Medical Research Archives

European Society of Medicine

OPEN ACCESS

Published: November 30, 2023

Citation: Inbar, O., et al., 2023. A Machine Learning Approach to Classify Exercise Limitation Severity Using Cardiopulmonary Exercise Testing - Development and Validation. Medical Research Archives, [online] 11(11). https://doi.org/10.18103/mra.v11i11.4721

Copyright: © 2023 European Society of Medicine. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

DOI: https://doi.org/10.18103/mra.v11i11.4721

ISSN: 2375-1924

RESEARCH ARTICLE

A Machine Learning Approach to Classify Exercise Limitation Severity Using Cardiopulmonary Exercise Testing - Development and Validation

Or Inbar[1], Omri Inbar[2*], R. Reuveny[3], M. Segel[2,3], H. Greenspan[1] and M. Scheinowitz[1,2]

[1]Department of Biomedical Engineering, Tel-Aviv University, Israel
[2]Faculty of Medicine, Tel-Aviv University, Tel-Aviv, Israel
[3]Pulmonary Institute, Sheba Medical Center Tel-Hashomer, Israel

[*]inbar@wincol.ac.il

ABSTRACT

Background: There is no agreement on the best strategy for interpreting/ analyzing cardiopulmonary exercise test (CPET) results.

Aims: This study aims to evaluate the feasibility of using computer-aided algorithms to evaluate CPET data to identify the exercise limitation/ intolerance level.

Methods: This study used 206 retrospective CPET files from the Pulmonary Institute and the Cardiology Rehabilitation Center at the Sheba Medical Center and 50 from the exercise physiology laboratory at the Washington-Hill College, both in Israel. Eighty patients with confirmed primary cardiovascular-related exercise limitation, seventy-six with ventilatory-related exercise limitation, and fifty healthy (none or very mild exercise restraints) subjects comprised the pool of examined CPET data. Support Vector Machine (SVM) Learning was performed on 150 (50 in each group) of the 206 CPETs, while model validation was performed on the remaining 56 CPET files. By applying the K-means clustering method, distribution analysis was used to compare the SVM interpretive module's performance to that of senior cardiologists, pulmonologists, and expert exercise physiologists.

Results: Overall, the proposed interpretive model has a predictive power of between 78% and 100%, as shown by its ability to correctly classify the degree of exercise limitation.

Conclusions: The proposed machine-learning CPET interpretive module was highly sensitive and specific in identifying patients with mild, moderate, or severe cardiovascular- or ventilatory-related exercise limitations/ intolerance. Comparable modules may be applied to additional (kinds of exercise limitations) and larger populations, making this tool powerful and clinically applicable.

Keywords: Cardiopulmonary exercise testing (CPET), Support Vector Machine, machine learning, exercise limitation severity, cardiovascular, ventilatory, data analysis.

Introduction

Cardiopulmonary exercise testing (CPET) is a valuable tool used in various fields of medicine to assess an individual's exercise capacity and diagnose exercise limitations[1,2,3]. It provides crucial insights into the functioning of the cardiovascular, respiratory, metabolic, and gas exchange systems during exercise, enabling clinicians to identify the underlying causes of exercise intolerance and prescribe appropriate interventions[2,3,4]. Interpretation of CPET is a multivariate time series problem involving simultaneous manual assessment of generated heart rate, ventilation, gas exchange (oxygen uptake), and carbon dioxide output. These time series' manual evaluations and traditional analytics are simplified to peak values, summary indices, and slopes[2,4]. A computer-aided module can automate the analysis process, ensuring accuracy, consistency, and saving time. It can perform complex calculations and algorithms more quickly and consistently than manual analysis, reducing the risk of human errors and providing reliable results.

In recent years, machine learning (ML) has emerged as the new flexible learning framework in which automatic extraction of relevant features happens based on the learning objective. This cutting-edge approach capitalizes on the power of machine learning algorithms, including Support Vector Machines (SVM), to derive meaningful patterns and correlations from complex CPET data[5,6]. SVMs, known for creating accurate decision boundaries between different data classes, are particularly advantageous in classifying and categorizing CPET results based on factors such as heart rate, oxygen consumption, pulmonary function, and gas exchange[7]. Machine learning algorithms have shown promising potential in extracting meaningful patterns from large and complex datasets, leading to more accurate and efficient diagnosis and treatment decisions. In line with this progress, researchers and clinicians have begun exploring the application of ML in the analysis of CPET data[5,6,7,8].

Only a few studies have applied ML to identify clinically relevant phenogroups from CPET data.

By harnessing the power of advanced algorithms, this methodology seeks to overcome the limitations of traditional subjective interpretations, offering a more objective, standardized, and time-saving approach to CPET analysis.

The present study continues our recently published article[5], which presents a novel approach using machine learning algorithms to identify individuals suffering from chronic heart failure (CHF) and chronic obstructive pulmonary disease (COPD) or is considered healthy[5]. The present study aims to present a novel machine learning analysis approach for classifying severity levels of exercise limitations (Cardiovascular-related exercise limitations (CREL) and ventilatory-related exercise limitations (VREL) using CPET data. Notably, in the majority of referrals for CPETs, the goal is **not** to diagnose a specific pathology (there are other ways to do this (though more expensive and time-consuming) but rather to identify the cause and degree of exercise intolerance - or to locate the reason(s) of unexplained dyspnea.

Materials and Methods

PARTICIPANTS - The data set in the present study consists of anonymized results from

