BMJ Open

Reference values for peak oxygen uptake: cross-sectional analysis of cycle ergometry-based cardiopulmonary exercise tests of 10090 adult German volunteers from the Prevention First Registry

Daniel Rapp,[1] Jürgen Scharhag,[2,3] Stefan Wagenpfeil,[1] Johannes Scholl[4]

[1]Institute for Medical Biometry, Epidemiology and Medical Informatics, Saarland University, Saarbrücken, Saarland, Germany
[2]Institute of Sports and Preventive Medicine, Saarland University, Saarbrücken, Saarland, Germany
[3]Department of Prevention and Sports Medicine, Klinikum rechts der Isar, Technical University Munich (TUM), Munich, Germany
[4]Dr. Scholl Prevention First GmbH, Rüdesheim am Rhein, Hessen, Germany

Correspondence to
Daniel Rapp;
daniel.rapp@gmx.net

ABSTRACT

Objective This study aims to construct quantile reference values for peak oxygen uptake ($\dot{V}O_{2peak}$) measured by cycle ergometry-based incremental cardiopulmonary exercise tests.

Design Cross-sectional study using quantile regressions to fit sex-specific and age-specific quantile curves. Exercise tests were conducted using cycle ergometry. Maximal effort in the exercise tests was assumed when respiratory exchange ratio ≥1.1 or lactate ≥8 mmol/L or maximal heart rate ≥90% of the age-predicted maximal heart rate. This was assessed retrospectively for a random subsample with an a priori calculated sample size of n=252 participants.

Setting A network of private outpatient clinics in three German cities recorded the results of cycle ergometry-based cardiopulmonary exercise tests to a central database (Prevention First Registry) from 2001 to 2015.

Participants 10090 participants (6462 men, 3628 women) from more than 100 local companies volunteered in workplace health promotion programmes. Participants were aged 21 to 83 years, were free of acute complaints and had primarily sedentary working environments.

Main outcome measure Peak oxygen uptake was measured as absolute $\dot{V}O_{2peak}$ in litres of oxygen per minute and relative $\dot{V}O_{2peak}$ in millilitres of oxygen per kilogram of body mass per minute.

Results The mean age for both men and women was 46 years. Median relative $\dot{V}O_{2peak}$ was 36 and 30 mL/kg/min at 40 to 49 years, as well as 32 and 26 mL/kg/min at 50 to 59 years for men and women, respectively. An estimated proportion of 97% of the participants performed the exercise test until exertion.

Conclusions Reference values and nomograms for $\dot{V}O_{2peak}$ were derived from a large sample of preventive healthcare examinations of healthy white-collar workers. The presented results can be applied to participants of exercise tests using cycle ergometry who are part of a population that is comparable to this study.

Strengths and limitations of this study

► This study is based on more than 10000 participants which is one of the largest samples in its field.
► Reference values were plotted as nomograms using quantile curves, and an interactive web application (www.uks.eu/vo2peak) was created to facilitate the interpretation.
► Subgroup analyses can be performed interactively, and CIs are given for the reference values.
► Criteria for maximal effort were recorded in medical records but not in the main study data base. Therefore, maximal effort was validated only in a random sample of n=252 participants.
► As in other reference values, differences between the present sample and its population might indicate selection bias.

INTRODUCTION

A comprehensive body of evidence shows that a low cardiorespiratory fitness (CRF) is a strong, independent and modifiable risk factor for a plethora of health threats such as premature death, cardiovascular disease,[1 2] diabetes mellitus[3] and neoplasia.[4] On the other hand, CRF can be improved by physical activity and exercise, which makes it a crucial target for health interventions.[5] Hence, the assessment of CRF should be a key component of clinical practice in preventive healthcare check-ups.[6] Despite its high predictive power, CRF has not been included in widely used cardiovascular risk models such as Framingham,[7] European Systematic Coronary Risk Evaluation (SCORE),[8] Joint British Society 3 (JBS3)[9] or Prospective Cardiovascular Munster Score (PROCAM).[10] Therefore, it is particularly important to incorporate the

measurement of CRF in preventive medicine beyond commonly used risk factors such as tobacco smoking or diabetes mellitus.

CRF is typically represented by the peak oxygen uptake ($\dot{V}O_{2peak}$) during incremental exercise tests. The participant is exposed to an incremental work rate until exhaustion is reached at the point of maximal volitional work rate. $\dot{V}O_{2peak}$ can then be estimated by prediction equations.[6 11] Still, the gold standard remains cardiopulmonary exercise testing using respiratory gas exchange measurement.[6 12 13]

CRF has been shown to be strongly dependent on sex and age.[12] For this reason, the exercise test result of an individual is not meaningful per se but must be interpreted using sex-specific and age-specific reference values.

The goal of the present analysis was to generate sex-specific and age-specific reference values for cycle ergometry-based $\dot{V}O_{2peak}$, based on a sample of more than 10 000 participants from primary preventive healthcare check-ups in three German cities. We constructed nomograms and created an interactive web application for visualisation.

METHODS
Study design and participants
A network of private outpatient clinics (Prevention First) recorded the results of preventive healthcare check-ups for participants aged 21–83 years old in three German cities (Rüdesheim, Frankfurt, Munich) from 2001 to 2015. A proportion of 95% of the participants were acquired in the course of workplace health promotion programmes. These programmes were performed in cooperation with more than 100 local companies from the private sector such as mid-sized companies, banks, insurance companies or business consulting. No participants in this group were civil servants or worked in the public sector. The other 5% comprised direct payers or persons with private health insurance. Overall, the majority of this study population consisted of white-collar workers and employees with office jobs and a primarily sedentary working environment.

Exercise tests were performed according to guidelines.[13–15] All participants were evaluated prior to exercise by an experienced physician. Pre-exercise evaluation included anamnesis, physical examination, resting ECG and laboratory tests. If a participant had no contraindications such as hypertensive crisis, acute infections or orthopaedic impairments,[15] the exercise test was performed with the goal of reaching exhaustion.

Pseudonymised data were recorded in a central database (Prevention First Registry). Participants were only included if they provided informed consent to use their data for scientific purposes. For the present cross-sectional analysis, only the first contact of a participant was considered. Follow-up examinations have not been included in the data.

Measurement of peak oxygen uptake
We performed incremental maximal exercise tests to assess $\dot{V}O_{2peak}$ using calibrated, electronically broke cycle ergometers. Gas exchange measurement was conducted through breath-by-breath analysis using the Ganshorn PowerCube system (Ganshorn Medizin Electronic, Niederlauer, Germany). We analysed and recorded the results with Ganshorn LF8 V.8.5 and the previous versions of this software. All three study sites performed daily calibration of their gas exchange measurement systems according to the manufacturer's instructions. Cycle ergometers were calibrated once a year and met the German directives for medical devices. Three to four exercise tests were performed each day using the same calibrated system. The calibration process was certified by a DIN EN ISO 9001 quality management standard.

