O'Donovan, G. et al. (2016) Non-exercise equations to estimate fitness in white European and South Asian men. *Medicine and Science in Sports and Exercise*, 48(5), pp. 854-859. (doi:10.1249/MSS.0000000000000836)

This is the author's final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

http://eprints.gla.ac.uk/113163/

Deposited on: 07 December 2015

Enlighten – Research publications by members of the University of Glasgow
http://eprints.gla.ac.uk

Non-exercise equations to estimate fitness in white European and South Asian men

Gary O'Donovan,[1] Kishan Bakrania,[1] Nazim Ghouri,[2] Thomas Yates,[1] Laura J Gray,[3] Mark Hamer,[4] Emmanuel Stamatakis,[5] Kamlesh Khunti,[1] Melanie Davies,[1] Naveed Sattar,[2] and Jason M. R. Gill[2]

[1]*University of Leicester, Diabetes Research Centre, Leicester Diabetes Centre, Leicester General Hospital, Leicester LE5 4PW, United Kingdom.* [2]*BHF Glasgow Cardiovascular Research Centre, Institute of Cardiovascular and Medical Sciences, College of Medical, Veterinary, and Life Sciences, University of Glasgow, Glasgow G12 8TA, United Kingdom.* [3]*University of Leicester, Department of Health Sciences, Leicester LE1 7RH, United Kingdom;* [4]*National Centre for Sport & Exercise Medicine – East Midlands, Loughborough University, Loughborough LE11 3TU, United Kingdom.* [5]*Charles Perkins Centre, Faculty of Health Sciences, University of Sydney, Sydney, Australia.*

Correspondence: Dr Gary O'Donovan, Leicester Diabetes Centre, Leicester General Hospital, Leicester LE5 4PW, United Kingdom. Phone: +44(0)116 258 8630. Email: go36@le.ac.uk

Abstract

Purpose: Cardiorespiratory fitness is a strong, independent predictor of health, whether it is measured in an exercise test or estimated in an equation. The purpose of this study was to develop and validate equations to estimate fitness in middle-aged white European and South Asian men.

Methods: Multiple linear regression models (n=168, including 83 white European and 85 South Asian men) were created using variables that are thought to be important in predicting fitness (VO_2 max, $mL{\cdot}kg^{-1}{\cdot}min^{-1}$): age (years); BMI ($kg{\cdot}m^{-2}$); resting heart rate ($beats{\cdot}min^{-1}$); smoking status (0=never smoked, 1=ex or current smoker); physical activity expressed as quintiles (0=quintile 1, 1=quintile 2, 2=quintile 3, 3=quintile 4, 4=quintile 5), categories of moderate- to vigorous-intensity physical activity (0=<75 $min{\cdot}wk^{-1}$, 1=75-150 $min{\cdot}wk^{-1}$, 2=>150-225 $min{\cdot}wk^{-1}$, 3=>225-300 $min{\cdot}wk^{-1}$, 4=>300 $min{\cdot}wk^{-1}$), or minutes of moderate- to vigorous-intensity physical activity ($min{\cdot}wk^{-1}$); and, ethnicity (0=South Asian, 1=white). The leave-one-out-cross-validation procedure was used to assess the generalizability and the bootstrap and jackknife resampling techniques were used to estimate the variance and bias of the models.

Results: Around 70% of the variance in fitness was explained in models with an ethnicity variable, such as: VO_2 max = 77.409 - (age*0.374) – (BMI*0.906) – (ex or current smoker*1.976) + (physical activity quintile coefficient) – (resting heart rate*0.066) + (white ethnicity*8.032), where physical activity quintile 1 is 1, 2 is 1.127, 3 is 1.869, 4 is 3.793, and 5 is 3.029. Only around 50% of the variance was explained in models without an ethnicity variable. All models with an ethnicity variable were generalizable and had low variance and bias.

Conclusion: These data demonstrate the importance of incorporating ethnicity in non-exercise equations to estimate cardiorespiratory fitness in multi-ethnic populations.

Key words: Physical fitness; exercise test; linear models; validation studies.