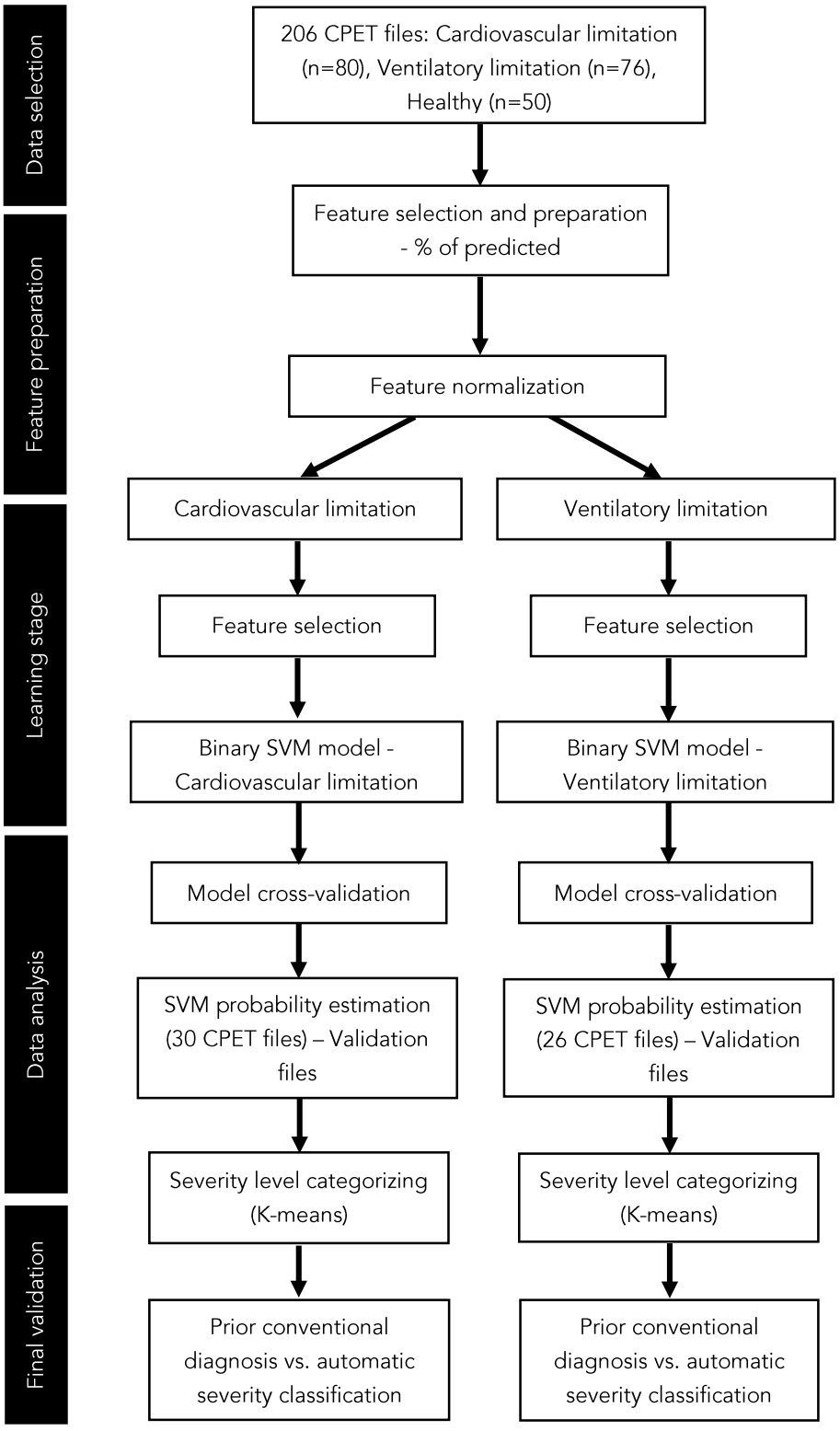
retrospective cardiopulmonary exercise tests (CPETs) of 156 patients with two mains clinically diagnosed conditions: cardiovascular-related (n=80) and ventilatory-related (n=76) exercise limitations (CREL and VREL, respectively). It should be pointed out that some of the studied patients presented with the coexistence of CREL and VREL, and their final group assignments were based on the most prominent CPET findings and symptoms (primary or secondary).

An additional 50 CPET files were obtained from studies carried out on healthy individuals (with no apparent particular exercise-related weakness) (H) at the exercise physiology laboratory of the Washington-Hill College in Yavne, Israel.

Each CPET file (including PFTs) was reviewed by a senior cardiologist, pulmonologist, and exercise physiologist well-versed in CPET studies using a standardized manual interpretation strategy[9,10]. Their ultimate verdict regarding the severity of exercise limitation was used as the "gold standard" to validate and compare the proposed model's results to that of the "gold standard".

THE EXERCISE PROTOCOL – All the CPET protocols in the study used cycle ergometers (Ergoselect 1200, Germany) with three phases: rest, test, and recovery. The CPETs of the "patients" were performed at the exercise physiology laboratory of the Pulmonary Institute at the Sheba Medical Center in Ramat-Gan and those of the healthy participants at the exercise physiology laboratory of the Washington-Hill College in Yavne, both in Israel. The CPET protocols and metabolic carts were the same in the two laboratories. The primary criteria for inclusion in the study cohort were: technically sound CPET, technically good resting pulmonary function test (PFT), maximal effort or symptom-limited CPETs (respiratory exchange ratio RER ≥ 1.00; test duration ≥ 6 min), and age ≥ 35 years old). The Institutional Review Board (IRB) of the Sheba Medical Center approved the protocol (No. 1730-14-SMC). Informed consent was not required due to the observational and retrospective nature of the study design.

Workflow of the study procedures



Before performing the CPETs, all study participants completed a pulmonary function test per the American Thoracic Society (ATS) guidelines[11]. An incremental symptom-limited maximal exercise test was performed after a 3-minute rest and 3 min of unloaded pedaling. Expired O_2 and CO_2 gases and the airflow rate were measured breath-by-breath

through a facemask connected to a metabolic cart (all from COSMED, Italy). Gas analyzers (O_2 and CO_2) were calibrated before each test. The airflow sensor was calibrated daily. The exercise protocols ensured that subjects reached volitional exhaustion within 8-12 minutes of incremental exercise. Work rate increments were individually adjusted, ranging from 5 to 25 watts.min^{-1}. Before entering the CPET data into the selected SVM learning and the respective validation processes, maximal and submaximal values of each CPET file were obtained using conventional algorithms embedded in the metabolic carts (COSMED, Italy). The CPET protocols and metabolic carts were the same in the two laboratories. Then, the relations of those measured values to their corresponding normal (predicted) values were calculated and assigned as **% of predicted**. The predicted "normal" values were based on Inbar et al.[12] and Wasserman et al.[10] CPET's reference values. Using % predicted values as input data for the SVM module ensured unbiased physiological feature comparisons (peak and submaximal) across various test procedures, ergometers, and populations with varying physical, physiological, and pathological characteristics.

NORMALIZING RANGES OF % OF PREDICTED VALUES (80%-100%) - During CPET, assorted physiological variables are measured with their widely spread "normal" peak values. To overcome the above problem and standardize the CPET predicted normal ranges, we rescaled the original boundaries of all expected normal ranges into equal limits of 80% to 100% of predicted normal (commonly used in medical sciences).[5] This procedure was accomplished by using three points to apply a linear regression equation for each CPET variable: The lower limit of the predicted normal range was set to 80% of normal, the average of the expected normal range to 90% of normal, and the upper limit to 100% of normal. Feature scaling is mapping the feature values of a dataset into the same range and is crucial for machine learning algorithms such as the SVM[13,14]. It should be stated that training an SVM classifier includes deciding on a boundary between classes. This boundary is known to have the maximum distance from the nearest point on each data class and differs for nonscaled and scaled cases. Also, the linear scaling of the input data in our study was done to avoid attributes with greater numeric ranges that could dominate those with smaller ones[15].

FEATURE SELECTION - Each "limitation" (CREL or VREL) was classified separately. A feature selection algorithm [SVM-recursive feature elimination (SVM-RFE)] with a correlation bias reduction process (CBR) was employed to identify the relevant features in the CPET dataset. The algorithm ranks the features based on their impact on the classification process[15]. As indicated above, the feature reduction algorithm was implemented separately for patients with moderate-to-severe ventilatory-related and cardiovascular-related limitations.

Feature selection can be achieved by choosing a group of top-ranked features. SVM-RFE is an application of RFE using the weight magnitude as a ranking criterion.

A linear SVM model was trained in each iteration of the recursive feature elimination (RFE). The feature with the minimal ranking criterion was removed since it has the least

impact on classification. The remaining features were kept for the SVM model in the next iteration. This process was repeated until all the features had been removed. Then, the features were sorted according to the order of removal. Since SVM-RFE includes inter-correlated features, which could incorrectly estimate the feature's importance, an additional algorithm was implemented: The Correlation Bias Reduction module (CBR)[15].

The SVM-RFE+CBR procedures produced a list of selected exercise limitation-related CPET features ranked in terms of their impact on the quality and reliability of the respective predictive model. (see Tables 5 and 6).

The SVM model implementations in this study were carried out using MATLAB R2013b, applying the Library for Support Vector Machines (LIBSVM) toolbox[16].

