Reference Value Prediction Model for Peak Oxygen Uptake in the Dutch Population Measured by Incremental Cardiopulmonary Exercise Cycle Ergometry

Masterthesis

Physiotherapy Science

Program in Clinical Health Sciences

Utrecht University

Name student:	C.F. (Caspar) Mijlius
Student number:	5506417
Date:	30-06-2017
Internship supervisor(s):	Prof. Dr. C. (Cees) van der Schans
Internship institute:	Hanze university of applied sciences, Groningen
Lecturer/supervisor Utrecht University:	Dr. J. (Janjaap) van der Net

"ONDERGETEKENDE

Caspar Frederik Mijlius,

bevestigt hierbij dat de onderhavige verhandeling mag worden geraadpleegd en vrij mag worden gefotokopieerd. Bij het citeren moet steeds de titel en de auteur van de verhandeling worden vermeld."

Examiner

Dr. M.F. Pisters

Assessors:

Prof. Dr. C. van der Schans
Dr. M. Pisters

Masterthesis, Physical Therapy Sciences, Program in Clinical Health Sciences, Utrecht University, Utrecht, 2017

ABSTRACT

Maximal oxygen uptake represents the functional limit of the body's ability to deliver and extract oxygen to meet the metabolic demands of vigorous exercise. It is recognized as the gold-standard for aerobic fitness and as an important component in estimating health risk. Therefore, it is essential to have accurate reference values to know what constitutes a healthy value. Currently, there is no existing consensus in scientific literature for the trajectory of the VO_{2peak} development across a person's entire lifespan. Therefore, there is a risk miss interpreting risk stratification if inadequate varieties of regression models are utilized.

Aim: To determine the nature of the regression model for the VO_{2peak} with the progression of age across a person's entire lifespan in a cohort of healthy Dutch participants.

Methods: A multi-center cross-sectional study was conducted by using the Low-lands Fitness Registry. The database contains measurements collected between January 2010 and December 2016 across 11 healthcare centers in the Netherlands. Generalized Additive Models are used as a semi-parametric way to identify the regression model for the VO_{2peak}.

Results: An additive model including weight, height and gender plus an interaction between age and gender best suited the trajectory of VO_{2peak} across a person's lifespan. This model yields an adjusted R^2=0.641, Akaike information criterion=42386.51, Bayesian information criterion=42478.83.

Conclusion: The trajectory of VO_{2peak} across a lifespan, can be modeled by semi-parametric regression without testing for various polynomial transformations of age.

Clinical Relevance: Representative values of aerobic fitness are sincerely needed since the current lack of data in the Dutch population. The prediction equation presented in this study can be used to determine reference values for the Dutch population. In future research aimed at determining reference value prediction equations for other nationalities, the type of regression equation fitted to the data should be modeled by semi-parametric regression.

Keywords:

Maximal oxygen uptake - Cardiopulmonary exercise testing - Regression model - Generalized additive model - reference value prediction equation

INTRODUCTION

Maximal oxygen uptake (VO_{2max}) represents the functional limit of the body's ability to deliver and extract VO_2 to meet the metabolic demands of vigorous exercise, it is recognized as the best expression of aerobic fitness.[1] Aerobic fitness takes an increasingly central role in healthcare because of the inverse association with type 2 diabetes, cardiovascular disease, and overall mortality in both a-symptomatic and symptomatic people.[1-5] The clinical decision making process critically depends upon reference value (RV) prediction equations given the importance of VO_{2max} in estimating health risk. It is essential to have accurate RV to know what constitutes a "normal" value. A person's VO_{2max} is influenced by multiple characteristics such as age, gender, height, weight, ethnicity, conditioning status, and the presence of disease or medications.[6-13] It is well known that VO_{2max} declines with age during adulthood and that higher values are observed in male compared with females.[1,6,14,15] The status of an individual's highest measured oxygen uptake (VO_{2peak}) can be evaluated through comparison with population RV from a prediction equation. These RV are obtained from healthy subjects with similar characteristics.[16] In 2003, the American Thoracic Society/American College of Chest Physicians (ATS/ACCP) published a statement containing guidelines for Cardiopulmonary Exercise Testing (CPET), with the aim to facilitate the interpretation and clinical application.[6] This statement recognized that RV were critical for any interpretative schema.[6] Knowing a person's VO_{2peak} relative to their peers does not only help to optimize risk stratification, but can also facilitate the clinical reasoning proces.[6,17]

Currently, there is no consensus in scientific literature about the type of regression model for the trajectory of the VO_{2peak} across a person's lifespan depending on age.[20] Several studies recommended a variety of both linear and non-linear VO_{2peak} RV prediction equations for the pediatric and adult population.[18-20] A 2014 systematic literature review by Paap et al. on CPET RV in predominantly adults, included nine linear regression equations and one non-linear prediction equation.[8] Publications including both the pediatric and adult population present the relation between VO_{2peak} and age in a linear model, this does not take the increasing VO_{2peak} in the pediatric population and a decreasing VO_{2peak} in the adult and elderly population in consideration.[6,8,17,21,22] A non-linear and inter-relation fashion of more than one independent variable has been hypothesized in both the pediatric and adult population.[14,21]

There is a risk of inadequate health evaluation and risk stratification if inadequate regression models are used in determining prediction equations. Therefore, determining a more adequate regression model for the VO_{2peak} RV prediction equation for the Dutch population is of importance. The aim of this study is to determine the nature of the regression model for the VO_{2peak} from the explanatory variables relation to age, weight and height in the healthy Dutch population with the progression of age across a person's entire lifespan.

METHODS

Study design

A multi-center cross-sectional study was conducted by using the Low-lands Fitness Registry, a database established with the primary aim to establish CPET RV for the Dutch population. The database contains measurements collected between January 2010 and December 2016 of patients from 11 of subjects from healthcare centers in the Netherlands. (See Acknowledgements) Institutes included in the database [1] met the ATS/ACCP statement equipment requirements to perform an incremental CPET using an electromagnetically braked cycle ergometry test utilizing gas exchange analysis by bag collection, mixing chamber or breath by breath analysis based upon averaging the values measured during last 30 to 60 seconds of the test;[6] and [2] perform equipment quality control in accordance with the ATS/ACCP-statement.[6] Subjects who underwent an individualized incremental CPET cycle ergometry test for multiple reasons were eligible for inclusion. These test reasons consisted of a test: initiated by a healthcare professional; work and sports related (mandatory) annual health checks; participation in scientific studies; or based on personal motivation like exercise response evaluation for the aid of a trainings scheme. Every institute provided anonymized, coded data to the data coordinator at the Wilhemina Kinderziekenhuis, Utrecht. All records were screened for measurement failures. If there were doubts, the testing institute was contacted to ensure the correct data was communicated.

