

## RESEARCH ARTICLE

# Toward characterizing cardiovascular fitness using machine learning based on unobtrusive data

Maria Cecília Moraes Frade<sup>1</sup>, Thomas Beltrame<sup>1,2\*</sup>, Mariana de Oliveira Gois<sup>1</sup>, Allan Pinto<sup>3</sup>, Silvia Cristina Garcia de Moura Tonello<sup>1</sup>, Ricardo da Silva Torres<sup>4</sup>, Aparecida Maria Catai<sup>1</sup>

**1** Department of Physical Therapy, Federal University of São Carlos, São Carlos, São Paulo, Brazil, **2** Samsung R&D Institute Brazil—SRBR, Campinas, São Paulo, Brazil, **3** Brazilian Synchrotron Light Laboratory (LNLS), Brazilian Center for Research in Energy and Materials (CNPEM), Campinas, São Paulo, Brazil, **4** Department of ICT and Natural Sciences, Faculty of Information Technology and Electrical Engineering, NTNU—Norwegian University of Science and Technology, Ålesund, Norway

\* [beltramethomas@gmail.com](mailto:beltramethomas@gmail.com)



## OPEN ACCESS

**Citation:** Frade MCM, Beltrame T, Gois MdO, Pinto A, Tonello SCGdM, Torres RdS, et al. (2023) Toward characterizing cardiovascular fitness using machine learning based on unobtrusive data. PLoS ONE 18(3): e0282398. <https://doi.org/10.1371/journal.pone.0282398>

**Editor:** Zulkarnain Jaafar, Universiti Malaya, MALAYSIA

**Received:** May 11, 2022

**Accepted:** February 14, 2023

**Published:** March 2, 2023

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0282398>

**Copyright:** © 2023 Frade et al. 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 author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and it will be in [Supporting Information](#) files.

## Abstract

Cardiopulmonary exercise testing (CPET) is a non-invasive approach to measure the maximum oxygen uptake ( $\dot{V}O_{2-max}$ ), which is an index to assess cardiovascular fitness (CF). However, CPET is not available to all populations and cannot be obtained continuously. Thus, wearable sensors are associated with machine learning (ML) algorithms to investigate CF. Therefore, this study aimed to predict CF by using ML algorithms using data obtained by wearable technologies. For this purpose, 43 volunteers with different levels of aerobic power, who wore a wearable device to collect unobtrusive data for 7 days, were evaluated by CPET. Eleven inputs (sex, age, weight, height, and body mass index, breathing rate, minute ventilation, total hip acceleration, walking cadence, heart rate, and tidal volume) were used to predict the  $\dot{V}O_{2-max}$  by support vector regression (SVR). Afterward, the SHapley Additive exPlanations (SHAP) method was used to explain their results. SVR was able to predict the CF, and the SHAP method showed that the inputs related to hemodynamic and anthropometric domains were the most important ones to predict the CF. Therefore, we conclude that the cardiovascular fitness can be predicted by wearable technologies associated with machine learning during unsupervised activities of daily living.

## Introduction

Noncommunicable chronic diseases (NCDs) are mainly responsible for all causes of death and illness among adults aged between 35–70 years, and cardiovascular diseases are accountable for the main cause of mortality in the world [1]. There are some modifiable risk factors associated with NCDs, such as high systolic arterial pressure, high fasting plasma glucose, as well as low physical activity [2, 3].

It is known that the cardiovascular diseases and their modifiable risk factors lead to a reduction in cardiovascular fitness (CF) [4, 5]. Moreover, higher CF levels have a protective effect

**Funding:** The research was supported by Coordination for the Improvement of Higher Education Personnel grants (CAPES) 001 and (CAPES) 88887.362954/2019-00 awarded to MOG. This work was also supported by São Paulo Research Foundation (FAPESP) grant FAPESP 2016/22215-7 awarded to AMC, FAPESP 2017/09639-5 and FAPESP 2018/19016-8 awarded to TB; FAPESP 2018/22818-9 awarded to MCMF; and FAPESP 2019/16253-1 awarded to AP. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

against cardiovascular diseases and all-cause mortality in varied populations [4, 6, 7]. Thus, due to the considerable relevance of increasing population lifetime, the continuous measurement of CF could be considered a vital sign, and thus, it should be a priority in public health [8]; however, the definition and ways of evaluation of the CF are contradictory [9–11].

CF is commonly evaluated by measuring the maximum oxygen uptake ( $\dot{V}O_{2-max}$ ), as the index of maximal aerobic power, obtained during cardiopulmonary exercise testing (CPET) [11–13]. The  $\dot{V}O_{2-max}$  reflects the maximal capacity of the pulmonary, cardiovascular, and metabolic systems to capture, transport, and utilize oxygen, respectively, which is directly influenced by the CF [13, 14]. However, the  $\dot{V}O_{2-max}$  measurement during the CPET requires trained professionals and expensive equipment [15–17], and is rarely used as a prevention tool in the general population. For this reason, the CF assessed by  $\dot{V}O_{2-max}$  during CPET is not available to all populations and cannot be obtained continuously.

Therefore, considering the difficulties of performing the CPET, but given the high clinical value to assess cardiovascular fitness, new methods for continuous assessment of CF are needed. These methods could be more realistic, unobstructive, and accessible to all populations if performed outside laboratory settings, during unsupervised activities of daily living (ADL) [18]. Wearable sensors and vital signal fusion might represent a unique possibility to infer CF continuously, allowing the use of this technology in the future for pre-symptomatic detection of NCDs, especially cardiovascular diseases [6, 7].

Furthermore, there is an increasing number of studies that have combined the use of wearables and machine learning techniques for monitoring patients with NCDs, especially in the cardiorespiratory field [19, 20]. In fact, longitudinal data from wearables seem to contain enough information to predict CF of healthy volunteers during unsupervised ADL from complex machine learning algorithms [21–25].

However, despite the great potential of the combination between wearables and machine learning, there is still a lack of evidence for using these technologies to predict CF in patients with NCDs, especially in diabetes mellitus, chronic pulmonary disease, and cardiovascular diseases. Furthermore, understanding how these models, trained from machine learning algorithms, can transform vital signals into  $\dot{V}O_{2-max}$  may provide complex mechanistic insights regarding the differences in CF between volunteers. Due to the complexity of the  $\dot{V}O_{2-max}$  prediction algorithms based on features obtained from wearable technologies [25], the interpretability of how longitudinal vital signals are being transformed into  $\dot{V}O_{2-max}$  is exceptionally low [26] because of the expected trade-off between the interpretability of a given model, and its performance to predict health outcomes [27].

Recently, explainable models have been used in medical science to better justify decision-making of the prediction models [26]. It is known that wearable sensors are useful for the continuous biological data acquisition that can be associated with machine learning techniques, such as Random Forest Regression, Neural Network and Support Vector Regression Machines to predict CF [21, 25]. Thus, understanding these models might also indicate how the human “black box” physiological systems interact with the environment, approximating the explainability of these complex algorithms to what we experience when using simpler methods, such as in linear regression models.

The SHapley Additive exPlanations (SHAP) is a valuable approach derived from the coalitional game theory, which can be used to interpret complex models built from supervised machine learning methods obtained from biological data [26, 28]. In this paper, we investigate the use of Shapley to assess the importance of features in the CF prediction problem. The main motivation for its use relies on (1) its ability to be model agnostic (i.e., method for explainability associated with any model to extract extra information about the prediction procedure

[26]. In this case, we can simply replace linear models with complex models without losing much interpretability; (2) to produce interpretations for a single data point; and (3) to produce human-friendly explanations for linear regression results when we deal with multiple regression problems. Furthermore, Shapley-based methods can produce visual interpretations in which we can easily visualize the global or local feature contributions [29].

Therefore, our main objective is to predict CF by using machine learning algorithms from data obtained by wearable technologies of volunteers with a broad spectrum of maximal aerobic power (or  $\dot{V}O_{2-max}$ , as an index of CF level). Afterward, an explainable artificial intelligence (AI) method will be used to investigate “how” CF can be estimated from the longitudinal signals acquired by wearables during unobtrusive experimental protocols. Our hypothesis is that machine learning algorithms can provide suitable models to predict  $\dot{V}O_{2-max}$  when trained with vital signals collected by wearables, and explainable AI methods can interpret the prediction models’ output of these algorithms. By doing so, we have a better understanding of how longitudinal signals during ADL are related with  $\dot{V}O_{2-max}$ , which has clinical implications for CF. Consequently, this study will demonstrate an innovative approach to predicting the onset of NCDs in future studies by continuously evaluating  $\dot{V}O_{2-max}$ , in addition to explaining the differences in CF among volunteers through explainable methods.

## Materials and methods

### Study design

This longitudinal study was approved by the Federal University of Sao Carlos Ethics Committee (CAAE: 80459817.5.1001.5504), and it was conducted in the Cardiovascular Physiotherapy Laboratory at the Federal University of Sao Carlos (UFSCar). All procedures followed the Helsinki declaration, and all volunteers signed a free and informed consent form in accordance with Resolution 466/2012 of the National Health Council.

