which leads to a reference value shift from the paediatric to the adult population. Additionally, such an age stratification implies more prediction uncertainty as accuracy is lowest at the boundary of the sample age scale. Estimation of an up-to-date general prediction model across the paediatric as well as the adult population would facilitate a smooth transition into adult care. Therefore, the aim of this study is to determine the best-fitting regression model for $V'O_{2peak}$ in the healthy Dutch paediatric and adult populations in relation to age.

Methods

This retrospective multicentre study was conducted using the LowLands Fitness Registry [6], a primary dataset of 8900 subjects from 11 healthcare centres in the Netherlands that was aggregated with the aim of establishing CPET reference values for the Dutch population. Additionally, to determine the external and predictive validity of the reference value prediction model, a cross-validation procedure was performed on an independent sample as recommended by the American Thoracic Society/American College of Chest Physicians (ATS/ACCP) [8]. Specifically, the cross-validation in this study was performed against an additional dataset obtained from the Diving Medical Center (Den Helder, The Netherlands) and the Wilhelmina Children's Hospital (Utrecht, The Netherlands). The cross-validation dataset contained 4536 subjects that were not included in the primary dataset. Both datasets contain incremental CPET measurements collected between January 2010 and December 2016. Institutes that were included satisfied the following criteria: 1) to meet the ATS/ACCP statement equipment requirements to perform an incremental CPET using an electromagnetically braked cycle ergometry test utilising gas exchange analysis by bag collection, mixing chamber or breath-by-breath analysis based upon averaging the values measured during the last 30–60 s of the test [8] and 2) to perform equipment quality control in accordance with the ATS/ACCP statement [8].

Subjects included in both datasets underwent an individualised incremental CPET cycle ergometry test for multiple reasons, including: initiated by a healthcare professional, work- and sports-related (mandatory) annual health checks, participation in scientific studies or based on personal motivation (*e.g.* an exercise response evaluation for the aid of a training scheme). Every institute provided anonymised, coded patient information to the data coordinator at the University Medical Center Utrecht (Utrecht, The Netherlands). All records were previously screened for measurement failures. If there were any uncertainties, the testing institute was contacted to ensure the communication of correct data. It has been confirmed by the medical ethical research committee of the University Medical Center Utrecht that the Dutch Medical Research Involving Human Subjects (WMO) Act does not apply to the current study.

Study sample

All of the subjects included in the study were Dutch residents, aged ⩽65 years. The status "healthy" was defined as the absence of any reported somatic signs of disease and the exclusion of registered available

risk factors [22]. Therefore, subjects were excluded if they reported somatic diseases at the time of testing or showed ECG irregularities prior to testing. Additionally, subjects were excluded from further analysis if they included a missing predictor or outcome values. To ensure subjects reached their maximal measurement (*i.e.* $V'O_2peak$), subjects were excluded if they did not reach a respiratory exchange peak ratio of at least 1.0 [23] or did not reach a minimum of 85% of the age-predicted maximum heart rate (beats·min^{-1}) determined as 208−(0.7×age) [24]. Furthermore, due to the abnormal working capacity and cardiovascular responses to exercise in underweight patients and the recognition of obesity as a disease by the World Health Organization, subjects who had a body mass index (BMI) ⩾30 kg·m^{-2} [25] or, in adult subjects, ⩽18.5 kg·m^{-2} [26] were excluded. Due to the decrease in $V'O_2peak$ associated with smoking, subjects who actively smoked at the time of the test were excluded [27]. Lastly, professional athletes were excluded because they were considered as not representative for the average Dutch population due to the positive effects of exercise training on $V'O_2peak$ [28]. The exclusion criteria were applied in both the primary and cross-validation datasets.

Statistical analyses

Statistical analyses were performed using R version 3.4.4 [29]. A p-value ⩽0.05 was considered significant. Continuous data were summarised as mean with standard deviation and categorical data as frequencies (percentage). The variables sex, age, weight and height were included in the analyses as these are commonly used as a basis for $V'O_2peak$ reference value prediction equations [9].

Generalised additive models (GAMs) were utilised to semiparametrically find the most appropriate fitting regression model [30, 31]. To determine the model best fitting the data, criteria such as the adjusted R^2, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were used [32, 33]. A higher adjusted R^2 and a lower AIC and BIC were considered as improving the fit. In cases of inconsistency between these, the BIC criterion was taken as the most decisive. The interpretation of the BIC score was: 0–2 as "minimal" improvement, 2–6 as "positive" improvement, 6–10 as "strong" improvement and >10 as "very strong" improvement [34].