$\dot{V}O_{2peak}$ was defined as the mean $\dot{V}O_2$ over the last 10 s of the exercise test. Absolute $\dot{V}O_{2peak}$ was measured in litres of oxygen per minute and relative $\dot{V}O_{2peak}$ in millilitres of oxygen per kilogram of body mass per minute.

We used ramp protocols and multistage exercise protocols depending on the request for capillary lactate measurement. Multistage protocols were used if the assessment of blood lactate levels was needed and ramp protocols were used if there was no assessment of lactate levels. The cycle ergometers' work load was independent from pedalling frequency in a range between 60 and 100/min. The participants were therefore instructed to maintain a pedalling frequency higher than 70/min. The increments of work rate were adjusted to the participant's estimated level of fitness. When ramp protocols were applied, participants typically reached exhaustion after 12–18 min of incremental work rate. In multistage protocols, we aimed for 4–6 stages lasting 3 min each. Prior to each exercise test, participants completed a low-pedalling phase lasting 3 min. It was aimed to continue the exercise test until exhaustion at the maximal volitional work rate, unless there were medical indications for termination.[11] Yet, we did not verify a maximal effort using any end criteria in the main sample.

Criteria for maximal effort were recorded in medical records but not in the main study database. Therefore, a subsequent data acquisition was performed for a random sample of n=252 participants. The sample size of this random sample was estimated based on CIs for a single proportion of 80% reaching exhaustion using a two-sided alpha level of 0.05 and 1-β of 0.8. The result was then adjusted for the proportion of missing cases in the variables of interest. The random sample was used to acquire (1) end criteria for maximal effort, (2) type of exercise protocol and (3) maximal heart rate during the exercise test.

Maximal effort of the participant was defined when one of the following criteria was met: (1) capillary lactate levels were ≥8 mmol/L; (2) respiratory exchange ratio ≥1.1; (3) maximal heart rate ≥90% of the age-predicted maximal heart rate.[16 17] Age-predicted maximal heart rate was estimated using the equation 208–0.7*age of the participant

in years.[18] Participants who did not continue the exercise test until exhaustion were not excluded from the analysis. Therefore, we used the random sample to estimate if the inclusion of those cases significantly altered the reference values. We fitted median regression models in the random sample for both, absolute and relative $\dot{V}O_{2peak}$ as dependent variables and age, a binary variable for exhaustion (yes/no) and an interaction term of both as independent variables.

Measurement of sample characteristics

Body weight and body fat were measured using the body composition analysers Tanita TBF 410 (Tanita, Tokyo, Japan) and Lange Skinfold Calipers (Beta Technology, Cambridge, Maryland, USA). Skinfold thickness was assessed at three sites (males: chest, abdomen, thigh; females: triceps, supprailium, thigh) according to Jackson and Pollock (1985).[19] After at least 5 min of rest,

blood pressure was assessed in a sitting position by averaging two measurements at the upper right arm using certified manometric blood pressure devices (BOSO, Jungingen, Germany) and measuring to the nearest 2 mm Hg. Following an overnight fast, blood samples were drawn in the morning prior to the exercise test and were analysed by an accredited laboratory (study centres Rüdesheim and Frankfurt: Labor Dr. Riegel, Wiesbaden, Germany; study centre Munich: Synlab, Augsburg, Germany).

Statistical methods

All statistical analyses were performed using R software V.3.4.0.[20] We used the package compareGroups for analysing descriptive statistics.[21] Continuous variables were described by mean (SD) and categorical variables by absolute and relative frequencies. Comparisons between two groups were performed using Student's t-test for
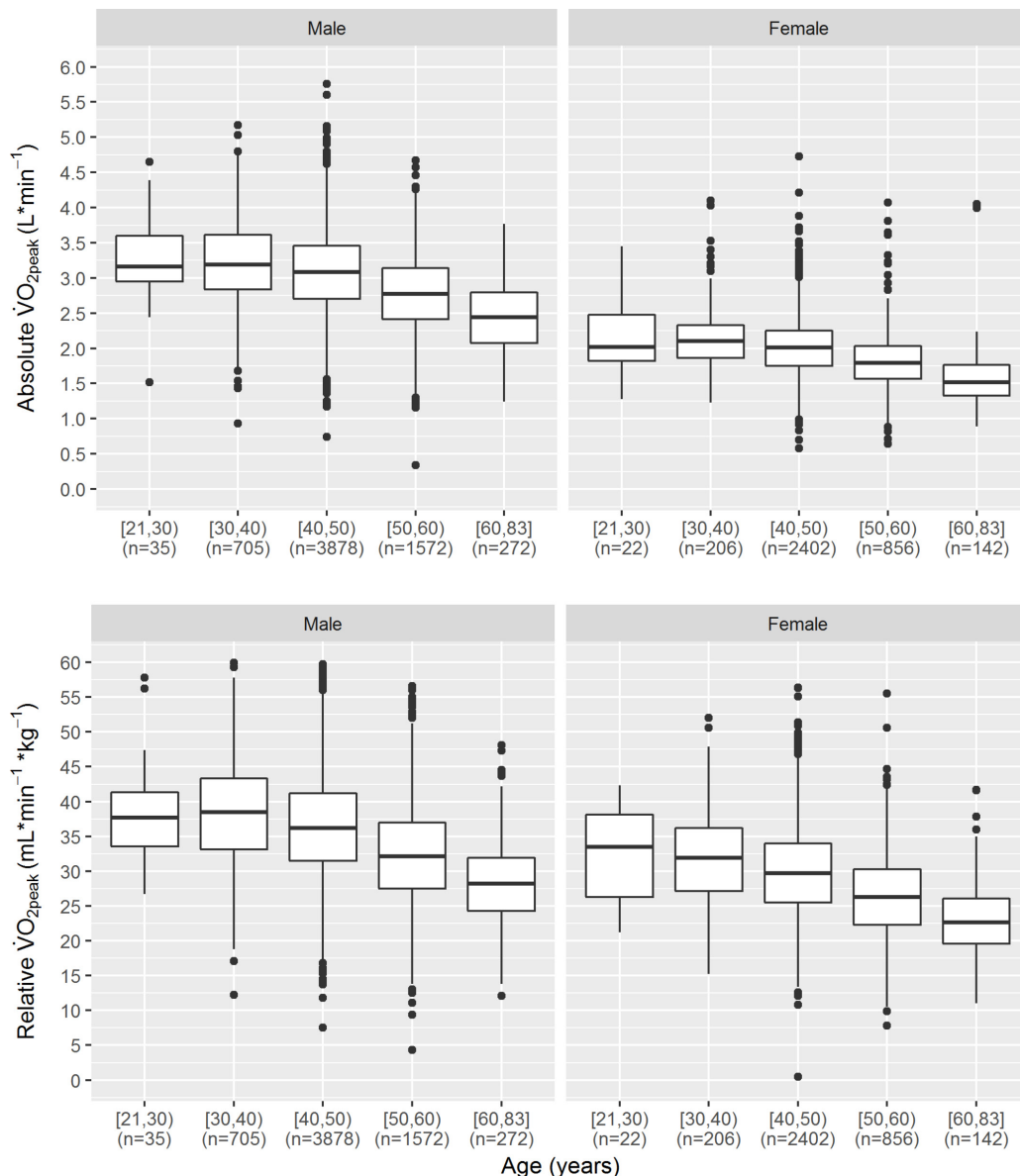

Figure 1 Box plots of absolute and relative $\dot{V}O_{2peak}$ by gender and age group. $\dot{V}O_{2peak}$, peak oxygen uptake.