The inclusion criteria were both sexes, ages of 18–80 years, and different levels of aerobic power (including apparently healthy volunteers, with risk factors to develop NCDs, or with type 2 diabetes mellitus, chronic obstructive pulmonary disease, or coronary arterial disease). All volunteer clinical conditions were validated from a medical diagnosis. Volunteers were excluded by orthopedic or neurological limitations; associated uncontrolled heart diseases, abnormality in the resting or exercising ECG response (infra-uneven ST-segment > 2 mm, unsustainable atrial tachycardia, atrial fibrillation or atrioventricular blocks, ventricular or supraventricular arrhythmias) that would prevent them from following the proposed protocol. Volunteers that wore the t-shirts for less than 5 days or less than 6 hours per day were also excluded. For the volunteers with NCDs, the pulse saturation (SpO<sub>2</sub>) was verified at resting by pulse oximeter (sense 10, ALFAMED, Lagoa do Sino, Brazil) for safety. The experimental protocol comprised three main steps.

### Step one

During the first laboratory visit, volunteers were questioned about their health, lifestyle habits, exercise practice, and disease conditions (if present). Afterwards, they performed a physical evaluation comprising measurements such as weight, height, thorax, and abdominal circumference, and resting vital signals (heart rate, breathing rate, and arterial pressure). Finally, the researcher explained about using the wearable device, including how to wear it properly, remove and wash it (smart shirt will be further explained in the text in the future) and how to charge the battery.

## Step two

On the second day in the laboratory, volunteers performed the CPET on a cycle ergometer (Quinton Corival® 400, Seattle, USA) with a ramp-type protocol to assess the CF. The power increment was calculated using the formula described by Wasserman, considering height, age, and sex [14]. The test consisted of: (1) five minutes at rest, (2) three minutes unloaded warm-up, (3)  $9.6 \pm 1.4$  minutes ramp protocol (with  $20.7 \pm 7.1$  watts per minute increment), and (4) six minutes unloaded cycling for active recovery. All volunteers were encouraged to keep a constant cycling of 60 to 65 rpm and were stimulated to continue the CPET until volitional fatigue. The oxygen uptake and minute ventilation were measured breath-by-breath by a metabolic system (Vmax29c, Sensor Medics, Yorba Linda, CA, USA) calibrated before each experiment, according to the manufacturer's manual. Moreover, heart rate (HR) was calculated during the exercise based on a single lead ECG system (BioAmp FE132, ADInstruments, Australia).

The interruption criteria were according to previous work [16], and just one volunteer had the CPET interrupted due to the unexpected (excessively high) increase in arterial pressure. One volunteer was also excluded from this study due to oxygen desaturation on resting immediately before the CPET.

## Step three

The last step took seven days, where the vital signals from the wearable sensors were collected during unsupervised ADL, in an unobtrusive way, where participants maintained their daily routine. The volunteers were instructed to use the smart shirt for seven days, at least for eight active hours per day, except during showering and water activities. The smart shirt has three embedded sensors, and the raw signals were used to obtain biological and environmental data by a previously validated [30] proprietary algorithm. The HR data was measured by an ECG system (one-lead ECG channel, frequency: 256 Hz and with 12 bits resolution), and an algorithm that filters and averages the HR over the last 16 heartbeats. The breathing rate (BR), minute ventilation ( $V_e$ ), and tidal volume ( $V_t$ ) were estimated by the thoracic and abdominal belts. The  $V_t$  variable was obtained by dividing  $V_e$  by BR ( $V_t = V_e/BR$ ). The respiratory belts were based on inductance plethysmography (sampled at 128 Hz with 16 bits resolution), and BR,  $V_e$ , and  $V_t$  were averaged over the last seven respiration cycles. The environmental data from total hip acceleration (Acc), and walking cadence (Cad) were based on triaxial accelerometer signals located at the right side of the hip. It was collected at 64 Hz, with a 13 bits resolution and a range of 16 g (with 0.004 g of resolution step). All data were resampled at 1 Hz.

## Data analysis

During the CPET, the following metabolic variables were measured: oxygen uptake ( $\dot{V}O_2$ ); carbon dioxide output ( $\dot{V}CO_2$ ); and respiratory exchange ratio (RER, or  $\dot{V}CO_2/\dot{V}O_2$ ). Data were pre-processed in MatLab routine where the data were interpolated at 1 Hz, and then, the metabolic data were synchronized with HR, which was also used as a secondary criterion to confirm  $\dot{V}O_{2-max}$ , as described in Table 1 [31]. For each variable, the maximum (including the  $\dot{V}O_{2-max}$ ,  $RER_{max}$ , and  $HR_{max}$ ) was considered as the average of the last 20 s of the exercise protocol, before the CPET interruption. The  $\dot{V}O_{2-max}$  (as a surrogate for CF) was considered as the ground truth for the machine learning algorithms training that should predict the  $\dot{V}O_{2-max}$  ( $pred \dot{V}O_{2-max}$ ) based on the inputs from the wearables and volunteer's personal information.

Another MatLab routine was used to process the unobtrusive longitudinal data from the smart shirt. Initially, the 1-Hz variables were downloaded from the Hexoskin's dashboard (please check the documentation at <https://www.hexoskin.com/pages/hexoskin-connected->

**Table 1. Characteristics of volunteers, peak variables obtained during the cardiopulmonary exercise testing, and the mean response of the variables obtained by the wearable.**

Characteristics of Subjects (n = 43)	
<b>Anthropometric</b>	
Sex (M/F)	32/11
Age (years)	37.50(25.00–55.00)
Weight (kg)	75.41±13.12
Height (m)	1.74±0.09
BMI (kg/m <sup>2</sup> )	24.84±3.12
<b>CPET Peak</b>	
$\dot{V}O_{2-max}$ (l/min)	2.42±0.80
RER <sub>max</sub>	1.28±0.10
HR <sub>max</sub> (bpm)	169.27(155.71–183.50)
<b>Wearable</b>	
μBR (rpm)	17.92(16.48–19.77)
μV <sub>E</sub> (l/min)	15.86±4.71
μAcc (g)	0.06±0.03
μHR (bpm)	84.09±8.57
Cad (spm)	5.61(3.47–8.38)
μVt (ml/min)	896.80±251.03

M: male; F: female; BMI: body mass index; CPET: cardiopulmonary exercise testing;  $\dot{V}O_{2-max}$ : maximum oxygen uptake; RER<sub>max</sub>: maximum respiratory exchange ratio; HR<sub>max</sub>: maximum heart rate; BR: breathing rate; V<sub>E</sub>: minute ventilation; Acc (g): total hip acceleration, by gravity; Cad (spm): walking cadence, by steps per minute; μ: mean response.

<https://doi.org/10.1371/journal.pone.0282398.t001>

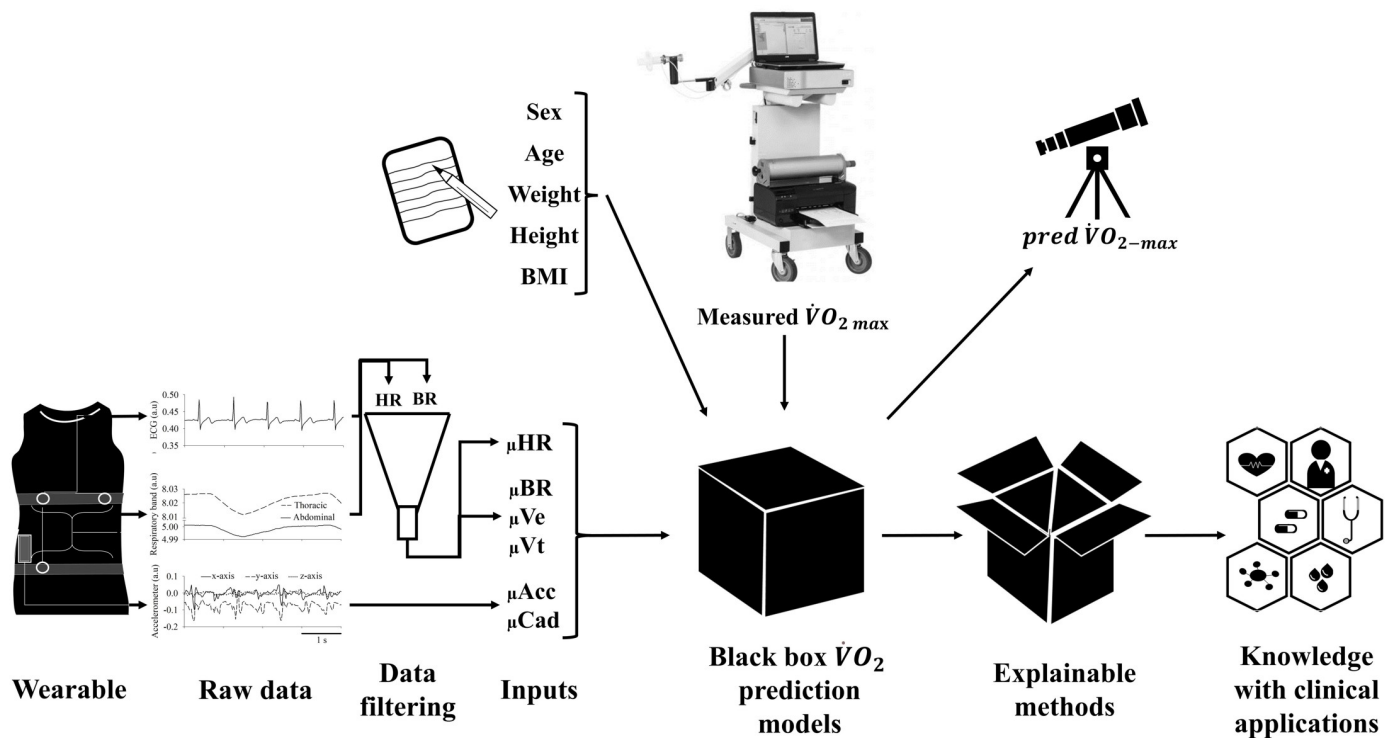
[health-platform](#) –as of August 2021). Each downloaded dataset (~7 days) was combined into a single dataset consisting of 65±13 hours; For HR, beats/min lower than 30 and higher than 220 were excluded. For respiratory measurements, BR values lower than 3 and higher than 79 were used as a reference to exclude data also from V<sub>E</sub> and V<sub>t</sub> variables, beyond BR. Finally, the average response for all variables (μHR, μBR, μV<sub>E</sub>, μV<sub>t</sub>, μAcc, μCad) was computed and used as the inputs for the machine learning algorithms.