Study sample

Data of healthy Dutch subjects, without age restrictions, was considered eligible for this study. The status of healthy was warranted through exclusion, a subject was excluded from the sample if the subject; [1] had a diagnosed illness at the time of testing or a diagnosis resulting from test results; [2] showed irregularities on the electrocardiogram (ECG) prior to testing; [3] was a professional athlete; [4] actively smoked at the time of the test or five years prior to the test; [5] had a body mass index (BMI) value ≥30, [6] did not perform a maximum test. Criterion six was determined as either a minimum of 85% of the age-predicted maximum heart rate (HR_{peak})[23] or not reaching a respiratory exchange peak ratio (RER_{peak}) of 1.0.[24] Before the CPET, information about each participant's demographic, medical and smoking history was gathered by the test leader. Participants were verbally asked about their smoking and athletics status as well as whether they had any diagnosed illness. Height and body weight were measured to the nearest 0.5 cm and 0.1 kg, respectively. BMI was calculated using weight divided by height squared. Age was calculated from day of testing minus date of birth. The WR_{peak}, RER_{peak}, HR_{peak} and VO_{2peak} was determined using the ergometry and gas exchange analysis based upon averaging the values measured during last 30 to 60 seconds of the test.

Statistical analyses

Statistical analyses are performed using R version 3.2.1, released 2015.[25] Throughout, a probability ≤.05 is considered significant. Continues data is summarized as mean (SD), categorical data as frequencies (percentage). Variables were considered eligible for inclusion in the model if they could be determined prior to testing.

Due to the lack of a consensus in scientific literature for the VO_{2peak} with the progression of age, Generalized Additive Models (GAM) are used as a semi-parametric way to identify a regression model for the VO_{2peak} depending on age.[26-28] The variables weight and height are included because of the common use in prediction equations in scientific literature and influence the development of VO_{2peak}.[8,26-28] The model fits the data though a cubic type of splines with smoothness determined by generalized cross-validation (GCV) embedded in GAM estimation procedures.[30] The fit of the model to the data was evaluated by comparing the adjusted R^2, Akaike information criterion (AIC)[30] and Bayesian information criterion (BIC)[31] for several model specifications. A higher adjusted R^2 and a lower AIC and BIC were considered as improved fit to the sample. If there was inconsistence in these scores, the BIC rating provided decisive. The interpretation of the BIC was 0 to 2 - minimal improvement, 2 to 6 positive improvement, 6 - 10 strong improvement and a >10 score as a very strong improvement.[32] To test for non-linearity various models were considered. We started with linear regression as it is the most basic and frequently used. Secondly, based upon inspection of the scatterplots of each explanatory variable with VO_{2peak}, an additive model is fitted with a spline transformation for the variable age and linear terms for height, weight and gender. Based upon the same scatterplots and because the hypothesized age and gender-dependent dynamics for $VO_{2peak,}$[20] an additional model with interaction terms between age and gender is fitted to the data to account for different VO_{2peak} levels of male and female subjects.

External validation

The external validation yielded a model from the ``train data set'' of which the predictive validity was tested in a cross-validation procedure on an independent test sample as recommended by the ATS/ACCP statement.[6] The cross-validation was performed against an independent data set from the Diving Medical Center from Den Helder, the Netherlands. To determine if the same type of model is the best fit for the Diving Medical Center sample, similar steps as used with the primary cohort are performed. The fit of each regression model to this sample is indicated by the adjusted R^2, AIC and BIC. Similar to the primary analysis, a higher adjusted R^2 and lower AIC and BIC were considered as improved fit. The BIC was decisive if inconsistent scores were noted.

RESULTS

The initial sample consisted of 8353 subjects. After applying the exclusion criteria 2,777 subjects remained, 2,386 males and 391 females with age ranging from 7.9 to 76 years. The sample is stratified by gender and per two years until the age of 20 years and per decade for 20 years and older. The descriptive characteristics of the sample are presented in **Table 1**. The descriptive CPET results of the sample are presented in **Table 2.** The development of VO_{2peak} with respectively age, weight and height amongst the male and female participants are shown in **Figure 1,2** and **3.**

Table 1. Study sample characteristics in mean and standard deviation. Age: years per decimal, Height: centimeters, Weight: kilograms, Body mass index(BMI): kilogram/meter²

Age group	Gender	Subjects % (No)	Age ±SD (year)	Weight ±SD (kg)	Height ±SD (cm)	BMI ±SD (kg/m²)
7-<10	Female	27 (0.97%)	9.0 (0.65)	32.7 (5.25)	138.9 (5.82)	16.8 (1.80)
	Male	19 (0.68%)	8.9 (0.68)	32.9 (4.87)	138.3 (6.25)	17.2 (2.12)
10-<12	Female	31 (1.11%)	11.1 (0.61)	39.5 (6.51)	150.0 (6.61)	17.4 (2.08)
	Male	29 (1.04%)	10.8 (0.61)	39.8 (6.76)	150.9 (7.93)	17.4 (2.06)
12-<14	Female	26 (0.93%)	12.8 (0.48)	44.9 (7.71)	158.7 (5.45)	17.7 (2.84)
	Male	28 (1.00%)	13.1 (0.61)	51.8 (11.07)	163.8 (8.17)	19.2 (3.31)
14-<16	Female	15 (0.54%)	14.7 (0.63)	57.9 (9.34)	173.0 (8.39)	19.2 (2.14)
	Male	14 (0.50%)	15.1 (0.70)	56.8 (7.74)	169.0 (7.29)	19.8 (2.16)
16-<18	Female	26 (0.93%)	17.0 (0.52)	65.3 (7.74)	177.1 (8.06)	20.7 (1.57)
	Male	46 (1.65%)	16.9 (0.41)	70.3 (9.68)	177.8 (5.70)	22.2 (2.76)
18-<20	Female	10 (0.36%)	18.5 (0.44)	65.8 (11.42)	175.0 (11.52)	21.4 (2.21)
	Male	92 (3.31%)	18,4 (0.50)	71.6 (10.33)	179.6 (6.66)	22.1 (2.60)
20-<30	Female	96 (3.45%)	24.5 (2.72)	68.6 (7.35)	172.1 (6.17)	23.1 (2.20)
	Male	921 (33.16%)	24.7 (2.80)	80.9 (9.49)	182.5 (6.74)	24.2 (2.29)
30-<40	Female	62 (2.23%)	34.7 (2.80)	65.4 (8.86)	171.6 (6.57)	22.1 (2.39)
	Male	708 (25.49%)	33.9 (2.83)	84.0 (9.52)	182.9 (6.68)	25.0 (2.21)
40-<50	Female	62 (2.23%)	44.5 (2.92)	68.2 (8.63)	170.7 (5.51)	23.3 (2.64)
	Male	395 (14.22%)	43.9 (2.88)	85.9 (8.83)	182.7 (6.79)	25.7 (2.24)
50-<60	Female	30 (1.08%)	54.1 (2.83)	70.6 (6.99)	169.0 (4.53)	24.7 (2.53)
	Male	98 (3.52%)	53.1 (2.66)	85.2 (8.97)	182.3 (5.82)	25.6 (2.17)
60-<70	Female	6 (0.21%)	62.9 (2.18)	70.3 (3.59)	166.1 (6.37)	25.5 (1.40)
	Male	26 (0.93%)	64.3 (2.66)	78.9 (8.84)	178.0 (6.66)	24.9 (2.21)
70-<80	Female	0 (0%)	-	-	-	-
	Male	10 (0.36%)	72.4 (2.32)	75.2 (3.70)	178.0 (6.46)	23.8 (1.77)
Total	Female	391 (14.07%)	28.2 (14.48)	61.1 (14.33)	167.0 (12.02)	21.6 (3.38)
	Male	2386 (85.92%)	31.5 (10.77)	80.9 (12.35)	181.3 (8.80)	24.4 (2.72)
	All	2777 (100%)	31.1 (11.42)	78.1 (14.38)	179.36 (10.57)	24.0 (2.99)