THE SVM ALGORITHM - is a supervised (or unsupervised) machine learning technique widely used in pattern recognition and classification problems. It includes a set of supervised (or unsupervised) learning methods developed in the 1990s[17,18] and is used to solve classification and regression problems. SVM is one of the most popular techniques for supervised classification[19], built on the structural risk minimization (SRM) induction principle and has found success in a variety of applications[20]. Machine learning approaches can be valuable for interpreting cardiopulmonary exercise tests (CPETs) by leveraging the power of data analysis and pattern recognition. However, the success of many applications using the SVM depends on the initial manual choice of features. As mentioned above, the present study looked at individuals of different genders, ages, weights, heights, levels of physical fitness, and kinds and degrees of exercise limitation.

THE LEARNING STAGE - Included in the exercise limitation severity (ELS**)** classification were patients demonstrating moderate to severe **ventilatory**-related exercise limitation (VREL; N=50), **cardiovascular**-related exercise limitation (CREL; N=50), and **healthy** participants (*H*; N=50).

We used the LIBSVM binary classifier for the ELS classification.

Two SVM binary models (one for ventilatory and one for cardiovascular) were separately applied following input of selected CPET features (normalized % of predicted).

CROSS-VALIDATION OF SVM MODELS – Following the SVM binary model learning stage, cross-validation was conducted for both (ventilatory- and cardiovascular-related) binary models.

Two cross-validation methods were carried out: leave-one-out and cross-validation repeated random sub-sampling.

FINAL VALIDATION OF THE EXERCISE LIMITATION SEVERITY CLASSIFICATION - 56 CPET files, not used for the learning stages, were added for the final confirmation of the SVM classification module. Of those patients, 26 were initially classified as suffering from ventilatory-related and 30 from cardiovascular-related exercise limitations. These patients showing various degrees of exercise limitations (mild, moderate, severe) were used for the final validation stage of the proposed classification module. Validation of the Model ELS classification (ventilatory- or cardiovascular-related) was based on the SVM probability estimation. The K-means

clustering method categorized the SVM probability values into three severity levels (mild, moderate, and severe)[21,22]. The categorization was carried out by minimizing the sum of squares of distances between the objects and the corresponding cluster or class centroid. Results from the SVM and K-means classifications of the exercise limitation severity were compared with those initially assigned by the attending physicians (all experts in assessing CPET results).

Statistical analyses

PARTICIPANTS' PHYSICAL CHARACTERISTICS AND CPET RESULTS - Discrete values (participant physical characteristics and CPET peak values) were calculated and are presented as means ± Standard deviation (SD).

A one-way analysis of variance (ANOVA) with a multiple comparison test was conducted to compare the three study groups' CPET peak and sub-peak values. A p-value of ≤0.05 was considered indicative of statistical significance.

PERFORMANCE QUANTIFICATION - The model's cross-validation outcomes (probability estimates) are presented as means±SD. The SVM classification outcomes versus prior physicians' clinical diagnoses were compared regarding test sensitivity, specificity, accuracy, and overall precision and presented in confusion tables. The confusion matrix contains information about actual and predicted classifications.

Sensitivity, specificity, accuracy, and overall precision were calculated based on the following formulas:

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

$$Percision = \frac{TP}{TP + FP}$$

TP, FP, TN, and FN represent the number of true positives, false positives, true negatives, and false negatives. A p value of ≤ 0.05 was considered statistically significant

Results

PARTICIPANTS - Tables 1 and 2 show the participants' physical characteristics during the learning and validation stages.

Table 1. Physical characteristics of the participants in the learning stage (means±SD)

Variable	CREL (N=50)	VREL (N=50)	Healthy (N=50)
Age [yr]	52.2 ±13.3	64.4±10.2	45.7±9.3
Height [cm]	172.3±6.2	169.0±6.7	173.0±4.5
Weight [kg]	79.4±11.7	70.7±13.3	76.6±5.6

CREL=Cardiovascular-related limitations; VREL=Ventilatory-related exercise limitations.

Table 2. Physical characteristics of the study participants of the validation stage (means±SD)

Variable	CREL (N=26) (18 males and 8 females)	VREL (N=30) (17 males and 13 females)
Age [yr]	54.3 ±13.5	65.6±7.3
Height [cm]	171.6±6.6	167.8±6.1
Weight [kg]	80.7±15.0	75.6±14.6

CREL=Cardiovascular-related limitation; VREL=Ventilatory-related exercise limitation.

CPET RESULTS - Tables 3 and 4 in the supplementary materials present the CPET raw data and selected resting pulmonary characteristics measured during the learning and validation stages, respectively. The tables detail the patient's measured test results (peak and submaximal values) and the % of the measured values of the corresponding predicted (expected) values of all CPET parameters. In addition, the tables show the statistical differences between the three groups.

Table 3. Comparisons of CPET actual and % of predicted values among the three studied groups- learning stage (means ± SD).

Variables	CREL (n=50)		VREL (n=50)		H (n=50)	
	Measured[a]	% of pred.[d]	Measured[b]	% of pred.[e]	Measured [c]	% of pred.[f]
Time [min]	9.3±2.4[b]	90.4±9.5	10.9±2.4[a]	95.7±16.5	9.9±1.8	91.8±8.1
Peak WR [watt]	83.5±35.0[bc]	43.8±13.0[f]	57.7±19.3[ac]	42.6±10.8[f]	180.2±27.1[ab]	90.7±6.8[de]
Peak VO_2 [l/min]	1.1±0.3[c]	45.2±13.1[ef]	1.2±0.4[c]	60.0±15.7[df]	2.5±0.2[ab]	92.2±9.8[de]
Peak VCO_2 [l/min]	1.1±0.4[c]	35.8±10.6[ef]	1.1±0.5[c]	43.5±15.0[df]	3.1±0.5[ab]	88.9±7.1[de]
RER	1.08±0.06[c]	85.0±4.3[f]	1.08±0.07[c]	85.0±4.8[f]	1.19±0.09[ab]	92.7±6.6[de]
Peak VO_2/kg [ml/kg/min]	13.3±4.2[bc]	42.2±9.7[ef]	16.9±3.7[ac]	64.3±10.1[df]	32.4±4.7[ab]	91.0±6.5[de]
VO_2/WR slope	7.6±1.0[bc]	71.1±8.0[ef]	10.4±1.1[a]	90.4±6.0[d]	10.2±1.2[a]	89.2±7.4[d]
Peak HR [beat/min]	103.7±13.1[bc]	50.8±6.1[ef]	114.0±14.3[ac]	58.8±6.4[df]	175.1±10.2[ab]	91.1±7.0[de]
Peak O_2Pulse [(ml/kg/beat)x100]	12.1±1.9[bc]	65.9±8.4[ef]	13.8±2.5[ac]	80.4±11.2[df]	19.0±2.1[ab]	93.5±6.7[de]
BR [l]	51.8±19.4[bc]	109.4±15.6[ef]	-1.6±13.3[ac]	66.7±10.7[df]	27.4±10.7[ab]	89.7±7.1[de]
Peak VE [l/min]	51.5±15.8[bc]	51.8±13.1[f]	44.4±10.1[ac]	55.9±10.4[f]	98.5±13.0[ab]	90.1±6.1[de]
Peak Vt [l/min]	1.3±0.3[bc]	48.0±9.5[f]	1.2±0.2[ac]	46.5±8.0[f]	2.6±0.2[ab]	90.6±5.6[de]
Peak Bf [1/min]	35.2±5.5[bc]	69.9±10.0[ef]	31.5±3.9[ac]	63.0±7.7[df]	47.6±5.2[ab]	90.1±7.4[de]
Peak VE/VO_2	42.5±6.2[c]	96.7±11.7[ef]	40.1±6.0	89.2±10.1[d]	39.5±4.7[a]	91.8±8.4[d]
Peak VE/VCO_2	42.0±5.4[bc]	125.1±16.9[ef]	37.6±4.9[ac]	108.0±14.1[df]	28.4±2.7[ab]	87.6±7.1[de]
Peak $PETO_2$ [mmHg]	113.8±4.8[bc]	85.6±5.7[ef]	108.3±5.7[ac]	79.6±5.3[df]	118.5±6.8[ab]	90.8±7.2[de]
Peak $PETCO_2$ [mmHg]	34.8±3.7[b]	92.0±9.6[e]	38.6±5.1[ac]	101.8±13.1[df]	33.5±2.6[b]	88.8±6.6[e]
VAT % of pred. VO_2/kg [%]	34.3±5.0[bc]	60.9±8.9[ef]	48.4±9.8[ac]	80.1±9.1[df]	60.2±9.2[ab]	90.1±6.8[de]
ECG grading (%)	N/A	79.6±12.2[ef]	N/A	99.2±3.4[df]	N/A	89.8±7.6[de]
O_2Pulse response grading [%]	N/A	74.4±12.2[ef]	N/A	86.2±6.6[df]	N/A	91.4±6.9[de]
SaO_2 [%]	97.8±1.5[bc]	98.7±2.6[ef]	88.5±4.9[ac]	80.1±9.1[df]	93.5±3.4[ab]	89.0±7.0[de]
VE/VCO_2 slope	38.1±3.8[bc]	130.3±11.3[ef]	31.9±5.1[ac]	105.6±16.6[df]	25.4±2.4[ab]	90.2±7.4[de]
FVC [l]	3.6±0.8[bc]	80.7±10.3[ef]	2.5±0.5[ac]	67.7±8.8[df]	4.5±0.7[ab]	90.2±7.5[de]
FEV1 [l]	2.8±0.6[bc]	78.3±10.3[ef]	1.2±0.2[ac]	43.0±6.0[df]	3.7±0.5[ab]	90.7±6.3[de]
FEV1/FVC [%]	88.2±8.8[bc]	97.3±6.1[ef]	49.0±4.1[ac]	64.8±5.2[df]	77.7±10.9[ab]	89.1±7.3[de]

