

Original Paper

# Estimating Cardiorespiratory Fitness Without Exercise Testing or Physical Activity Status in Healthy Adults: Regression Model Development and Validation

Robert Sloan<sup>1</sup>, PhD; Marco Visentini-Scarzanella<sup>1</sup>, PhD; Susumu Sawada<sup>2</sup>, PhD; Xuemei Sui<sup>3</sup>, MD, PhD; Jonathan Myers<sup>4</sup>, PhD

<sup>1</sup>Department of Social and Behavioral Medicine, Kagoshima University Graduate Medical School, Kagoshima, Japan

<sup>2</sup>Faculty of Sport Sciences, Waseda University, Saitama, Japan

<sup>3</sup>Department of Exercise Science, Arnold School of Public Health, University of South Carolina, Columbia, SC, United States

<sup>4</sup>Division of Cardiovascular Medicine, Veterans Affairs Palo Alto Health Care System, Stanford University, Palo Alto, CA, United States

**Corresponding Author:**

Robert Sloan, PhD

Department of Social and Behavioral Medicine

Kagoshima University Graduate Medical School

8-35-1 Sakuragaoka

Kagoshima

Japan

Phone: 81 99 275 5751

Email: [rsloan@m.kufm.kagoshima-u.ac.jp](mailto:rsloan@m.kufm.kagoshima-u.ac.jp)

## Abstract

**Background:** Low cardiorespiratory fitness (CRF) is an independent predictor of morbidity and mortality. Most health care settings use some type of electronic health record (EHR) system. However, many EHRs do not have CRF or physical activity data collected, thereby limiting the types of investigations and analyses that can be done.

**Objective:** This study aims to develop a nonexercise equation to estimate and classify CRF (in metabolic equivalent tasks) using variables commonly available in EHRs.

**Methods:** Participants were 42,676 healthy adults (female participants: n=9146, 21.4%) from the Aerobics Center Longitudinal Study examined from 1974 to 2005. The nonexercise estimated CRF was based on sex, age, measured BMI, measured resting heart rate, measured resting blood pressure, and smoking status. A maximal treadmill test measured CRF.

**Results:** After conducting nonlinear feature augmentation, separate linear regression models were used for male and female participants to calculate correlation and regression coefficients. Cross-classification of actual and estimated CRF was performed using low CRF categories (lowest quintile, lowest quartile, and lowest tertile). The multiple correlation coefficient (*R*) was 0.70 (mean deviation 1.33) for male participants and 0.65 (mean deviation 1.23) for female participants. The models explained 48.4% (SE estimate 1.70) and 41.9% (SE estimate 1.56) of the variance in CRF for male and female participants, respectively. Correct category classification for low CRF (lowest tertile) was found in 77.2% (n=25,885) of male participants and 74.9% (n=6,850) of female participants.

**Conclusions:** The regression models developed in this study provided useful estimation and classification of CRF in a large population of male and female participants. The models may provide a practical method for estimating CRF derived from EHRs for population health research.

(*JMIR Public Health Surveill* 2022;8(7):e34717) doi: [10.2196/34717](https://doi.org/10.2196/34717)

**KEYWORDS**

nonexercise estimated cardiorespiratory fitness; public health; surveillance; epidemiology; electronic health record; EHR; fitness; cardiorespiratory; physical activity; regression model; nonexercise equation

## Introduction

### Background

The use of data from electronic health records (EHRs) beyond day-to-day medical management is rapidly emerging in the fields of digital health, public health, and epidemiology [1,2]. However, access to cardiorespiratory fitness (CRF), a valuable health metric, is limited. This limitation is primarily due to the medical service (cardiopulmonary stress test) being costly, time-consuming, and generally focused on cardiac patients [3-5]. CRF is a comprehensive measure of one's functional capacity ( $\text{mL O}_2 \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$ ) driven by the combination of heart, lung, and muscle function [6]. It is an important marker of health status in the general adult population [3,6]. Further demonstrating the importance of CRF, the American Heart Association released a scientific statement proposing that CRF be considered a clinical vital sign. The scientific rationale behind the statement is driven by the voluminous evidence that demonstrates that low CRF is a strong independent predictor of adverse health outcomes (ie, all-cause mortality, cancer, stroke, heart disease, and diabetes incidence) [3].