Table 2. Study characteristics; peak RER: highest measured respiratory exchange rate ratio, peak HR: highest measured heart rate beats per minute, peak WR: highest measured wattage, peak VO2: highest measured oxygen uptake

Age group	Gender	Subjects % (No)	Peak RER ±SD	Peak HR (bpm) ±SD	Peak WR (wattage) ±SD	Peak VO_2 ±SD	Peak VO_2 (ml·min-1·kg-1) ±SD
7-<10	Female	27 (0.97%)	1.15 (0.06)	187.0 (9.03)	114.0 (25.01)	1538.8 (277.19)	47.17 (5.85)
	Male	19 (0.68%)	1.12 (0.07)	188.2 (7.74)	104.7 (19.63)	1392.4 (213.95)	42.75 (7.57)
10-<12	Female	31 (1.11%)	1.14 (0.07)	187.8 (9.01)	146.5 (27.00)	1887.6 (236.32)	48.37 (5.90)
	Male	29 (1.04%)	1.19 (0.07)	191.6 (9.78)	128.4 (23.61)	1606.5 (263.66)	40.67 (5.16)
12-<14	Female	26 (0.93%)	1.16 (0.06)	190.6 (8.76)	179.7 (33.91)	2194.1 (368.15)	49.35 (6.92)
	Male	28 (1.00%)	1.16 (0.09)	192.1 (8.17)	187.3 (52.28)	2376.5 (702.97)	46.27 (11.55)
14-<16	Female	15 (0.54%)	1.13 (0.07)	194.0 (10.43)	236.6 (42.75)	2907.1 (567.28)	48.77 (6.09)
	Male	14 (0.50%)	1.14 (0.07)	190.2 (7.80)	210.5 (52.87)	2461.7 (578.30)	43.39 (8.45)
16-<18	Female	26 (0.93%)	1.16 (0.07)	192.0 (10.87)	262.1 (43.16)	3173.0 (466.08)	48.77 (6.09)
	Male	46 (1.65%)	1.17 (0.08)	195.2 (9.94)	292.2 (58.45)	3484.0 (656.26)	49.80 (9.33)
18-<20	Female	10 (0.36%)	1.17 (0.08)	188.1 (4.48)	232.4 (77.11)	2575.5 (772.57)	38.81 (7.67)
	Male	92 (3.31%)	1.17 (0.08)	196.4 (7.28)	310.8 (45.99)	3734.2 (557.41)	52.40 (6.40)
20-<30	Female	96 (3.45%)	1.18 (0.07)	189.0 (7.75)	247.9 (44.81)	2782.0 (500.53)	40.82 (7.69)
	Male	921 (33.16%)	1.19 (0.08)	190.9 (9.17)	332.0 (48.77)	3904.9 (558.82)	48.46 (6.26)
30-<40	Female	62 (2.23%)	1.20 (0.08)	182.2 (10.27)	229.1 (47.17)	2509.1 (500.58)	38.92 (8.83)
	Male	708 (25.49%)	1.20 (0.07)	186.0 (9.25)	336.9 (49.46)	3893.5 (560.76)	46.64 (6.72)
40-<50	Female	62 (2.23%)	1.19 (0.08)	176.1 (9.65)	223.8 (44.99)	2377.3 (449.77)	35.36 (7.85)
	Male	395 (14.22%)	1.19 (0.07)	180.5 (9.80)	334.1 (48.30)	3824.0 (598.19)	44.71 (7.04)
50-<60	Female	30 (1.08%)	1.19 (0.08)	167.1 (13.74)	199.9 (41.25)	2164.4 (377.33)	31.13 (7.37)
	Male	98 (3.52%)	1.19 (0.07)	171.9 (11.71)	311.9 (58.04)	3496.4 (652.56)	41.32 (8.25)
60-<70	Female	6 (0.21%)	1.14 (0.05)	161.5 (7.89)	162.1 (35.72)	1752.1 (250.95)	25.03 (4.34)
	Male	26 (0.93%)	1.19 (0.08)	161.6 (11.29)	252.1 (71.82)	2789.7 (789.68)	35.98 (11.73)
70-<80	Female	0 (0%)	-	-	-	-	
	Male	10 (0.36%)	1.17 (0.09)	162.6 (12.64)	220.4 (65.37)	2676.2 (699.67)	35.39 (8.55)
Total	Female	391 (14.07%)	1.17 (0.08)	184.0 (12.19)	214.6 (59.00)	2441.0 (610.81)	41.11 (9.50)
	Male	2386 (85.92%)	1.19 (0.08)	186.8 (11.19)	323.3 (61.67)	3765.0 (702.89)	46.80 (7.34)
	All	2777 (100%)	1.19 (0.08)	186.4 (11.38)	308.1 (71.99)	3578.6 (830.07)	46.00 (7.93)