Table 1 Characteristics of maximal exercise test participants

	Male, n=6462	Female, n=3628	P value
Age (years)	46.0 (7.09)	46.1 (6.87)	0.620
$\dot{V}O_{2peak}$			
Relative (mL/kg/min/)	35.4 (7.68)	28.9 (6.66)	<0.001
Absolute (L/min)	3.00 (0.60)	1.96 (0.43)	<0.001
Weight (kg)	85.9 (13.2)	69.1 (13.5)	<0.001
Height (cm)	181 (7.14)	168 (6.87)	<0.001
BMI (kg/m^2)	26.2 (3.63)	24.6 (4.67)	<0.001
Body fat (calliper) (%)	23.1 (6.27)	30.8 (7.50)	<0.001
Blood pressure			
Systolic (mm Hg)	128 (15.4)	120 (16.6)	<0.001
Diastolic (mm Hg)	83.1 (8.95)	78.3 (9.28)	<0.001
Blood glucose (mg/dL)	97.1 (14.5)	92.2 (14.5)	<0.001
HbA1c (%)	5.42 (1.43)	5.41 (1.52)	0.756
Total cholesterol (mg/dL)	216 (39.1)	211 (38.4)	<0.001
HDL cholesterol (mg/dL)	54.1 (12.4)	67.6 (15.1)	<0.001
LDL cholesterol (mg/dL)	136 (35.9)	124 (41.6)	<0.001
Triglycerides (mg/dL)	135 (88.1)	100 (63.7)	<0.001
No (%) of smoker	867 (13.5)	522 (14.5)	0.182
No (%) of former smoker	1634 (29.6)	910 (29.7)	0.928

Values are mean (SD) unless stated otherwise. P values were calculated using Student's t-test for two independent samples or χ^2 test as appropriate.
BMI, body mass index; HbA1c, glycated haemoglobin; HDL, high-density lipoprotein; LDL, low-density lipoprotein; $\dot{V}O_{2peak}$, peak oxygen uptake.

means and Pearson's χ^2 test for proportions. All 95% CIs were calculated using 10 000 bootstrap samples.

To describe the generalisability of our sample, we compared eligible characteristics of our population with the results of a study which represents the German population (DEGS1).[22–25] To ascertain comparability of the present study and the DEGS1, we performed direct age standardisation of our variables using the R package epitools.[26] The German population from 2011 was selected as the standard population.[27]

We computed quantile curves for nomograms by fitting quantile regressions using the R package quantreg.[28] Absolute or relative $\dot{V}O_{2peak}$ was used as the dependent variable and the age in years of the participant was used as the independent variable. Quantile regression was selected because it allows the estimation of conditional quantiles. In comparison to ordinary least square regression, which estimates conditional means, it was possible to provide detailed information about the distribution of peak oxygen uptake for a given age and to calculate CIs for the predicted values. Another approach would have been to calculate sample quantiles, but in this case, age would have to be aggregated into classes.

We also created an interactive web application to visualise the reference values with the use of the R package shiny.[29]

Our analyses were stratified for sex and adjusted for age, if appropriate. Statistical significance was assumed for p values of less than 0.05 or if 95% CIs did not overlap using a two-sided significance level of α=5%. In this epidemiological, explorative analysis, there was no correction of p values for multiple testing.

RESULTS
Description of study population
Overall, the results of 10 090 (6462 men, 3628 women) healthy participants from preventive healthcare examinations were eligible for the analysis and provided plausible, non-missing values for age and peak oxygen uptake.

The mean age was 46 years for both men and women. There were large differences between numbers of cases of different age classes (figure 1, online supplementary table 1). High numbers were observed in the age classes from 30 to 59 years and lower numbers in the marginal age classes. The highest number of cases was in the age group 40–49 years. In this particular age group, 3878 and 2402 participants were recorded for men and women, respectively. For the age group 21–29 years, there were comparatively few cases.

Peak oxygen uptake was significantly higher in men. Mean relative $\dot{V}O_{2peak}$ was 35 and 29 mL/kg/min for men and women, respectively (table 1). Furthermore, a decline in peak oxygen uptake among older participants was observed (figure 1).

There were also significant differences between our study population and the German population. In men, the proportions of smokers, overweight and obese participants were significantly lower compared with the DEGS1 study (table 2). Likewise, in women, the proportions of smokers, overweight, obese and hypertensive participants were significantly lower compared with the DEGS1 study.

The bivariate distributions of absolute as well as relative $\dot{V}O_{2peak}$ and age class are displayed in figure 1 and online supplementary tables 1 and 2. Figures 2 and 3 are nomograms including percentile curves. The nomograms can also be accessed as an interactive web application at www.uks.eu/vo2peak.

Description of exercise test modalities from random sample
Within our entire study population, we also drew a random sample of n=252. These participants did not

Table 2 Comparing characteristics of present study to results representing German population (DEGS1)

	Males		Females	
	PF Registry	DEGS1	PF Registry	DEGS1
Smoker	**14.5 (12.5 to 16.5)**	**26.1 (24.0 to 28.2)**	**15.1 (12.6 to 17.9)**	**21.4 (19.7 to 23.1)**
Former smoker	29.3 (26.9 to 31.9)	33.7 (31.9 to 35.5)	26.9 (23.9 to 30.2)	22.8 (21.4 to 24.2)
Overweight	**60.5 (57.9 to 63.0)**	**67.1 (65.0 to 69.2)**	**37.9 (34.6 to 41.3)**	**53.0 (50.8 to 55.1)**
Obese	**13.7 (12.0 to 15.6)**	**23.3 (21.2 to 25.4)**	**12.5 (10.3 to 15.0)**	**23.9 (22.0 to 25.9)**
Hypertension	35.7 (33.2 to 38.3)	33.3 (31.1 to 35.6)	**20.9 (18.2 to 23.8)**	**29.9 (28.1 to 31.9)**

Values are in per cent (95% CI). Results were directly age standardised using German population of 2011 to ascertain comparability.
Characteristics were defined according to DEGS1: Overweight, BMI ≥25 kg/m^2, obesity=BMI ≥30 kg/m^2, hypertension=systolic blood pressure ≥140 mm Hg or diastolic blood pressure ≥90 mm Hg.
BMI, body mass index; DEGS1, Studie zur Gesundheit Erwachsener in Deutschland[23–25]; PF Registry, Prevention First Registry.
Significant results are printed bold.

differ significantly from the entire study population for the variables sex, age, peak oxygen uptake, body mass index (BMI) and smoking status.