Introduction

Cardiorespiratory fitness reflects the ability of the lungs and cardiovascular system to transport oxygen and the ability of the tissues and organs to extract and use oxygen during sustained exercise. Cardiorespiratory fitness is a strong, independent predictor of morbidity and mortality when measured during a maximal exercise test (4, 13). However, the measurement of cardiorespiratory fitness is not deemed feasible in many healthcare settings because it is expensive, time-consuming, and requires trained personnel (12, 16, 20). Cardiorespiratory fitness can be estimated using equations containing variables that might be readily available, such as age, body mass index (BMI), resting heart rate, and physical activity (12, 15). When cardiorespiratory fitness is estimated from non-exercise testing equations, it is associated with all-cause and cardiovascular disease mortality (16, 20). The existing non-exercise testing equations were developed in samples of predominantly white men and women and may not be generalizable to others (12, 15). Cardiorespiratory fitness is significantly different when measured during a maximal exercise test in white men and men of South Asian descent matched for age and BMI; and, the lower cardiorespiratory fitness of South Asian men cannot be explained by their lower physical activity (7). The purpose of this study was to develop and validate equations to estimate cardiorespiratory fitness in white men and South Asian men.

Methods

Data source

Non-exercise testing cardiorespiratory fitness models were developed using data from a cross-sectional study of 100 white European and 100 South Asian men matched for age and BMI living in Glasgow, UK (6). Volunteers were without coronary heart disease, cerebrovascular disease, peripheral vascular disease, or known diabetes. Whites reported having two parents of white European origin and South Asians reporting having two parents of Indian, Pakistani, Bangladeshi, or Sri Lankan origin. Age and smoking status were reported by the participant, BMI was determined by a trained investigator, resting heart rate was assessed by ECG, and physical activity was assessed by accelerometer (GT3X+ or ActiTrainer, ActiGraph, Florida, USA). Cardiorespiratory fitness was assessed during a maximal exercise, which included a continuous incremental uphill walking protocol (21). Respiratory gasses were measured using the Douglas bag technique and cardiorespiratory fitness was expressed as maximal oxygen consumption (VO_2 max, $mL \cdot kg^{-1} \cdot min^{-1}$) (6). The study was approved by West of Scotland Research Ethics Committee and all participants gave written informed consent.

Model development

A listwise deletion approach was used for missing data. Multiple linear regression models were created using variables that are thought to be important in predicting cardiorespiratory fitness (6, 16, 20): age (years); BMI ($kg \cdot m^{-2}$); resting heart rate ($beats \cdot min^{-1}$); smoking status (0=never smoked, 1=ex or current smoker); physical activity expressed as quintiles (0=quintile 1, 1=quintile 2, 2=quintile 3, 3=quintile 4, 4=quintile 5), categories of moderate- to vigorous-intensity physical activity (0=<75 $min \cdot wk^{-1}$, 1=75-150 $min \cdot wk^{-1}$, 2=>150-225 $min \cdot wk^{-1}$, 3=>225-300 $min \cdot wk^{-1}$, 4=>300 $min \cdot wk^{-1}$), or minutes of moderate- to vigorous-intensity physical activity ($min \cdot wk^{-1}$); and, ethnicity (0=South Asian, 1=white). The outcome variable was VO_2 max ($mL \cdot kg^{-1} \cdot min^{-1}$). Three models were created without an ethnicity variable: VO_2 max = age + BMI + heart rate + smoking status + activity quintile (model 1); VO_2 max = age + BMI+ heart rate + smoking status + activity category (model 2); VO_2 max = age + BMI + heart rate + smoking status + activity minutes (model 3). Three models were created with an ethnicity variable: VO_2 max = age + BMI + heart rate + smoking status + activity quintile + ethnicity (model 4); VO_2 max = age + BMI + heart rate + smoking status + activity category + ethnicity (model 5); VO_2 max = age + BMI + heart rate + smoking status + activity minutes + ethnicity (model 6). Physical activity was expressed as a categorical variable or a continuous variable in the equations because physical activity is usually expressed as a categorical variable or a continuous variable in observational and experimental studies (10). The regression coefficients are reported along with their corresponding standard errors, 95% confidence intervals, and P values. The R^2 statistic and the adjusted R^2 value (which accounts for the number of parameters in the model) are reported in order to indicate the proportion of total variance in VO_2 max explained by the model. The root mean squared error (RMSE) is also reported because it provides a measure of the error between the observed values and the predicted values.