Data presented as mean ± SD.

WR = work rate; VO_2 = Oxygen Consumption; VCO_2 = Carbon Dioxide Production; RER = Respiratory Exchange Ratio; HR = Heart Rate; O_2Pulse = Oxygen Puls; BR = Breathing Reserve; VE = Minute Ventilation; Vt = Tidal Volume; Bf = Breathing Frequency; $PETO_2$ = End-Tidal Oxygen tension; $PETCO_2$ = End-Tidal Carbon Dioxide tension; VAT = Ventilatory Anaerobic Threshold; % predicted = percent of predicted normal value; ECG = Electrocardiography; SaO_2 = Oxygen Saturation; FVC = Forced Vital Capacity; FEV1 = Forced Expiratory Volume in 1 second; CHF = Chronic Heart Failure; CREL=patients diagnosed with cardiovascular-related limitations; VREL= patients diagnosed with ventilatory-related exercise limitations; H=healthy (normal) patients.

[a] Letters a, b, and c represent significant differences ($P < 0.05$) related to measured values between the specified groups.

[b] Letters d, e, and f, represent significant differences ($P < 0.05$) related to % of predicted values between the specified groups.

O_2pulse response grading (% of predicted normal)[23,24]: Up-sloping 90%, flat 50%, down-sloping 30%.

ECG (%) changes in ECG tracings were classified based on clinical severity: Normal 100%, nonspecific changes 80%, specific T-wave changes 75%, ventricular conduction defects 70%, atrial arrhythmia 60%, ST depression (>2 mm) 50%, ventricular arrhythmia 40%, ST elevation (>2 mm) 30%.

Table 4. Comparisons of CPETs actual and % of pred. values among the three studied groups - validation stage (means ± SD).

Variables	CREL (n=26)		VREL (n=30)	
	Measured[a]	% of pred.[c]	Measured[b]	% of pred.[d]
Time [min]	8.2±1.7[b]	85.8±5.1[d]	12.5±2.2[a]	102.2±14.5[c]
Peak WR [watt]	74.3±37.8	41.3±16.5	58.6±20.3	43.4±18.3
Peak VO_2 [l/min]	1.0±0.2	40.5±10.5	1.0±0.2	53.8±14.7
Peak VCO_2 [l/min]	1.1±0.3	31.7±12.7	1.0±0.2	42.9±13.6
RER	1.06±0.05	84.1±4.3	1.07±0.06	84.3±3.8
Peak VO_2/kg [ml/kg/min]	13.7±5.0	44.7±17.4[d]	19.2±5.9	68.7±19.2[c]
Slope VO_2/WR	8.9±2.8	80.5±22.0	10.7±1.7	85.7±12.0
Peak HR [beat/min]	99.1±21.2[b]	49.3±9.5[d]	116.5±16.2[a]	59.9±8.6[c]
Peak O_2Pulse [(ml/kg/beat)x100]	12.3±4.5	67.4±25.4	14.1±4.1	82.9±21.2
BR [l]	41.8±18.2[b]	101.0±14.5[d]	17.2±11.0[a]	81.8±8.6[c]
Peak VE [l/min]	45.5±16.6	47.2±17.8	47.3±10.5	58.5±17.2
Peak Vt [l/min]	1.4±0.4	50.5±32.8	1.2±0.3	51.9±11.2
Peak Bf [1/min]	32.9±8.7	66.1±16.4	32.1±7.2	62.1±13.2
Peak VE/VO_2	48.4±12.3[b]	110.0±26.0[d]	40.7±7.3[a]	89.0±13.5[c]
Peak VE/VCO_2	46.2±10.1[b]	136.2±29.6[d]	39.7±7.2[a]	113.0±19.4[c]
Peak $PETO_2$ [mmHg]	115.0±5.5	86.1±6.3	111.3±5.2	82.1±6.2
Peak $PETCO_2$ [mmHg]	33.8±4.7	89.2±12.2	35.5±4.3	93.6±11.0
VAT % of pred. VO_2/kg [%]	33.7±11.6[b]	57.5±17.1[d]	49.1±13.6[a]	79.3±14.7[c]
ECG grading [%]	N/A	80.1±18.7[d]	N/A	98.1±6.3[c]
O_2Pulse response grading [%]	N/A	71.1±21.8	N/A	77.2±19.5
SaO_2 [%]	98.1±1.3[b]	98.4±3.3[d]	92.5±3.4[a]	88.3±7.5[c]
Slope VE/VCO_2	46.0±17.7[b]	157.0±55.3[d]	38.0±11.2[a]	124.6±37.0[c]
FVC [l]	3.8±1.3	84.2±19.4	3.6±1.0	87.2±22.7
FEV1 [l]	3.1±1.3[b]	81.6±24.7[d]	1.2±0.4[a]	52.8±19.5[c]
FEV1/FVC [%]	85.9±19.4[b]	97.1±18.1[d]	54.8±16.5[a]	70.3±17.6[c]

Data presented as mean ± SD.