### Prior Work

To help increase accessibility to CRF data, researchers have developed an array of nonexercise estimated CRF (NEECRF) equations to estimate CRF [7-9]. NEECRF equations commonly include age, gender, resting heart rate, smoking status, BMI, and self-reported physical activity status (PAS) [8,10-12]. Studies have shown NEECRF to predict all-cause and cardiovascular disease mortality on par with measured CRF [13,14]. However, the assessment of PAS required to calculate NEECRF in patients is not typically conducted or documented in health care settings [15]. Therefore, using an NEECRF model without PAS (non-PAS) may be more feasible.

In a 2019 comprehensive NEECRF review, a few peer-reviewed non-PAS NEECRF equations were identified in adult populations reporting correlations and SE estimates [8]. Most of the equations were developed using only age, height, and weight combinations, and some used variables not commonly found in EHRs (eg, waist girth, predicted/ideal weight, or exercise mode).

Though correlations were moderate to high, the sample populations were too small to determine the classification accuracy of low CRF. Accurate classification of low CRF is essential for large-scale investigations [3,16]. Low CRF (lowest tertile) classification was recently investigated by Peterman et al [9], who found Baynard's [7] simplified non-PAS NEECRF equation to have poor classification ability in a sizable ( $n=4871$ ) adult population. Other researchers have also suggested that simple non-PAS NEECRF formulas have low validity, and nuanced approaches are warranted [16,17]. Investigators have also noted that valid non-PAS NEECRF models may have broad applications for public health, epidemiology, surveillance, practice, and research [3,8,9,16,18].

### Goal of This Study

Because it is standard practice to assess and document resting heart rate, blood pressure, BMI, and smoking status during a typical clinic visit, the primary aim of this study was to develop new models for NEECRF that could potentially be used in large-scale population health investigations using variables commonly found in EHRs. To accomplish this, we compared a non-PAS NEECRF equation to clinically measured CRF and evaluated its ability to estimate and classify CRF.

## Methods

### Study Sample

The Aerobics Center Longitudinal Study (ACLS) is a prospective epidemiological investigation of participants that began in 1970 [19]. The original data set for this investigation included 43,257 healthy adults who voluntarily participated in a comprehensive preventive medical examination at the Cooper Clinic in Dallas, Texas between 1974 and 2005. At baseline, all participants were free of diabetes, heart disease, stroke, cancer, positive electrocardiograms, and completed a maximal graded exercise test. Each participant gave informed consent to join the longitudinal study. The research population demographic primarily consists of Caucasian college-educated adults of middle to high socioeconomic status with an average age at baseline of 43.5 (range 20-79) years.

### Ethical Approval

The study was reviewed and approved annually by the Cooper Institute Institutional Review Board, and all participants provided written informed consent.

### Measurements

Predictor variables were assessed during a preventive health examination that included objective measurements of age, BMI, resting heart rate, systolic blood pressure, diastolic blood pressure, and self-reported smoking status between 1974 and 2005 at the Cooper Clinic. Age was verified at the time of the examination. Height and weight were measured on a calibrated scale using US customary units and converted to metric scales for this investigation. BMI was calculated from measured height and weight as  $\text{kg}/\text{m}^2$ . Manual auscultation was used to measure resting blood pressure while seated. Resting heart rate was calculated using the R-R interval on an electrocardiogram while seated. CRF was expressed as absolute metabolic equivalent tasks (METs;  $1 \text{ MET} = 3.5 \text{ mL O}_2 \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$ ) based on the total duration of a symptom-limited maximal Balke graded exercise test [6]. Following American College for Sports Medicine Guidelines, patients were encouraged to give maximal effort, and the test end point was volitional exhaustion or termination by the physician for medical reasons [6]. METs were calculated based on the final treadmill speed and grade [6]. The Balke graded exercise test is highly correlated ( $r=0.94$ ) with maximal graded cardiopulmonary exercise testing [20,21]. A standardized medical questionnaire was used to ascertain demographic information, lifestyle habits, and chronic disease status. More detailed information on the preventive health examination is available in prior ACLS publications [10,13].

## Statistical Analysis

### Data Exclusion

Data were first examined for outliers and skewness. We removed the outliers from the data set by removing participants with predictor variables (continuous) beyond the  $\pm 3 \sigma$  interval. To do this, we calculated the mean and SD of each predictor variable, excluding smoking (categorical). Any participant with at least one predictor variable above 3 SDs or below 3 SDs was flagged as an outlier. After removing outliers and incomplete entries ( $n=581$ ), the data set comprised 33,530 male participants and 9146 female participants, 98.7% of the original 43,257 participants.