Checking assumptions, collinearity and goodness-of-fit

The assumption of constancy of variance was checked by plotting the standardised residuals against the predicted values. Constancy of variance, or homoscedasticity, was assumed to exist if the spread of residuals was constant. The assumption of normality was checked using normal probability plots. Normality was assumed to exist if the standardized residuals were normally distributed. The assumption of linearity was checked by plotting the standardised residuals against each of the covariates; and, polynomial terms up to the order of 3 were manually added to the models and tested for significance following the principles of parsimony (data and graphs not shown). The variance inflation factor (VIF) was used to check for collinearity between any of the covariates in a model. Goodness-of-fit of the model was checked by plotting the observed VO_2 max values against the predicted ones.

Model validation

Cross-validated models tend to exhibit far better generalizability (out-of-sample performance) than conventionally fitted models (18). In the present study, the leave-one-out-cross-validation procedure was used to assess the predictive performance of the models on 'unseen' data: a training (n-1 observations) and testing analysis (1 observation) were implemented (n times, with a different observation left out each time). The R^2 and RMSE of the original models were considered alongside the leave-one-out-cross-validation R^2 and RMSE. Bootstrapping is an appropriate way to validate a model in the absence of a large second dataset (5). In the present study, the bootstrap and jackknife resampling techniques were used to estimate the variance and bias of the models (5). The models were bootstrapped with the use of Monte Carlo simulations. A bootstrap program was written to resample observations with replacement from the data. The number of samples was increased until convergence was seen. The regression coefficients, R^2 statistic and RMSE of the original models were considered alongside the corresponding bootstrapped standard errors and non-parametric 95% confidence intervals based on percentiles. The regression coefficients, standard errors and 95% confidence intervals of the original models were considered alongside the corresponding jackknife standard errors and confidence intervals. Stata (version 13.1., Stata Corp.) was used for model development, model checking, and model validation.

Results

Physical activity was missing in 31 men and cardiorespiratory fitness was missing in one man. Therefore, participants in the present study were 83 white European men and 85 South Asian men. Age was 50±7 (40-69) years in the present sample and the original sample (mean±SD(range)). Body mass index was 27.1±4.0 (17.7-46.7) $kg \cdot m^{-2}$ in the present sample and 27.0±4.1 (17.8-46.7) $kg \cdot m^{-2}$ in the original sample. Resting heart rate was 64±9 (45-94) $beats \cdot min^{-1}$ in the present sample and 64±9 (41-94) $beats \cdot min^{-1}$ in the original sample. Sixty eight per cent reported never smoking in the present sample and the original sample. Moderate- to vigorous-intensity physical activity was 238±160 (14-817) $min \cdot wk^{-1}$ in the present sample and the original sample. Table 1 shows the characteristics of participants in the present sample by ethnicity.

No assumptions were violated, there was no evidence of collinearity, and all models were a good fit. Figure 1 shows the fit of a model without an ethnicity variable and Figure 2 shows the fit of the same model with an ethnicity variable. Around fifty per cent of the variance in cardiorespiratory fitness was explained by the models without an ethnicity variable (n=168 in all models). Fifty-four per cent of the variance in cardiorespiratory fitness was explained by model 1, the model containing age, BMI, smoking status, resting heart rate, and physical activity quintile (Table 2). Fifty four per cent of the variance in cardiorespiratory fitness was explained by model 2, a similar model containing physical activity category (Table S1, Supplemental Digital Content). Fifty three per cent of the variance in cardiorespiratory fitness was explained by model 3, a similar model in which physical activity was expressed as minutes (Table S2, Supplemental Digital Content).

Around seventy per cent of the variance in cardiorespiratory fitness was explained by the models with an ethnicity variable (n=168 in all models). Seventy-one per cent of the variance in cardiorespiratory fitness was explained by model 4, the model containing age, BMI, smoking status, resting heart rate, physical activity quintile, and ethnicity (Table 3). Similar amounts of variance were explained in models 5 and 6, similar models in which physical activity was expressed as categories or minutes (Tables S3 and S4, Supplemental Digital Content).