In the random sample, exercise tests were performed using the multistage protocols in 130/243 (54%, 95% CI 47% to 60%) cases, and using the ramp protocols in 113/243 (47%, 95% CI 40% to 53%). Nine cases did not provide any information about the protocol that was used. Mean maximal heart rate was 172 (15)/min in both men and women.

Maximal exhaustion was reached in 239/247 (97%, 95% CI 94% to 99%) participants. This proportion was 150/155 (97%, 95% CI 94% to 99%) in men and 89/92 (97%, 95% CI 93% to 100%) in women. Reasons for termination prior to exhaustion were mainly orthopaedic impairments, anxiety from wearing the mask for gas exchange measurement or muscular exhaustion due to a low level of fitness. We also tested whether median regressions were significantly altered by inclusion or exclusion of participants who did not meet the criteria for exhaustion. However, these median regressions could not be calculated for women due to low numbers of observations with no exhaustion. In men, all p values for the variable exhaustion (yes/no) were higher than 0.05. This means that inclusion of participants with poor effort did not alter the regression significantly in the random sample. The results of this regression have to be interpreted with caution because only five men met the criterion 'no exhaustion'.

Quantile regressions and nomograms

We fitted quantile regressions to our data to derive quantile reference curves. Peak oxygen uptake served as the dependent variable while age in years was the independent variable. Second-degree polynomial quantile regressions showed a better goodness of fit compared with age as a linear predictor. Therefore, quantile regressions were formulated as:

$$\dot{V}O_{2peak\ i,\ \tau} = \hat{\beta}_0(\tau) + \hat{\beta}_1(\tau) * age_i + \hat{\beta}_2(\tau) * age_i^2 + e_{i,\ \tau}$$

where τ denotes the τ-quantile, i=1,…,10 090 the i-th participant, age the age in years of the

participant, $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\beta}_2$ regression coefficients to be estimated and e the error term, respectively.

DISCUSSION

The presented quantile reference values for $\dot{V}O_{2peak}$ were derived from a sample of more than 10 000 participants who volunteered in preventive healthcare check-ups, primarily in the course of workplace health promotion programmes. To our knowledge, this currently constitutes the largest sample for $\dot{V}O_{2peak}$ reference values using cycle ergometry-based exercise tests. Data were acquired in three different German cities. All study centres adhered to predefined quality standards and performed daily calibration of the exercise test devices. The goal was to continue each exercise test until exhaustion, and this criterion was met by 97% of the men and 97% of the women in the subgroup (n=252). The mean maximal heart rate in this sample was 172/min for men and women. This value appears to be high in light of the mean age calculated in the whole sample (46 years) and might support the assumption that participants in this study exerted particularly high effort.

Other reference values

The reference values published by the Cooper Clinic (Dallas, Texas, USA)[14] are among the most commonly used and comprehensive reference values for $\dot{V}O_{2peak}$. The exercise tests to acquire those reference values were performed using treadmill ergometers and an indirect measurement of $\dot{V}O_{2peak}$ using prediction equations based on the achieved treadmill time.[30] Earlier publications by Hansen et al[31] as well as Jones et al[32] were based on rather low numbers of observations, which led to imprecise estimations.[12] Furthermore, Hansen et al used a sample of shipyard workers, which is a highly selected population.[31] An early systematic review collected and arranged normal standards that were published before 1990.[33] However, those results are now only relevant from a historical perspective.[12]