All models fitted to the data included an age by sex interaction term to account for the different $V'O_2peak$ levels between male and female subjects [11]. In order to compare with a best-performing polynomial regression model, each predictive variable was modelled using linear, quadratic and cubic effects by stepwise minimum BIC procedures [35]. The resulting model was employed to represent the polynomial model type in the model fitting procedures. Additionally, based upon the hypothesised nonlinear age dynamics for $V'O_2peak$, an additive model with a smooth spline type of transformation for age was included [17–20, 31].

To determine the fit of the models in the separate paediatric and adult populations, the predictive accuracy of the models was measured using stratified age groups by comparing the residual standard error of the estimate (SEE). The groups were stratified by ⩽20 and >20 years of age. The better the predictive fit of either of the three types of models, the less variability there is and the smaller the standard error of the estimate [36].

Models are of little clinical value unless these have predictive accuracy for independent samples. A cross-validation procedure was performed using each identified model per type (linear, polynomial and GAM) against a cross-validation dataset. Similar to criteria for the primary analysis, the model performance was evaluated by a larger adjusted R^2 and a smaller standard error of the estimate.

For the purpose of illustration, examples of $V'O_2peak$ predictions are reported using the best-performing regression model. For these examples, cases with an increase of 5 years per paediatric case and 10 years per adult case are used; corresponding average height and weight were used determined by data provided by Statistics Netherlands (www.cbs.nl). The 2.5th, 5th, 10th, 25th, 50th, 75th, 90th, 95th and 97.5th prediction percentile intervals are reported.

Results

The complete registry consisted of 8900 cases (1641 females); after applying the exclusion criteria for missing values (n=2674), nonmaximal tests (n=480), BMI ⩾30 kg·m^{-2} or, in adults, ⩽18.5 kg·m^{-2} (n=324), smokers (n=881) and professional athletes (n=64), a sample of 4477 cases labelled as "healthy" remained (907 females) with age ranging from 7.9 to 65.0 years. The cross-validation sample contained 4536 subjects; after applying the exclusion criteria for missing values (n=0), nonmaximal tests (n=64), BMI ⩾30 kg·m^{-2} or, in adults, ⩽18.5 kg·m^{-2} (n=260), smoking (n=694) and professional athletes (n=0), a sample of 3518 subjects (170 females) with an age range from 6.8 to 59.0 years remained. Table 1 summarises the characteristics of both samples. Figure 1 shows the age distribution of the primary sample.

TABLE 1 Sample characteristics

	Primary sample					Cross-validation sample				
	Subjects	Age years per decimal	Weight kg	Height cm	BMI kg·m^{-2}	Subjects	Age years per decimal	Weight kg	Height cm	BMI kg·m^{-2}
Female	907	32.2±12.8	64.3±11.8	168.6±9.3	22.5±3.0	170	23.0±9.8	61.5±14.5	166.6±12.5	21.8±3.4
Male	3570	34.6±11.5	81.6±11.6	181.7±8.1	24.6±2.6	3348	33.9±10.0	84.3±11.6	182.7±8.2	25.1±2.4
All	4477	34.1±11.8	78.1±13.6	179.1±9.9	24.2±2.9	3518	33.4±10.2	83.2±12.7	182.0±9.1	25.0±2.6

Data are presented as n or mean±SD.

The best-performing polynomial regression model that was found *via* stepwise minimum BIC was: $V'O_{2peak}$ (mL·min^{-1})=−1469+(673.00×sex)+(16.87×age)+(−0.47×age^2)+(0.07×height^2)+(39.70×weight)+(−0.16×weight^2) (adjusted R^2=0.56, AIC=69 480.15, BIC=69 531.40), where male=1 and female=0, age is in years, height is in centimetres, and weight is in kilograms.

Table 2 summarises various estimated models and their fit measures. The best-fitting model to the dataset was the additive model that includes a smooth spline transformation for age and an interaction term between age and sex plus linear terms for weight and height. The fit of the model yields an adjusted R^2=0.57, AIC=69 342.81 and BIC=69 449.50. This additive model demonstrates "very strong" improvements [34] compared with both the linear model (BIC difference=170.34) and the polynomial model (BIC difference=81.9). The age-dependent transformations of $V'O_{2peak}$ are shown in figure 2. Additionally, the linear dependencies of weight and length are shown in figures 3 and 4.