The leave-one-out-cross-validation procedure showed small differences in the observed and cross-validated sets in models 4, 5 and 6, suggesting that these models are generalizable (Table 4). For example, the observed R^2 value was 0.71, the leave-one-out-cross-validation R^2 value was 0.68, and the R^2 difference was -5.4% in model 4. The bootstrapping program showed that there was low variance and low bias in model 4, the model containing age, BMI, smoking status, resting heart rate, physical activity quintile, and ethnicity (Table S5, Supplemental Digital Content). There was evidence of convergence with replication. For example, the R^2 value was 0.71 in the original model (Table 3) and the standard error was 0.035 (95% confidence interval: 0.661, 0.792) with 500 replications, 0.036 (0.654, 0.795) with 10,000 replications, and 0.036 (0.654, 0.795) with 40,000 replications (Table S5, Supplemental Digital Content). The bootstrapping program showed that there was also low variance and low bias in models 5 and 6, similar models in which physical activity was expressed as categories or minutes (data not shown). The jackknife standard errors were observed to be similar to the original standard errors and the jackknife 95% confidence intervals did not change the importance or statistical significance of any of the variables in the models, which suggests that there was no over-fitting of the data and that the models with an ethnicity variable are reliable (Table S6, Supplemental Digital Content, shows the results for model 4, for example).

Discussion

The present study demonstrates the importance of incorporating ethnicity in non-exercise equations to estimate cardiorespiratory fitness in multi-ethnic populations. The equations containing age, BMI, smoking status, resting heart rate, physical activity and ethnicity explain much of the variance in cardiorespiratory fitness and are generalizable and reliable.

Existing non-exercise testing equations were developed in samples of predominantly white men and women and may not be generalizable to others (12, 15). The equations in the present study were developed using data from a study of white European and South Asian men living in the UK (6). In that study, Ghouri and colleagues (6) measured a difference in VO_2 max in whites and South Asians of 8.24 $mL \cdot kg^{-1} \cdot min^{-1}$ (95% confidence interval: 6.33, 10.15). Ghouri and colleagues (6) concluded that the lower VO_2 max values of South Asians could not be explained by their lower physical activity levels. Ghouri and colleagues (6) noted that South Asians had lower fitness levels at all activity levels and they suggested that there might be innate differences in fitness between whites and South Asians. In the present study, equations containing age, BMI, smoking status, resting heart rate and physical activity explained around 50% of the variance in VO_2 max and similar equations with an ethnicity variable explained around 70% of the variance in VO_2 max. These data demonstrate the importance of using an ethnicity variable to help estimate VO_2 max in non-exercise testing equations.

The non-exercise testing equations with an ethnicity variable perform favourably in comparison to existing equations. Consider model 4 in the present study, for example. R-squared was 0.72 in model 4, 0.65 in the NASA equation, 0.60 in the ACLS equation, and 0.58 in the ADNFS equation (NASA is National Aeronautics and Space Administration; ACLS is Aerobics Centre Longitudinal Study; ADNFS is Allied Dunbar National Fitness Survey) (12). The cross-validated R^2 value indicates how well an equation might predict fitness in an unseen dataset. Importantly, the cross-validated R^2 was 0.67 in model 4 in the present study, 0.58 and 0.56 in the NASA equation, 0.64 and 0.55 in the ACLS equation, and 0.58 and 0.52 in the ADNFS equation (these values were obtained by squaring the cross-validity correlations reported in Table 6 of Jurca and colleagues' paper (12)). These data suggest that the equations with an ethnicity variable would better predict cardiorespiratory fitness in an unseen dataset than existing equations.

Wier and colleagues (22) added an ethnicity variable to an early version of the NASA equation. They found that the addition of a nominal ethnicity variable raised R^2 by 0.008 in a model containing age, gender, physical activity and waist girth, by 0.006 in a model containing age, gender, physical activity and per cent body fat, and by 0.004 in a model containing age, gender, physical activity and BMI (all $P<0.001$). Wier and colleagues (22) suggested that ethnicity explained little more of the variance in cardiorespiratory fitness because the sample of 140 non-whites was considerably smaller than the sample of 2417 whites. It is also possible that the blacks, Hispanics, and Asian-Pacific Islanders who made up the non-whites were different ethnic groups. The families of the non-whites in the present study were all from South Asia, a region of high diabetes prevalence (11).

It is clear that the best performing of the equations in the present study were those containing an ethnicity variable. These equations contain other variables that might also be readily available in many healthcare settings: age, BMI, smoking status, resting heart rate, and physical activity. There are many ways of expressing physical activity outcomes in observational and experimental studies and it is noteworthy that these equations explained much of the variance in cardiorespiratory fitness, whether physical activity was expressed as quintiles, categories or minutes. The leave-one-out-cross-validation procedure showed that these equations are generalizable and the bootstrap and jackknife resampling techniques showed that there is low variance and bias in these equations.

