


RESEARCH ARTICLE

Open Access



# Utilizing machine learning dimensionality reduction for risk stratification of chest pain patients in the emergency department

Nan Liu<sup>1,2,3\*</sup> , Marcel Lucas Chee<sup>4</sup>, Zhi Xiong Koh<sup>5</sup>, Su Li Leow<sup>1</sup>, Andrew Fu Wah Ho<sup>1,5</sup>, Dagang Guo<sup>6</sup> and Marcus Eng Hock Ong<sup>1,2,5</sup>

## Abstract

**Background:** Chest pain is among the most common presenting complaints in the emergency department (ED). Swift and accurate risk stratification of chest pain patients in the ED may improve patient outcomes and reduce unnecessary costs. Traditional logistic regression with stepwise variable selection has been used to build risk prediction models for ED chest pain patients. In this study, we aimed to investigate if machine learning dimensionality reduction methods can improve performance in deriving risk stratification models.

**Methods:** A retrospective analysis was conducted on the data of patients > 20 years old who presented to the ED of Singapore General Hospital with chest pain between September 2010 and July 2015. Variables used included demographics, medical history, laboratory findings, heart rate variability (HRV), and heart rate n-variability (HRnV) parameters calculated from five to six-minute electrocardiograms (ECGs). The primary outcome was 30-day major adverse cardiac events (MACE), which included death, acute myocardial infarction, and revascularization within 30 days of ED presentation. We used eight machine learning dimensionality reduction methods and logistic regression to create different prediction models. We further excluded cardiac troponin from candidate variables and derived a separate set of models to evaluate the performance of models without using laboratory tests. Receiver operating characteristic (ROC) and calibration analysis was used to compare model performance.

**Results:** Seven hundred ninety-five patients were included in the analysis, of which 247 (31%) met the primary outcome of 30-day MACE. Patients with MACE were older and more likely to be male. All eight dimensionality reduction methods achieved comparable performance with the traditional stepwise variable selection; The multidimensional scaling algorithm performed the best with an area under the curve of 0.901. All prediction models generated in this study outperformed several existing clinical scores in ROC analysis.

**Conclusions:** Dimensionality reduction models showed marginal value in improving the prediction of 30-day MACE for ED chest pain patients. Moreover, they are black box models, making them difficult to explain and interpret in clinical practice.

**Keywords:** Machine learning, Dimensionality reduction, Heart rate n-variability (HRnV), Heart rate variability (HRV), Chest pain, Emergency department

\* Correspondence: [liu.nan@duke-nus.edu.sg](mailto:liu.nan@duke-nus.edu.sg)

<sup>1</sup>Duke-NUS Medical School, National University of Singapore, Singapore, Singapore

<sup>2</sup>Health Services Research Centre, Singapore Health Services, Singapore, Singapore

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

Chest pain is among the most common chief complaints presenting to the emergency department (ED) [1–3]. The assessment of chest pain patients poses a diagnostic challenge in balancing risk and cost. Inadvertent discharge of acute coronary syndrome (ACS) patients is associated with higher mortality rates while inappropriate admission of patients with more benign conditions increases health service costs [4, 5]. Hence, the challenge lies in recognizing low-risk chest pain patients for safe and early discharge from the ED. There has been increasing focus on the development of risk stratification scores. Initially, risk scores such as the Thrombolysis in Myocardial Infarction (TIMI) score [6, 7] and the Global Registry of Acute Coronary Events (GRACE) score [8] were developed from post-ACS patients to estimate short-term mortality and recurrence of myocardial infarction. The History, Electrocardiogram (ECG), Age, Risk factors, and initial Troponin (HEART) score was subsequently designed for ED chest pain patients [9], which demonstrated superior performance in many comparative studies on the identification of low-risk chest pain patients [10–17]. Nonetheless, the HEART score has its disadvantages. Many potential factors can affect its diagnostic and prognostic accuracy, such as variation in patient populations, provider determination of low-risk heart score criteria, specific troponin reagent used, all of which contribute to clinical heterogeneity [18–21]. In addition, most risk scores still require variables that may not be available during the initial presentation of the patient to the ED such as troponin. There remains a need for a more efficient risk stratification tool.

We had previously developed a heart rate variability (HRV) prediction model using readily available variables at the ED, in an attempt to reduce both diagnostic time and subjective components [22]. HRV characterizes beat-to-beat variation using time, frequency domain, and nonlinear analysis [23] and has proven to be a good predictor of major adverse cardiac events (MACE) [22, 24, 25]. Most HRV-based scores were reported to be superior to TIMI and GRACE scores while achieving comparable performance with HEART score [17, 24, 26, 27]. Recently, we established a new representation of beat-to-beat variation in ECGs, the heart rate  $n$ -variability (HRnV) [28]. HRnV utilizes variation in sampling RR-intervals and overlapping RR-intervals to derive additional parameters from a single strip of ECG reading. As an extension to HRV, HRnV potentially supplements additional information about adverse cardiac events while reducing unwanted noise caused by abnormal heartbeats. Moreover, HRV is a special case of HRnV when  $n = 1$ . The HRnV prediction model, developed from multivariable stepwise logistic regression,

outperformed the HEART, TIMI, and GRACE scores in predicting 30-day MACE [28]. Nevertheless, multicollinearity is a common problem in logistic regression models where supposedly independent predictor variables are correlated. They tend to overestimate the variance of regression parameters and hinder the determination of the exact effect of each parameter, which could potentially result in inaccurate identification of significant predictors [29, 30]. In the paper, 115 HRnV parameters were derived but only seven variables were left in the final prediction model, and this implies the possible elimination of relevant information [28].