See Table 4 for abbreviations.

[a] Letters a, and b, represent significant differences ($P < 0.05$) related to measured values between the specified groups.

[b] Letters c, and d, represent significant differences ($P < 0.05$) related to % of predicted values between the specified groups.

O2pulse response grading (% of normal)[23,24]: Up-sloping 90%, flat 50%, down-sloping 30%.

ECG grading (% of normal), changes in ECG tracings were classified based on clinical severity: Normal 100%, nonspecific changes 80%, specific T-wave changes 75%, ventricular conduction defects 70%, atrial arrhythmia 60%, ST depression (>2 mm) 50%, ventricular arrhythmia 40%, ST elevation (>2 mm) 30%.

FEATURE SELECTION - A feature reduction algorithm (SVM-RFE+CBR) was implemented to identify the most contributory CPET features for each level of exercise limitation (ventilatory, cardiovascular, or healthy).

Table 5 and 6 present the relative contribution of CPET parameters to the respective cardiovascular-related and ventilatory-related ELS.

Table 5. CPET feature contribution to correct classification of ELS in order of impact – Cardiovascular-related exercise limitation severity classification.

Parameter	Sn (%)	Sp (%)	Acc (%)	Pr (%)
Peak VO_2/kg	83	92	89	83
Peak HR	63	82	76	63
Peak O_{2Pulse}	80	90	87	80
VE/VCO_2 Slope	83	92	89	83
Peak VE	83	92	89	83
BR	**97**	**98**	**98**	**97**
Peak Vt	77	88	84	77
SaO_2	77	88	84	77
Peak VE/VCO_2	73	87	82	73
O_{2Pulse} response	73	87	82	73
Peak WR	77	88	84	77
VO2/WR Slope	80	90	87	80
Peak Bf	77	88	84	77
FEV1	77	88	84	77
VAT	77	88	84	77
FEV1/FVC	80	90	87	80
ECG	83	92	89	83
Peak $PETO_2$	83	92	89	83
Peak $PETCO_2$	83	92	89	83
Peak VE/VO_2	83	92	89	83
FVC	83	92	89	83

Sn=Sensitivity; Sp=Specificity; Acc=Accuracy; and Pr=precision.

Table 6. CPET feature contribution to correct classification of ELS in order of impact – Ventilatory-related exercise limitation severity classification.

Parameter	Sn (%)	Sp (%)	Acc (%)	Pr (%)
FEV1	73	87	82	73
Peak WR	77	88	85	77
Peak HR	85	92	90	85
Peak Vt	85	92	90	85
FEV1/FVC	65	83	77	65
Peak Bf	85	92	90	85
Peak VE	81	90	87	81
BR	**92**	**96**	**95**	**92**
FVC	58	79	72	58
ECG	81	90	87	81
Peak VO_2/kg	81	90	87	81
SaO_2	81	90	87	81
Peak $PETCO_2$	81	90	87	81
O_{2Pulse} response	77	88	85	77
Peak VE/VCO_2	77	88	85	77
Peak $PETO_2$	81	90	87	81
Peak O_{2Pulse}	77	88	85	77
VO2/WR Slope	77	88	85	77
VAT	77	88	85	77
Peak VE/VO_2	77	88	85	77
VE/VCO_2 Slope	77	88	85	77

Sn=Sensitivity; Sp =Specificity; Acc=Accuracy; and Pr=precision.

We selected the best combination of features (CPET parameters) based on feature reduction ranking, considering the sensitivity, specificity, accuracy, and precision for the SVM ELS classifications. Consequently, six and eight CPET parameters for cardiovascular-related and ventilatory-related limitations provided the best classification outcomes (pale gray features).

The selected features for the two SVM exercise limitation severity models, in order of impact, were:

For the cardiovascular-related limitations: Peak VO_2/kg, Peak HR, Peak O_{2Pulse}, VE/VCO_2 Slope, Peak VE, and BR.
For the ventilatory-related limitations: FEV1, Peak WR, Peak HR, Peak Vt, FEV1/FVC, Peak Bf, Peak VE, and BR.

CROSS-VALIDATION OF SVM MODELS – Table 7 and 8 present the results of the cross-validation process estimating the accuracy of the predictive binary models' performance for the ELS classifications. The results show significant separation between the two groups (healthy and exercise-limited) and high similarity within each group of the exercise-limited patients (low SD) in both binary models (cardiovascular and ventilatory - related).

Table 7. Results of the SVM binary model cross-validation – Cardiovascular-related exercise limitation severity classification.

Physician			SVM probability estimation average (%)	
Cross-Validation method	No. of iterations	Limitation	Cardiovascular limitation	Healthy patients
50% (training) – 50% (validation)	1250	Cardiovascular	**95±5**	4±3
		Healthy patients	5±5	**96±3**
70% (training) – 30% (validation)	450	Cardiovascular	**96±5**	4±5
		Healthy patients	4±2	**96±2**
80% (training) – 20% (validation)	200	Cardiovascular	**97±4**	3±4
		Healthy patients	3±3	**97±3**
Leave-One-Out	100	Cardiovascular	**98±3**	2±3
		Healthy patients	3±2	**97±2**

Values are means ± SD. Bold numbers denote the average probability estimates of the respective group.

Table 8. Results of the SVM binary model cross-validation – Ventilatory-related exercise limitation severity classification.

Physician			SVM probability estimation average (%)	
Cross-Validation method	No. of iterations	Limitation	Ventilatory limitation	Healthy patients
50% (training) – 50% (validation)	1250	Ventilatory	**96±3**	4±3
		Healthy patients	5±3	**95±3**
70% (training) – 30% (validation)	450	Ventilatory	**97±2**	3±2
		Healthy patients	4±2	**96±2**
80% (training) – 20% (validation)	200	Ventilatory	**97±2**	3±2
		Healthy patients	3±3	**97±3**
Leave-One-Out	100	Ventilatory	**98±2**	2±2
		Healthy patients	2±2	**98±2**

Values are means ± SD. Bold numbers denote the average probability estimates of the respective group.

Final validation of the SVM module as a classifier for exercise limitation severity.

CLUSTER ASSIGNMENT (K-MEANS) - The cluster assignments determined by the K-means method following the SVM probability estimation are depicted in Figures 1 and 2.

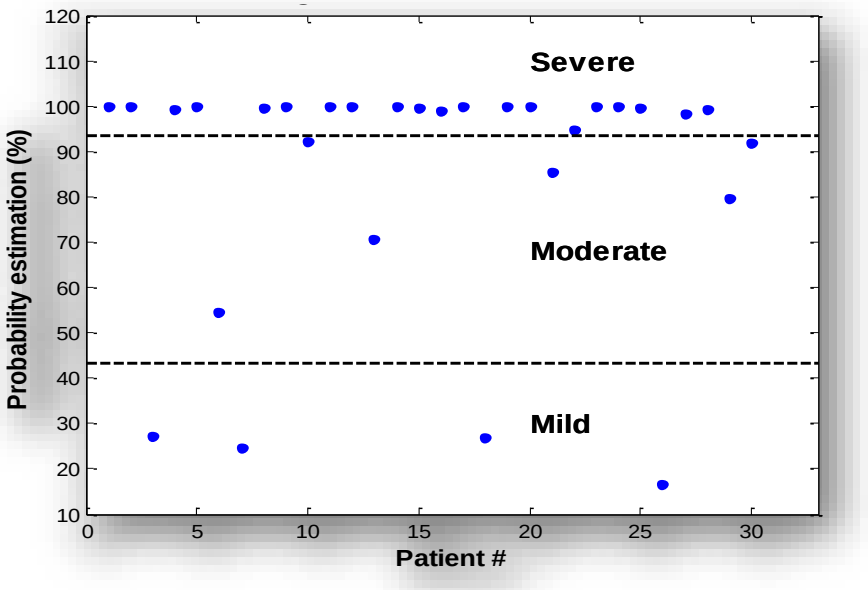


Figure 1. K-means cluster assignment for exercise limitation severity levels - **cardiovascular-related**.