### Regression

The main analysis was based on apparently healthy adults at baseline. Using a supervised machine learning technique, we conducted separate linear regression analyses for men and women to predict non-PAS NEECRF based on nonlinear augmentation of the predictor variables [1]. We also considered advanced machine learning models but did not find them advantageous. The male and female non-PAS NEECRF prediction equations were formulated to minimize the average mean squared error, where  $N$  is the number of samples in our data set:

$$MSE = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2$$

The prediction equation used age, height (Ht), weight (Wt), BMI, resting heart rate (rHR), systolic blood pressure (SBP), diastolic blood pressure (DBP), and smoking. All variables were continuous except for smoking status (nonsmoker=0, current smoker=1). Data were standardized by subtracting the mean and dividing by the SD for each variable. Next, separate models were then trained for male and female participants. We augmented the original 8 variables with second order and interaction terms, and regressed them linearly to the dependent variable for training. In this way, the nonlinearity was transferred from the regressor to the independent variables, while the model's overall interpretability was maintained. The augmentation procedure added the following 28 second order and interaction terms:  $Wt^2$ ,  $Wt \times Ht$ ,  $Wt \times Age$ ,  $Wt \times rHR$ ,  $Wt \times SBP$ ,  $Wt \times DBP$ ,  $Wt \times BMI$ ,  $Ht^2$ ,  $Ht \times Age$ ,  $Ht \times rHR$ ,  $Ht \times SBP$ ,  $Ht \times DBP$ ,  $Ht \times BMI$ ,  $Age^2$ ,  $Age \times rHR$ ,  $Age \times SBP$ ,  $Age \times DBP$ ,  $Age \times BMI$ ,  $rHR^2$ ,  $Ht \times SBP$ ,  $Ht \times DBP$ ,  $rHR \times BMI$ ,  $SBP^2$ ,  $SBP \times DBP$ ,  $SBP \times BMI$ ,  $DBP^2$ , and  $DBP \times BMI$  (for a total of 36 variables). Because smoking status was a categorical variable, it was not used to create the additional variables. The augmented data set was input into an elastic net linear regressor and trained and evaluated via 10-fold cross-validation [22]. Optimal model hyperparameters were calculated for the male ( $\alpha=.001$ ,  $\lambda=1.0$ ) and female ( $\alpha=.004$ ,  $\lambda=1.0$ ) data sets through the cross-validation procedure. Pearson correlation coefficients were then calculated using the non-PAS NEECRF equations

for the male and female data sets (shown in [Multimedia Appendix 1](#)). Lastly, for comparison, we cross-validated Baynard's [7] simplified non-PAS NEECRF equation ( $77.96 - 10.35 (\text{sex}; M=0, F=1) - 0.92 (\text{BMI}) - 0.32 (\text{age})$ ) [9].

### Classification Accuracy

We cross-classified non-PAS NEECRF and CRF for three specified cut points (lowest quintile, lowest quartile, and lowest tertile) commonly used in epidemiological investigations to define low CRF [3]. Next, CRF distributions were then grouped by males and females where  $\alpha$  served as the value of the  $\alpha$ -th percentile using the calculation non-PAS NEECRF  $> \alpha$ . After classification, we determined the non-PAS NEECRF accuracy, sensitivity, positive predictive value, and  $F_1$  score. The reference standard was the measured CRF, and low CRF was defined as a positive test. All analyses were performed in scikit-learn version 0.22.2 (NumFOCUS).