Within the general medical literature, machine learning dimensionality reduction methods are uncommon and limited to a few specific areas, such as bioinformatics studies on genetics [31, 32] and diagnostic radiological imaging [33, 34]. Despite this, dimensionality reduction in HRV has been investigated and shown to effectively compress multidimensional HRV data for the assessment of cardiac autonomic neuropathy [35]. In this paper, we attempted to investigate several machine learning dimensionality reduction algorithms in building predictive models, hypothesizing that these algorithms could be useful in preserving useful information while improving prediction performance. We aimed to compare the performance of the dimensionality reduction models against the traditional stepwise logistic regression model [28] and conventional risk stratification tools such as the HEART, TIMI, and GRACE scores, in the prediction of 30-day MACE in chest pain patients presenting to the ED.

## Methods

### Study design and clinical setting

A retrospective analysis was conducted on data collected from patients > 20 years old who presented to Singapore General Hospital ED with chest pain between September 2010 to July 2015. These patients were triaged using the Patient Acuity Category Scale (PACS) and those with PACS 1 or 2 were included in the study. Patients were excluded if they were lost to the 30-day follow-up or if they presented with ST-elevation myocardial infarction (STEMI) or non-cardiac etiology chest pain such as pneumothorax, pneumonia, and trauma as diagnosed by the ED physician. Patients with ECG findings that precluded quality HRnV analysis such as artifacts, ectopic beats, paced or non-sinus rhythm were also excluded.

### Data collection

For each patient, HRV and HRnV parameters were calculated using HRnV-Calc software suite [28, 36] from a five to six-minute single-lead (lead II) ECG performed via the X-series Monitor (ZOLL Medical, Corporation, Chelmsford, MA). Table 1 shows the full list of HRV and HRnV parameters used in this study. Besides, the

**Table 1** List of traditional heart rate variability (HRV) and novel heart rate n-variability (HRnV) parameters used in this study. HRnV is a new representation of beat-to-beat variation in ECGs and parameter “n” controls the formation of new RR-intervals that are used for parameter calculation. Details of HRnV definition can be found in [28]