The fit of the models compared in the combined and separate paediatric and adult populations is shown in table 3. The additive model provides the largest predictive accuracy overall with an adjusted R^2=0.57 and SEE=556.50 mL·min^{-1} in the entire primary sample; the polynomial and the additive models performed equally against the cross-validation sample, specifically R^2=0.57 compared with the linear model with R^2=0.55. Additionally, the additive model also provided the smallest standard error of the estimate in the stratified age groups in both samples, *i.e.* SEE=495.18 and 420.72 mL·min^{-1} in those ⩽20 years old and SEE=563.82 and 476.92 mL·min^{-1} in those >20 years old. The largest improvement between models in both samples occurred in the ⩽20-year-old age group. In this age group, the additive model has a better fit than both the linear and polynomial models, with an equal adjusted R^2 difference=0.05. Similar improvements are discerned in the standard error of the estimate between the additive and the linear and polynomial models of SEE=65.47 and 53.62 mL·min^{-1} in the primary sample and SEE=108.14 and 35.81 mL·min^{-1} in the cross-validation sample, respectively.

Reference values with corresponding prediction intervals are constructed using average weight and height per sex and age provided by Statistics Netherlands. Table 4 shows the predictions for the female cases and table 5 shows the predictions for the male cases. In both sexes, the 2.5th and 97.5th prediction interval in the 60-year-old cases is the largest: 352 mL·min^{-1} for the female cases and 213 mL·min^{-1} for the male cases. The 2.5th and 97.5th prediction interval of the 30-year-old cases is the smallest: 131 and 78 mL·min^{-1}, respectively. Both sexes have increasing $V'O_{2peak}$ prediction until the age of 20 years followed by a decline.

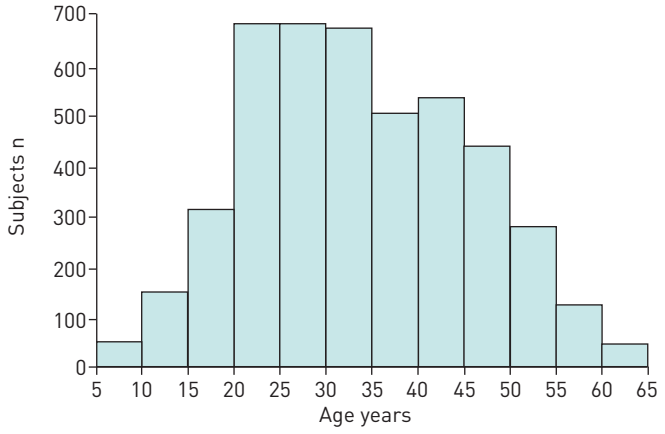

FIGURE 1 Age distribution of the primary sample.

TABLE 2 Fitting comparison by regression model type

	Estimate	Standard error	T-value	p-value	Adjusted R^2	AIC	BIC
Linear model							
Intercept	−3039.01	206.02	−14.75	<0.001	0.55	69 581.40	69 619.84
Sex	634.32	25.75	24.63	<0.001			
Age	−16.50	0.79	−20.66	<0.001			
Height	29.22	1.46	19.95	<0.001			
Weight	16.17	1.11	14.48	<0.001			
Polynomial model							
Intercept	−1469.00	158.80	−9.25	<0.001	0.56	69 480.15	69 531.40
Sex	673.00	25.89	25.99	<0.001			
Age	16.87	4.81	3.50	<0.001			
Age^2	−0.47	0.06	−7.31	<0.001			
$Height^2$	0.07	<0.01	16.52	<0.001			
Weight	39.70	5.17	7.67	<0.001			
$Weight^2$	−0.16	0.03	−5.05	<0.001			
Additive model							
Intercept	−2537.29	224.98	−11.28	<0.001	0.57	69 342.81	69 449.50
Sex	743.35	26.30	28.26	<0.001			
Height	24.30	1.52	15.91	<0.001			
Weight	12.57	1.12	11.21	<0.001			
S(age): male	4.263[#]	5.260[¶]	22.59[+]	<0.001			
S(age): female	7.391[#]	8.288[¶]	70.38[+]	<0.001			

AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion; S(age): spline function for age per sex. Sex: 0=female, 1=male; age=years; height=centimetres; weight=kilograms. [#]: effective degrees of freedom; [¶]: reference number of degrees of freedom; [+]: F-value.

Discussion

The aim of this study was to determine reference values for $V'O_{2peak}$ based upon an optimal regression model in healthy Dutch paediatric and adult populations. Based on adjusted R^2, AIC, BIC and standard error of the estimate, the additive model was the best fitting with the largest predictive accuracy. From the model, it can be concluded that $V'O_{2peak}$ is sex specific and depends nonlinearly on years of age.