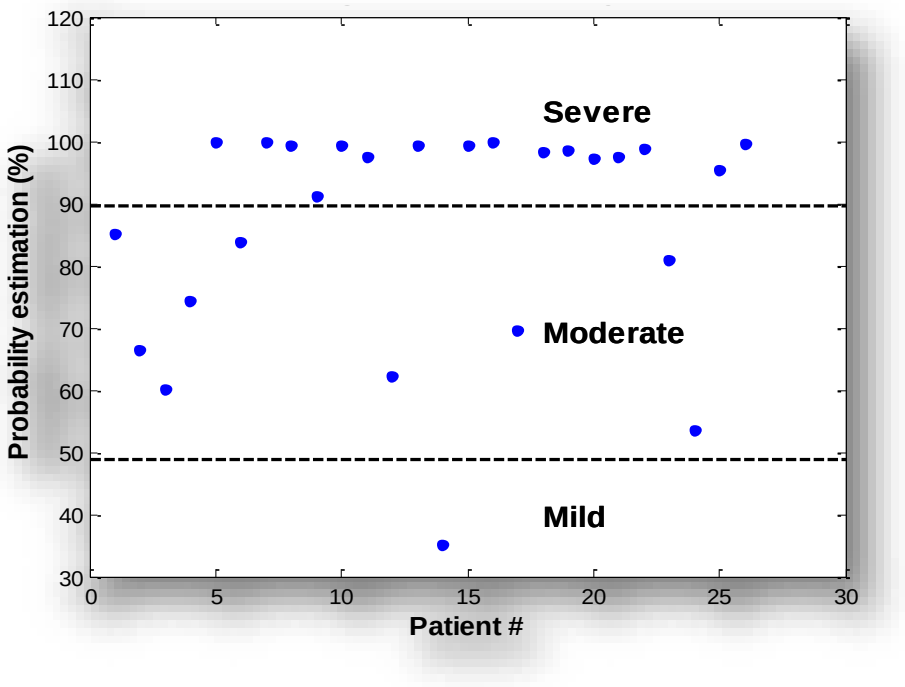


Figure 2. K-means cluster assignment for exercise limitation severity levels-**Ventilatory-related**.

Each dot represents the SVM model probability estimation (%).

The dashed lines show probability estimation thresholds between levels of limitation severity (mild, moderate, and severe).

Table 9 presents the distribution of the SVM ELS classification compared with prior physician evaluation.

Table 9. Confusion matrix – exercise limitation severities

Limitation	*Severity*	*TP*	*FN*	*FP*	*TN*
Ventilatory	mild	1	2	0	23
	moderate	7	0	2	17
	severe	16	0	0	10
Cardiovascular	mild	4	0	0	26
	moderate	5	0	1	24
	severe	20	1	0	9

TP, FN, FP, and TN denote true positive, false negative, false positive, and true negative, respectively.

The SVM models' ability to correctly predict the original classification level of ELS is shown in Table 10. Fifty-three (53) of the validated patients (56) were correctly classified (95%).

Table 10. Comparison between the original physicians' and SVM models' exercise limitation severity classification.

PHYSICIAN			SVM				
Limitation	Severity level	Count		Mild	Moderate	Severe	Total
Ventilatory	Mild	3	Count	1	2	0	3
			% within "true"	**33%**	67%	0%	
	Moderate	7	Count	0	7	0	7
			% within "true"	0%	**100%**	0%	
	Severe	16	Count	0	0	16	16
			% within "true"	0%	0%	**100%**	
Cardiovascular	Mild	4	Count	4	0	0	4
			% within "true"	**100%**	0%	0%	
	Moderate	5	Count	0	5	0	5
			% within "true"	0%	**100%**	0%	
	Severe	21	Count	0	1	20	21
			% within "true"	0%	5%	**95%**	

The count in the Physician section represents the number of study participants in each group. The count in the SVM section represents the classification results of the SVM model. In addition, % within "true" (bold percentage) denotes the percent of correct classifications of the entire respective group.

The sensitivity, specificity, accuracy, and precision of the SVM binary models' classifications are displayed in Tables 11 and 12. The results show that the SVM model classification had an overall predictive power of 83% to 100% for cardiovascular-related exercise limitations and 78% to 100% for respiratory-related exercise limitations. These numbers point to the high accuracy of the machine learning SVM module in predicting exercise limitations. The sensitivity of the mild ventilatory-related ELS was an exceptionally low finding (33%) and was linked to the small number of patients in this category.

Table 11 - Validation of the cardiovascular-relate exercise limitation severity classification.

Severity	Sn (%)	Sp (%)	Acc (%)	Pr (%)
Mild	100	100	100	100
Moderate	100	96	97	83
Severe	95	100	97	100

Sn-sensitivity; Sp-specificity; Acc-accuracy; Pr-precision.

Table 12. Validation of the ventilatory-related exercise limitation severity classification.

Severity	Sn (%)	Sp (%)	Acc (%)	Pr (%)
Mild	33*	100	92	100
Moderate	100	89	92	78
Severe	100	100	100	100

Sn-sensitivity; Sp- specificity; Acc-accuracy; Pr-precision. ***N=3.**

Discussion

Interpreting and analyzing cardiopulmonary exercise test (CPET) results remains a topic of debate in the medical field, with no clear consensus on the best strategy. Traditionally, all CPET interpretive strategies are performed manually following expert-based guidelines[1,4,10,11]. These interpretation strategies, including flow charts, graphs, and tables, are time-consuming and require extensive knowledge and understanding of the meaning and implications of all CPET physiological variables.[2] As such, the potential exists for subjective and sometimes inaccurate interpretation of CPET results.[1,14,25] Moreover, there is no consensus on any presently applied interpretation strategies for CPET data analysis, thus reducing the objectivity and consistency of interpretation and its reporting[26,27,28].

Recently, more sophisticated statistics and machine learning techniques have been applied in the context of CPET[28,29,30].

They might prompt a revolution in how machines will support experts in analyzing CPET results and the degree of exercise limitation/intolerance. In the present study, a novel automatic (computerized) CPET interpretation strategy using SVM models was developed and validated using a supervised machine learning procedure. The SVM was trained to model three groups of patients (CREL, VREL, and Healthy) and to discriminate between their respective exercise limitation severities.

Our recently published study[5] looked at how well computer-aided algorithms could be used to evaluate CPET data to find individuals suffering from chronic heart failure (CHF), chronic obstructive pulmonary disease (COPD), and healthy people. The present study aimed to expand the paradigm using CPET data and SVM algorithms to determine exercise limitations' severity (mild, moderate, severe).