It has been argued that models should include variables that are thought to be important from the literature, whether or not they reach statistical significance in a particular dataset (2). The literature suggests age, BMI, smoking status, resting heart rate, physical activity and ethnicity are important predictors of cardiorespiratory fitness (6, 16, 20). Accordingly, these variables were retained in models in the present study. Resting heart rate was a statistically significant predictor of cardiorespiratory fitness in model 1, model 2, model 3, and model 6 (all $P<0.05$). Resting heart rate was not a statistically significant predictor in model 4 ($P=0.091$) or model 5 ($P=0.064$). Resting heart rate was not removed from any model because the removal of important variables might lead to unreliable models (2).

Non-exercise testing equations developed in white men and women (12, 15) have been shown to predict cardiovascular and all-cause mortality in whites (16, 20). In a study of 32,319 adults aged 35 to 70 years at baseline, Stamatakis and colleagues (20) found that non-exercise testing cardiorespiratory fitness predicted cardiovascular and all-cause mortality during nine years of follow up after adjustment for potential confounders. In a study of 20,112 adults aged 20 to 60 years at baseline, Nes and colleagues (16) found that non-exercise testing cardiorespiratory fitness predicted cardiovascular and all-cause mortality during 24 years of follow up after adjustment for potential confounders. Further research is required to determine whether non-exercise testing cardiorespiratory fitness predicts mortality in multi-ethnic populations, such as South Asians. Insulin resistance is higher in South Asian men than white men and it was recently reported that directly measured cardiorespiratory fitness explained more than two thirds of the ethnic difference in insulin resistance (lower cardiorespiratory fitness explained 68%, lower physical activity explained 29%, and greater adiposity explained 52% of the ethnic variance in HOMA) (6). These data suggest that it may be particularly important to measure or to estimate cardiorespiratory fitness in South Asian men.

This study has some limitations. The development and validation dataset was drawn from a convenience sample that may not be representative (6); however, physical activity and cardiorespiratory fitness levels were similar in other studies of white European and South Asian men (7). Questionnaires and accelerometers have their advantages and disadvantages (17), but questionnaires were not used in the original study of Ghouri and colleagues (6). Questionnaires might be more feasible in many healthcare settings and model 1 (without an ethnicity variable) and model 4 (with an ethnicity variable) might be used with questionnaires that provide a summary measure of physical activity that can be divided into quintiles, such as the Baecke questionnaire (kcal·week^{-1}) (1), the Five-City Project 7-Day Physical Activity Recall Questionnaire (kcal·day^{-1}) (19), or the ACLS 7-Day Recall questionnaire (hr·wk^{-1}) (14). The dataset was made up of men and more research is required to develop and validate equations for women. The non-exercise testing cardiorespiratory fitness equations were developed and validated in the same dataset; however, bootstrapping is an appropriate way to validate a model in the absence of a large second dataset (5). The cardiorespiratory fitness of white and South Asian men has only been reported in three other studies, in which sample size ranged from 40 to 92 (3, 8, 9).

To our knowledge, this is the first study to develop and validate equations to estimate cardiorespiratory fitness in white men and Sough Asian men. This study shows the importance of incorporating ethnicity in non-exercise equations to estimate cardiorespiratory fitness. The equations contain readily available variables, they explain much of the variance in cardiorespiratory fitness, and they are generalizable and reliable.

List of Supplemental Digital Content

The following tables appear in the Word document entitled, Supplemental Digital Content, O'Donovan and colleagues.

Table S1. A non-exercise testing regression model for estimating cardiorespiratory fitness in whites and South Asians using physical activity categories without an ethnicity variable

Table S2. A non-exercise testing regression model for estimating cardiorespiratory fitness in whites and South Asians using physical activity minutes per week without an ethnicity variable

Table S3. A non-exercise testing regression model for estimating cardiorespiratory fitness in whites and South Asians using physical activity categories with an ethnicity variable

Table S4. A non-exercise testing regression model for estimating cardiorespiratory fitness in whites and South Asians using physical activity minutes per week with an ethnicity variable

Table S5. Bootstrapping data for the model containing age, BMI, smoking status, resting heart rate, physical activity quintile, and ethnicity

Table S6. Jackknife data for the model containing age, BMI, smoking status, resting heart rate, physical activity quintile, and ethnicity

Acknowledgements

Ghouri was supported by a Fellowship from Chest, Heart and Stroke Scotland. The research was supported by The National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care – East Midlands (NIHR CLAHRC – EM), the Leicester Clinical Trials Unit and the NIHR Leicester-Loughborough Diet, Lifestyle and Physical Activity Biomedical Research Unit which is a partnership between University Hospitals of Leicester NHS Trust, Loughborough University and the University of Leicester. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Conflicts of interest

The authors have no conflicts of interest. The results of the present study do not constitute endorsement by ACSM.