We determined that the additive model results in a smaller standard error of the estimate especially in the ⩽20-year-old subjects because, in contrast to the linear model, the additive model is able to adjust for age-related transformations such as the increase in $V'O_{2peak}$ associated with the growth-related weight and

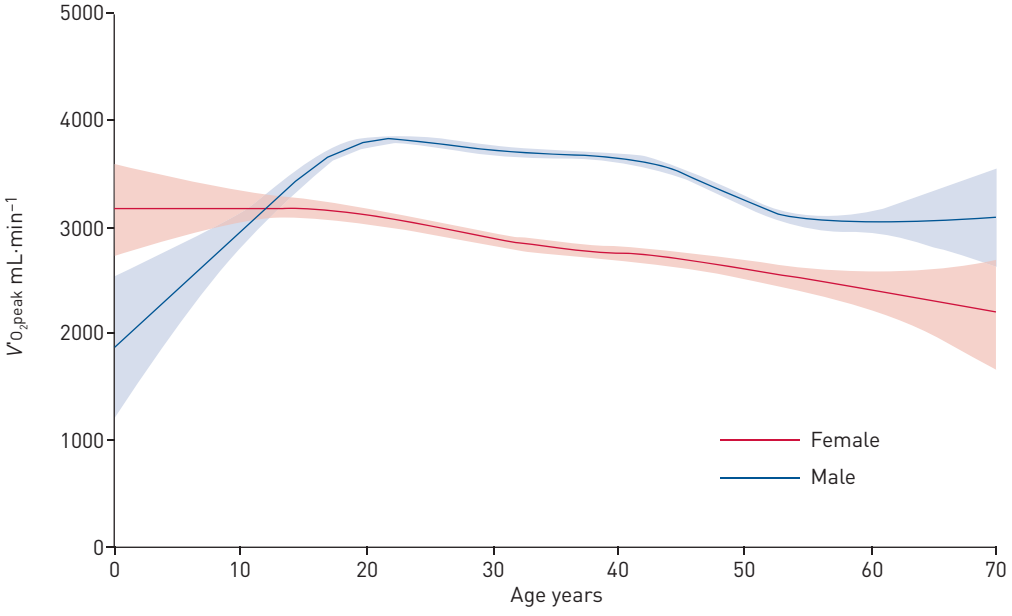

FIGURE 2 Age-dependent transformation of mean peak oxygen uptake ($V'O_{2peak}$). Shading represents the pointwise 95% confidence interval.

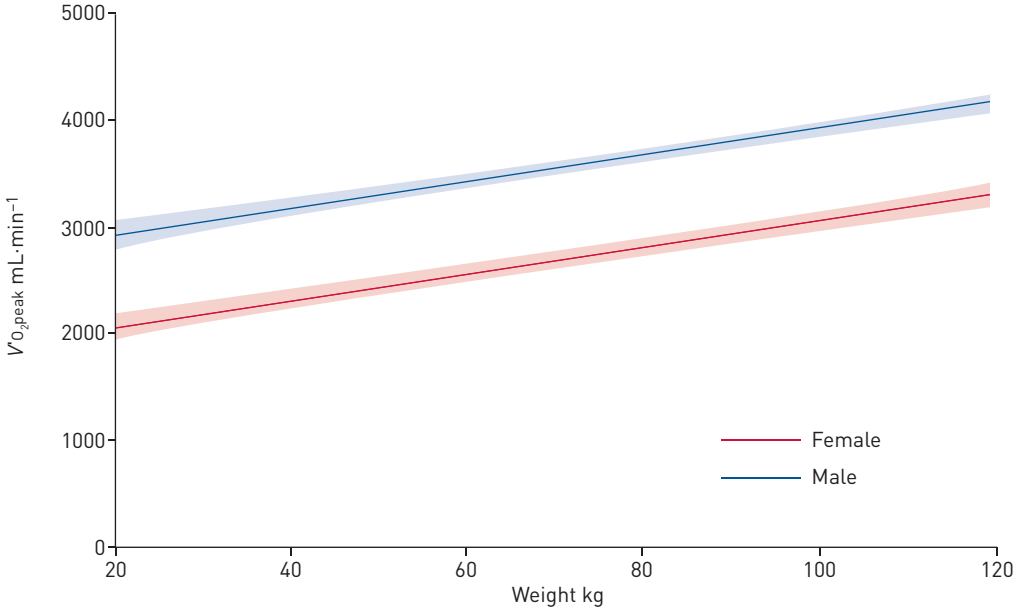

FIGURE 3 Relationship between mean peak oxygen uptake ($V'O_2$peak) and weight. Shading represents the pointwise 95% confidence interval.

height gain during childhood and adolescence. The increase in skeletal muscle mass during this life phase accounts for the majority of weight gained [37]. As skeletal muscle mass is responsible for the majority of oxygen utilised during exercise, the increase in skeletal muscle mass associated with increasing age in ⩽20-year-old subjects partially explains the increase in $V'O_2$peak during this life phase [38]. During adulthood, the increase in skeletal muscle mass and height are limited. $V'O_2$peak decreases during adulthood because of a decrease in muscle mass and a loss of chronotropic competence [24, 39].