## Framework

As described before, the MatLab scripts were used to calculate the biological signal-derived inputs and the output data ( $\dot{V}O_{2-max}$ ). Beyond the inputs from wearables (i.e., μHR, μBR, μV<sub>E</sub>, μV<sub>t</sub>, μAcc, and μCad), the age, sex, weight, height, and body mass index (BMI, weight/height<sup>2</sup>) of each volunteer were also used as inputs to estimate the *pred*  $\dot{V}O_{2-max}$  by machine learning algorithms (further explained). These steps are illustrated in [Fig 1](#).

## Machine learning algorithm

Support vector machine (SVM) comprises a set of supervised learning algorithms used for classification and regression analysis. Introduced by Cortes and Vapnik [32], SVM is one of the most robust and flexible machine learning algorithms that has been successfully applied to several different problems [32]. In short, SVM algorithms build a model by finding a hyperplane in an n-dimensional space in which data points could be distinctly classified. Differently from other linear regression methods, the SVM algorithm creates a safety boundary from both sides of the hyperplane (known as margins), which is paramount information for better



**Fig 1. The wearable system has embedded cardiac, respiratory, and movement sensors that measure unsupervised and unobtrusive biological data.** These raw data are processed, filtered, and averaged. Mean response to heart rate ( $\mu\text{HR}$ ), breathing rate ( $\mu\text{BR}$ ), minute ventilation ( $\mu\text{VE}$ ), tidal volume ( $\mu\text{Vt}$ ), total hip acceleration ( $\mu\text{Acc}$ ), and walking cadence ( $\mu\text{Cad}$ ), as well as sex, age, weight, height, and body mass index (BMI), were used as inputs to predict the maximum oxygen uptake ( $\text{pred } \dot{V}O_{2\max}$ ). The resultant prediction model is a black box due to its high complexity and low explainability; therefore, explainable methods are necessary to extract meaningful knowledge that might have clinical applications.

<https://doi.org/10.1371/journal.pone.0282398.g001>

modelling the uncertainties in the decision boundary zone considering two-class distribution. Thus, the SVM algorithm maximizes the separation (margin) between two classes in a higher-dimensional space from the input features. In the context of regression analysis, the SVM algorithm aims to find a linear function  $f(x)$  under the condition that  $f(x)$  is within a required accuracy epsilon from the  $y(x)$  of every data point, i.e.,  $|y(x) - f(x)| \leq \epsilon$  where  $\epsilon$  is the distance between observed and predicted values for each data point. This work adopted the use of Support Vector Regression (SVR), an SVM formulation for regression problems [33] operating with a radial basis function as Kernel.

In turn, the SVM algorithms can produce more accurate and flexible models, despite the lack of some interpretability. SVM models can be considered partially interpretable models as we can determine which training data point is relevant for the prediction (i.e., the support vectors). On the other hand, it is hard to infer the contribution of features to the model's output as the input data points are projected into a higher-dimensional space to decide the predicted output value. To have accurate, and yet interpretable models, our methodology considers the use of a specific method for interpretable machine learning methods. In this work, we adopted the use of the SHapley Additive exPlanations (SHAP) method to estimate a local surrogate model, which was used to explain individual predictions. Thus, we can better manage the trade-off between accuracy and interpretability by taking advantage of robust regression algorithms and the SHAP method, which is described in the next section.

## Explainable methods

The ability to explain and interpret the prediction model's output is essential since the understanding of these models might also indicate how the human physiological systems interact with the environment. EXplainable Artificial Intelligence (XAI) is a growing research topic in the machine learning community and several methods have been proposed recently [34]. We can categorize the current approach for XAI as global and local explainable methods. While local methods provide explanations for each data point individually [35], global methods are able to provide explanations that make the entire model easier to understand, in addition to providing the rationale for the models to produce all possible results [34].

The SHapley Additive exPlanations (SHAP) approach [36] aims to explain the prediction of a given data point by computing the Shapley value for each feature input, which represents how much the features contribute to the model's prediction value. The concept of the Shapley value was originally introduced by Lloyd in the context of the cooperative game theory [37] that involves a fair distribution of both gains and costs to several players acting in coalition. Thus, Shapley value tries to ensure that each actor gains as much or more as they would have from acting independently. In explainable machine learning, the Shapley value of a feature input comprises its contribution to the model's prediction value, weighted and summed over all possible feature input combinations.

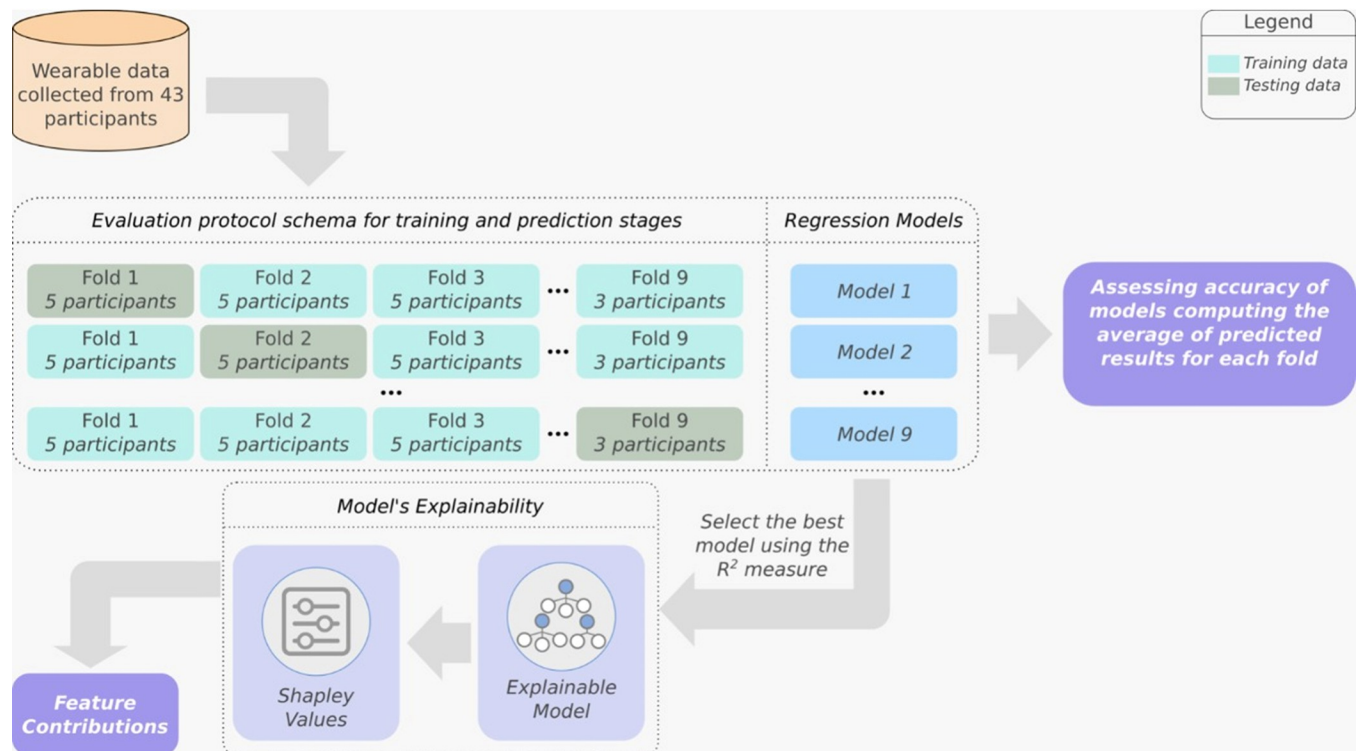
To evaluate the robustness and reliability of the regression models built in this study, we adopted the use of k-fold cross-validation. The evaluation protocol adopted in this study was designed to have predictions for each participant. To do this, we split the data into k folds ( $k = 9$ ) disjoint among participants, which means that we do not have the same participant in two or more folds. Fig 2 illustrates the methodology adopted in this study. Typically, k-fold implementations available in well-known software packages fill the gaps by duplicating some arbitrary data points when there is no integer division between the data point and the number of folds. We decided to avoid this duplication due to the possibility of biasing our results. This strategy resulted in 9 values of R-squared, Mean Absolute Error (MAE) and Pearson correlation coefficient (R), and we assess the overall performance of our algorithm by computing the average of these metrics. Then, we estimated a regression model for each fold by using the k-th fold for validation purposes and the remaining folds for training purposes. To evaluate the effectiveness of built models, we computed the average of MAE and R between the observed ( $\dot{V}O_{2-max}$ ) and predicted ( $pred \dot{V}O_{2-max}$ ) value of  $\dot{V}O_{2-max}$ . For the MAE metrics low values are better, while for the Pearson correlation coefficient, values near 1.0 indicate a near-perfect correlation.