HRV	HR <sub>2</sub> V	HR <sub>2</sub> V <sub>1</sub>	HR <sub>3</sub> V	HR <sub>3</sub> V <sub>1</sub>	HR <sub>3</sub> V <sub>2</sub>
Mean NN	HR <sub>2</sub> V Mean NN	HR <sub>2</sub> V <sub>1</sub> Mean NN	HR <sub>3</sub> V Mean NN	HR <sub>3</sub> V <sub>1</sub> Mean NN	HR <sub>3</sub> V <sub>2</sub> Mean NN
SDNN	HR <sub>2</sub> V SDNN	HR <sub>2</sub> V <sub>1</sub> SDNN	HR <sub>3</sub> V SDNN	HR <sub>3</sub> V <sub>1</sub> SDNN	HR <sub>3</sub> V <sub>2</sub> SDNN
RMSSD	HR <sub>2</sub> V RMSSD	HR <sub>2</sub> V <sub>1</sub> RMSSD	HR <sub>3</sub> V RMSSD	HR <sub>3</sub> V <sub>1</sub> RMSSD	HR <sub>3</sub> V <sub>2</sub> RMSSD
Skewness	HR <sub>2</sub> V Skewness	HR <sub>2</sub> V <sub>1</sub> Skewness	HR <sub>3</sub> V Skewness	HR <sub>3</sub> V <sub>1</sub> Skewness	HR <sub>3</sub> V <sub>2</sub> Skewness
Kurtosis	HR <sub>2</sub> V Kurtosis	HR <sub>2</sub> V <sub>1</sub> Kurtosis	HR <sub>3</sub> V Kurtosis	HR <sub>3</sub> V <sub>1</sub> Kurtosis	HR <sub>3</sub> V <sub>2</sub> Kurtosis
Triangular index	HR <sub>2</sub> V Triangular index	HR <sub>2</sub> V <sub>1</sub> Triangular index	HR <sub>3</sub> V Triangular index	HR <sub>3</sub> V <sub>1</sub> Triangular index	HR <sub>3</sub> V <sub>2</sub> Triangular index
NN50	HR <sub>2</sub> V NN50	HR <sub>2</sub> V <sub>1</sub> NN50	HR <sub>3</sub> V NN50	HR <sub>3</sub> V <sub>1</sub> NN50	HR <sub>3</sub> V <sub>2</sub> NN50
pNN50	HR <sub>2</sub> V pNN50	HR <sub>2</sub> V <sub>1</sub> pNN50	HR <sub>3</sub> V pNN50	HR <sub>3</sub> V <sub>1</sub> pNN50	HR <sub>3</sub> V <sub>2</sub> pNN50
–	HR <sub>2</sub> V NN50n	HR <sub>2</sub> V <sub>1</sub> NN50n	HR <sub>3</sub> V NN50n	HR <sub>3</sub> V <sub>1</sub> NN50n	HR <sub>3</sub> V <sub>2</sub> NN50n
–	HR <sub>2</sub> V pNN50n	HR <sub>2</sub> V <sub>1</sub> pNN50n	HR <sub>3</sub> V pNN50n	HR <sub>3</sub> V <sub>1</sub> pNN50n	HR <sub>3</sub> V <sub>2</sub> pNN50n
Total power <sup>a</sup>	HR <sub>2</sub> V Total power	HR <sub>2</sub> V <sub>1</sub> Total power	HR <sub>3</sub> V Total power	HR <sub>3</sub> V <sub>1</sub> Total power	HR <sub>3</sub> V <sub>2</sub> Total power
VLF power	HR <sub>2</sub> V VLF power	HR <sub>2</sub> V <sub>1</sub> VLF power	HR <sub>3</sub> V VLF power	HR <sub>3</sub> V <sub>1</sub> VLF power	HR <sub>3</sub> V <sub>2</sub> VLF power
LF power	HR <sub>2</sub> V LF power	HR <sub>2</sub> V <sub>1</sub> LF power	HR <sub>3</sub> V LF power	HR <sub>3</sub> V <sub>1</sub> LF power	HR <sub>3</sub> V <sub>2</sub> LF power
HF power	HR <sub>2</sub> V HF power	HR <sub>2</sub> V <sub>1</sub> HF power	HR <sub>3</sub> V HF power	HR <sub>3</sub> V <sub>1</sub> HF power	HR <sub>3</sub> V <sub>2</sub> HF power
LF power norm	HR <sub>2</sub> V LF power norm	HR <sub>2</sub> V <sub>1</sub> LF power norm	HR <sub>3</sub> V LF power norm	HR <sub>3</sub> V <sub>1</sub> LF power norm	HR <sub>3</sub> V <sub>2</sub> LF power norm
HF power norm	HR <sub>2</sub> V HF power norm	HR <sub>2</sub> V <sub>1</sub> HF power norm	HR <sub>3</sub> V HF power norm	HR <sub>3</sub> V <sub>1</sub> HF power norm	HR <sub>3</sub> V <sub>2</sub> HF power norm
LF/HF	HR <sub>2</sub> V LF/HF	HR <sub>2</sub> V <sub>1</sub> LF/HF	HR <sub>3</sub> V LF/HF	HR <sub>3</sub> V <sub>1</sub> LF/HF	HR <sub>3</sub> V <sub>2</sub> LF/HF
Poincaré SD1	HR <sub>2</sub> V Poincaré SD1	HR <sub>2</sub> V <sub>1</sub> Poincaré SD1	HR <sub>3</sub> V Poincaré SD1	HR <sub>3</sub> V <sub>1</sub> Poincaré SD1	HR <sub>3</sub> V <sub>2</sub> Poincaré SD1
Poincaré SD2	HR <sub>2</sub> V Poincaré SD2	HR <sub>2</sub> V <sub>1</sub> Poincaré SD2	HR <sub>3</sub> V Poincaré SD2	HR <sub>3</sub> V <sub>1</sub> Poincaré SD2	HR <sub>3</sub> V <sub>2</sub> Poincaré SD2
Poincaré SD1/SD2 ratio	HR <sub>2</sub> V Poincaré SD1/SD2	HR <sub>2</sub> V <sub>1</sub> Poincaré SD1/SD2	HR <sub>3</sub> V Poincaré SD1/SD2	HR <sub>3</sub> V <sub>1</sub> Poincaré SD1/SD2	HR <sub>3</sub> V <sub>2</sub> Poincaré SD1/SD2
SampEn	HR <sub>2</sub> V SampEn	HR <sub>2</sub> V <sub>1</sub> SampEn	HR <sub>3</sub> V SampEn	HR <sub>3</sub> V <sub>1</sub> SampEn	HR <sub>3</sub> V <sub>2</sub> SampEn
ApEn	HR <sub>2</sub> V ApEn	HR <sub>2</sub> V <sub>1</sub> ApEn	HR <sub>3</sub> V ApEn	HR <sub>3</sub> V <sub>1</sub> ApEn	HR <sub>3</sub> V <sub>2</sub> ApEn
DFA, $\alpha_1$	HR <sub>2</sub> V DFA, $\alpha_1$	HR <sub>2</sub> V <sub>1</sub> DFA, $\alpha_1$	HR <sub>3</sub> V DFA, $\alpha_1$	HR <sub>3</sub> V <sub>1</sub> DFA, $\alpha_1$	HR <sub>3</sub> V <sub>2</sub> DFA, $\alpha_1$
DFA, $\alpha_2$	HR <sub>2</sub> V DFA, $\alpha_2$	HR <sub>2</sub> V <sub>1</sub> DFA, $\alpha_2$	HR <sub>3</sub> V DFA, $\alpha_2$	HR <sub>3</sub> V <sub>1</sub> DFA, $\alpha_2$	HR <sub>3</sub> V <sub>2</sub> DFA, $\alpha_2$