Our additive regression model differs from previously utilised linear and polynomial regression models [9, 40, 41]. The use of the advanced statistical analysis method, GAM, in the current study makes it possible to determine the best-fitting regression model for the combined paediatric and adult populations. This method fits the data through cubic-type splines with the degree of smoothness determined by generalised

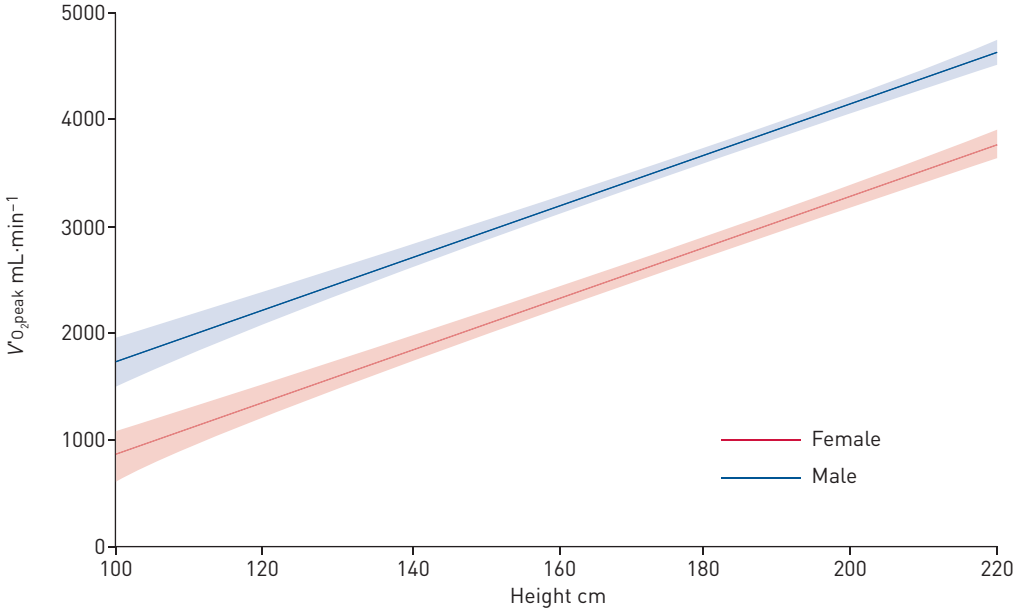

FIGURE 4 Relationship between mean peak oxygen uptake ($V'O_2$peak) and height. Shading represents the pointwise 95% confidence interval.

TABLE 3 Fit of model type per age group and sample set

Model and age group	Primary sample		Cross-validation sample	
	Adjusted R^2	SEE mL·min^{-1}	Adjusted R^2	SEE mL·min^{-1}
Linear model				
All	0.55	572.89	0.54	487.84
⩽20 years	0.71	560.65	0.81	528.86
>20 years	0.50	574.43	0.37	484.56
Polynomial model				
All	0.56	566.20	0.57	476.72
⩽20 years	0.71	548.80	0.82	456.03
>20 years	0.51	568.37	0.38	478.26
Additive model				
All	0.57	556.50	0.57	473.15
⩽20 years	0.76	495.18	0.84	420.72
>20 years	0.52	563.82	0.38	476.92

SEE: standard error of the estimate.

cross-validation, which facilitates combining the previously hypothesised nonlinear and interrelated fashion of more than one independent variable in the paediatric and adolescent populations, and the curvilinear decline with age in the adult population [7, 17–20]. This method results in an improved fit across the entire population [17, 18, 20]. Therefore, prediction of $V'O_{2peak}$ in the transition group between adolescents and adulthood is more precise when using the additive model.

In comparison with the prediction models currently utilised in the Dutch clinical settings, the additive model improves the fit in both the adult and paediatric populations. The linear prediction model for adults provided by Jones *et al.* [14] yields R^2=0.41 to the primary sample and R^2=0.33 to the cross-validation sample compared with R^2=0.52 and 0.38, respectively, in the additive model. The linear prediction equation provided by Ten Harkel *et al.* [41] is most frequently used in the Dutch paediatric population, this equation yields R^2=0.58 and 0.73 compared with R^2=0.76 and 0.84 in the primary and

TABLE 4 Additive model peak oxygen uptake ($V'O_{2peak}$) prediction percentiles: female cases