The above-mentioned shortcomings have been raised by the American Thoracic Society and by the American
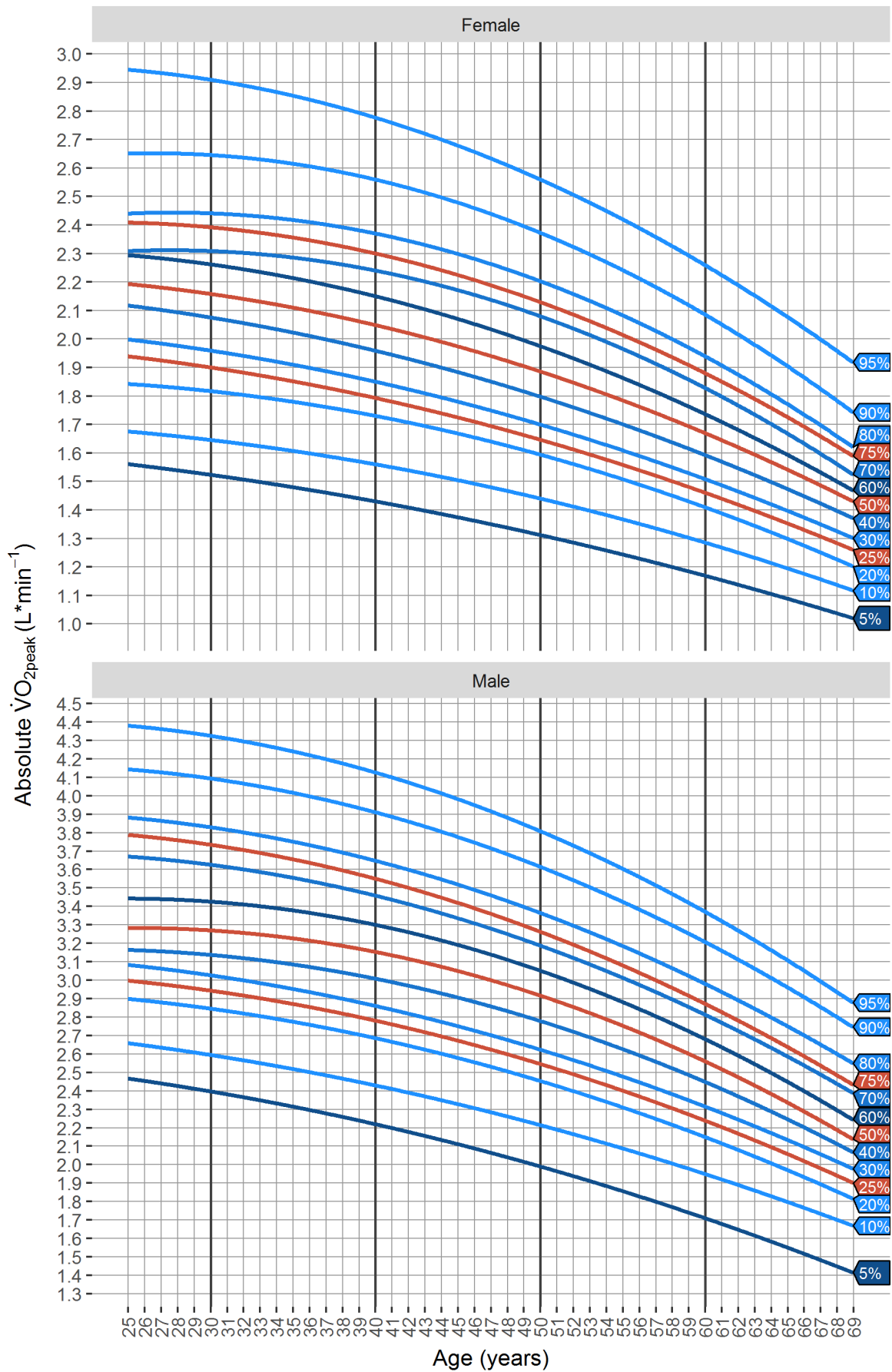
Figure 2 Nomograms of percentile reference values of absolute $\dot{V}O_{2peak}$ by gender. $\dot{V}O_{2peak}$, peak oxygen uptake.
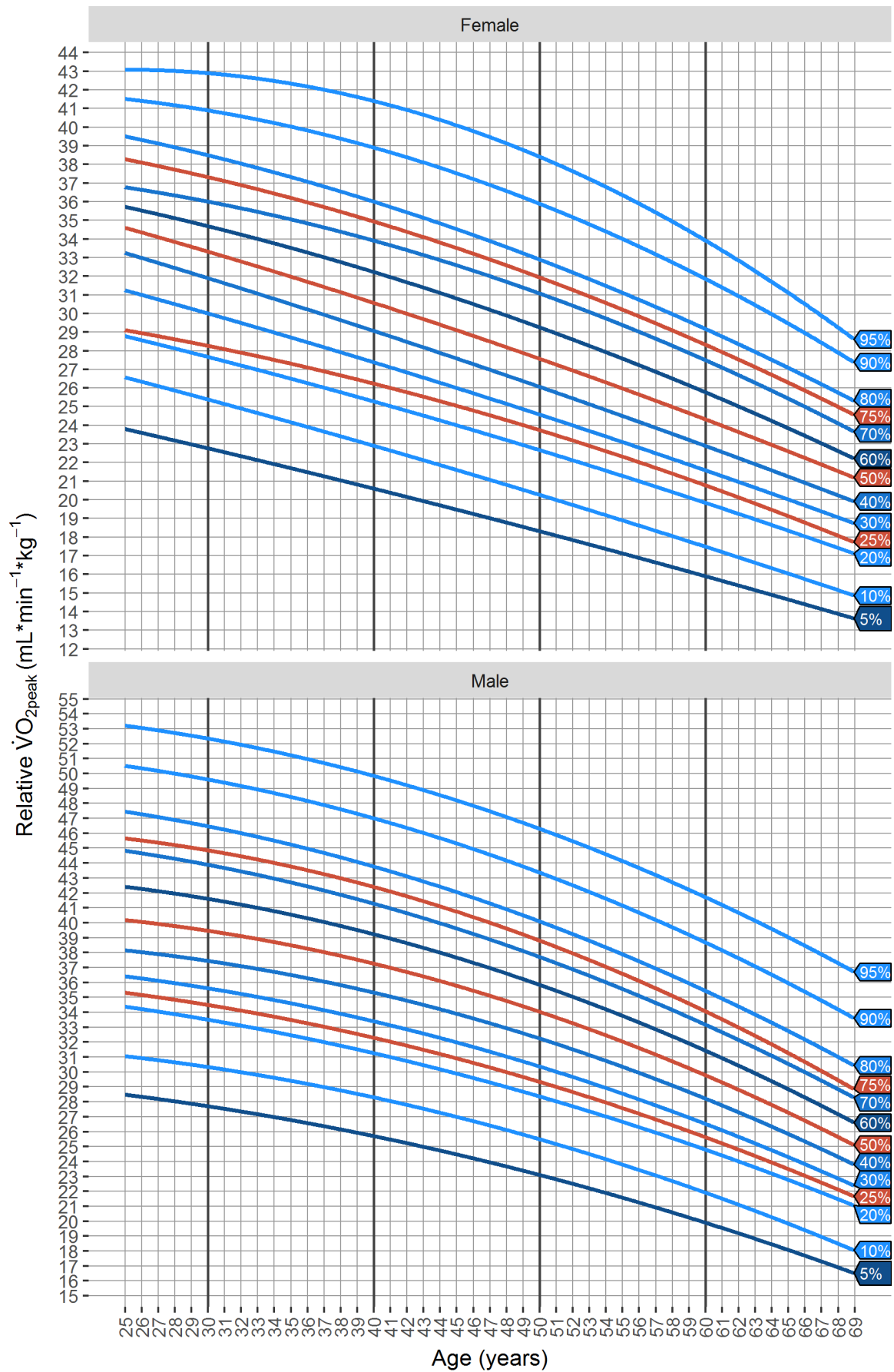
Figure 3 Nomograms of percentile reference values of relative $\dot{V}O_{2peak}$ by gender. $\dot{V}O_{2peak}$, peak oxygen uptake.
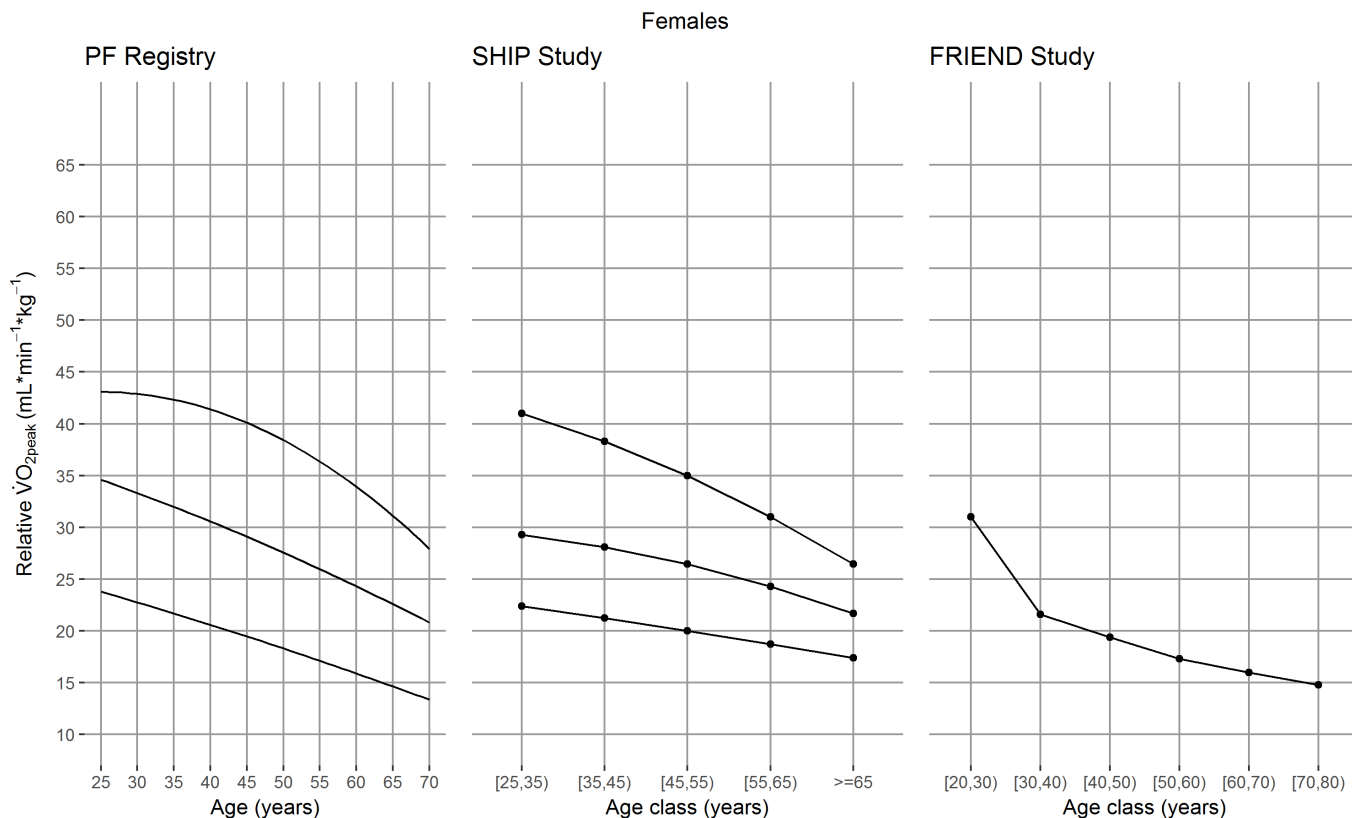
Figure 4 Cycle ergometry-based reference values for relative $\dot{V}O_{2peak}$ of females compared between three studies. PF Registry, SHIP Study[36] and FRIEND Study.[35] 25th, 50th and 75th percentiles are plotted. For the FRIEND Study, only the 50th percentile is plotted. FRIEND, Fitness Registry and the Importance of Exercise National Database; PF, Prevention First; SHIP, Study of Health in Pomerania; $\dot{V}O_{2peak}$, peak oxygen uptake.