*Mean NN* average of R-R intervals, *SDNN* standard deviation of R-R intervals, *RMSSD* square root of the mean squared differences between R-R intervals, *NN50* the number of times that the absolute difference between 2 successive R-R intervals exceeds 50 ms, *pNN50*, *NN50* divided by the total number of R-R intervals, *NN50n* the number of times that the absolute difference between 2 successive  $RR_n/RR_{n+m}$  sequences exceeds  $50 \times n$  ms, *pNN50n* *NN50n* divided by the total number of  $RR_n/RR_{n+m}$  sequences, *VLF* very low frequency, *LF* low frequency, *HF* high frequency, *SD* standard deviation, *SampEn* sample entropy, *ApEn* approximate entropy, *DFA* detrended fluctuation analysis

<sup>a</sup>In frequency domain analysis, the power of spectral components is the area below the relevant frequencies presented in absolute units (square milliseconds)

first 12-lead ECGs taken during patients’ presentation to the ED were interpreted by two independent clinical reviewers and any pathological ST changes, T wave inversions, and Q-waves were noted. Patients’ demographics, medical history, first set of vital signs, and troponin-T values were obtained from the hospital’s electronic health records (EHR). In this study, high-sensitivity troponin-T was selected as the cardiac biomarker and an abnormal value was defined as > 0.03 ng/mL.

The primary outcome measured was any MACE within 30 days, including acute myocardial infarction, emergent revascularization procedures such as percutaneous coronary intervention (PCI) or coronary artery bypass graft (CABG), or death. The primary outcome was captured through a retrospective review of patients’ EHR.

### Machine learning dimensionality reduction

Dimensionality reduction in machine learning and data mining [37] refers to the process of transforming high-dimensional data into lower dimensions such that fewer features are selected or extracted while preserving essential information of the original data. Two types of dimensionality reduction approaches are available, namely variable selection and feature extraction. Variable selection methods generally reduce data dimensionality by choosing a subset of variables, while feature extraction methods transform the original feature space into lower-dimensional space through linear or nonlinear feature projection. In clinical predictive modeling, variable selection techniques such as stepwise logistic regression are popular for constructing prediction models [38]. In contrast, feature extraction approaches [39] are less

commonly used in medical research, although they have been widely used in computational biology [40], image analysis [41, 42], physiological signal analysis [43], among others. In this study, we investigated the implementation of eight feature extraction algorithms and evaluated their contributions to prediction performance in risk stratification of ED chest pain patients. We also compared them with a prediction model that was built using conventional stepwise variable selection [28]. Henceforth, we use the terms “dimensionality reduction” and “feature extraction” interchangeably.

Given that there were  $n$  samples  $(x_i, y_i)$ ,  $i = 1, 2, \dots, n$ , in the dataset  $(X, y)$ , where each sample  $x_i$  had original  $D$  features and its label  $y_i = 1$  or  $0$ , with  $1$  indicating a positive primary outcome, i.e., MACE within 30 days. We applied dimensionality reduction algorithms to project  $x_i$  into a  $d$ -dimensional space ( $d < D$ ). As a result, the original dataset  $X \in \mathbb{R}^{n \times D}$  became  $\hat{X} \in \mathbb{R}^{n \times d}$ . There was a total of  $D = 174$  candidate variables in this study. As suggested in Liu et al. [28], some variables were less statistically significant in terms of contributions to the prediction performance. Thus, we conducted univariable analysis and preselected a subset of  $\tilde{D}$  variables if their  $p < \tilde{P}$ . In this study, we determined  $\tilde{P}$  by running principal component analysis (PCA) [44] and logistic regression through 5-fold cross-validation; we plotted a curve to visualize the choice of a threshold and its impact on predictive performance. PCA was used for demonstration because of its simplicity and fast running speed. Other than PCA, we also implemented seven dimensionality reduction algorithms, including kernel PCA (KPCA) [45] with polynomial kernel function, latent semantic analysis (LSA) [46], Gaussian random projection (GRP) [47], sparse random projection (SRP) [48], multidimensional scaling (MDS) [49], Isomap [50], and locally linear

embedding (LLE) [51]. All these algorithms are unsupervised learning methods, meaning the transformation of feature space does not rely on sample labels  $y$ . Among the eight methods, MDS, Isomap, and LLE are manifold learning-based techniques for nonlinear dimensionality reduction. Table 2 gives a brief introduction to these eight methods.