Age decile years	Height cm	Weight kg	$V'O_{2peak}$ mL·min^{-1} prediction percentile								
			2.5th	5th	10th	25th	50th	75th	90th	95th	97.5th
10	143.0	33.5	1581	1601	1624	1662	1704	1746	1784	1807	1826
15	164.0	52.0	2345	2359	2375	2401	2429	2458	2484	2500	2513
20	168.8	63.2	2543	2556	2570	2593	2619	2645	2668	2682	2694
30	169.3	68.5	2415	2426	2438	2458	2481	2503	2523	2536	2546
40	169.3	70.3	2298	2309	2322	2343	2367	2391	2412	2425	2436
50	167.7	70.5	2089	2104	2120	2148	2180	2211	2239	2255	2270
60	166.6	71.6	1793	1821	1854	1908	1969	2029	2084	2117	2145

TABLE 5 Additive model peak oxygen uptake ($V'O_{2peak}$) prediction percentiles: male cases

Age decile years	Height cm	Weight kg	$V'O_{2peak}$ mL·min^{-1} prediction percentile								
			2.5th	5th	10th	25th	50th	75th	90th	95th	97.5th
10	143.0	34.0	1329	1351	1377	1421	1469	1518	1561	1587	1610
15	168.0	53.0	2775	2788	2804	2831	2860	2889	2916	2932	2945
20	183.5	78.1	3808	3816	3825	3841	3858	3875	3891	3900	3908
30	183.7	83.3	3818	3825	3832	3844	3857	3870	3882	3889	3896
40	182.4	85.1	3718	3725	3733	3747	3763	3778	3792	3800	3808
50	181.3	86.4	3292	3301	3311	3327	3346	3364	3381	3391	3399
60	179.2	84.4	2969	2986	3006	3039	3076	3112	3145	3165	3182

cross-validation sample, respectively. These improved fits make the additive model provided by the current study a more adequate reference prediction equation to utilise in both the paediatric and adult populations.

Primary sample analysis and cross-validation showed consistent results, specifically a stronger predictive accuracy in ⩽20-year-old subjects, and accuracy improvement in >20-year-old subjects and the entire sample. This consistent increase in predictive accuracy indicates a good generalisability to the Dutch population. This is underlined by the fit of R^2=0.54 (SEE=556.55 mL·min^{-1}) of the additive model to the whole sample, including smokers, all BMI values and athletes. The somewhat lower adjusted R^2 of the additive model obtained in the cross-validation >20-year-old subgroup suggests a difference from the primary sample analysis. This is possibly caused by the use of a variety of more institutions providing >20-year-old subjects in the cross-validation sample. Every subject >20 years old in this sample was tested at a single institute aimed at test indications such as sport- and work-related (mandatory) annual health checks. The underrepresentation of tests initiated by a healthcare professional results in a cross-validation sample with higher aerobic fitness compared with the more heterogeneous primary sample (healthy workers effect).

The strength of our study is the wide age range of 7.9–65.0 years. The LowLands Fitness Registry that we used in our study is a reasonable representation of the Dutch population. Additionally, the utilisation of a diverse variety of healthcare centres, including hospitals, sports medicine clinics and occupational medicine clinics, ensures representation of every conditioning status. The familiarity of the Dutch population with cycling and the low risk of injury during testing ensures this method of measurement is fitting for the population and participants of all ages [8].

Study results are limited by the retrospective and institution-based nature of the study. Preferably, $V'O_2peak$ reference value research should be performed using a prospective community-based method [8], since a retrospective study design has potential data quality issues. Although every institution used measurement methods and equipment described by the ACCP/ATS statement [8], the exclusion of 4364 subjects emphasises the variety of data quality in the primary sample. The majority of excluded subjects were due to missing values, accounting for 2674 excluded subjects. It is of primary importance that CPET instructors increase their skills and knowledge, and stringently apply the test guidelines provided by the ATS/ACCP statement in order to facilitate data harmonisation [8].

Representative reference $V'O_2peak$ values are genuinely needed because of the current lack of reference data in the Dutch population. The currently employed North American reference values from 1985 may plausibly underestimate the aerobic fitness for the Dutch population; hence, subjects are misclassified as having normal aerobic fitness. The additive regression equation presented in the current study can be used to determine a reference value for the Dutch population. In future research aimed at determining reference value prediction equations, the type of regression model fitted to the data may be conveniently modelled by semiparametric regression. This research can best be performed in a prospective, community-based setting with emphasis on the inclusion of sufficient numbers of female participants.

Conclusion

In conclusion, this study has provided a robust additive regression model for $V'O_2peak$ in the Dutch population. $V'O_2peak$ is sex specific and has a nonlinear relationship with age. Publicly usable reference values can be conveniently obtained by suitable software implementation.