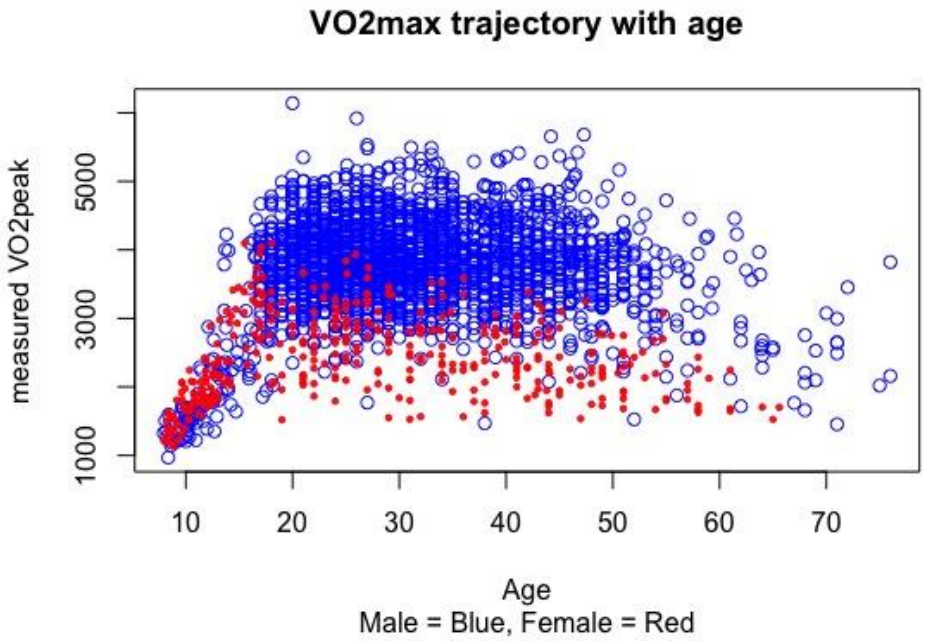

Figure 1. VO2peak trajectory with age 1

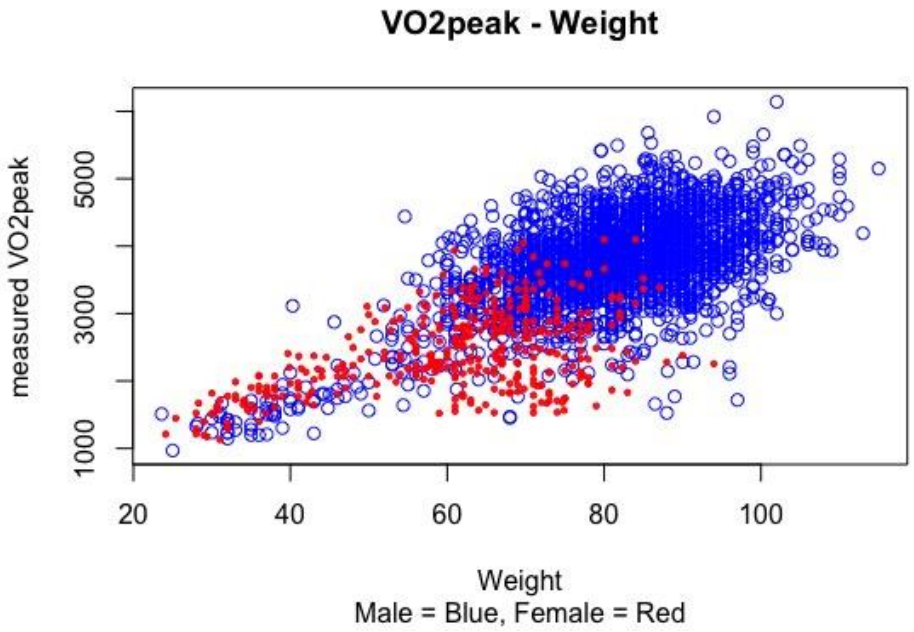

Figure 2. VO2peak related to weight

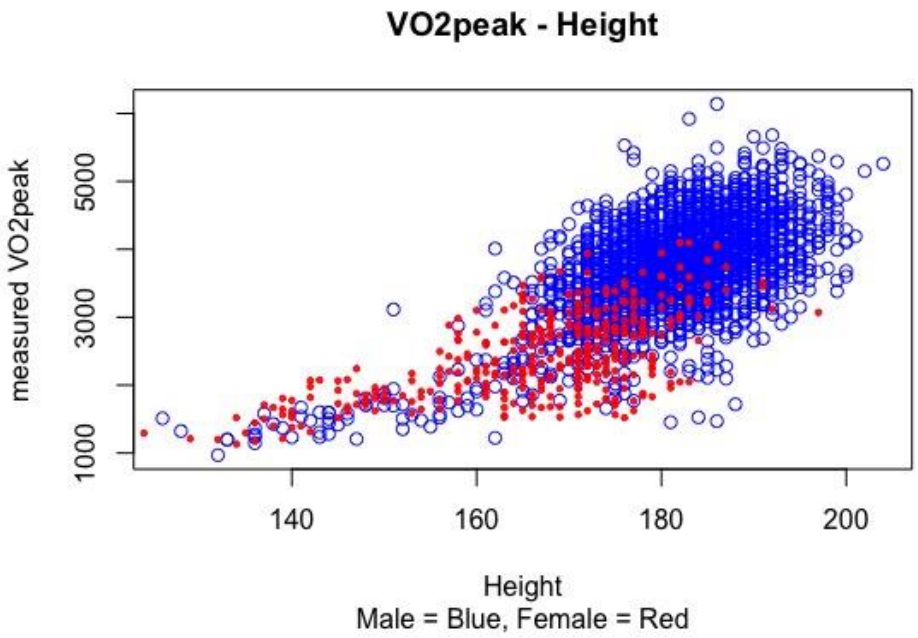

Figure 3. VO2peak trajectory with height

Generalized regression model

Each consecutive fitted model type showed an increasing improved fit to the data. The linear regression model with terms weight, height, gender and age yield an adjusted R^2=0.607, AIC=42621.97, BIC=42657.55, whilst the additive model with a non-linear term for age yields adjusted R^2=0.622, AIC=42519.52, BIC=42584.04. The additive model with an additional interaction between age and gender performed the best with adjusted R^2=0.641, AIC=42386.51, BIC=42478.83. The models and corresponding adjusted R^2, AIC, BIC and GCV are displayed in **Table 3.**

Table 3. Model type fitting - Gender: 0=male, 1=female, Age= years, height= centimeters, weight= kilograms