References

1. Baecke JAH, Burema J, Frijters JER. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. *Am J Clin Nutr*. 1982;36:936-42.
2. Collins GS, Mallett S, Omar O, Yu LM. Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. *BMC medicine*. 2011;9:103.
3. Davey G, Roberts J, Patel S et al. Effects of exercise on insulin resistance in South Asians and Europeans. *J Exerc Physiol*. 2000;3(2):6-11.
4. Fogelholm M. Physical activity, fitness and fatness: relations to mortality, morbidity and disease risk factors. A systematic review. *Obes Rev*. 2010;11(3):202-21.
5. Fox J. Bootstrapping regression models. In. *Applied regression analysis and generalized linear models*. London: Sage; 2015, pp. 587-606.
6. Ghouri N, Purves D, McConnachie A, Wilson J, Gill JM, Sattar N. Lower cardiorespiratory fitness contributes to increased insulin resistance and fasting glycaemia in middle-aged South Asian compared with European men living in the UK. *Diabetologia*. 2013;56(10):2238-49.
7. Gill JM, Celis-Morales CA, Ghouri N. Physical activity, ethnicity and cardio-metabolic health: does one size fit all? *Atherosclerosis*. 2014;232(2):319-33.
8. Hall LM, Moran CN, Milne GR et al. Fat oxidation, fitness and skeletal muscle expression of oxidative/lipid metabolism genes in South Asians: implications for insulin resistance? *PLoS One*. 2010;5(12):e14197.
9. Hardy CP, Eston RG. Aerobic fitness of Anglo-Saxon and Indian students. *Br J Sports Med*. 1985;19(4):217-8.

10. Helmerhorst HJ, Brage S, Warren J, Besson H, Ekelund U. A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. *The international journal of behavioral nutrition and physical activity*. 2012;9:103.

11. International Diabetes Federation. *IDF Diabetes Atlas*. Brussels, Belgium: International Diabetes Federation2013. Available from: International Diabetes Federation.

12. Jurca R, Jackson AS, LaMonte MJ et al. Assessing cardiorespiratory fitness without performing exercise testing. *Am J Prev Med*. 2005;29(3):185-93.

13. 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(19):2024-35.

14. Kohl HW, Blair SN, Paffenbarger RS, Jr., Macera CA, Kronenfeld JJ. A mail survey of physical activity habits as related to measured physical fitness. *Am J Epidemiol*. 1988;127(6):1228-39.

15. Nes BM, Janszky I, Vatten LJ, Nilsen TI, Aspenes ST, Wisloff U. Estimating V.O 2peak from a nonexercise prediction model: the HUNT Study, Norway. *Med Sci Sports Exerc*. 2011;43(11):2024-30.

16. Nes BM, Vatten LJ, Nauman J, Janszky I, Wisloff U. A simple nonexercise model of cardiorespiratory fitness predicts long-term mortality. *Med Sci Sports Exerc*. 2014;46(6):1159-65.

17. Pedisic Z, Bauman A. Accelerometer-based measures in physical activity surveillance: current practices and issues. *Br J Sports Med*. 2015;49(4):219-23.

18. Porta M. *A dictionary of epidemiology*. 5th ed. Oxford: Oxford University Press; 2008.

19. Sallis JF, Haskell WL, Wood PD et al. Physical activity assessment methodology in the Five-City Project. *Am J Epidemiol*. 1985;121(1):91-106.

20. Stamatakis E, Hamer M, O'Donovan G, Batty GD, Kivimaki M. A non-exercise testing method for estimating cardiorespiratory fitness: associations with all-cause and cardiovascular mortality in a pooled analysis of eight population-based cohorts. *Eur Heart J*. 2013;34(10):750-8.

21. Taylor HL, Buskirk E, Henschel A. Maximal oxygen intake as an objective measure of cardio-respiratory performance. *J Appl Physiol*. 1955;8(1):73-80.

22. Wier LT, Jackson AS, Ayers GW, Arenare B. Nonexercise models for estimating VO2max with waist girth, percent fat, or BMI. *Med Sci Sports Exerc*. 2006;38(3):555-61.