Acknowledgements: We would like to thank the collaborating centres for providing the anonymous exercise testing data.

The LowLands Fitness Registry Study Group: Harriet Wittink (University of Applied Science Utrecht, Utrecht, The Netherlands), Marcel Schmitz (In2motion, Roermond, The Netherlands), Pieter-Jan van Ooi (The Royal Netherlands Navy's Diving Medical Center, Den Helder, The Netherlands), Geert van Beek (AZ Jan Portaels, Vilvoorde, Belgium), Leendert van Galen (InspanningLoont, Utrecht, The Netherlands), Jeroen Molinger (Belife Human Performance Lab, Rotterdam, The Netherlands), Robert Rozenberg (Erasmus Medical Center, Rotterdam, The Netherlands), Marieke van den Oord (The Center for Man and Aviation, Soesterberg, The Netherlands), Yvonne Hartman (Radboud University Medical Center, Nijmegen, The Netherlands), Nicolle Verbaarschot (St-Anna Hospital, Geldrop, The Netherlands), Aernout Snoek (Isala Hospital, Zwolle, The Netherlands), Jaap Stomphorst (Isala Hospital) and Joep van Kesteren (Sport Medical Center Sportmax, Veldhoven, The Netherlands) .

Conflict of interest: None declared.

References

1 Myers J, McAuley P, Lavie C, *et al.* Physical activity and cardiorespiratory fitness as major markers of cardiovascular risk: their independent and interwoven importance to health status. *Prog Cardiovasc Dis* 2015; 57: 306–314.