A total of 206 retrospective CPET files from two medical centers: The Pulmonary Institute and the Cardiology Rehabilitation Center at the Sheba Medical Center in Israel and from the exercise physiology laboratory at the Washington-Hill College. Among these files, 80 belonged to patients with confirmed primary **cardiovascular-related** exercise limitations, 76 patients with **ventilatory-related** exercise limitations, and 50 to healthy subjects with no apparent exercise limitation, forming the data pool for analysis. We employed SVM Learning on 150 CPET files (50 from each group) for model training, while the remaining 56 CPET files were used for model validation. Additionally, distribution analysis was conducted to compare the performance of the SVM interpretive module with that of senior cardiologists, pulmonologists, and expert exercise physiologists. The study's results demonstrated that the proposed interpretive model based on machine learning (SVM) exhibited high predictive power, ranging between 78% and 100%. This indicates its ability to accurately classify the degree of exercise limitation/intolerance.

Accurately identifying exercise limitations makes the proposed platform a potent and clinically applicable solution. Furthermore, we suggest that similar modules could be developed and applied to other exercise limitations and larger populations.

Unlike the **disease classification** study[5], the exercise **limitation severity** classification

consisted of two separate SVM binary models, one for each limitation (cardiovascular-related or ventilatory-related). Each exercise limitation was characterized by selected limitation-specific physiological responses (features). Hence, a linear SVM-RFE+CBR algorithm for the feature reduction process was employed. This algorithm ranks the CPET features' impact on the exercise limitation classification process in descending order. Various combinations of features were examined to determine the most powerful (i.e., accurate) combination of CPET features for predicting ventilatory- or cardiovascular-related exercise limitation severity (see Tables 5 and 6).

Surprisingly, a few features (variables) contributed significantly to classifying cardiovascular-related exercise limitations, which are truly ventilatory and vice versa (i.e., VE, VE/VCO_2 slope, and BR for the cardiovascular-related limitation and HR for the ventilatory-related limitation). This outcome may be partially attributable to many cardiovascular-related patients suffering from pulmonary edema, among other deficiencies. As such, this may cause, among other effects, right heart overload and malfunction[29,30]. Such malfunctions, no doubt, could affect the pulmonary gas exchange and breathing patterns (Vt and Bf) and, therefore, their resulting minute ventilation (VE) and VA/Q mismatch (VE/VCO_2 slope)[4,31,32,33]. These functions are truly pulmonary-related attributes. Similarly, many pulmonary-related patients suffer peripheral (muscular-mitochondrial) abnormalities that, if severe enough, could hamper metabolic energy production and, thereby, limit exercise tolerance. Furthermore, due to their ventilatory deficiencies, many of these patients reach their ventilatory potential (MVV) well before that of their cardiovascular system[2,4,12]. These health-related factors could be responsible for the observed relatively low peak HR (cardiovascular-related feature) and exercise intolerance. The presence of overlapping characteristics among different limitations and intolerances may contribute to the complexities and discrepancies encountered when analyzing results from cardiopulmonary exercise testing (CPET). Interestingly, human vision can be approximated by computer vision[34]. As a result, the use of machine learning algorithms designed to find discriminatory features in patterns of variables is most likely the closest artificial representation of what a human expert does when manually analyzing/interpreting CPET results. By examining multiple variables concurrently, an SVM can detect differences between patterns less likely to be distorted by internal and/or external "noise" than conventional manual interpretation or linear regression algorithms operating on single variables at a time[35,36,37]. Therefore, the implementation of an ML algorithm, which is naturally conceived to find discriminatory features in patterns of variables, is most likely the closest artificial representation of what the human expert is doing when analyzing/ interpreting CPET results. Accordingly, computer-aided algorithms to interpret CPET results have several potential advantages. It can improve analysis efficiency, consistency, and objectivity, reducing inter-observer variability. Additionally, such algorithms can aid decision-making and treatment planning by providing objective and standardized assessments. Our approach may enable a more comprehensive, faster, and observer-

friendly way to evaluate CPET results than current approaches, or it may be a complementary tool. Nonetheless, prospective studies with larger and more diverse populations are necessary to validate the findings further and ensure the generalizability of the SVM interpretive modules. Additionally, the present study focused on a specific set of exercise limitations, and further research is needed to explore the module's performance in different clinical contexts.

Conclusions

CPET-based clustering utilizing supervised machine learning algorithms identified integrative CPET profiles, stratifying a heterogeneous patient sample according to **CREL** and **VREL**. Such phenogrouping may enable an integrative interpretation of CPET results, facilitate a more comprehensive assessment of CPET results, and complement risk stratification strategies.

The proposed new **CREL** and **VREL** modules for CPET analysis are among the few recently proposed computer-aided interpretation algorithms for CPET interpretation[5,35,36,37]. Both our modules were found to be highly sensitive and specific. Using them should reduce complexity, increase objectivity, and economize on the time of CPET interpretation in clinical settings. These encouraging results suggest that the proposed classification method should be extended to additional clinical conditions and exercise limitations. Lastly, we believe that the phenogrouping model used in this study has high clinical potential. However, a stringent development procedure for the software to be used as a medical device is still required before it can be used as a real-life medical application.

Conflict of Interest Statement:

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgement Statement:

The authors acknowledge Dr. Ron Dlin's assistance in editing this work.

Funding Statement:

None

Authors' Contributions:

Or I, Omri I, Mickey S, and Hayit G conceived and designed the research. Or I, Omri I, Ronen R, and Michael S analyzed the data. Or I and Omri I wrote the manuscript. All authors read and approved the manuscript.

Future research directions:

With the study's promising outcomes, it is self-evident that the proposed diagnostic strategies should be extended to other pathologies and exercise limitations/ intolerances.

These may include abnormal O2 delivery, muscle metabolic dysfunction, pulmonary gas-exchange abnormalities, chronotropic incompetence, overweight, and more.

Study limitations:

As with most research, some shortcomings are to be expected:

1. This study was conducted using retrospective CPET data, which may have inherent biases and limitations associated with data collection and documentation.
2. Only three sample populations (CREL, VREL, and healthy) were selected for the computer-aided exercise limitation severity classification, undermining universal generalizability and limiting clinical applicability.
3. Even though prior clinical diagnosis and ELS were corroborated by experienced physicians and exercise physiologists using "gold standards" manual CPET interpretation guidelines, no manual interpretative process is entirely free of potential error and inaccuracy.

Abbreviations:

ATS: American Thoracic Society
CPET: Cardiopulmonary exercise testing
ECG: Electrocardiogram
SD: Standard deviation
ML: Machine learning
SVM: Support Vector Machine.
LIBSVM: Library for Support Vector Machines
ELS: Exercise limitation severity
CREL: Cardiovascular-related exercise limitation
VREL: Ventilatory-related exercise limitation
CHF: Chronic heart failure
COPD: Chronic obstructive pulmonary disease

References:

1. Sietsema KE, Stringer WW, Sue DY, Ward S. Wasserman & Whipp: Principles of Exercise Testing and Interpretation: Including Pathophysiology and Clinical Applications. Philadelphia: Lippincott Williams & Wilkins; 2020.

2. Stickland MK, Neder A, Guenette JA, et al. Using Cardiopulmonary Exercise Testing to Understand Dyspnea and Exercise Intolerance in Respiratory Disease. *CHEST*. 2022; 161(6): 1505-1516.

3. Guazzi M, Bandera F, Ozemek C, Systrom D, Arena R. Cardiopulmonary Exercise Testing: What Is its Value? *J Am Coll Cardiol*. 2017; 70(13):1618-1636. doi: 10.1016/j.jacc. 2017.08.012.