Model type	Prediction equation	Adjusted R^2	AIC	BIC	GCV
Linear	VO_{2peak} = -2220.71 - (555.97 * gender) - (13.11 * age) + (25.09 * height) + (22.85 * weight)	0.607	42621.97	42657.55	2.70
Additive	VO_{2peak} = -1577.67 - (576.19 * gender) + (20.59 * height) + (19.76 * weight) * (s (age))	0.622	42519.52	42584.04	2.61
Additive with interaction	VO_{2peak} = -1252.01 - (688.08 * gender) + (19.47 * height) + (18.05 * weight) + (s (age, by gender))	0.641	42386.51	42478.83	2.48

Cross-validation

The minimized GCV of the additive with interaction model is 2.48. The external-validation is made with an independent sample to determine the external validity of the new model. This independent sample contains 3.747 unique subjects, 3568 males and 179 females with an age range between 6.8 and 59.0 years (mean 33.74±10.26). The fit of each model to this sample is shown in **Table 4**. Similar to the primary sample, the additive model including an interaction between age and gender showed the best fit to the sample (R^2=0.583, AIC=56352.29, BIC=56437.2).

Table 4. Model type fitting to cross validation sample - Gender: 0=male, 1=female, Age= years, height= centimeters, weight= kilograms

Model type	Prediction equation	Adjusted R^2	AIC	BIC
Linear	VO_{2peak} = -2326.01 - (646.51 * gender) - (11.94 * age) + (27.24 * height) + (19.44 * weight)	0.545	56672.42	56709.79
Additive	VO_{2peak} = -1337.49 - (601.82 * gender) + (20.68 * height) + (17.08 * weight) + (s (age))	0.58	56386.66	56467.70
Additive with interaction	VO_{2peak} = -1180.75 - (565.36 * gender) + (19.90 * height) + (16.90 * weight) + (s (age, by gender))	0.583	56352.29	56437.20

Predictions

Through the utilization of the newly formulated model, reference value predictions can be formulated. For instance, a 40.0-year-old female (subject 1) with a body composition of height 185.0 centimeters and 110.0 kilograms has a predicted VO_{2peak} value of 2869,09. A 16.6-year-old male (subject 2) with a body height of 176.0 centimeters and 54,6 kilograms has a predicted VO_{2peak} of 3095.06. The corresponding percentiles are displayed in **Table 5.**

Table 5. VO_{2peak} prediction examples

Percentiles Subject	3%	10%	25%	50%	75%	90%	97%
Subject 1	2813.32	2839.692	2866.409	2896.094	2925.778	2952.495	2978.868
Subject 2	3019.76	3043.758	3068.06	3095.062	3122.065	3146.367	3170.357

DISCUSSION

The aim of this study was to determine the nature of the regression model for the VO_{2peak} with the progression of age across a person's entire lifespan in a cohort of healthy Dutch participants. We tested several generalized regression models and mathematical transformations to obtain prediction curves that minimized residual association. The best fitting model includes terms for weight, height, gender and a spline interaction between age and gender. The new model type has an improved fit with both the primary sample (adjusted R^2=0.641, AIC=42519.52, BIC=42584.04) and the cross-validation sample (adjusted R^2=0.583, AIC=56352.29, BIC=56437.2) compared to the frequently used linear model type (primary sample: adjusted R^2=0.607, AIC=42621.97, BIC=42657.55 and cross-validation sample: adjusted R^2=0.545, AIC=56672,42, BIC=56709,79). In both cohorts, the BIC noted a very strong improvement per model. Between the linear and additive model the corresponding BIC improvements were 73,51 in the primary cohort and 285,76 in the cross-validation cohort. Between the additive and additive with an interaction model, respectively BIC improvements were 105,21 in the primary cohort and 30,5 in the cross-validation cohort.[32] It can be concluded that the VO_{2peak} development across a person's lifespan, can be modeled by semi-parametric regression without testing for various polynomial transformations of age.

The development of GAM provides an improved method to determine the best type of regression model. This method is relatively unknown in VO_{2peak} reference value prediction research.[26] The use of this advanced method makes it possible to determine a regression model for the development of VO2peak across a person's entire lifespan. Earlier research stratified the sample which minimized the age effect. The current tendency to differentiate between the adult and pediatric population is the causes of an unrealistic transition from pediatric care to an adult hospital. The newly defined model will facilitate a smooth transition to adult care. This is important since many patients with congenital disease will reach into adulthood nowadays because of the improvements in care.

The current finding of an additive model including an interaction between age and gender, merge the consistent findings in scientific literature of an increasing VO_{2peak} in the pediatric population, followed by a decreasing VO_{2peak} in the adult population.[8,18] Nonetheless, it is in contrast with the linear prediction model presented by Jones et al.,[33] a study from 1989 including both the pediatric and adult population in the sample with an age range from 15 till 71 years.[33] Compared to the current cohort, the prediction models of Jones et al. yield a correlation of r=0.44 for the female specific prediction equation and r=0.49 for the males.[33] Similar to the prediction equation of Jones et al., the prediction equation provided by Wasserman et al. is commonly used in the Dutch adult population.[16,33] In comparison with the current sample, this linear prediction equation yields a correlation of r=0.54 for both male and female adults.[16] The linear prediction equation of Ten Harkel et al.[22] including the term age, is commonly used in the Dutch pediatric care. When applied to the same age group of the current cohort, the prediction equation yields a correlation of r=0.83. The main reason

behind this result is the inclusion of the Ten Harkel et al.[22] sample in the sample of this current research.

The large sample of the healthy Dutch population makes it possible to provide a robust regression model for VO_{2peak} prediction. Additionally, 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 are represented in this study.[6] Nonetheless, study results are limited by the retrospective and institution based nature of the study. Institution based research is at an increased risk of inclusion bias. Preferably, VO_{2peak} RV research is performed in a prospective community based method to ensure a good representation of the population.[6] The effect of institution based tests is minimized by including a high amount and large variety of test locations and test reasons in the sample. This includes both voluntary and mandatory workforce health checks, sports, recreational and health related tests. This varied selection procedure limits the risk of an overrepresentation of persons with higher or lower aerobic fitness. The representativeness of the sample is underlined by the close approximation of the samples mean height and weight to the mean height in Dutch population in people over the age of 20 years old, presented by Statistics Netherlands (CBS).[34] The mean height of the males included in the sample rate 181.39cm and females 171.17cm, compared to 180.90cm and 167.5cm presented by the CBS.[34] The mean weight in the sample, 81.50kg for male and 68.08kg for female participants, is lower compared to the data presented by the CBS of 84Kg for male and 70kg for female participants. This difference is explained by the exclusion of subject with a BMI >30 in the current study. These subjects are excluded because the World Health Organization labels a BMI >30 as a disease.[35] The CBS mean weight is 84.0kg for males includes 10.2% subjects with a BMI rating >30 and the 70kg for females include 30.3% with a BMI >30. Although every institution used measurement methods and equipment described by the ACCP/ATS statement,[6] the variety in equipment and the large number of instructors involved in the process causes an increased risk for measurement deviations.