## Statistical analysis

We calculated the R and the Bland Altman plot to further investigate the agreement level between the  $\dot{V}O_{2-max}$  and the  $pred \dot{V}O_{2-max}$  for each volunteer. The Bland-Altman plot was done in Microsoft Excel (Office package 365, Microsoft Corporation, Redmond, WA), and the prediction quality was classified as "valid" when the R-value was higher than 0.7 [38].

The source of the data variability of the Shapley value of each input feature (i.e., their contribution for the predicted value of  $pred \dot{V}O_{2-max}$ ) originated from the cross-validation method described above. The Shapley data normality was tested by the Shapiro Wilk test, and all the Shapley values presented a non-normal distribution. Thus, Friedman repeated-measures analysis along with the post-hoc Tukey Test was used to compare the final Shapley values (obtained from SHAP) between all inputs, since the Shapley values are normalized between the inputs.

In addition, the Shapley values for each eleven inputs were grouped into four domains: Anthropometric (age, weight, height, sex, and BMI), Hemodynamic (HR), Physical Activity



**Fig 2. Evaluation protocol adopted for this study.** Given a dataset containing wearable data from 43 participants, we first use the k-fold cross-validation ( $k = 9$ ) to evaluate generalization aspects of regression models. Then, we use the R-squared measure to select the best model, which is used to build an explainable model via Shapley values to assess the feature contributions. We also computed the average Mean Absolute Error (MAE), and Pearson correlation to assess the accuracy of regression models.

<https://doi.org/10.1371/journal.pone.0282398.g002>

(Acc and Cad), and Pulmonary (BR, Ve, and Vt). For each input, nine Shapley values were computed from the cross-validation. Afterward, within each domain, the Shapley values of the inputs were summed and divided by the number of inputs for this domain. Moreover, Friedman repeated-measures analysis or One Way Repeated Measures Analysis (depending on the data distribution tested by Shapiro-Wilk), with the post-hoc Tukey Test, were used to compare the final domain importance level between the four domains. Finally, the Spearman correlation (as the data were non-normally distributed) was used to verify the correlation level between all the Shapley values of the inputs (i.e., sex, age, weight, height, and BMI, BR, Ve, Acc, HR, Cad, and Vt).

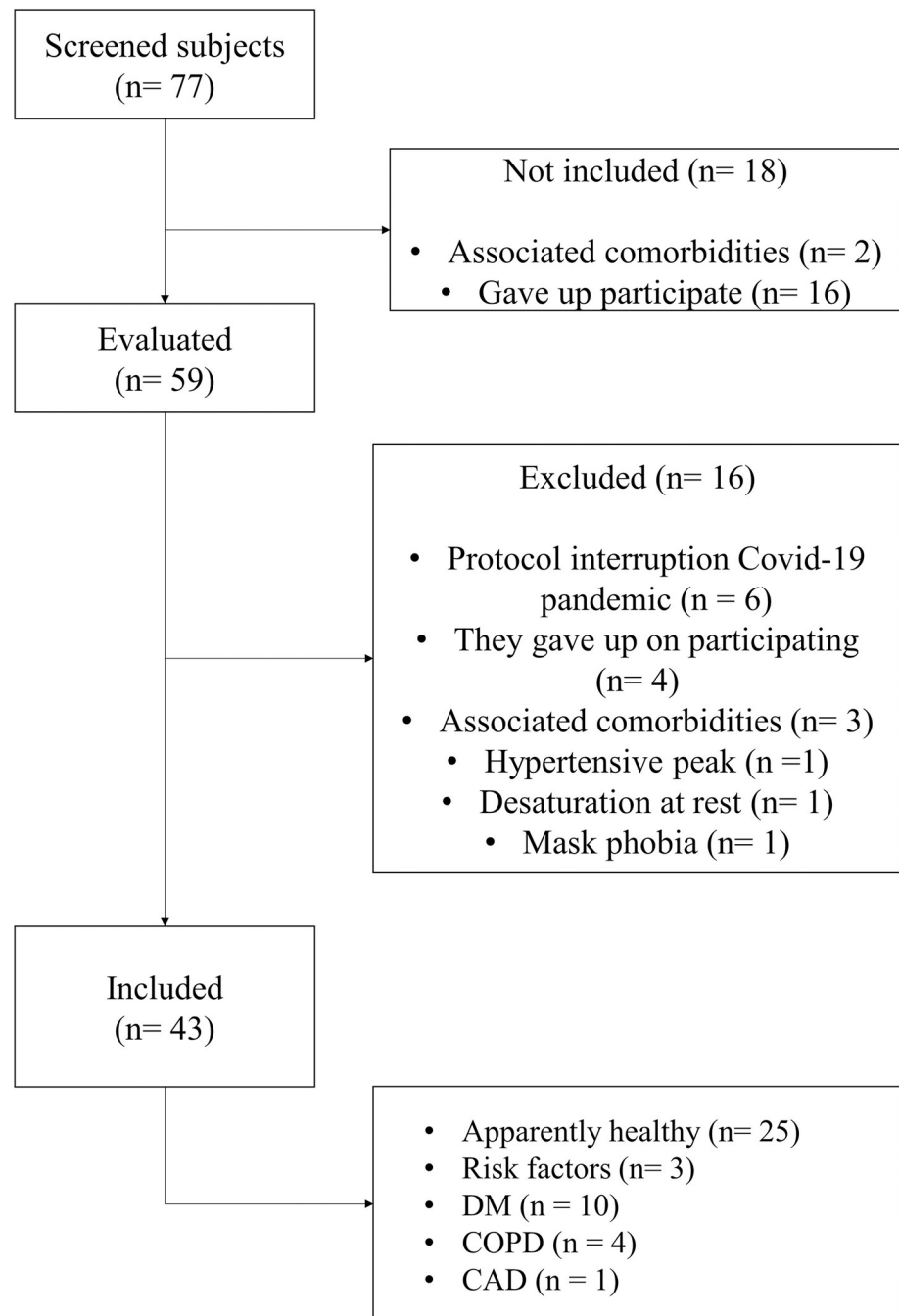
Statistical analyses and graphs were done in Sigma Plot 14.0 (Systat Software Inc, Chicago, 2018). The statistical significance level ( $p$ ) was set at 0.05.

## Results

As presented in Fig 3, 43 volunteers were included in the statistical analysis.

The demographic and anthropometric characteristics of the volunteers are presented in Table 1. From the 43 volunteers, 74.4% were men, and the age ranged from 19 to 72 years. According to a previous publication [39]; 5%, 21%, 51%, 21%, and 2% of the volunteers were classified as very low, low, fair, good, and excellent aerobic power, respectively accordingly their  $\dot{V}O_{2-max}$  in relation to the participant's weight, which indicates a broad aerobic power spectrum. All CPET were interrupted due to volitional fatigue where the mean of maximum respiratory exchange ratio ( $RER_{max}$ ) among all participants was higher than 1.1. In addition,