### Predictive and statistical analysis

In this study, we chose logistic regression as the classification algorithm to predict the MACE outcome. As described earlier, we determined the threshold  $\tilde{P}$  to preselect a subset of  $\tilde{D}$  variables, ensuring the removal of less significant variables as indicated by univariable analysis, after which  $X \in \mathbb{R}^{n \times D}$  became  $\tilde{X} \in \mathbb{R}^{n \times \tilde{D}}$ . In summary, the inputs to all dimensionality reduction algorithms were in  $\tilde{D}$ -dimensional space. Subsequently, conventional logistic regression was implemented to take  $d$ -dimensional  $\hat{X}$  to predict  $y$ , where 5-fold cross-validation was used.

We compared the models built with machine learning dimensionality reduction with our previous stepwise model [28], in which the following 16 variables were used: age, diastolic blood pressure, pain score, ST-elevation, ST-depression, Q wave, cardiac history (the “History” component in the HEART score), troponin, HRV NN50, HR<sub>2</sub>V skewness, HR<sub>2</sub>V SampEn, HR<sub>2</sub>V ApEn, HR<sub>2</sub>V<sub>1</sub> ApEn, HR<sub>3</sub>V RMSSD, HR<sub>3</sub>V skewness, and HR<sub>3</sub>V<sub>2</sub> HF power. As described in [28], we selected candidate variables with  $p < 0.2$  in univariable analysis and subsequently conducted multivariable analysis using backward stepwise logistic regression. In the current study, we further built eight dimensionality reduction models without using the cardiac troponin and compared them with the stepwise model without the

**Table 2** Summary of machine learning dimensionality reduction methods used in this study

Methods	Descriptions
Principal component analysis (PCA) [44]	PCA decomposes data into a set of successive orthogonal components that explain a maximum amount of the variance
Kernel PCA (KPCA) [45]	KPCA extends PCA by using kernel functions to achieve non-linear dimensionality reduction
Latent semantic analysis (LSA) [46]	LSA is similar to PCA but differs in that the data matrix does not need to be centered
Gaussian random projection (GRP) [47]	GRP projects the original input features onto a randomly generated matrix where components are drawn from a Gaussian distribution
Sparse random projection (SRP) [48]	SRP projects the original input features onto a sparse random matrix, which is an alternative to dense Gaussian random projection matrix
Multidimensional scaling (MDS) [49]	MDS is a technique used for analyzing similarity or dissimilarity data, seeking a low-dimensional representation of the data in which the distances respect well the distances in the original high-dimensional space
Isomap [50]	Isomap is a manifold learning algorithm, seeking a lower-dimensional embedding that maintains geodesic distances between all points
Locally linear embedding (LLE) [51]	LLE projects the original input features to a lower-dimensional space by preserving distances within local neighborhoods

troponin component. This analysis enabled us to check the feasibility of avoiding the use of laboratory results for quick risk stratification.

In evaluating the modeling performance, we performed the receiver operating characteristic (ROC) curve analysis and reported the corresponding area under the curve (AUC), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) measures. Moreover, we generated the calibration plots for prediction models. In describing the data, we reported continuous variables as the median and interquartile range (IQR) and statistical significance using two-sample t-test. We reported categorical variables as frequency and percentage and statistical significance using chi-square test. All analyses were conducted in Python version 3.8.0 (Python Software Foundation, Delaware, USA).

## Results

We included 795 chest pain patients in this study, of which 247 (31%) patients had MACE within 30 days of presentation to the ED. Table 3 presents the baseline characteristics of the patient cohort. Patients with MACE were older (median age 61 years vs. 59 years,  $p = 0.035$ ) and more likely to be male (76.1% vs. 64.6%,  $p = 0.002$ ). History of diabetes, current smoking status, and pathological ECG changes such as ST elevation, ST depression, T wave inversion, pathological Q waves, and QTc prolongation were significantly more prevalent in patients with the primary outcome. Troponin-T and creatine kinase-MB levels were also significantly elevated in patients with the primary outcome. There was no statistically significant difference in patient ethnicity between MACE and non-MACE groups.