2 Wasserman K, Hansen JE, Sue D, *et al.* Principles of exercise testing and interpretation: pathophysiology and clinical applications. *Shock* 2005; 14: 1–180.
3 Older P, Levett D. Cardiopulmonary exercise testing and surgery. *Ann Am Thorac Soc* 2017; 14: S74–S83.
4 Puente-Maestu L, Palange P, Casaburi R, *et al.* Use of exercise testing in the evaluation of interventional efficacy: an official ERS statement. *Eur Respir J* 2016; 47: 429–460.
5 Guazzi M, Arena R, Halle M, *et al.* 2016 focused update: clinical recommendations for cardiopulmonary exercise testing data assessment in specific patient populations. *Circulation* 2016; 133: e694–e711.
6 Van de Poppe DJ, Hulzebos E, Takken T, *et al.* Reference values for maximum work rate in apparently healthy Dutch/Flemish adults: data from the LowLands fitness registry. *Acta Cardiol* 2018; in press [http://dx.doi.org/10.1080/00015385.2018.1478763].
7 Fleg JL, Morrell CH, Bos AG, *et al.* Accelerated longitudinal decline of aerobic capacity in healthy older adults. *Circulation* 2005; 112: 674–682.
8 Ross R, Beck K, Casaburi R, *et al.* ATS/ACCP statement on cardiopulmonary exercise testing. *Am J Respir Crit Care Med* 2003; 167: 1451.
9 Paap D, Takken T. Reference values for cardiopulmonary exercise testing in healthy adults: a systematic review. *Expert Rev Cardiovasc Ther* 2014; 12: 1439–1453.
10 Kaminsky LA, Arena R, Beckie TM, *et al.* The importance of cardiorespiratory fitness in the United States: the need for a national registry: a policy statement from the American Heart Association. *Circulation* 2013; 127: 652–662.
11 Kokkinos P, Kaminsky LA, Arena R, *et al.* A new generalized cycle ergometry equation for predicting maximal oxygen uptake: the Fitness Registry and the Importance of Exercise National Database (FRIEND). *Eur J Prev Cardiol* 2018; 25: 1077–1082.
12 Kaminsky L, Arena R, Myers J. Reference standards for cardiorespiratory fitness measured with cardiopulmonary exercise testing. *Mayo Clin Proc* 2015; 90: 1515–1523.
13 Kaminsky LA, Imboden MT, Arena R, *et al.* Reference standards for cardiorespiratory fitness measured with cardiopulmonary exercise testing using cycle ergometry: data from the Fitness Registry and the Importance of Exercise National Database (FRIEND) registry. *Mayo Clin Proc* 2017; 92: 228–233.
14 Jones N, Makrides L, Hitchkock C, *et al.* Normal standards for an incremental progressive cycle ergometer test. *Am Rev Respir Dis* 1985; 131: 700–708.
15 de Oliveira RJ, Bottaro M, Mota AM, *et al.* Association between sarcopenia-related phenotypes and aerobic capacity indexes of older women. *J Sport Sci Med* 2009; 8: 337–343.
16 Graves R, Mahnken J, Perea R, *et al.* Modeling percentile rank of cardiorespiratory fitness across the lifespan. *Cardiopulm Phys Ther J* 2015; 26: 108–113.
17 Hawkins SA, Wiswell RA. Rate and mechanism of maximal oxygen consumption decline with aging: implications for exercise training. *Sport Med* 2003; 33: 877–888.
18 Blais S, Berbari J, Counil F-P, *et al.* A systematic review of reference values in pediatric cardiopulmonary exercise testing. *Pediatr Cardiol* 2015; 36: 1553–1564.
19 Bouaziz W, Kanagaratnam L, Vogel T, *et al.* Effect of aerobic training on peak oxygen uptake among seniors aged 70 or older: a meta-analysis of randomized controlled trials. *Rejuvenation Res* 2018; 21: 341–349.
20 Nevill A, Cooke C. The dangers of estimating VO_{2max} using linear, nonexercise prediction models. *Med Sci Sports Exerc* 2017; 49: 1036–1042.
21 Graves L, Batterham A, Foweather L, *et al.* Scaling of peak oxygen uptake in children: a comparison of three body size index models. *Med Sci Sports Exerc* 2013; 45: 2341–2345.
22 Gräsbeck R. The evolution of the reference value concept. *Clin Chem Lab Med* 2004; 42: 692–697.
23 Guazzi M, Adams V, Conraads V, *et al.* Clinical recommendations for cardiopulmonary exercise testing data assessment in specific patient populations. *Eur Heart J* 2012; 33: 2917–2927.
24 Tanaka H, Monahan KD, Seals DR. Age-predicted maximal heart rate revisited. *J Am Coll Cardiol* 2001; 37: 153–156.
25 Allison DB, Downey M, Atkinson RL, *et al.* Obesity as a disease: a white paper on evidence and arguments commissioned by the council of the obesity society. *Obesity* 2008; 16: 1161–1177.
26 Biadi O, Rossini R, Musumeci G, *et al.* Cardiopulmonary exercise test in young women affected by anorexia nervosa. *Ital Heart J* 2001; 2: 462–467.
27 Suminski R, Wier L, Poston W, *et al.* The effect of habitual smoking on measured and predicted VO_2max. *J Phys Act Health* 2009; 6: 667–673.
28 Bovens AMPM, Van Baak MA, Vrencken JGPM, *et al.* Maximal aerobic power in cycle ergometry in middle-aged men and women, active in sports, in relation to age and physical activity. *Int J Sports Med* 1993; 14: 66–71.
29 R Development Core Team. R: A Language and Environment for Statistical Computing. Vienna, R Foundation for Statistical Computing, 2008.
30 Wood SN. Generalized additive models: an introduction with R. *J R Stat Soc A* 2006; 170: 388.
31 Wood SN. Stable and efficient multiple smoothing parameter estimation for generalized additive models. *J Am Stat Assoc* 2004; 99: 673–686.
32 Akaike H. A new look at the statistical model identification. *IEEE Trans Autom Control* 1974; 19: 716–723.
33 Schwarz G. Estimating the dimension of a model. *Ann Stat* 1978; 6: 461–464.
34 Kass R, Raftery A. Bayes factors. *J Am Stat Assoc* 1995; 90: 773–795.
35 Zhang Z. Variable selection with stepwise and best subset approaches. *Ann Transl Med* 2016; 4: 136–136.
36 Portney G, Watkins P. Foundations of Clinical Research: Applications to Practice. 3rd Edn. Upper Saddle River, Prentice Hall, 2009.
37 Weber DR, Leonard MB, Zemel BS. Body composition analysis in the pediatric population. *Pediatr Endocrinol Rev* 2012; 10: 130–139.
38 Loomba-Albrecht L, Styne D. Effect of puberty on body composition. *Curr Opin Endocrinol Diabetes Obes* 2009; 16: 10–15.
39 Köhler A, King R, Bahls M, *et al.* Cardiopulmonary fitness is strongly associated with body cell mass and fat-free mass. *Scand J Med Sci Sport* 2018; 28: 1628–1635.

40 Rapp D, Scharhag J, Wagenpfeil S, *et al.* Reference values for peak oxygen uptake: cross-sectional analysis of cycle ergometry-based cardiopulmonary exercise tests of 10 090 adult German volunteers from the Prevention First Registry. *BMJ Open* 2018; 8: e018697.
41 Ten Harkel AD, Takken T, Van Osch-Gevers M, *et al.* Normal values for cardiopulmonary exercise testing in children. *Eur J Prev Cardiol* 2011; 18: 48–54.