The underrepresentation of females in both the primary sample and cross validation sample is a limiting factor for generalization to the female population. Historically, females are understudied in VO_{2peak} RV studies, measurements of 34% fewer female subjects were included in all the studies included in the systematic review by Paap et al.[8] Similar underrepresentation is found in the clinical setting within the Netherlands, females less commonly participate in CPET testing.

Representative norms of aerobic fitness are sincerely needed since the current lack of data in the Dutch population. Currently employed RVs might under estimate the aerobic fitness for the Dutch population and hence subjects are misclassified as having a normal aerobic fitness. The prediction equation presented in this study can be used to determine RV for the Dutch population. In future research aimed at determining RV prediction equations, the type of regression equation fitted to the data should be modeled by semi-parametric regression. This

research should be performed in a prospective, community based setting with emphasize on the inclusion of female participants and taking physical activity into consideration.[36] Also, given the large number of excluded subjects, there is a need for data harmonization amongst institutions in the Netherlands**.**

CONCLUSION

In conclusion, this study has provided a robust semi-parametric regression model type for aerobic fitness in the Dutch population, in future VO_{2peak} RV prediction equation research, there is no need to test for various polynomial transformations of age. VO_{2max} develops in a gender-specific non-linear fashion over time best expressed by a spline model.

REFERENCES

1. Myers J, McAuley P, Lavie C, Despres JP, Arena R, Kokkinos P. Physical activity and cardiorespiratory fitness as major markers of cardiovascular risk: their independent and interwoven importance to health status. Prog Cardiovasc Dis. 2015; 57(4):306–315.

2. Kokkinos P Myers J. Exercise and Physical Activity: Clinical Outcomes and Applications. Circulation. 2010;122(16):1637-1648

3. Swift D, Lavie C, Johannsen N, Arena R, Earnest C, O Keefe J Et al. Physical Activity, Cardiorespiratory Fitness, and Exercise Training in Primary and Secondary Coronary Prevention. Circulation Journal. 2013;77(2):281-292

4. Myers J, Prakash M, Froelicher V, Do D, Partington S, Atwood J. Exercise capacity and mortality among men referred for exercise testing. ACC Current Journal Review. 2002;11(4):33-34

5. Balady G, Larson M, Vasan R. Usefulness of exercise testing in the prediction of coronary disease risk among asymptomatic persons as a function of the Framingham Risk Score. ACC Current Journal Review. 2005;14(2):30

6. Ross R. ATS/ACCP statement on cardiopulmonary exercise testing. Am J Respir Crit Care Med. 2003;167(10):211-277

7. Arena R, Myers J, Williams M, Gulati M, Kligfield P, Balady G Et al. Assessment of Functional Capacity in Clinical and Research Settings: A Scientific Statement from the American Heart Association Committee on Exercise, Rehabilitation, and Prevention of the Council on Clinical Cardiology and the Council on Cardiovascular Nursing. Circulation. 2007;116(3):329-343.

8. Paap D Takken T. Reference values for cardiopulmonary exercise testing in healthy adults: a systematic review. Expert Rev Cardiovasc Ther. 2014;12(12):1439-1453

9. Loe H, Rognmo Ø, Saltin B, Wisløff U. Aerobic Capacity Reference Data in 3816 Healthy Men and Women 20–90 Years. PLoS ONE. 2013;8(5): 64319

10. Edvardsen E, Hansen B, Holme I, Dyrstad S, Anderssen S. Reference Values for Cardiorespiratory Response and Fitness on the Treadmill in a 20- to 85-Year-Old Population. Chest. 2013;144(1):241-248

11. Kaminsky L, Arena R, Myers J. Reference Standards for Cardiorespiratory Fitness Measured with Cardiopulmonary Exercise Testing. Mayo Clin Proc. 2015;90(11):1515-1523

12. Eriksen L, Grønbaek M, Helge J, Tolstrup J. Cardiorespiratory fitness in 16 025 adults aged 18-91 years and associations with physical activity and sitting time. Scand J Med Sci Sports 2015; doi:10.1111/sms.12608

13. Jones NL, Summers E, Killian KJ. Influence of age and stature on exercise capacity during incremental cycle ergometry in men and women. Am Rev Respir Dis 1989; 140(5):1373-80

14. Hawkins S, Wiswell R. Rate and Mechanism of Maximal Oxygen Consumption Decline with Aging. Sports Medicine. 2003;33(12):877-888

15. Koch B, Schaper C, Ittermann T, Spielhagen T, Dorr M, Volzke H. et al. Reference values for cardiopulmonary exercise testing in healthy volunteers: the SHIP study. Eur Res J 2009;33(2):389-397

16. Wasserman K, Hansen J, Sue D, Whipp B, Froelicher V. Principles of Exercise Testing and Interpretation. J Cardiopulm Rehabil. 1987;7(4):189.

17. Palange P, Ward SA, Carlsen KH, Casaburi R, Gallagher CG, Gosselink R, et al. Recommendations on the use of exercise testing in clinical practice. Eur Respir J 2007; 29:185–209.

18. Jones NL, Makrides L, Hitchcock C, Chypchar T, Mccartney N. Normal standards for an incremental progressive cycle ergometer test. Am Rev Respir Dis 1985;131(5):700-708

19. Singh R, Singh HJ, Sirisinghe RG. Cardiopulmonary fitness in a sample of Malaysian population. JPJ J Physiol 1989; 39:475-485

20. Dubowy KO, Baden W, Bernitzki S, Peters B. A practical and transferable new protocol for treadmill testing of children and adults. Cardiol Young 2008;18(6): 615-23

21. Blais S, Berbari J, Counil FP, Dallaire F. A systematic review of reference values in pediatric cardiopulmonary exercise testing. Pediatr Cardiol 2015;36(8):1553-64.