Figure captions

Figure 1. Observed and predicted maximal oxygen consumption (VO_2 max, $mL \cdot kg^{-1} \cdot min^{-1}$) in model 1, the model containing age, BMI, smoking status, resting heart rate, and physical activity quintile.

Figure 2. Observed and predicted maximal oxygen consumption (VO_2 max, $mL \cdot kg^{-1} \cdot min^{-1}$) in model 4, the model containing age, BMI, smoking status, resting heart rate, physical activity quintile, and ethnicity.

Table 1. Participants' characteristics by ethnicity*

	White European men, n=83	South Asian men, n=85
Age, years	50±7 (40-69)	49±7 (40-69)
Body mass index, $kg \cdot m^{-2}$	27.2±4.5 (17.7-46.7)	27.0±3.6 (19.8-38.9)
Resting heart rate, $beats \cdot min^{-1}$	61±9 (45-83)	67±9 (49-94)
Smoking status, n (%)		
Never smoked	46 (55 %)	69 (81 %)
Ex-smoker	29 (35 %)	6 (7 %)
Current smoker	8 (10 %)	10 (12 %)
MVPA, $min \cdot wk^{-1}$	298±175 (49-817)	179±117 (14-599)
VO_2 max, $mL \cdot kg^{-1} \cdot min^{-1}$	39.68±7.71 (19.50-63.70)	31.31±5.90 (17.23-47.90)

*Values are mean±SD (range) unless stated otherwise. MVPA is moderate- to vigorous-intensity physical activity. VO_2 max is maximal oxygen consumption.

Table 2

Table 2. A non-exercise testing regression model for estimating cardiorespiratory fitness in whites and South Asians using physical activity quintiles without an ethnicity variable*

Variable	Observed coefficient	Standard error	95% CI	p-value
Constant	81.030	5.218	70.724, 91.337	<0.001
Age, y	-0.326	0.062	-0.449, -0.203	<0.001
BMI, $kg{\cdot}m^{-2}$	-0.813	0.112	-1.034, -0.593	<0.001
Smoking status				
Never smoked	0, reference	-	-	-
Ex or current	0.540	0.946	-1.328, 2.408	0.569
Physical activity quintile†				
1	1, reference	-	-	-
2	0.768	1.377	-1.952, 3.487	0.578
3	3.646	1.420	0.841, 6.451	0.011
4	5.353	1.394	2.600, 8.106	<0.001
5	7.064	1.408	4.284, 9.844	<0.001
Resting heart rate	-0.172	0.050	-0.270, -0.074	0.001

* This model is referred to as model 1 in the text. Model R^2 value was 0.5421, adjusted R^2 was 0.5190, and root-mean-square error was 5.5635. N=168.

Table 3. A non-exercise testing regression model for estimating cardiorespiratory fitness in whites and South Asians using physical activity quintiles with an ethnicity variable*

Variable	Observed coefficient	Standard error	95% CI	p-value
Constant	77.409	4.148	69.217, 85.601	<0.001
Age, y	-0.374	0.050	-0.472, -0.276	<0.001
BMI, kg·m^{-2}	-0.906	0.089	-1.082, -0.731	<0.001
Smoking status				
Never smoked	0, reference	-	-	-
Ex or current	-1.976	0.792	-3.540, -0.413	0.014
Physical activity quintile†				
1	1, reference	-	-	-
2	1.127	1.091	-1.028, 3.281	0.303
3	1.869	1.139	-0.381, 4.118	0.103
4	3.793	1.115	1.591, 5.996	0.001
5	3.029	1.188	0.682, 5.376	0.012
Resting heart rate	-0.066	0.041	-0.146, 0.015	0.108
Ethnicity				
South Asian	0, reference	-	-	-
White	8.032	0.821	6.410, 9.654	<0.001

*This model is referred to as model 4 in the text. Model R^2 value was 0.7148, adjusted R^2 was 0.6986, and root-mean-square error was 4.4046. N=168.

Table 4. Leave-one-out-cross-validation for the models with an ethnicity variable

Model	Observed R-squared	LOOCV R-squared	R-squared % difference	Observed RMSE	LOOCV RMSE	RMSE, % difference
4*	0.7148	0.6760	- 5.4%	4.4046	4.5565	+ 3.4%
5†	0.7170	0.6791	- 5.3%	4.3879	4.5343	+ 3.3%
6‡	0.7051	0.6772	- 4.0%	4.4372	4.5463	+ 2.5%

*This model is shown in Table 3 in the manuscript. †This model is shown in Table S3 in the supplement. ‡This model is shown in Table S4 in the supplement. LOOCV is leave-one-out-cross-validation. RMSE is root mean squared error.