## Results

Descriptive statistics are provided in [Table 1](#). Correlation coefficients between each independent variable and CRF are presented in [Multimedia Appendix 1](#). The multiple Rs and mean deviations for non-PAS NEECRF (in METs) were high at 0.70 (mean deviation 1.33) for male participants and moderate at 0.65 (mean deviation 1.23) for female participants. The models explained 48.4% (SE estimate 1.70, 95% CI 0.05-3.97) of the variance in CRF for male participants and 41.9% (SE estimate 1.56, 95% CI 0.05-3.48) for female participants. [Multimedia Appendix 1](#) provides a simple independent variable input Google Sheet for researchers and data scientists to easily calculate NEECRF. [Table 2](#) provides the findings regarding the accuracy, positive predictive, and sensitivity values using the lowest quintile, quartile, and tertile to classify low CRF for male and female participants. While overall classification accuracy was meaningful for a nondiagnostic test across the three models, the optimal model was the lowest tertile. Combined male and female positive predictive value were 0.60, sensitivity 0.67, and  $F_1$  score 0.63 [23,24]. The  $F_1$  score is the best practice summary metric consisting of the harmonic mean of positive predictive value and sensitivity for classification (0=low, 1=high) [23]. Assuming a balanced data set ( $n=2529$ ) by Peterman et al [16], we calculated the  $F_1$  scores from their reported findings for equations applicable to EHRs [16]. We found  $F_1$  scores ranging from 0.04 to 0.56. Based on a residual plot ([Multimedia Appendix 1](#)), we found the model was most accurate for CRF MET values in the 7.5 to 12.5 METs range but tended to underestimate MET values  $>12.5$  METs and overestimate MET values  $<7.5$  METs. Notably, this is a common finding in non-PAS NEECRF studies [3]. Similar to Peterman et al [9], we cross-validated Baynard's [7] non-PAS NEECRF equation with our reference CRF data set and found low positive correlations for male ( $r=0.49$ , mean deviation 1.60) and female ( $r=0.43$ , mean deviation 1.46) participants.

**Table 1.** Baseline characteristics of participants.

	All (N=42,676)	Male participants (n=33,530)	Female participants (n=9146)
Age (years), mean (SD)	44.1 (9.6)	44.1 (9.5)	44.1 (10.2)
BMI (kg/m <sup>2</sup> ), mean (SD)	25.8 (3.8)	26.5 (3.5)	23.3 (3.7)
Resting heart rate (bpm), mean (SD)	60.9 (10.5)	60.1 (10.4)	64.0 (10.1)
Systolic blood pressure (mmHg), mean (SD)	118.9 (13.4)	120.7 (12.7)	112.4 (13.9)
Diastolic blood pressure (mmHg), mean (SD)	80.0 (9.6)	81.1 (9.3)	76.0 (9.3)
Smoker, n (%)	6361 (14.9)	5569 (16.7)	792 (9.7)
Measured CRF <sup>a</sup> (mL/kg/min), mean (SD)	11.5 (2.5)	12.0 (2.4)	9.8 (2.0)

<sup>a</sup>CRF: cardiorespiratory fitness (maximal oxygen consumption).

**Table 2.** Predictive accuracy of nonexercise estimated CRF classification of lowest cardiovascular fitness against the reference CRF.

Group	Low CRF <sup>a</sup> (lowest quintile)			Low CRF (lowest quartile)			Low CRF (lowest tertile)		
	ACC <sup>b</sup> (%)	PPV <sup>c</sup> (%)	SEN <sup>d</sup> (%)	ACC (%)	PPV (%)	SEN (%)	ACC (%)	PPV (%)	SEN (%)
Male participants	81.5	44.7	69.2	79.5	50.9	67.4	77.2	59.9	66.7
Female participants	80.6	43.9	57.8	77.9	55.4	55.6	74.9	60.4	66.9

<sup>a</sup>CRF: cardiorespiratory fitness.

<sup>b</sup>ACC: accuracy.

<sup>c</sup>PPV: positive predictive value.

<sup>d</sup>SEN: sensitivity.

## Discussion

### Principal Findings

Using a greater combination of clinical measures commonly found in EHRs, this study compared non-PAS NEECRF with objectively measured CRF in the largest population to date [10,11,25-27]. Overall, our model may provide a more applicable method for estimating and classifying CRF than previous methods [8,9,16]. Moreover, because the vital signs and medical information used to calculate non-PAS NEECRF are routinely captured during health care visits, our approach places nominal demand on health care staff and patients for collecting data. From a public health perspective, a moderate positive predictive value is practical, given that non-PAS NEECRF is a nondiagnostic test that is no cost and easily accessible [24]. Likely, some individuals classified with low fitness may be at the lower end of the fit spectrum and benefit from health promotion [24]. From a clinical perspective and considering moderate sensitivity, we concur with previous investigators that, while estimation equations are applicable for epidemiological investigations, they should not replace clinical exercise testing for patient diagnosis and management [3,16]. Our findings show that the ACLS non-PAS NEECRF may provide a useful assessment of CRF to conduct population health research.