College of Chest Physicians in their comprehensive statement on exercise tests in 2003.[15] They emphasised that valid and representative reference values were critical for the interpretation of CRF, but reliable reference values were lacking at that time in the USA. This issue has recently been addressed by an initiative that recorded data from several laboratories in the USA to a registry (Fitness Registry and the Importance of Exercise National Database (FRIEND)).[34 35] Reference quantiles were obtained for treadmill ergometry-based[34] as well as cycle ergometry-based[35] exercise tests using the results of 7783 and 4494 participants, respectively. The exercise tests that were recorded by the FRIEND study were performed in the course of exercise programmes or research studies.

Reference values for a German population were published in 2009 using data from a prospective, population-based study (SHIP (Study of Health in Pomerania) study).[36] A representative sample of 7008 adults was drawn from a north-eastern region of Germany. Due to non-responders and rigorous exclusion of smokers, obese participants and other factors, the final sample yielded 534 participants (253 men, 281 women) who were eligible for the exercise tests. Measures of exhaustion were not published in the described study.

A comprehensive systematic review that summarises past reference values for $\dot{V}O_{2peak}$ was conducted in 2014.[12] The authors established a grading system for critical evaluation of the quality of reference values. They emphasised that only 4/35 (11%) studies reached high quality, and only two of these were conducted after the year 2000.

Comparison of reference values

The exercise tests in our study were performed using cycle ergometers. Cycle and treadmill ergometers were also the most common choices in past studies. However, to select appropriate reference values, it has to be considered that the choice of ergometer has a large impact on the obtained reference values. Peak oxygen uptake measured by treadmill ergometers was assumed to be 5% to 10% higher compared with cycle ergometers as a larger muscle mass is involved in treadmill ergometry and cycle ergometry is often terminated due to localised muscle fatigue.[15] This effect appears to be even stronger when compared with the results of the FRIEND study. A 35-year-old man showed a median relative $\dot{V}O_{2peak}$ of 42 mL/kg/min using treadmill ergometer and 30 mL/kg/min using cycle ergometer.[34 35] Therefore, reference values should only be considered for interpretation if the type of ergometer in the performed exercise test is identical to those used for calculating the reference values.