Figure 1

Click here to download Figure Figure 1.tif

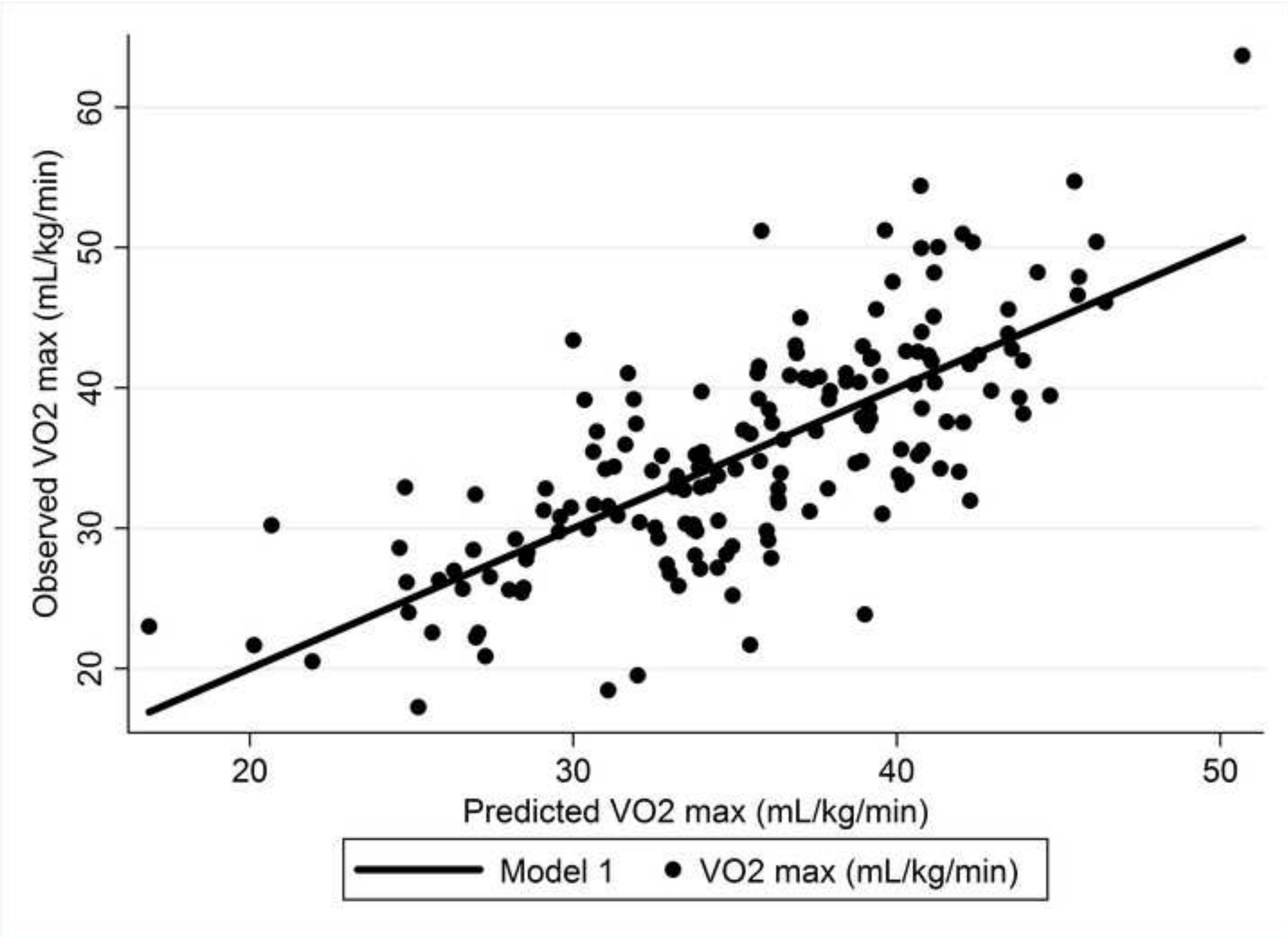

Figure 2

Click here to download Figure Figure 2.tif