### Comparison to Prior Work

Comparatively, PAS-based NEECRF models have demonstrated higher positive correlation values (0.71-0.93) along with a higher degree (~90%) of correct classification accuracy for low CRF than non-PAS NEECRF models [3]. Because physical activity

is a key contributor to CRF, including a PAS variable in an NEECRF model improves accuracy [3,8,16]. However, a recent review of distinct EHRs across 20 countries found that only 18.8% of family practice clinics had structured PAS questionnaires embedded within the EHR, with documented PAS in the EHR ranging from 10% to 86% [15]. Notably, no validated questionnaires designed for PAS-based NEECRF calculations were used. Therefore, the ability to conduct large-scale studies aggregating existing EHR data across local, domestic, or international systems to predict CRF is unlikely. Conversely, our model may provide a global approach to aggregating EHR data across systems to predict CRF and conduct analyses.

In 2019, Wang et al [8] provided a comprehensive list of peer-reviewed non-PAS NEECRF models that used some combination of age, BMI, or gender to predict CRF [8]. Samples were from small populations, and moderate to high correlations were reported generally without SE estimates [8]. Notably, the validity and usefulness of these simplified types of equations have been called into question [8,9,16,17]. We found that the findings from such investigations may be limited because the studies only reported correlations and lacked sufficient sample sizes to calculate prediction values for low CRF. It is also important to note that high correlation values do not necessarily result in a more accurate classification of low CRF [9,16].

Recently, Peterman et al [16] determined the ability of 7 non-PAS NEECRF equations to accurately classify low NEECRF (tertile) compared to measured CRF in a demographically comparable cohort (n=2529) to ACLS. Only 3 of the 7 equations apply because of variables not commonly

found in EHRs. On balance, the classification accuracy of low CRF (tertile) appeared to be better for ACLS. In a separate investigation, Peterman et al [9] also tested Baynard's [7] simplified non-PAS NEECRF using the Ball State cohort (n=4871) to assess classification accuracy. The equation had an  $r$  of 0.76; however, there was poor accuracy for detecting individuals positive for low CRF (37%). We also cross-validated Baynard's [7] equation with our data set and found low correlations, thus did not attempt to determine accuracy [9]. Although the ACLS and Ball State cohorts are demographically similar, our finding is expected to some degree because the ACLS equation is specifically trained from the ACLS data set [9,16]. Moreover, these equations need to be tested in epidemiological investigations with EHR data to see how well they predict health outcomes.

### Limitations

This study is not without limitations. Our study's primary limitation regarding correlation was that the measured reference CRF was conducted using a Balke graded maximal exercise test that estimates absolute METs. This testing strongly correlates with adults' maximal graded cardiopulmonary

exercise testing and is routinely used for clinical and epidemiological purposes [3,20,21]. Though the ACLS data set provides the largest known healthy female data set for clinically measured CRF, a larger female sample size may have provided a slightly better predictive model [3,27]. Nonetheless, our analyses demonstrated reasonable correlation and classification. Our analyses are based on a large predominantly Caucasian cohort; it is unknown if the results generalize to other ethnic groups. Notably, the homogeneity of the ACLS cohort may have strengthened the internal validity of our results by limiting possible confounders. The main strength of this investigation is that it was conducted on the largest cohort to date with a larger number of objectively measured predictive variables to estimate non-PAS NEECRF.

### Conclusions

The ACLS non-PAS NEECRF equation may provide a useful population health metric for CRF. More work should be conducted regarding diverse populations, the incidence of chronic conditions, and longitudinal repeated measures analyses toward improving public health and surveillance capability.

### Acknowledgments

The authors are grateful to the Cooper Clinic physicians and technicians for collecting the baseline data and the Cooper Institute staff for data entry and management. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the Japan Society for the Promotion of Science.

This research was supported by the Japan Society for the Promotion of Science KAKENHI Grant 19K19437.

### Authors' Contributions

RS conceptualized the study. RS and MVS designed the methodology. MVS used the software. RS and MVS validated the study. MVS conducted the formal analysis. RS conducted the investigation and obtained the resources. XS curated the data. RS and MVS wrote and prepared the original draft. RS, MVS, JM, XS, and SS reviewed and edited the manuscript. RS supervised the study and acquired the funding. All authors have read and agreed to the published version of the manuscript.

### Conflicts of Interest

None declared.