4. Balady GJ, Arena R, Sietsema K, et al. Clinician's Guide to Cardiopulmonary Exercise Testing in Adults: a scientific statement from the American Heart Association: A scientific statement from the American Heart Association. *Circulation*. 2010; 122(2): 191-225. oi:10.1161/CIR.0b013 e3181e52e69

5. Inbar O, Inbar O, Reuveny R, Segel MJ, Greenspan H, Scheinowitz M. A machine learning approach to the interpretation of cardiopulmonary exercise tests: Development and validation. *Pulm Med*. 2021; 2021: 5516248. doi:10.1155/2021/5516248.

6. Akay MF, Zayid EIM, Aktürk E, George JD. Artificial neural network-based model for predicting VO2max from a submaximal exercise test. *Expert Syst Appl*. 2011; 38(3):2007-2010. doi: 10.1016/j.eswa.2010. 07.135.

7. Sakr S, Elshawi R, Ahmed A, 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. doi: 10.1371/journal. pone.0195344.

8. Portella JJ, Andonian BJ, Brown DE, et al. Using machine learning to identify organ system specific limitations to exercise via cardiopulmonary exercise Testing. *IEEE J Biomed Health Inform*. 2022;26(8):4228-4237. doi:10.1109/JBHI.2022.3163402.

9. Inbar O, Katz U, Dubnov RG, Epstein Y. "Guidelines for Cardiopulmonary Exercise Tests (CPET) Results". *Medicine (Harefuah),* 157 (9); 1-7, 2018. (Hebrew).

10. Wasserman K, Hansen JE, Sue DY, Stringer WW, Whipp BJ. Principles of exercise testing and interpretation: including pathophysiology and clinical applications. Medicine and Science in Sports and Exercise. 2005;37(7).

11. American Thoracic Society, American College of Chest Physicians. ATS/ACCP Statement on cardiopulmonary exercise testing. *Am J Respir Crit Care Med*. 2003; 167(2):211-277. doi:10.1164/rccm.167.2.211.

12. Inbar O, Oren A, Scheinowitz M, Rotstein A, Dlin R, Casaburi R. Normal cardiopulmonary responses during incremental exercise in 20- to 70-year-old men. *Med Sci Sports Exer*. 1994; 26(5):538-546. doi:10.1249/00005768-199405000-00003.

13. Hsu CW, Chang CC, Lin CJ. ACM transactions on intelligent systems and technology, IEEE Transactions on Neural Networks and Learning Systems. Published online 2003:1396-1400.

14. Chacey M, Goodman LF, Kynyk JA, Russell C, Magalang U, Asif MJ. Analysis of interpretation practices of cardiopulmonary

exercise tests: a quality initiative. American Journal of Respiratory and Critical Care Medicine. 2012;185.

15. Yan K, Zhang D. Feature selection and analysis on correlated gas sensor data with recursive feature elimination. Sens Actuators B Chem. 2015; 212:353-363. doi: 10.1016/j.snb.2015.02.025.

16. Chang CC, Lin CJ. LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology. 2011;2(3).

17. Vapnik VN. An overview of statistical learning theory. IEEE Trans Neural Netw. 1999; 10(5):988-999. doi:10.1109/72.788640.

18. Vapnik VN. The Nature of Statistical Learning Theory. 840 Springer-Verlag New York. Vol 840. Springer-Verlag New York." Inc; 1995.

19. Awad M, Khan L, Bastani F, Yen IL. An effective support vector machines (SVMs) performance using hierarchical clustering. In: 16th IEEE International Conference on Tools with Artificial Intelligence. IEEE Comput. Soc; 2005.

20. Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. In: Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory; 1992:144-152.

21. Hartigan JA, Wong MA. Algorithm AS 136: A K-means clustering algorithm. J R Stat Soc Ser C Appl Stat. 1979; 28(1):100. doi: 10.2307/2346830.

22. Macqueen J. Some methods for classification and analysis of multivariate observations. Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. 1967; 1:281-297.

23. Inbar O, Yamin C, Bar-On I, Nice S, David D. Transluminal Coronary Angioplasty (PTCA) on cardiopulmonary responses during exercise. J Sports Med Phys Fitness. 2008; 48(2):235-245.

24. Klainman E, Fink G, Lebzelter J, Zafrir N. Assessment of functional results after percutaneous transluminal coronary angioplasty by cardiopulmonary exercise test. Cardiology. 1998; 89(4):257-262. doi:10.1159/000006797

25. Schmid A, Schilter D, Fengels I, et al. Design and validation of an interpretative strategy for cardiopulmonary exercise tests. Respirology. 2007; 12(6):916-923. doi:10.1111/j.1440-1843.2007.01197. x.

26. Ashley E, Myers J, Froelicher V. Exercise testing scores as an example of better decisions through science. Med Sci Sports Exer. 2002; 34(8):1391-1398. doi:10.1097/00005768-200208000-00023.

27. Swets JA, Dawes RM, Monahan J. Better decisions through science. Sci Am. 2000; 283(4):82-87. doi:10.1038/scientificamerican1000-82.

28. Zile MR, Brutsaert DL. New concepts in diastolic dysfunction and diastolic heart failure: Part I: diagnosis, prognosis, and measurements of diastolic function: Diagnosis, prognosis, and measurements of diastolic function. Circulation. 2002; 105(11):1387-1393. doi:10.1161/hc1102. 105289.

29. Brown DE, Sharma S, Jablonski JA, Weltman A. Neural network methods for diagnosing patient conditions from cardiopulmonary exercise testing data. BioData Min. 2022; 15(1). doi:10.1186/s13040-022-00299-6.

30. Cauwenberghs N, Sente J, Van Criekinge H, et al. Integrative interpretation of

cardiopulmonary exercise tests for cardiovascular outcome prediction: A machine learning approach. Diagnostics (Basel). 2023; 13(12). doi:10.3390/diagnostics 13122051.

31. Solin P, Bergin P, Richardson M, Kaye DM, Walters EH, Naughton MT. Influence of pulmonary capillary wedge pressure on central apnea in heart failure. Circulation. 1999; 99(12):1574-1579. doi: 10.1161/01.cir. 99.12.1574.

32. Writing Committee, EACPR, M. Guazzi, et al. Clinical recommendations for cardiopulmonary exercise testing data assessment in specific patient populations," European Heart Journal, vol. 33, no. 23, pp. 2917–2927, 2012.

33. Weisman IM, Zeballos RJ. Clinical exercise testing. Clin Chest Med. 2001; 22(4):679-701. doi:10.1016/s0272-5231(05) 70060-5.

34. Portella JJ, Andonian BJ, Brown DE, Mansur J, Wales D, West VL, Kraus WE, Hammond WE. Using machine learning to identify organ system specific limitations to exercise via cardiopulmonary exercise testing. IEEE J. Biomed. Health Inform. 2022, 26, 4228–4237.

35. Zignoli A. Machine learning models for the automatic detection of exercise thresholds in cardiopulmonary exercising tests: From regression to generation to explanation. Sensors (Basel). 2023; 23(2):826. doi:10.3390/s23020826.

36.Ismail Fawaz H, Forestier G, Weber J. Idoumghar Muller PA. Deep learning for time series classification: A review. Data Min. Knowl. Discov. 2019, 33, 917–963.

37. Längkvist M, Karlsson, Loutfi A. A review of unsupervised feature learning and deep learning for time-series modeling. Pattern Recognition Letters. 2014; 42 (1): 11-24.