Figure 1a depicts the PCA-based predictive performance versus the threshold  $\tilde{P}$  (for preselection of variables) and Fig. 1b shows the number of preselected variables versus threshold  $\tilde{P}$ . The predictive performance peaked at  $\tilde{P} = 0.02$ , where a total of 30 variables were preselected, including gender, diastolic blood pressure, pain score, ST-elevation, ST-depression, T-wave inversion, Q wave, cardiac history, EKG, and risk factor components of the HEART score, troponin, HRV RMSSD, HRV NN50, HRV pNN50, HRV HF power, HRV Poincaré SD1, HR<sub>2</sub>V RMSSD, HR<sub>2</sub>V NN50, HR<sub>2</sub>V pNN50, HR<sub>2</sub>V HF power, HR<sub>2</sub>V Poincaré SD1, HR<sub>2</sub>V<sub>1</sub> RMSSD, HR<sub>2</sub>V<sub>1</sub> NN50, HR<sub>2</sub>V<sub>1</sub> HF power, HR<sub>2</sub>V<sub>1</sub> Poincaré SD1, HR<sub>3</sub>V<sub>1</sub> RMSSD, HR<sub>3</sub>V<sub>1</sub> HF power, HR<sub>3</sub>V<sub>1</sub> Poincaré SD1, HR<sub>3</sub>V<sub>2</sub> RMSSD, and HR<sub>3</sub>V<sub>2</sub> Poincaré SD1. These were used as inputs to all dimensionality reduction algorithms whose outputs were linear or nonlinear combinations of these 30 variables.

Figure 2 shows the predictive performance (in terms of AUC value) versus feature dimension (i.e., number of “principal components”) for all eight dimensionality reduction algorithms. The AUC values of GRP, SRP, and KPCA gradually increased with the increment of feature dimension, while the AUC values of PCA, LSA, MDS, Isomap, and LLE drastically jumped to more than 0.8 when feature dimension  $d \geq 3$  and plateaued in the curves when  $d \geq 15$ . The highest AUC values of PCA, KPCA, LSA, GRP, SRP, MDS, Isomap, and LLE were 0.899, 0.896, 0.899, 0.896, 0.898, 0.901, 0.888, and 0.898, achieved with feature dimensions of 15, 30, 15, 22, 20, 27, 23, and 30, respectively.

Figure 3 shows the ROC curves of the eight dimensionality reduction algorithms, the stepwise logistic regression [28], and three clinical scores. All eight dimensionality reduction methods performed comparably with the stepwise variable selection, and MDS achieved the highest AUC of 0.901. Table 4 presents ROC analysis results of all 12 methods/scores where sensitivity, specificity, PPV, and NPV are reported with 95% confidence intervals (CIs), noting that the performance of the stepwise model in this paper was slightly different from that reported in [28] due to the choice of cross-validation scheme, i.e., 5-fold (AUC of 0.887) versus leave-one-out (AUC of 0.888). Figure 4 presents the calibration curves of predictions by all methods/scores. The stepwise model and seven dimensionality reduction models (PCA, KPCA, LSA, GRP, SRP, MDS, and Isomap) showed reasonable model calibrations, in which their curves fluctuated along the diagonal line, meaning these models only slightly overestimated or underestimated the predicted probability of 30-day MACE. The LLE model was unable to achieve good calibration. In comparison, all three clinical scores (HEART, TIMI, and GRACE) generally underpredicted the probability of 30-day MACE.

Figure 5 shows the ROC curves of prediction models without using cardiac troponin. At feature dimensions of 13, 21, 13, 29, 24, 17, 18, and 18, the highest AUC values of PCA, KPCA, LSA, GRP, SRP, MDS, Isomap, and LLE were 0.852, 0.852, 0.852, 0.852, 0.851, 0.852, 0.845, and 0.849, respectively. The stepwise model without troponin yielded an AUC of 0.834 compared to 0.887 with troponin. All prediction models outperformed both the TIMI and GRACE scores while achieving comparable results with the HEART score.

## Discussion

In this study, we showed that machine learning dimensionality reduction yielded only marginal, non-significant improvements compared to stepwise model in predicting the risk of 30-day MACE among chest pain patients in the ED. This corroborates with similar observations that