### Multimedia Appendix 1

Supplementary material.

[\[PDF File \(Adobe PDF File\), 123 KB-Multimedia Appendix 1\]](#)

### References

1. Wiemken TL, Kelley RR. Machine learning in epidemiology and health outcomes research. *Annu Rev Public Health* 2020 Apr 02;41:21-36. [doi: [10.1146/annurev-publhealth-040119-094437](https://doi.org/10.1146/annurev-publhealth-040119-094437)] [Medline: [31577910](https://pubmed.ncbi.nlm.nih.gov/31577910/)]
2. Kim J. Big data, health informatics, and the future of cardiovascular medicine. *J Am Coll Cardiol* 2017 Feb 21;69(7):899-902 [FREE Full text] [doi: [10.1016/j.jacc.2017.01.006](https://doi.org/10.1016/j.jacc.2017.01.006)] [Medline: [28209228](https://pubmed.ncbi.nlm.nih.gov/28209228/)]
3. Ross R, Blair SN, Arena R, Church TS, Després JP, Franklin BA, American Heart Association Physical Activity Committee of the Council on Lifestyle and Cardiometabolic Health, Council on Clinical Cardiology, Council on Epidemiology and Prevention, Council on Cardiovascular and Stroke Nursing, Council on Functional Genomics and Translational Biology, Stroke Council. 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 Dec 13;134(24):e653-e699. [doi: [10.1161/CIR.0000000000000461](https://doi.org/10.1161/CIR.0000000000000461)] [Medline: [27881567](https://pubmed.ncbi.nlm.nih.gov/27881567/)]
4. Sakr S, Elshawi R, Ahmed A, Qureshi WT, Brawner C, Keteyian S, et al. Using machine learning on cardiorespiratory fitness data for predicting hypertension: The Henry Ford Exercise Testing (FIT) Project. *PLoS One* 2018;13(4):e0195344 [FREE Full text] [doi: [10.1371/journal.pone.0195344](https://doi.org/10.1371/journal.pone.0195344)] [Medline: [29668729](https://pubmed.ncbi.nlm.nih.gov/29668729/)]