22. Ten Harkel ADJ, Takken T, Van Osch-Gevers M, Helbing WA. Normal values for cardiopulmonary exercise testing in children. Eur J Cardiovasc Prev Rehabil. 2011;18(1):48-54

23. Tanaka H, Monahan KD, Seals DR. Age-predicted maximal heart rate revisited. J Am Coll Cardiol. 2001;37(1):153-156.

24. Guazzi M, Adams V, Conraads V, Halle M, Mezzani A, Vanhees L, et al. EACPR/AHA Joint Scientific Statement. Clinical recommendations for cardiopulmonary exercise testing data assessment in specific patient populations. Eur Heart J. 2012;33(23):2917-2927.

25. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2007 URL https://www.R-project.org/.

26. Wood SN, Stable and efficient multiple smoothing parameter estimation for generalized additive models 2004, J. Amer. Statist. Ass. 99:673-686.

27. Wood SN. Generalized Additive Models: An Introduction with R. 2006b Chapman and Hall/CRC Press, p199-138

28. Wahba G. Spline Models of Observational Data. SIAM 1990, p177

29. Wood SN. Stable and efficient multiple smoothing parameter estimation for generalized additive models. J. Amer. Statist. Ass. 2004; 99:673-686.

30. Akaike H, A new look at the statistical model identification 1974, IEEE Transactions on Automatic Control, 19 (6): 716–723.

31. Schwarz G. Estimating the dimension of a model. Ann Stat 1978, 6:461 – 464.

32. Kass, Robert E.; Raftery, Adrian E. (1995), "Bayes Factors", Journal of the American Statistical Association, 90 (430): 773–795

33. Jones NL, Makrides L, Hitchcock C, Chypchar T, Mccartney N. Normal standards for an incremental progressive cycle ergometer test. Am Rev Respir Dis 1985;131(5):700-708

34. Statistics Netherlands (2011) Lengte en gewicht van personen, ondergewicht en overgewicht; vanaf 1981 [Height and body weight of individuals, underweight and overweight; from 1981]. Statistics Netherlands. http://statline.cbs.nl/StatWeb/publication/default.aspx?VW=T&DM=SLNL&PA=81565NED&D1=2-6&D2=1-2&D3=5&D4=0&D5=0-30&HD=121122-1317&HDR=G1%2cT&STB=G2%2cG3%2cG4. Accessed 11 Jun 2017

35. Pollack A. A.M.A. Recognizes obesity as a disease. [Internet]. 2013 [cited 2016 Nov 20]. Available from: http://www.nytimes.com/2013/06/19/business/ama-recognizes-obesity-as-a-disease.html.

36. Bacon, A. P., Carter, R. E., Ogle, E. A., & Joyner, M. J. (2013). VO_2max Trainability and High Intensity Interval Training in Humans: A Meta-Analysis. *PLoS ONE*, *8*(9)

APPENDIX

Syntex R

```
library(MASS);library(haven);library(nlme);library(mgcv);library(gdata);library(MASS);library(reshape)
;library(psych);attach(mtcars)
dfa <- subset(Data_referentiewaarden_VO2max_TOTAAL_832017)
dfa <- with(dfa,data.frame(cbind(VO2max,Gender,Genderr,Age,Length,Weight,VO2maxKg,BMI)))
dfa$BMI <- (dfa$Weight) / ((dfa$Length/100)*(dfa$Length/100))
attach(dfa)
dfa$Agecat[dfa$Age < 10.00] <- "1"
dfa$Agecat[dfa$Age >= 10.00 & Age < 12.00] <- "2"
dfa$Agecat[dfa$Age >= 12.00 & Age < 14.00] <- "3"
dfa$Agecat[dfa$Age >= 14.00 & Age < 16.00] <- "4"
dfa$Agecat[dfa$Age >= 16.00 & Age < 18.00] <- "5"
dfa$Agecat[dfa$Age >= 18.00 & Age < 20.00] <- "6"
dfa$Agecat[dfa$Age >= 20.00 & Age < 30.00] <- "7"
dfa$Agecat[dfa$Age >= 30.00 & Age < 40.00] <- "8"
dfa$Agecat[dfa$Age >= 40.00 & Age < 50.00] <- "9"
dfa$Agecat[dfa$Age >= 50.00 & Age < 60.00] <- "10"
dfa$Agecat[dfa$Age >= 60.00 & Age < 70.00] <- "11"
dfa$Agecat[dfa$Age >= 70.00] <- "12"
detach(dfa)
females<-subset(dfa,subset = Gender=="1")
males<-subset(dfa,subset = Gender=="0")
dfa$Genderr <- with(dfa,factor(ifelse(Gender == 0, "male", "female")))

describe(dfa[1-10])
describe(males[1-10])
describe(females[1-10])
describeBy(dfa[1-10], list(dfa$Agecat))
describeBy(males[1-10], list(males$Agecat))
describeBy(females[1-10], list(females$Agecat))

plot(males$Weight, xlab= "Weight",males$VO2max, ylab = "measured VO2peak", pch=1,
col=rgb(0,0,1,alpha=0.9))
points(females$Weight,xlab= "Weight",females$VO2max, ylab = "measured VO2peak", pch=20,
col=rgb(1,0,0,alpha=0.9), cex=0.7)
title(main="VO2peak - Weight", sub ="Male = Blue, Female = Red")
plot(males$Age, xlab= "Age",males$VO2max, ylab = "measured VO2peak", pch=1,
col=rgb(0,0,1,alpha=0.9))
points(females$Age,xlab= "Age",females$VO2max, ylab = "measured VO2peak",
pch=20,col=rgb(1,0,0,alpha=0.9), cex=0.7)
title(main="VO2peak trajectory with Age", sub ="Male = Blue, Female = Red")
plot(males$Length, xlab= "Height",males$VO2max, ylab = "measured
VO2peak",pch=1,col=rgb(0,0,1,alpha=0.9))
```

```
points(females$Length,xlab= "Height",females$VO2max, ylab = "measured VO2peak", pch=20,
col=rgb(1,0,0,alpha=0.9), cex=0.7)
title(main="VO2peak - Height", sub ="Male = Blue, Female = Red")

linmod <- gam(VO2max~Genderr+Age+Length+Weight,data=dfa)
summary(linmod)
splinemod <- gam(VO2max~Gender + s(Age)+Length+Weight,data=dfa)
summary(splinemod)
Defmod <- gam(VO2max~Gender + s(Age,by=Genderr) + Length + Weight ,data=dfa)
summary(Defmod)
cbind(AIC(linmod,splinemod,Defmod),BIC(linmod,splinemod,Defmod))

dfacv <- bestand_crossvalidatie_MCAS
dfacv$Genderr <- with(dfacv,factor(ifelse(Gender == 0, "male", "female")))
dfacv$VO2maxKg <- dfacv$VO2max/ dfacv$Weight
dfacv$BMI <- (dfacv$Weight) / ((dfacv$Length/100)*(dfacv$Length/100))
dfacv <- with(dfacv,data.frame(cbind(VO2max,Genderr,Gender,Age,Length,Weight,VO2maxKg,BMI)))
attach(dfacv)
dfacv$Agecat[dfacv$Age < 10.00] <- "1"
dfacv$Agecat[dfacv$Age >= 10.00 & Age < 12.00] <- "2"
dfacv$Agecat[dfacv$Age >= 12.00 & Age < 14.00] <- "3"
dfacv$Agecat[dfacv$Age >= 14.00 & Age < 16.00] <- "4"
dfacv$Agecat[dfacv$Age >= 16.00 & Age < 18.00] <- "5"
dfacv$Agecat[dfacv$Age >= 18.00 & Age < 20.00] <- "6"
dfacv$Agecat[dfacv$Age >= 20.00 & Age < 30.00] <- "7"
dfacv$Agecat[dfacv$Age >= 30.00 & Age < 40.00] <- "8"
dfacv$Agecat[dfacv$Age >= 40.00 & Age < 50.00] <- "9"
dfacv$Agecat[dfacv$Age >= 50.00 & Age < 60.00] <- "10"
dfacv$Agecat[dfacv$Age >= 60.00 & Age < 70.00] <- "11"
dfacv$Agecat[dfacv$Age >= 70.00] <- "12"
detach(dfacv)
femalescv<-subset(dfacv,subset = Gender=="1")
malescv<-subset(dfacv,subset = Gender=="0")

describe(dfacv[1-6])
describe(malescv[1-6])
describe(femalescv[1-6])
describeBy(dfacv[1-6], list(dfacv$Agecat))
describeBy(malescv[1-6], list(malescv$Agecat))
describeBy(femalescv[1-6], list(femalescv$Agecat))

linmodcv <- gam(VO2max~Gender+Age+Length+Weight,data=dfacv)
summary(linmodcv)
splinemodcv <- gam(VO2max~Gender + s(Age)+Length+Weight,data=dfacv)
summary(splinemodcv)
Defmodcv <- gam(VO2max~Gender + s(Age,by=Genderr)+Length+Weight,data=dfacv)
```

```
summary(Defmodcv)
cbind(AIC(linmodcv,splinemodcv,Defmodcv),BIC(linmodcv,splinemodcv,Defmodcv))

predictive_explanatory_variables <- dfa[,2:6]
yhat <- predict(Defmod,type="response", se.fit=TRUE,newdata=predictive_explanatory_variables)
predicted <- data.frame(yhat$fit,yhat$se)
p <- c(0.03,0.10, 0.25,0.50,0.75, 0.90,0.97)
quantile <- qnorm(p)
x <- predictive_explanatory_variables[1:2,]
x[1,2:4] <- c(0,40,185,110)
yhat <- predict(Defmod,type="response", se.fit=TRUE,newdata=x)
predicted <- data.frame(yhat$fit,yhat$se)
predicted.percentiles <- matrix(NA,2,7)
for (i in 1:2) {predicted.percentiles[i,] <- predicted[i,1] + predicted[i,2] *  quantile}
colnames(predicted.percentiles) <- as.character(p)
```