Exercise test protocols of past studies were diverse or were not described.[14 35 36] However, the choice of exercise test protocol does not seem to have a strong effect on $\dot{V}O_{2peak}$.[37 38]
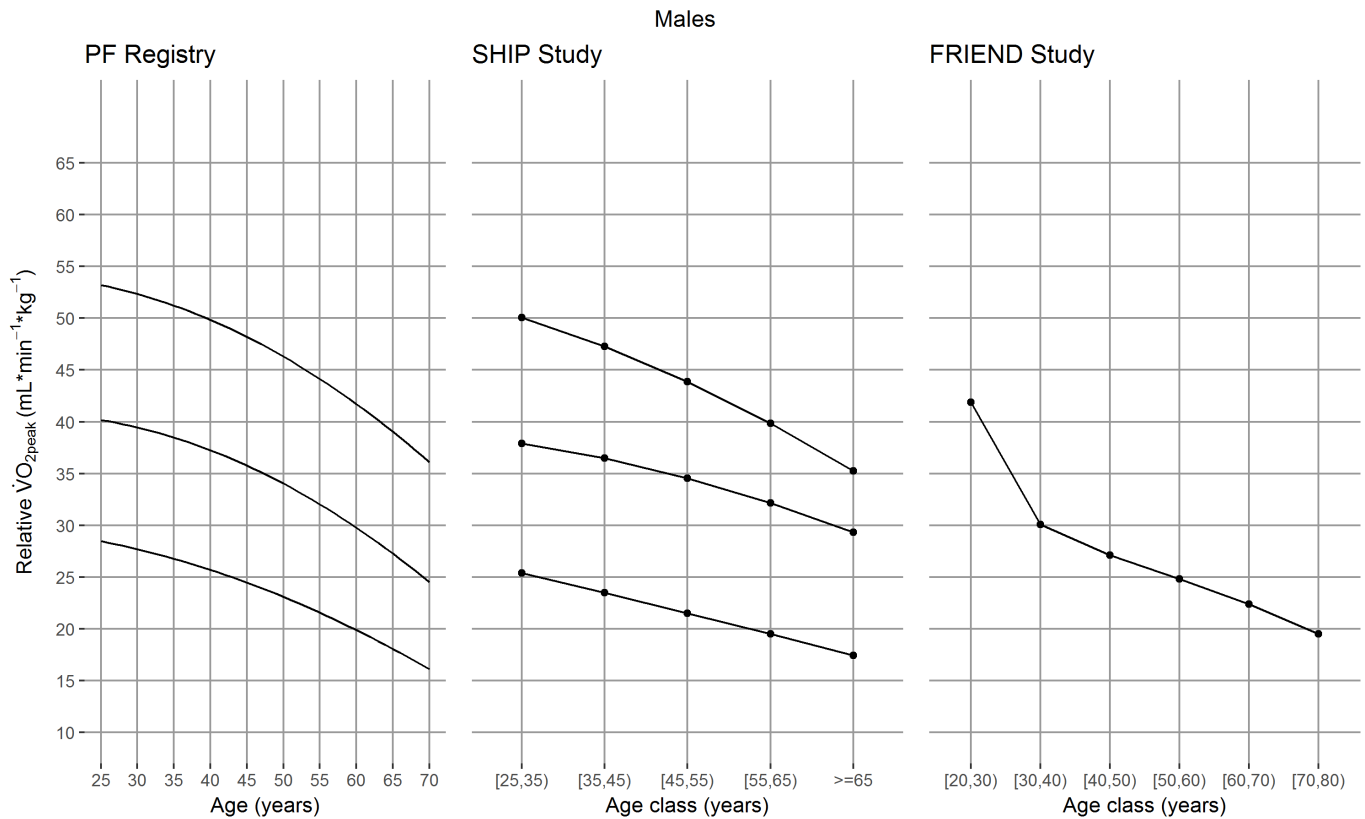
Figure 5 Cycle ergometry-based reference values for relative $\dot{V}O_{2peak}$ of males compared between three studies. PF Registry, SHIP Study[36] and FRIEND Study.[35] 25th, 50th and 75th percentiles are plotted. For the FRIEND Study, only the 50th percentile is plotted. FRIEND, Fitness Registry and the Importance of Exercise National Database; PF, Prevention First; SHIP, Study of Health in Pomerania; $\dot{V}O_{2peak}$, peak oxygen uptake.

Considering the methodological differences, our reference values were slightly higher compared with other cycle ergometry-based studies. The median relative $\dot{V}O_{2peak}$ of a 35-year-old man measured in mL/kg/min was 38 in our study, 36 in the SHIP study[36] (using BMI <25 kg/m²) and 30 in the FRIEND study.[35] The value presented by the Cooper Institute was 42, but this was obtained using treadmill ergometers and indirect measurement of $\dot{V}O_{2peak}$.[14] All mentioned studies treated age as a categorical predictor using 10-year age classes,[14 34–36] but $\dot{V}O_{2peak}$ changed distinctly within 10 years in other studies. This effect might be even stronger in higher age classes and can be observed in a Japanese sample in which relative $\dot{V}O_{2peak}$ of men between 50 and 59 years measured in mL/kg/min showed to be 29 and 26, respectively.[37] This effect was also observed in our study, where the estimated medians of men aged 50 and 59 years old were 34 and 30 mL/kg/min. Due to the differences within 10-year age groups, we considered age as a quantitative predictor measured in years. The reference values of the SHIP study, the cycle ergometry-based FRIEND study and our results are compared in figures 4 and 5.

Generalisability of the study sample
Our results were based on a sample of German white-collar workers with a predominantly sedentary working environment. This economic sector describes a large and increasing proportion of the population in Germany and in other industrialised countries.[39]

However, our study sample had significantly lower proportions of smokers, overweight and obese persons compared with the overall German population. These differences are likely due to our sample including primarily white-collar workers and also due to a selection of participants with a healthier lifestyle than the German population. A selection of healthy participants might yield reference values that are higher than in the whole population. As selection bias is hard to challenge and has also been observed in past population-based studies,[36 40] it is critical to quantify the extent of the selection. This might also be considered in future studies. In our study, a selection of individuals with normal weight and lower prevalence of smoking was addressed by conducting subgroup analyses with exclusion of smokers and obese subjects (www.uks.eu/vo2peak) and by reporting peak oxygen uptake standardised for body weight in mL/kg/min.

Strengths and weaknesses of the study
Strengths of the present study include the high number of observations from three different German cities. Based on this sample, it was possible to obtain reference values with high precision and narrow CIs. Exercise tests were performed by experienced personnel according to guidelines and predefined quality standards which yielded

reliable test results. In contrast to earlier studies that commonly used age in 10-year age classes, we used quantile regressions to create nomograms with age in years as an independent variable. Based on that, the exercise test results of an individual at a certain age can be interpreted more precisely and in light of the interindividual variability. Furthermore, nomograms and an interactive web application may help clinicians and participants of exercise tests to better understand the results.

A weakness of our study is that participants who terminated the exercise test prior to exhaustion were not excluded . This might decrease our reference values. It should also be acknowledged that $\dot{V}O_{2peak}$ was defined as the average over the last 10 s of the exercise test, which is lower than some recommendations for exercise testing.[12 13] This might be a source of inaccuracy in the assessment of $\dot{V}O_{2peak}$.

Furthermore, this study shares the limitations of earlier studies sourced from databases. Despite large numbers of observations, overall, the distribution of age was not uniform, as the majority of cases were in the age class of 30 to 59 years. Only 16% and 10% of the male and female participants, respectively, were <30 years or ≥60 years old. Consequently, these age groups showed lower precision and wider CIs.

Our sample was a selection of a homogenous group of workers and therefore was significantly different from the German population in some respects. It is likely that this sample was a selection of volunteers with a particularly healthy lifestyle, which perhaps increased the reference values. The question of how well the underlying sample represents a population has not been discussed extensively in the literature. However, this should be emphasised in future reference values.

CONCLUSIONS AND IMPLICATIONS FOR CLINICIANS

The reference values for peak oxygen uptake presented by this study may be used in populations that are comparable to our sample. Laboratories using cycle ergometry-based cardiopulmonary exercise tests can interpret their results precisely and with background information. The reference values have also been embedded into an interactive web application (www.uks.eu/vo2peak) with the goal of facilitating the interpretation of exercise tests in clinical practice and improving the communication of exercise test results to the participant.