5. Al-Mallah MH, Keteyian SJ, Brawner CA, Whelton S, Blaha MJ. Rationale and design of the Henry Ford Exercise Testing Project (the FIT project). *Clin Cardiol* 2014 Aug;37(8):456-461. [doi: [10.1002/clc.22302](https://doi.org/10.1002/clc.22302)] [Medline: [25138770](https://pubmed.ncbi.nlm.nih.gov/25138770/)]
6. American College of Sports Medicine. *ACSM's Guidelines for Exercise Testing and Prescription*. 10th ed. Philadelphia, PA: Wolters Kluwer Health; 2017.
7. Baynard T, Arena RA, Myers J, Kaminsky LA. The role of body habitus in predicting cardiorespiratory fitness: The FRIEND Registry. *Int J Sports Med* 2016 Oct;37(11):863-869. [doi: [10.1055/s-0042-110572](https://doi.org/10.1055/s-0042-110572)] [Medline: [27490112](https://pubmed.ncbi.nlm.nih.gov/27490112/)]
8. Wang Y, Chen S, Lavie CJ, Zhang J, Sui X. An overview of non-exercise estimated cardiorespiratory fitness: estimation equations, cross-validation and application. *J Sci Sport Exerc* 2019 Jun 10;1(1):38-53. [doi: [10.1007/s42978-019-0003-x](https://doi.org/10.1007/s42978-019-0003-x)]
9. Peterman JE, Harber MP, Imboden MT, Whaley MH, Fleenor BS, Myers J, et al. Accuracy of exercise-based equations for estimating cardiorespiratory fitness. *Med Sci Sports Exerc* 2021 Jan;53(1):74-82. [doi: [10.1249/MSS.0000000000002435](https://doi.org/10.1249/MSS.0000000000002435)] [Medline: [32694370](https://pubmed.ncbi.nlm.nih.gov/32694370/)]
10. Jackson AS, Sui X, O'Connor DP, Church TS, Lee D, Artero EG, et al. Longitudinal cardiorespiratory fitness algorithms for clinical settings. *Am J Prev Med* 2012 Nov;43(5):512-519. [doi: [10.1016/j.amepre.2012.06.032](https://doi.org/10.1016/j.amepre.2012.06.032)] [Medline: [23079174](https://pubmed.ncbi.nlm.nih.gov/23079174/)]
11. Jackson AS, Wier LT, Ayers GW, Beard EF, Stuteville JE, Blair SN. Changes in aerobic power of women, ages 20-64 yr. *Med Sci Sports Exerc* 1996 Jul;28(7):884-891. [doi: [10.1097/00005768-199607000-00016](https://doi.org/10.1097/00005768-199607000-00016)] [Medline: [8832543](https://pubmed.ncbi.nlm.nih.gov/8832543/)]
12. Jurca R, Jackson AS, LaMonte MJ, Morrow JR, Blair SN, Wareham NJ, et al. Assessing cardiorespiratory fitness without performing exercise testing. *Am J Prev Med* 2005 Oct;29(3):185-193. [doi: [10.1016/j.amepre.2005.06.004](https://doi.org/10.1016/j.amepre.2005.06.004)] [Medline: [16168867](https://pubmed.ncbi.nlm.nih.gov/16168867/)]
13. Artero EG, Jackson AS, Sui X, Lee D, O'Connor DP, Lavie CJ, et al. Longitudinal algorithms to estimate cardiorespiratory fitness: associations with nonfatal cardiovascular disease and disease-specific mortality. *J Am Coll Cardiol* 2014 Jun 03;63(21):2289-2296 [FREE Full text] [doi: [10.1016/j.jacc.2014.03.008](https://doi.org/10.1016/j.jacc.2014.03.008)] [Medline: [24703924](https://pubmed.ncbi.nlm.nih.gov/24703924/)]
14. 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 Mar;34(10):750-758. [doi: [10.1093/eurheartj/ehs097](https://doi.org/10.1093/eurheartj/ehs097)] [Medline: [22555215](https://pubmed.ncbi.nlm.nih.gov/22555215/)]
15. Lindeman C, McCurdy A, Lamboglia CG, Wohlers B, Pham ANQ, Sivak A, et al. The extent to which family physicians record their patients' exercise in medical records: a scoping review. *BMJ Open* 2020 Feb 12;10(2):e034542 [FREE Full text] [doi: [10.1136/bmjopen-2019-034542](https://doi.org/10.1136/bmjopen-2019-034542)] [Medline: [32054628](https://pubmed.ncbi.nlm.nih.gov/32054628/)]
16. Peterman JE, Whaley MH, Harber MP, Fleenor BS, Imboden MT, Myers J, et al. Comparison of non-exercise cardiorespiratory fitness prediction equations in apparently healthy adults. *Eur J Prev Cardiol* 2021 Apr 10;28(2):142-148. [doi: [10.1177/2047487319881242](https://doi.org/10.1177/2047487319881242)] [Medline: [33838037](https://pubmed.ncbi.nlm.nih.gov/33838037/)]
17. Malek MH, Berger DE, Housh TJ, Coburn JW, Beck TW. Validity of VO<sub>2</sub>max equations for aerobically trained males and females. *Med Sci Sports Exerc* 2004 Aug;36(8):1427-1432. [doi: [10.1249/01.mss.0000135795.60449.ce](https://doi.org/10.1249/01.mss.0000135795.60449.ce)] [Medline: [15292753](https://pubmed.ncbi.nlm.nih.gov/15292753/)]
18. Radford NB, DeFina LF, Leonard D, Barlow CE, Willis BL, Gibbons LW, et al. Cardiorespiratory fitness, coronary artery calcium, and cardiovascular disease events in a cohort of generally healthy middle-age men: results from the cooper center longitudinal study. *Circulation* 2018 May 01;137(18):1888-1895. [doi: [10.1161/CIRCULATIONAHA.117.032708](https://doi.org/10.1161/CIRCULATIONAHA.117.032708)] [Medline: [29343464](https://pubmed.ncbi.nlm.nih.gov/29343464/)]
19. Blair SN, Kohl HW, Paffenbarger RS, Clark DG, Cooper KH, Gibbons LW. Physical fitness and all-cause mortality. A prospective study of healthy men and women. *JAMA* 1989 Nov 03;262(17):2395-2401. [doi: [10.1001/jama.262.17.2395](https://doi.org/10.1001/jama.262.17.2395)] [Medline: [2795824](https://pubmed.ncbi.nlm.nih.gov/2795824/)]
20. Balke B, Ware RE. An experimental study of physical fitness of Air Force personnel. *U S Armed Forces Med J* 1959 Jun;10(6):675-688. [Medline: [13659732](https://pubmed.ncbi.nlm.nih.gov/13659732/)]
21. Pollock ML, Foster C, Schmidt D, Hellman C, Linnerud AC, Ward A. Comparative analysis of physiologic responses to three different maximal graded exercise test protocols in healthy women. *Am Heart J* 1982 Mar;103(3):363-373. [doi: [10.1016/0002-8703\(82\)90275-7](https://doi.org/10.1016/0002-8703(82)90275-7)] [Medline: [7064770](https://pubmed.ncbi.nlm.nih.gov/7064770/)]
22. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J Royal Stat Soc Ser B Stat Methodology* 2005 Apr;67(2):301-320. [doi: [10.1111/j.1467-9868.2005.00503.x](https://doi.org/10.1111/j.1467-9868.2005.00503.x)]
23. Xiang L, Deng K, Mei Q, Gao Z, Yang T, Wang A, et al. Population and age-based cardiorespiratory fitness level investigation and automatic prediction. *Front Cardiovasc Med* 2021;8:758589. [doi: [10.3389/fcvm.2021.758589](https://doi.org/10.3389/fcvm.2021.758589)] [Medline: [35071342](https://pubmed.ncbi.nlm.nih.gov/35071342/)]
24. Trevethan R. Sensitivity, specificity, and predictive values: foundations, pliabilities, and pitfalls in research and practice. *Front Public Health* 2017;5:307. [doi: [10.3389/fpubh.2017.00307](https://doi.org/10.3389/fpubh.2017.00307)] [Medline: [29209603](https://pubmed.ncbi.nlm.nih.gov/29209603/)]
25. Zeiher J, Ombrellaro KJ, Perumal N, Keil T, Mensink GBM, Finger JD. Correlates and determinants of cardiorespiratory fitness in adults: a systematic review. *Sports Med Open* 2019 Sep 03;5(1):39. [doi: [10.1186/s40798-019-0211-2](https://doi.org/10.1186/s40798-019-0211-2)] [Medline: [31482208](https://pubmed.ncbi.nlm.nih.gov/31482208/)]
26. Ahmadian HR, Sclafani JJ, Emmons EE, Morris MJ, Leclerc KM, Slim AM. Comparison of predicted exercise capacity equations and the effect of actual versus ideal body weight among subjects undergoing cardiopulmonary exercise testing. *Cardiol Res Pract* 2013;2013:940170. [doi: [10.1155/2013/940170](https://doi.org/10.1155/2013/940170)] [Medline: [23653881](https://pubmed.ncbi.nlm.nih.gov/23653881/)]
27. Rodgers JL. Degrees of freedom at the start of the second 100 years: a pedagogical treatise. *Adv Methods Practices Psychological Sci* 2019 Dec 05;2(4):396-405. [doi: [10.1177/2515245919882050](https://doi.org/10.1177/2515245919882050)]