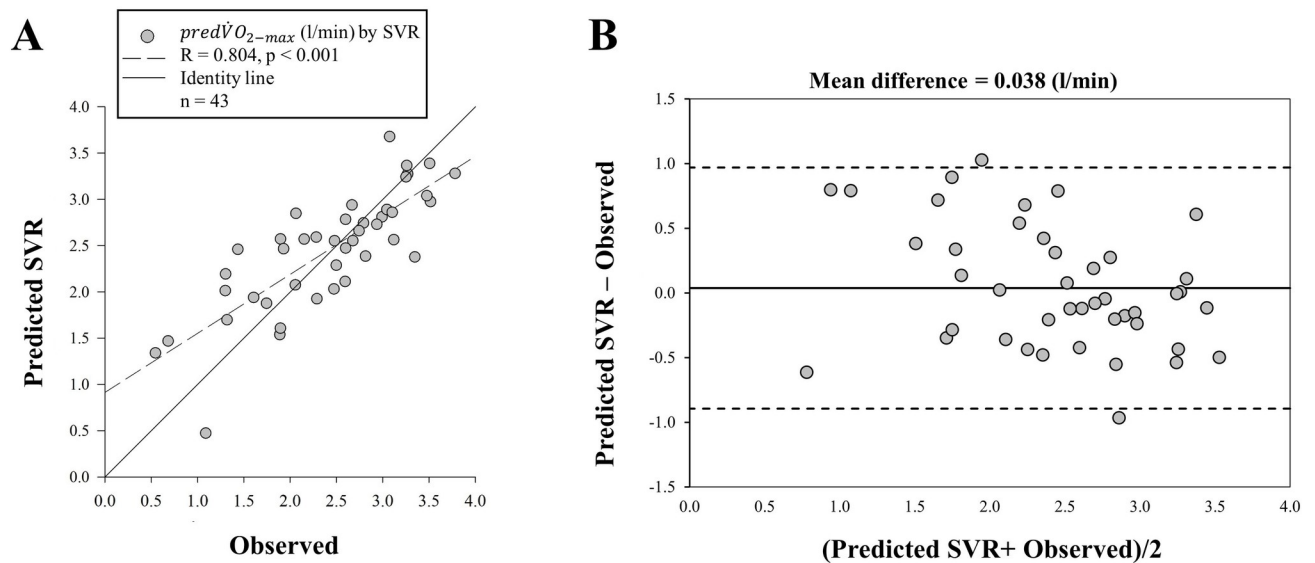




**Fig 3. Flowchart of screening, evaluation and inclusion and exclusion criteria for the study.** This flow diagram illustrates the sample size and the volunteer characteristics. DM: diabetes mellitus; COPD: chronic obstructive pulmonary disease; CAD: coronary artery disease.

<https://doi.org/10.1371/journal.pone.0282398.g003>

the mean of the maximum heart rate ( $HR_{max}$ , described in Table 1) reached during CPET (among all participants) represented 91.5% of the predicted heart rate by age ( $HR_{predicted} = 220 - \text{age}$ ) [31]. For each participant, the  $RER_{max}$  and  $HR_{max}$  were calculated as the mean value of the last 20 s of the incremental exercise. The number of days and time per day of the longitudinal data collection, as well as the average response for all variables ( $\mu BR$ ,  $\mu Ve$ ,  $\mu Vt$ ,  $\mu HR$ ,  $\mu Acc$  and  $\mu Cad$ ) are also shown in Table 1.



**Fig 4.** Linear correlation between maximum oxygen uptake during CPET and the predicted maximum oxygen uptake by machine learning technique (on letter A) and Bland-Altman plot of maximum oxygen uptake and prediction of the maximum oxygen uptake with the bias and the confidence interval (CI<sub>95</sub>) (on letter B). Support vector regression (SVR); Pearson coefficient (R).

<https://doi.org/10.1371/journal.pone.0282398.g004>

### General prediction model validation

The reproducibility and validity of the prediction of the  $\dot{V}O_{2-max}$  ( $pred \dot{V}O_{2-max}$ ) were tested using SVR, as described in Fig 4. The Bland-Altman analysis was used to verify the reproducibility between the  $\dot{V}O_{2-max}$  measured during the CPET and the  $pred \dot{V}O_{2-max}$ . We found that the mean differences between model and observations was low (0.038 l/min). Furthermore, the agreement level by the Pearson correlation coefficient ( $R = 0.804$ ,  $p < 0.001$ ) was high and positive between the  $\dot{V}O_{2-max}$  and the  $pred \dot{V}O_{2-max}$ .

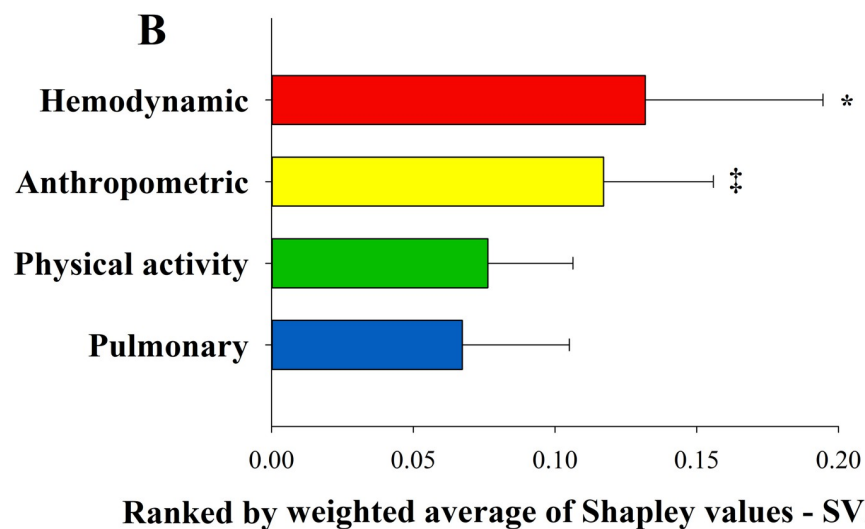
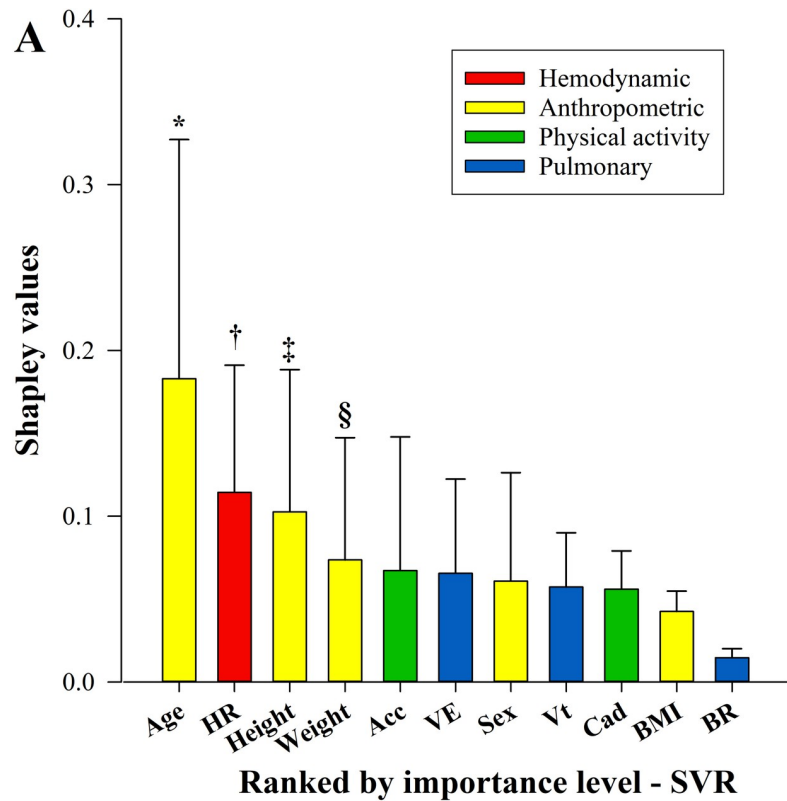
### Evaluation protocol

As mentioned before, we generalize aspects of the built models by adopting the k-fold cross-validation evaluation protocol (see Fig 2). To measure the effectiveness of models, we computed the average of Mean Absolute Error (MAE) and the Pearson correlation coefficient (R) between the observed and predicted  $\dot{V}O_{2-max}$  for each fold.

We observed a slight variability in the performance of the results achieved for the nine models. On average of MAE, the SVR regressors reached  $0.384 \pm 0.134$ . In addition, the Pearson coefficient was high and positive ( $R = 0.8$ ).

### Explainable models

According to the game theory [40], the resultant Shapley value from the above described SHAP method indicates the importance level of each input feature used to predict the variable  $pred \dot{V}O_{2-max}$ . Fig 5 shows the median Shapley values for each regression algorithm and the statistical differences between each input feature considered in this study. The feature age had the highest values, and the HR, height, weight, and Acc are ranked in the top-five list of the most important features. While the last four places are represented by respiratory measurements, such as the BR and Vt, Cad, and BMI. Moreover, when grouping the inputs into four domains (Anthropometric, Hemodynamic, Physical Activity, and Pulmonary), the



**Fig 5. Shapley values (importance level) of the inputs used to predict cardiovascular fitness.** A- Median and 25-75th percentile of Shapley values of the inputs from the Support Vector Regression (SVR). \* Significant difference between age and BR ( $p > 0.001$ ), between age and BMI ( $p > 0.001$ ), between age and Cad ( $p = 0.006$ ). † Significant difference between HR and BR ( $p = 0.003$ ). ‡ Significant difference between height and BR ( $p = 0.004$ ). § Significant difference between Weight and BR ( $p = 0.049$ ). B- Mean $\pm$ SD of Weighted average of Shapley values of the domains from the Support Vector Regression (SVR) model. \* Significant difference between Hemodynamic and Physical Activity ( $p = 0.010$ ), between Hemodynamic and Pulmonary ( $p = 0.003$ ). ‡ Significant difference between Anthropometric and Pulmonary ( $p = 0.023$ ). HR: heart rate, Acc: total hip acceleration; BMI: body mass index; BR: breathing rate, Vt: tidal volume; Cad: walking cadence, Ve: minute ventilation.