**Table 3** Baseline characteristics of patient cohorts

	Total (n = 795)	MACE (n = 247)	Non-MACE (n = 548)	p-value
<b>Age, median (IQR)</b>	60 (51–68)	61 (54–68)	59 (50–68)	0.035
<b>Male gender, n (%)</b>	542 (68.2)	188 (76.1)	354 (64.6)	0.002
<b>Race, n (%)</b>				0.623
Chinese	492 (61.9)	159 (64.4)	333 (60.8)	0.374
Indian	129 (16.2)	34 (13.8)	95 (17.3)	0.246
Malay	150 (18.9)	46 (18.6)	104 (19.0)	0.984
Other	24 (3.0)	8 (3.2)	16 (2.9)	0.984
<b>Vital signs, median (IQR)</b>				
Temperature (°C)	36.4 (36.0–36.7)	36.3 (36.0–36.7)	36.4 (36.0–36.7)	0.793
Heart rate (beats/min)	76 (67–89)	80 (69–92.5)	75 (66–87)	0.03
Respiratory rate (breaths/min)	18 (17–18)	18 (17–18)	18 (17–18)	0.716
Systolic blood pressure (mmHg)	138 (123.0–159.0)	142 (123.5–165.5)	137 (122.0–156.2)	0.037
Diastolic blood pressure (mmHg)	76.0 (68.0–86.0)	78.0 (70.0–89.0)	75.0 (67.0–84.0)	0.001
SpO <sub>2</sub> (%)	99.0 (97.0–100.0)	99.0 (97.0–100.0)	99.0 (97.0–100.0)	0.842
Pain score	2.0 (0.0–5.0)	2.0 (0.0–5.0)	2.0 (0.0–5.0)	0.008
Glasgow Coma Scale (GCS) score	15.0 (15.0–15.0)	15.0 (15.0–15.0)	15.0 (15.0–15.0)	0.121
<b>Medical history, n (%)</b>				
Ischaemic heart disease	343 (43.1)	115 (46.6)	228 (41.6)	0.22
Diabetes	278 (35.0)	106 (42.9)	172 (31.4)	0.002
Hypertension	509 (64.0)	161 (65.2)	348 (63.5)	0.707
Hypercholesterolaemia	476 (59.9)	151 (61.1)	325 (59.3)	0.683
Stroke	58 (7.3)	15 (6.1)	43 (7.8)	0.458
Cancer	29 (3.6)	7 (2.8)	22 (4.0)	0.537
Respiratory disease	31 (3.9)	5 (2.0)	26 (4.7)	0.102
Chronic kidney disease	87 (10.9)	26 (10.5)	61 (11.1)	0.896
Congestive heart failure	38 (4.8)	9 (3.6)	29 (5.3)	0.407
History of PCI	199 (25.0)	68 (27.5)	131 (23.9)	0.316
History of CABG	71 (8.9)	26 (10.5)	45 (8.2)	0.355
History of AMI	133 (16.7)	48 (19.4)	85 (15.5)	0.205
Active smoker	197 (24.8)	73 (29.6)	124 (22.6)	0.045
<b>ECG pathology, n (%)</b>				
ST elevation	65 (8.2)	48 (19.4)	17 (3.1)	< 0.001
ST depression	92 (11.6)	69 (27.9)	23 (4.2)	< 0.001
T wave inversion	209 (26.3)	86 (34.8)	123 (22.4)	< 0.001
Pathological Q wave	86 (10.8)	51 (20.6)	35 (6.4)	< 0.001
QTc prolongation	174 (21.9)	73 (29.6)	101 (18.4)	0.001
Left axis deviation	64 (8.1)	16 (6.5)	48 (8.8)	0.34
Right axis deviation	16 (2.0)	6 (2.4)	10 (1.8)	0.773
Left bundle branch block	8 (1.0)	3 (1.2)	5 (0.9)	0.991
Right bundle branch block	56 (7.0)	14 (5.7)	42 (7.7)	0.385
Interventricular conduction delay	30 (3.8)	13 (5.3)	17 (3.1)	0.201
Left atrial abnormality	12 (1.5)	4 (1.6)	8 (1.5)	0.886
Left ventricular hypertrophy	103 (13.0)	38 (15.4)	65 (11.9)	0.21
Right ventricular hypertrophy	6 (0.8)	1 (0.4)	5 (0.9)	0.747

**Table 3** Baseline characteristics of patient cohorts (Continued)

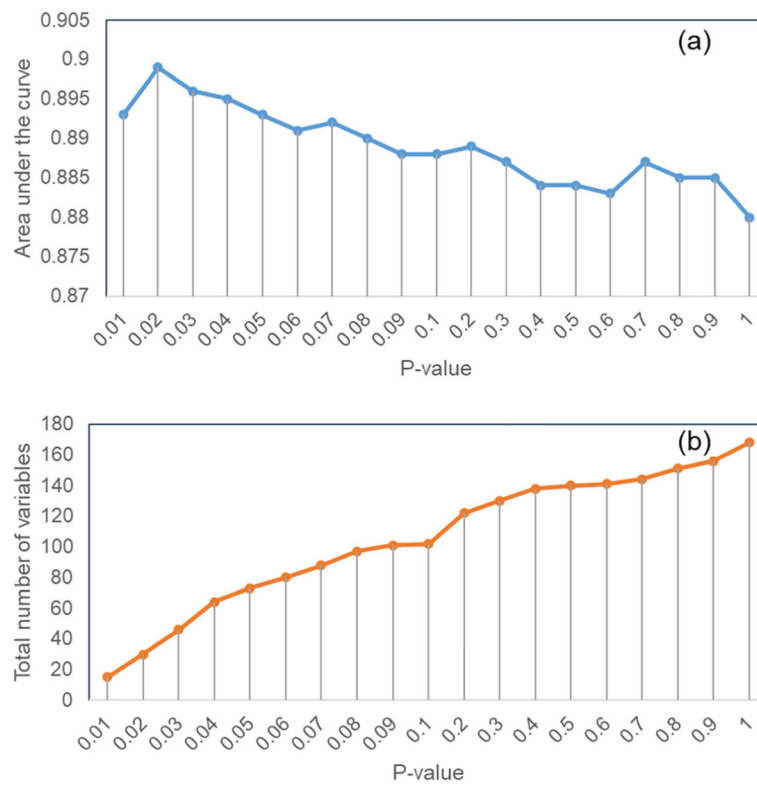
	Total (n = 795)	MACE (n = 247)	Non-MACE (n = 548)	p-value
<b>Laboratory findings, median (IQR)</b>				
Troponin (ng/L)	0 (0–39.5)	40 (10–170)	0 (0–15.2)	< 0.001
Creatine kinase-MB	2.4 (1.8–3.2)	2.7 (2.1–6.0)	2.4 (1.7–2.7)	< 0.001
<b>Clinical scores, median (IQR)</b>				
HEART	5.0 (4.0 to 7.0)	7.0 (6.0 to 8.0)	4.0 (3.0 to 6.0)	< 0.001
TIMI	2.0 (1.0 to 4.0)	3.0 (2.0 to 4.0)	2.0 (1.0 to 3.0)	< 0.001
GRACE	104.0 (83.5 to 128.0)	119.0 (97.0 to 139.0)	98.0 (78.0 to 125.0)	< 0.001