## Abbreviations

**ACLS:** Aerobics Center Longitudinal Study  
**CRF:** cardiorespiratory fitness  
**DBP:** diastolic blood pressure  
**EHR:** electronic health record  
**Ht:** height  
**MET:** metabolic equivalent task  
**NEECRF:** nonexercise estimated cardiorespiratory fitness  
**PAS:** physical activity status  
**rHR:** resting heart rate  
**SBP:** systolic blood pressure  
**Wt:** weight

*Edited by H Bradley; submitted 04.11.21; peer-reviewed by L Larsen, C Lavie, J Peterman; comments to author 24.02.22; revised version received 14.03.22; accepted 27.05.22; published 06.07.22*

*Please cite as:*

*Sloan R, Visentini-Scarzanella M, Sawada S, Sui X, Myers J*

*Estimating Cardiorespiratory Fitness Without Exercise Testing or Physical Activity Status in Healthy Adults: Regression Model Development and Validation*

*JMIR Public Health Surveill 2022;8(7):e34717*

*URL: <https://publichealth.jmir.org/2022/7/e34717>*

*doi: [10.2196/34717](https://doi.org/10.2196/34717)*

*PMID:*

©Robert Sloan, Marco Visentini-Scarzanella, Susumu Sawada, Xuemei Sui, Jonathan Myers. Originally published in JMIR Public Health and Surveillance (<https://publichealth.jmir.org>), 06.07.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.