<https://doi.org/10.1371/journal.pone.0282398.g005>

Hemodynamic domain presented statistically ( $p < 0.05$ ) higher importance to predict the  $\dot{V}O_{2-max}$  compared with the Physical Activity and Pulmonary domains. Moreover, the Anthropometric domain was statistically ( $p < 0.05$ ) higher than the Pulmonary domain. We did not find any evidence of statistically significant differences ( $p > 0.05$ ) between the Hemodynamic and Anthropometric domains. Similarly, we also did not find any evidence of statistically significant differences between the Physical Activity and Pulmonary domains.

Finally, the correlations between the Shapley values of the inputs were calculated. We found some statistically ( $p < 0.05$ ) high and positive correlations between the Shapley values. For the SVR model, there were two high and positive correlations between Acc and age ( $R = 0.817$ ,  $p = 0.004$ ); and between minute ventilation (Ve) and height ( $R = 0.767$ ,  $p = 0.012$ ).

## Discussion

The SVR method showed to be reliable to predict maximal oxygen uptake (cardiovascular fitness), with an average MAE of  $0.38 \pm 0.13$  l/min. Afterward, we identified the most important inputs (and their respective domains) modelled to predict the  $\dot{V}O_{2-max}$ , using the explainable model. Hemodynamic and Anthropometric domains were more important to predict cardiovascular fitness than the Physical Activity and Pulmonary domains.

As previously reported, the CF measurement by CPET was predicted by machine learning techniques [24, 25], however the  $\dot{V}O_{2-max}$  prediction from data exclusively obtained during unobtrusive ADL protocols is still under investigation. Our results corroborate with previous studies, as mentioned before as the  $pred \dot{V}O_{2-max}$  and  $\dot{V}O_{2-max}$  were statistically similar ( $p = 0.602$  for SVR) and the predictions were reliable as verified by the low MAE and high R. Moreover, our mean errors of the Bland-Altman were low (SVR =  $0.038$  l/min), and these results are very close to what were previously described in a study ( $0.22$  l/min) [41] that used the same wearables that we used. However, our agreement limits were higher ( $0.970$  to  $-0.894$  l/min) than Amelard, Hedge and Hughson, 2021 ( $0.218$  to  $-0.262$  l/min), although our data being within the agreement limits.

The SVR using a radial basis function (RBF) as Kernel, can estimate a non-linear machine learning model that takes a small number of critical boundary samples, called support vectors. These optimize the predictions compared with models limited to single-dimension linear boundaries, such as the linear regression [42]. Therefore, SVR should allow the expression of more complexities from the input-output relationships, improving the results of regressions. In addition, SVR has been used in Medical Sciences to predict coronary artery disease and stroke [19].

Similar to our study, previous reports [24] used SVR models to identify the activity levels of unsupervised ADL from HR and accelerometer data in healthy adults (both sexes,  $25.1 \pm 6.0$  years,  $22.7 \pm 2.5$  kg/m<sup>2</sup>). These authors used linear regression models to estimate the CF by wearables. In our study, we used unobtrusive longitudinal data collected by a wearable system that also considered more physiological inputs (such as the Ve and BR from the respiratory sensors). These signals were then used to train machine learning models to predict the  $\dot{V}O_{2-max}$  of volunteers with a broad spectrum of aerobic power, which included (contrary to previous publications) apparently healthy volunteers with risk factors for NCDs and with NCDs. Therefore, our results add to the current literature to further support the use of machine learning models to predict CF in the general population, including diseased groups.

Once the prediction models were validated for this broad spectrum of CF, the Shapley method was applied to check the importance level of these 11 measured inputs (i.e., sex, age, height, weight, BMI,  $\mu$ BR,  $\mu$ Ve,  $\mu$ Acc,  $\mu$ HR,  $\mu$ Cad, and  $\mu$ Vt). All inputs were considered as

“good players”, according to the Shapley value that represents the contribution of each input feature to predict CF [28, 43].

It is worth mentioning that the Shapley values are able to better isolate each input individual influence over the predictions, thus it is less influenced by expected multicollinearities between the input features [43, 44], as we expect in the relationship between age and CF, for example. The values were expressed as positive (all the cases in our study) or negative, which means that the contribution of a particular input led to a better output prediction, i.e.,  $\text{pred } \dot{V}O_{2-\max}$  [45].

Explainable methods are used as an explanatory tool for complex models, especially when the resultant model is derived from a machine learning approach with a considerable number of hyperparameters and weights or coefficients, such as the SVR method. Thus, the Shapley values provided an approximation for the global input importance in predictions of complex responses, as we expect from biological systems [28, 46]. In fact, the CF level depends on several factors [47–51] that can be tracked by some inputs used in our study, including those measured by the wearable sensors.

Among all the inputs, age was the most important variable for both models, which also corroborates with the model-agnostic characteristics of the SHAP method. In a systematic review that aimed to identify the determinants of CF, the authors found that more than 80% of the studies identified an inverse relationship between CF and age [51]. This relationship might be justified by the influences of the aging process over the aerobic response, including the reduction of lean mass. In addition, the decrease of the  $\dot{V}O_{2-\max}$  by aging, might be also related to the reduction of the activity levels and the diseased states [52, 53].

When comparing the Shapley values (importance levels), the feature age was highly (and positively) correlated with the Shapley value of feature Acc. Thus, when age was more important for the CF definition (i.e., prediction), Acc was also more important. In a longitudinal follow-up of 8 years, Katznel, Sorkin, and Fleg 2001, found that maintaining a high level of training is inversely related to the rate of decline of aerobic power, due to aging [54]. Thus, the SVR model was able to identify this expected observation. It is known that the  $\dot{V}O_{2-\max}$  during CPET on cycle ergometry is influenced by sex, age, weight, and height [55]. Thus, in our study, the Anthropometric domain of the inputs took second place for the CF prediction. In addition, this domain was not statistically different from the Hemodynamic domain, but statistically different from the Pulmonary domain.

Between the vital signals measured by the wearable system, the Hemodynamic domain (evaluated here by the  $\mu\text{HR}$ ) was the most important input for the  $\dot{V}O_{2-\max}$  prediction, which, to some extent, corroborates with previous literature [51] that associated resting HR with maximal aerobic power. High values of resting HR were related to low values of  $\dot{V}O_2$  and consequently, lower levels of CF [56]. Altini et al., 2016 [25] found that HR explained 64% of CF variability when including sex, weight, and age as predictors, and this percentage rise as the intensity of physical activities increased. In our results, the  $\mu\text{HR}$  was very crucial to our predictions, maybe because  $\mu\text{HR}$  also includes a great deal of information regarding the resting HR as it was calculated as the average HR response throughout 7 days that comprise many resting periods.

## Study limitations

Some limitations of the present study should be considered. As described before, the feature extraction method (simple average of the longitudinal signal) might have reduced the complexity of the data from the wearable system. Thus, more studies are necessary to develop new feature extraction methods for mining longitudinal data and extracting more complex and meaningful information. Although the volunteers wore the wearables for most of the day, they

did not use it full time, including during sleep, and two volunteers practiced swimming as their main sport activity. Collecting information during sleep and water activities might improve the understanding of the CF from longitudinal data obtained from wearable sensors. It is known that HR and mean blood pressure have been used for assessing hemodynamic conditions, however assessing blood pressure must be done invasively or indirectly (which depends on the HR to be calculated), [57] thus we only consider HR as the hemodynamic domain. In this study, we used the SVR with an RFB Kernel, which means that we have a two-dimensional parameter space. Due to the low complexity of searching hyperparameters in this parameter space, we adopted a simple, effective, and well-known method named Grid Search [58]. In short, given a set of values for the variables that comprise the model, i.e., the parameters C and Gamma, the Grid Search algorithm makes a direct search on a set of all trials, which is formed by assembling every possible combination of values for the C and Gamma. This approach does not guarantee the best values for the parameters, which could be a limitation of this study. However, it is one of the most widely used approaches for hyperparameter training in machine learning [58].

## Conclusion

Cardiovascular fitness can be predicted by wearable technologies associated with artificial intelligence. Explainable models can be used to extract clinical insights from these predictions. Therefore, the  $pred \dot{V}O_{2-max}$  showed to be reproducible and valid in volunteers apparently healthy, with risk factors to develop NCDs and with NCDs. Thus, the association between longitudinal and unobtrusive biological data from wearables, machine learning, and explainable models represents a unique framework in Health Science.

## Supporting information

**S1 Dataset.**  
(XLSX)

## Author Contributions

**Conceptualization:** Maria Cecília Moraes Frade, Thomas Beltrame, Mariana de Oliveira Gois, Aparecida Maria Catai.

**Data curation:** Maria Cecília Moraes Frade, Thomas Beltrame, Mariana de Oliveira Gois, Allan Pinto, Silvia Cristina Garcia de Moura Tonello, Ricardo da Silva Torres, Aparecida Maria Catai.

**Formal analysis:** Maria Cecília Moraes Frade, Thomas Beltrame, Mariana de Oliveira Gois, Allan Pinto, Silvia Cristina Garcia de Moura Tonello, Ricardo da Silva Torres, Aparecida Maria Catai.