REFERENCES

1. Kodama S, Saito K, Tanaka S, et al. Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women: a meta-analysis. *JAMA* 2009;301:2024–35.
2. Shah RV, Murthy VL, Colangelo LA, et al. Association of fitness in young adulthood with survival and cardiovascular risk: the coronary artery risk development in young adults (CARDIA) Study. *JAMA Intern Med* 2016;176:87–95.
3. Zaccardi F, O'Donovan G, Webb DR, et al. Cardiorespiratory fitness and risk of type 2 diabetes mellitus: a 23-year cohort study and a meta-analysis of prospective studies. *Atherosclerosis* 2015;243:131–7.
4. Schmid D, Leitzmann MF. Cardiorespiratory fitness as predictor of cancer mortality: a systematic review and meta-analysis. *Ann Oncol* 2015;26:272–8.
5. Kokkinos P. Physical activity, health benefits, and mortality risk. *ISRN Cardiol* 2012;2012:1–14.
6. Ross R, Blair SN, Arena R, et al. Importance of assessing cardiorespiratory fitness in clinical practice: a case for fitness as a clinical vital sign: a scientific statement from the American Heart Association. *Circulation* 2016;134:e653–99.
7. D'Agostino RB, Vasan RS, Pencina MJ, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. *Circulation* 2008;117:743–53.
8. Conroy RM, Pyörälä K, Fitzgerald AP, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J* 2003;24:987–1003.
9. JBS3 Board. Joint British Societies' consensus recommendations for the prevention of cardiovascular disease (JBS3). *Heart* 2014;100(Suppl 2):ii1–67.
10. Assmann G, Cullen P, Schulte H. Simple scoring scheme for calculating the risk of acute coronary events based on the 10-year follow-up of the prospective cardiovascular Münster (PROCAM) study. *Circulation* 2002;105:310–5.
11. Franklin BA, Balady GJ. American College of Sports Medicine. *ACSM's guidelines for exercise testing and prescription*. 6. ed. Philadelphia: Lippincott Williams & Wilkins, 2000. ISBN: 978-0-7817-2735-8.
12. Paap D, Takken T. Reference values for cardiopulmonary exercise testing in healthy adults: a systematic review. *Expert Rev Cardiovasc Ther* 2014;12:1439–53.
13. Myers J, Arena R, Franklin B, et al. Recommendations for clinical exercise laboratories: a scientific statement from the American Heart Association. *Circulation* 2009;119:3144–61.
14. Pescatello LS. American College of Sports Medicine, editors. *ACSM's guidelines for exercise testing and prescription*. 9th ed. Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins Health, 2014:ISBN: 978-1-60913-605-5.
15. American Thoracic SocietyAmerican College of Chest Physicians. ATS/ACCP Statement on cardiopulmonary exercise testing. *Am J Respir Crit Care Med* 2003;167:211–77.
16. Midgley AW, McNaughton LR, Polman R, et al. Criteria for determination of maximal oxygen uptake: a brief critique and recommendations for future research. *Sports Med* 2007;37:1019–28.
17. Scharhag-Rosenberger F, Schommer K. Die Spiroergometrie in der Sportmedizin. *Dtsch Z Für Sportmed* 2013;64:362–6.
18. Tanaka H, Monahan KD, Seals DR. Age-predicted maximal heart rate revisited. *J Am Coll Cardiol* 2001;37:153–6.
19. Jackson AS, Pollock ML. Practical assessment of body composition. *Phys Sportsmed* 1985;13:76–90.

20. R Development Core Team. R: a language and environment for statistical computing. 2008 http://www.R-project.org

21. Subirana I, Sanz H, Vila J. Building bivariate tables: the compareGroups Package for R. *J Stat Softw* 2014;57:1–16.

22. Kamtsiuris P, Lange M, Hoffmann R. Die erste Welle der Studie zur Gesundheit Erwachsener in Deutschland (DEGS1): Stichprobendesign, Response, Gewichtung und Repräsentativität. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz* 2013;56:620–30.

23. Lampert T, von der Lippe E, Müters S. Verbreitung des Rauchens in der Erwachsenenbevölkerung in Deutschland: Ergebnisse der Studie zur Gesundheit Erwachsener in Deutschland (DEGS1). *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz* 2013;56:802–8.

24. Mensink GBM, Schienkiewitz A, Haftenberger M. Übergewicht und Adipositas in Deutschland: Ergebnisse der Studie zur Gesundheit Erwachsener in Deutschland (DEGS1). *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz* 2013;56:786–94.

25. Neuhauser H, Thamm M, Ellert U. Blutdruck in Deutschland 2008–2011: Ergebnisse der Studie zur Gesundheit Erwachsener in Deutschland (DEGS1). *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz* 2013;56:795–801.

26. Aragon TJ. epitools: epidemiology tools. 2012. https://CRAN.R-project.org/package=epitools

27. Statistisches Bundesamt. Bevölkerung und Erwerbstätigkeit. Bevölkerungsfortschreibung auf Grundlage des Zensus 2011. 2016. https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/Bevoelkerungsstand/Bevoelkerungsfortschreibung2010130147004.pdf?__blob=publicationFile

28. Koenker R. quantreg: quantile regression. 2016. https://CRAN.R-project.org/package=quantreg

29. Chang W, Cheng J, Allaire J, *et al*. Shiny: web application framework for R. 2017. https://CRAN.R-project.org/package=shiny

30. Heyward VH. *Advanced fitness assessment and exercise prescription*. 5th ed. Champaign: IL: Human Kinetics, 2006. ISBN: 978-0-7360-5732-5.

31. Hansen JE, Sue DY, Wasserman K. Predicted values for clinical exercise testing. *Am Rev Respir Dis* 1984;129:S49–55.

32. Jones NL, Makrides L, Hitchcock C, *et al*. Normal standards for an incremental progressive cycle ergometer test. *Am Rev Respir Dis* 1985;131:700–8.

33. Shvartz E, Reibold RC. Aerobic fitness norms for males and females aged 6 to 75 years: a review. *Aviat Space Environ Med* 1990;61:3–11.

34. Kaminsky LA, Arena R, Myers J. Reference standards for cardiorespiratory fitness measured with cardiopulmonary exercise testing. *Mayo Clin Proc* 2015;90:1515–23.

35. 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;92:228–33.

36. Koch B, Schäper C, Ittermann T, *et al*. Reference values for cardiopulmonary exercise testing in healthy volunteers: the SHIP study. *Eur Respir J* 2009;33:389–97.

37. Itoh H, Ajisaka R, Koike A, *et al*. Heart rate and blood pressure response to ramp exercise and exercise capacity in relation to age, gender, and mode of exercise in a healthy population. *J Cardiol* 2013;61:71–8.

38. Yoon BK, Kravitz L, Robergs R. VO2max, protocol duration, and the VO2 plateau. *Med Sci Sports Exerc* 2007;39:1186–92.

39. Buckley JP, Hedge A, Yates T, *et al*. The sedentary office: an expert statement on the growing case for change towards better health and productivity. *Br J Sports Med* 2015;49:1357–62.

40. Edvardsen E, Hansen BH, Holme IM, *et al*. Reference values for cardiorespiratory response and fitness on the treadmill in a 20- to 85-year-old population. *Chest* 2013;144:241–8.