Acknowledgements

The Isala Sophia hospital, Zwolle the Netherlands
The Diving medical center, Den Helder the Netherlands
In2Motion Sports Bureau, Roermond the Netherlands.
University Medical center Amsterdam, the Netherlands
Erasmus Medical Center Rotterdam, the Netherlands
Ministry of defense testing center, Soesterberg the Netherlands
Wilhelmina Children's hospital, Utrecht the Netherlands
University of Applied Science Utrecht, the Netherlands
TOPsupport Geldrop, the Netherlands
InspanningLoont center, Utrecht the Netherlands
Maxima Medical Center Veldhoven, the Netherlands

SAMENVATTING

Een persoons maximale zuurstof opname vertegenwoordigt de functionele limiet van het lichaam om zuurstof op te nemen en af te leveren om aan de metabolische eisen van inspanning te voldoen. Deze waarde wordt erkent als de gouden-standaard voor aerobic fitheid en als een belangrijk component in het bepalen van gezondheidsrisico's. Voor deze risicostratificatie is het essentieel om accurate referentiewaarden te hebben van de gezonde populatie. Momenteel is er geen consensus in de wetenschappelijke literatuur wat betreft het de ontwikkeling van de maximale zuurstof opname gedurende iemands leven. Hierdoor ontstaat er een risico op inadequate risicostratificatie van een predictie formule.

Doelstelling

Het doel van de studie is om het type regressie model te bepalen voor de ontwikkeling van de maximale zuurstof opname met de toename van leeftijd, gedurende een persoons gehele leven binnen een cohort gezonde Nederlanders.

Methode

Het betreft een multicenter transversale studie uitgevoerd met de Low-lands Fitness Registery. Deze database bevat metingen van 11 instituten in Nederland, uitgevoerd tussen januari 2010 en december 2016. Generalized Additive Models zijn gebruikt als een semi-parametrische methode om het best passende regressie model te bepalen voor de maximale zuurstof opname met de variabelen leeftijd, gender, gewicht en lengte.

Resultaten

Een additief model met de variabelen gewicht, lengte en geslacht en een spline functie voor leeftijd met een interactie met geslacht is het best passende model voor de ontwikkeling van maximale zuurstof opname gedurende een persoons levensduur. Dit model heeft een adjusted R^2=0.641, Akaike Information Criterion=42386.51, Bayesian Information Criterion=42478.83.

Conclusie

De maximale zuurstof opname ontwikkeld zich gedurende een persoons levensduur middels een spline ontwikkeling. De ontwikkeling van maximale zuurstof opname gedurende iemands levensduur kan het beste gemodelleerd worden met een semi-parametrische regressie.

Klinische relevantie

Representatieve waarden voor aerobic fitness zijn nodig gezien het gebrek aan deze gegevens van de Nederlandse populatie. De momenteel gebruikte referentiewaarden kunnen mogelijk zorgen voor een onderschatting van de aerobic fitheid van de Nederlandse populatie. Hierdoor kunnen personen onterecht geclassificeerd worden als het hebben van een normale fitheid.

In toekomstig onderzoek, gericht op het opstellen van predictieformules voor maximale zuurstof opname, het gebruikte regressie model dient te bestaan uit een semi-parametrische model.