**Funding acquisition:** Maria Cecília Moraes Frade, Thomas Beltrame, Mariana de Oliveira Gois, Ricardo da Silva Torres, Aparecida Maria Catai.

**Investigation:** Maria Cecília Moraes Frade, Thomas Beltrame, Mariana de Oliveira Gois, Allan Pinto, Silvia Cristina Garcia de Moura Tonello, Ricardo da Silva Torres, Aparecida Maria Catai.

**Methodology:** Maria Cecília Moraes Frade, Thomas Beltrame, Mariana de Oliveira Gois, Allan Pinto, Ricardo da Silva Torres, Aparecida Maria Catai.

**Project administration:** Maria Cecília Moraes Frade, Thomas Beltrame, Mariana de Oliveira Gois, Aparecida Maria Catai.

**Resources:** Maria Cecília Moraes Frade, Thomas Beltrame, Mariana de Oliveira Gois, Ricardo da Silva Torres, Aparecida Maria Catai.

**Supervision:** Thomas Beltrame, Mariana de Oliveira Gois, Ricardo da Silva Torres, Aparecida Maria Catai.

**Validation:** Maria Cecília Moraes Frade, Thomas Beltrame, Mariana de Oliveira Gois, Aparecida Maria Catai.

**Visualization:** Maria Cecília Moraes Frade, Thomas Beltrame, Mariana de Oliveira Gois, Allan Pinto, Silvia Cristina Garcia de Moura Tonello, Ricardo da Silva Torres, Aparecida Maria Catai.

**Writing – original draft:** Maria Cecília Moraes Frade, Thomas Beltrame, Mariana de Oliveira Gois, Allan Pinto, Silvia Cristina Garcia de Moura Tonello, Ricardo da Silva Torres, Aparecida Maria Catai.

**Writing – review & editing:** Maria Cecília Moraes Frade, Thomas Beltrame, Mariana de Oliveira Gois, Allan Pinto, Silvia Cristina Garcia de Moura Tonello, Ricardo da Silva Torres, Aparecida Maria Catai.

## References

1. Dagenais GR, Leong DP, Rangarajan S, Lanas F, Lopez-Jaramillo P, Gupta R, et al. Variations in common diseases, hospital admissions, and deaths in middle-aged adults in 21 countries from five continents (PURE): a prospective cohort study. *The Lancet*. 2020; 395: 785–794. [https://doi.org/10.1016/S0140-6736\(19\)32007-0](https://doi.org/10.1016/S0140-6736(19)32007-0) PMID: 31492501
2. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, et al. Global Burden of Cardiovascular Diseases and Risk Factors, 1990–2019: Update From the GBD 2019 Study. *J Am Coll Cardiol*. 2020; 76: 2982–3021. <https://doi.org/10.1016/j.jacc.2020.11.010> PMID: 33309175
3. Budreviciute A, Damiati S, Sabir DK, Onder K, Schuller-Goetzburg P, Plakys G, et al. Management and Prevention Strategies for Non-communicable Diseases (NCDs) and Their Risk Factors. *Front Public Health*. 2020; 8: 788. <https://doi.org/10.3389/fpubh.2020.574111> PMID: 33324597
4. Lavie CJ, Ozemek C, Carbone S, Katzmarzyk PT, Blair SN. Sedentary Behavior, Exercise, and Cardiovascular Health. *Circ Res*. 2019; 124: 799–815. <https://doi.org/10.1161/CIRCRESAHA.118.312669> PMID: 30817262
5. Breneman CB, Polinski K, Sarzynski MA, Lavie CJ, Kokkinos PF, Ahmed A, et al. The Impact of Cardiorespiratory Fitness Levels on the Risk of Developing Atherogenic Dyslipidemia. *Am J Med*. 2016; 129: 1060–1066. <https://doi.org/10.1016/j.amjmed.2016.05.017> PMID: 27288861
6. Harber MP, Kaminsky LA, Arena R, Blair SN, Franklin BA, Myers J, et al. Impact of Cardiorespiratory Fitness on All-Cause and Disease-Specific Mortality: Advances Since 2009. *Prog Cardiovasc Dis*. 2017; 60: 11–20. <https://doi.org/10.1016/j.pcad.2017.03.001> PMID: 28286137
7. Blair SN. Influences of Cardiorespiratory Fitness and Other Precursors on Cardiovascular Disease and All-Cause Mortality in Men and Women. *JAMA: The Journal of the American Medical Association*. 1996; 276: 205. <https://doi.org/10.1001/jama.1996.03540030039029>
8. Després JP. Physical Activity, Sedentary Behaviours, and Cardiovascular Health: When Will Cardiorespiratory Fitness Become a Vital Sign? *Canadian Journal of Cardiology*. 2016; 32: 505–513. <https://doi.org/10.1016/j.cjca.2015.12.006> PMID: 26907579
9. Kinnunen H, Rantanen A, Kentt T, Koskim ki H. Feasible assessment of recovery and cardiovascular health: accuracy of nocturnal HR and HRV assessed via ring PPG in comparison to medical grade ECG. *Physiol Meas*. 2020; 41: 04NT01. <https://doi.org/10.1088/1361-6579/ab840a> PMID: 32217820
10. Gaye B, Tajeu GS, Vasan RS, Lassale C, Allen NB, Singh-Manoux A, et al. Association of Changes in Cardiovascular Health Metrics and Risk of Subsequent Cardiovascular Disease and Mortality. *J Am Heart Assoc*. 2020; 9. <https://doi.org/10.1161/JAHA.120.017458> PMID: 32985301

11. Kodama S, Saito K, Tanaka S, Maki M, Yachi Y, Asumi M, et al. CLINICIAN ' S CORNER Cardiorespiratory Fitness as a Quantitative Predictor of All-Cause Mortality and Cardiovascular Events. *JAMA*. 2009; 301: 2024–2035.
12. Beltrame T, Gois MO, Hoffmann U, Koschate J, Hughson RL, Frade MCM, et al. Relationship between maximal aerobic power with aerobic fitness as a function of signal-to-noise ratio. *J Appl Physiol*. 2020; 129: 522–532. <https://doi.org/10.1152/jappphysiol.00310.2020> PMID: 32730176
13. Poole DC, Jones AM. Measurement of the maximum oxygen uptake  $\dot{V}O_{2max}$  :  $\dot{V}O_{2peak}$  is no longer acceptable. *J Appl Physiol*. 2017; 122: 997–1002. <https://doi.org/10.1152/jappphysiol.01063.2016> PMID: 28153947
14. Wasserman K, Hansen J, Sue D, Stringer W, Whipp B. Principles of exercise testing and interpretation. Wilkins LW&, editor. Philadelphia; 1999.
15. Guazzi M, Arena R, Halle M, Piepoli MF, Myers J, Lavie CJ. 2016 focused update: Clinical recommendations for cardiopulmonary exercise testing data assessment in specific patient populations. *Circulation*. 2016; 133: e694–e711. <https://doi.org/10.1161/CIR.0000000000000406> PMID: 27143685
16. Weisman IM, Weisman IM, Marciniuk D, Martinez FJ, Sciruba F, Sue D, et al. ATS/ACCP Statement on cardiopulmonary exercise testing. *Am J Respir Crit Care Med*. 2003; 167: 211–277. <https://doi.org/10.1164/rccm.167.2.211> PMID: 12524257
17. Nelson N, Asplund CA. Exercise Testing: Who, When, and Why? *PM and R*. 2016; 8: S16–S23. <https://doi.org/10.1016/j.pmrj.2015.10.019> PMID: 26972264
18. Guo Y, Liu X, Peng S, Jiang X, Xu K, Chen C, et al. A review of wearable and unobtrusive sensing technologies for chronic disease management. *Computers in Biology and Medicine*. Elsevier Ltd; 2021. p. 104163. <https://doi.org/10.1016/j.compbiomed.2020.104163> PMID: 33348217
19. Krittanawong C, Virk HUH, Bangalore S, Wang Z, Johnson KW, Pinotti R, et al. Machine learning prediction in cardiovascular diseases: a meta-analysis. *Sci Rep*. 2020; 10: 16057. <https://doi.org/10.1038/s41598-020-72685-1> PMID: 32994452
20. Dunn J, Runge R, Snyder M. Wearables and the medical revolution. *Per Med*. 2018; 15: 429–448. <https://doi.org/10.2217/pme-2018-0044> PMID: 30259801
21. Beltrame T, Amelard R, Wong A, Hughson RL. Extracting aerobic system dynamics during unsupervised activities of daily living using wearable sensor machine learning models. *J Appl Physiol*. 2017; 129: 00299.2017. <https://doi.org/10.1152/jappphysiol.00299.2017> PMID: 28596271
22. Beltrame T, Amelard R, Villar R, Shafiee MJ, Wong A, Hughson RL. Estimating oxygen uptake and energy expenditure during treadmill walking by neural network analysis of easy-to-obtain inputs. *J Appl Physiol*. 2016; 121: 1226–1233. <https://doi.org/10.1152/jappphysiol.00600.2016> PMID: 27687561
23. Beltrame T, Amelard R, Wong A, Hughson RL. Prediction of oxygen uptake dynamics by machine learning analysis of wearable sensors during activities of daily living. *Sci Rep*. 2017; 7: 45738. <https://doi.org/10.1038/srep45738> PMID: 28378815
24. Altini M, Penders J, Amft O. Estimating Oxygen Uptake During Nonsteady-State Activities and Transitions Using Wearable Sensors. *IEEE J Biomed Health Inform*. 2016; 20: 469–475. <https://doi.org/10.1109/JBHI.2015.2390493> PMID: 25594986
25. Altini M, Casale P, Penders J, ten Velde G, Plasqui G, Amft O. Cardiorespiratory fitness estimation using wearable sensors: Laboratory and free-living analysis of context-specific submaximal heart rates. *J Appl Physiol*. 2016; 120: 1082–1096. <https://doi.org/10.1152/jappphysiol.00519.2015> PMID: 26940653
26. Barredo Arrieta A, Díaz-Rodríguez N, del Ser J, Bannetot A, Tabik S, Barbado A, et al. Explainable Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*. 2020; 58: 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>
27. Lundberg SM, Nair B, Vavilala MS, Horibe M, Eisses MJ, Adams T, et al. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nat Biomed Eng*. 2018; 2: 749–760. <https://doi.org/10.1038/s41551-018-0304-0> PMID: 31001455
28. Štrumbelj E, Kononenko I. Explaining prediction models and individual predictions with feature contributions. *Knowl Inf Syst*. 2014; 41: 647–665. <https://doi.org/10.1007/s10115-013-0679-x>
29. Molnar C. Interpretable Machine Learning. *Lulu.com*, editor. 2020.
30. Villar R, Beltrame T, Hughson RL. Validation of the Hexoskin wearable vest during lying, sitting, standing, and walking activities. *Applied Physiology, Nutrition and Metabolism*. 2015; 40: 1019–1024. <https://doi.org/10.1139/apnm-2015-0140> PMID: 26360814
31. Midgley AW, McNaughton LR, Polman R, Marchant D. Criteria for determination of maximal oxygen uptake: A brief critique and recommendations for future research. *Sports Medicine*. *Sports Med*; 2007. pp. 1019–1028. <https://doi.org/10.2165/00007256-200737120-00002> PMID: 18027991



32. Cortes C, Vapnik V. Support-vector networks. *Machine Learning* 1995 20:3. 1995; 20: 273–297. <https://doi.org/10.1007/BF00994018>
33. Drucker H., Burges C.J., Kaufman L., Smola A. and Vapnik V. Support Vector Regression Machines. *Adv Neural Inf Process Syst.* 1996 [cited 4 May 2022]. Available: <https://papers.nips.cc/paper/1996/hash/d38901788c533e8286cb6400b40b386d-Abstract.html>
34. Adadi A, Berrada M. Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access.* 2018; 6: 52138–52160. <https://doi.org/10.1109/ACCESS.2018.2870052>
35. Ribeiro MT, Singh S, Guestrin C. “Why should i trust you?” Explaining the predictions of any classifier. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* 2016;13-17-Augu: 1135–1144. <https://doi.org/10.1145/2939672.2939778>
36. Lundberg SM, Lee S-I. A Unified Approach to Interpreting Model Predictions. *Adv Neural Inf Process Syst.* 2017; 30.
37. Shapley L. Stochastic Games. *Proceedings of the National Academy of Sciences.* 1953; 39: 1095–1100. <https://doi.org/10.1073/PNAS.39.10.1095> PMID: 16589380
38. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007; 60: 34–42. <https://doi.org/10.1016/j.jclinepi.2006.03.012> PMID: 17161752
39. Herdy AH, Caixeta A. Brazilian Cardiorespiratory Fitness Classification Based on Maximum Oxygen Consumption. *Arq Bras Cardiol.* 2016; 389–395. <https://doi.org/10.5935/abc.20160070> PMID: 27305285
40. Lipovetsky S, Conklin M. Analysis of regression in game theory approach. *Appl Stoch Models Bus Ind.* 2001; 17: 319–330. <https://doi.org/10.1002/ASMB.446>
41. Amelard R, Hedge ET, Hughson RL. Temporal convolutional networks predict dynamic oxygen uptake response from wearable sensors across exercise intensities. *npj Digital Medicine* 2021 4:1. 2021; 4: 1–8. <https://doi.org/10.1038/s41746-021-00531-3> PMID: 34764446
42. Witten IH, FRANK E, Hall MA. *Data Mining: Practical Machine Learning Tools and Techniques.* 3rd ed. Kaufmann M, editor. Massachusetts; 2010.
43. Cesari G, Algaba E, Moretti S, Nepomuceno JA. An application of the Shapley value to the analysis of co-expression networks. *Appl Netw Sci.* 2018; 3: 3–35. <https://doi.org/10.1007/s41109-018-0095-y> PMID: 30839839
44. Li X, Dvornek NC, Zhou Y, Zhuang J, Ventola P, Duncan JS. Efficient Interpretation of Deep Learning Models Using Graph Structure and Cooperative Game Theory: Application to ASD Biomarker Discovery. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).* 2018; 11492 LNCS: 718–730. Available: <http://arxiv.org/abs/1812.06181>
45. Toba MN, Zavaglia M, Malherbe C, Moreau T, Rastelli F, Kaglik A, et al. Game theoretical mapping of white matter contributions to visuospatial attention in stroke patients with hemineglect. *Hum Brain Mapp.* 2020; 41: 2926–2950. <https://doi.org/10.1002/hbm.24987> PMID: 32243676
46. Orlenko A, Moore JH. A comparison of methods for interpreting random forest models of genetic association in the presence of non-additive interactions. *BioData Min.* 2021; 14: 9. <https://doi.org/10.1186/s13040-021-00243-0> PMID: 33514397
47. Benck L, Cuttica M, Colangelo L, Sidney S, Dransfield M, Mannino D, et al. Association between Cardiorespiratory Fitness and Lung Health from Young Adulthood to Middle Age. *Am J Respir Crit Care Med.* 2017; 195: 1236–1243. <https://doi.org/10.1164/rccm.201610-2089OC> PMID: 28248551
48. Kunutsor SK, Laukkanen T, Laukkanen JA. Cardiorespiratory Fitness is Associated with Reduced Risk of Respiratory Diseases in Middle-Aged Caucasian Men: A Long-Term Prospective Cohort Study. *Lung.* 2017; 195: 607–611. <https://doi.org/10.1007/s00408-017-0039-9> PMID: 28698945
49. Ross R, Blair S, Arena R, Church T, Després J, Franklin B, 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–e699. <https://doi.org/10.1161/CIR.0000000000000461> PMID: 27881567
50. Saxena A, Minton D, Lee D, Sui X, Fayad R, Lavie CJ, et al. Protective Role of Resting Heart Rate on All-Cause and Cardiovascular Disease Mortality. *Mayo Clin Proc.* 2013; 88: 1420. <https://doi.org/10.1016/j.mayocp.2013.09.011> PMID: 24290115
51. 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; 5. <https://doi.org/10.1186/s40798-019-0211-2> PMID: 31482208

52. Schwartz R, Buchner D. Exercise in the elderly: Physiologic and functional effects. 3rd ed. In: Hazzard WR, Blass JP, Ettinger WH, Halter JB, Ouslander JG, editors. *Principles of Geriatric Medicine and Gerontology*. 3rd ed. New York: Mc-Graw Hill; 1999. pp. 143–152.
53. Lakatta EG. Cardiovascular regulatory mechanisms in advanced age. *Physiological Reviews*. 1993. pp. 413–465. <https://doi.org/10.1152/physrev.1993.73.2.413> PMID: 8475195
54. Katzel LI, Sorkin JD, Fleg JL. A Comparison of Longitudinal Changes in Aerobic Fitness in Older Endurance Athletes and Sedentary Men. *J Am Geriatr Soc*. 2001; 49: 1657–1664. <https://doi.org/10.1046/j.1532-5415.2001.t01-1-49276.x> PMID: 11844000
55. Neder J, Nery L, Castelo A, Andreoni S, Lerario M, Sachs A, et al. Prediction of metabolic and cardiopulmonary responses to maximum cycle ergometry: a randomised study. *Eur Respir J*. 1999; 14: 1304–1313. <https://doi.org/10.1183/09031936.99.14613049> PMID: 10624759
56. Facioli TP, Philbois S v, Gastaldi AC, Almeida DS, Maida KD, Rodrigues JAL, et al. Study of heart rate recovery and cardiovascular autonomic modulation in healthy participants after submaximal exercise. *Scientific Reports* |. 2021; 11: 1–9. <https://doi.org/10.1038/s41598-021-83071-w> PMID: 33574441
57. Truijen J, van Lieshout JJ, Wesselink WA, Westerhof BE. Noninvasive continuous hemodynamic monitoring. *J Clin Monit Comput*. 2012; 26: 267. <https://doi.org/10.1007/s10877-012-9375-8> PMID: 22695821
58. Bergstra J, Yoshua B. Random Search for Hyper-Parameter Optimization Yoshua Bengio. *Journal of Machine Learning Research*. 2012; 13: 281–305. Available: <http://scikit-learn.sourceforge.net>.