IQR interquartile range, MACE major adverse cardiac events, PCI percutaneous coronary intervention, CABG coronary artery bypass graft, AMI acute myocardial infarction, HEART History, ECG, Age, Risk factors and Troponin, TIMI Thrombolysis in Myocardial Infarction, GRACE Global Registry of Acute Coronary Events

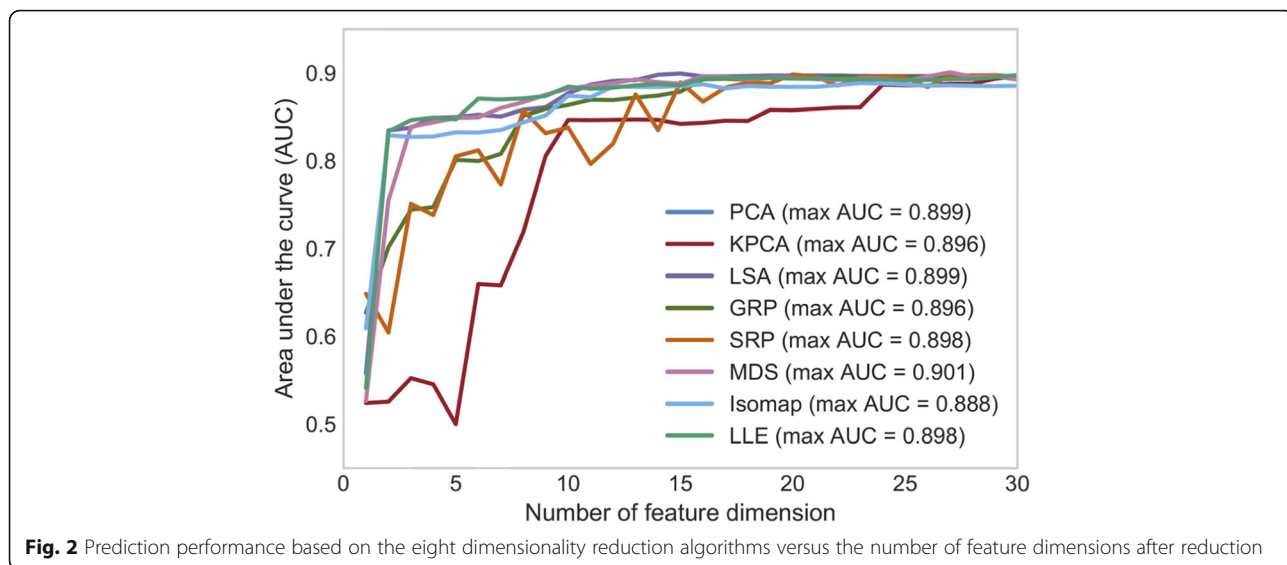
traditional statistical methods can perform comparably to machine learning algorithms [52, 53]. Among the dimensionality reduction models integrated with cardiac troponin, the MDS model had the highest discriminative performance (AUC of 0.901, 95% CI 0.874–0.928) but did not significantly outperformed the traditional stepwise model (AUC of 0.887, 95% CI 0.859–0.916). Among the models without using troponin, PCA, KPCA, LSA, GRP, and MDS performed equally well, achieving an AUC of 0.852, compared with the stepwise model without troponin which had an AUC of 0.834. In general, the traditional stepwise approach was proved to be comparable to machine learning dimensionality reduction

methods in risk prediction, while benefiting from model simplicity, transparency, and interpretability that are desired in real-world clinical practice.

High-dimensional data suffers from the curse of dimensionality, which refers to the exponentially increasing sparsity of data and sample size required to estimate a function to a given accuracy as dimensionality increases [54]. Dimensionality reduction has successfully mitigated the curse of dimensionality in the analysis of high-dimensional data in various domains such as computational biology and bioinformatics [31, 32]. However, clinical predictive modeling typically considers relatively few features, limiting the effects of the curse of



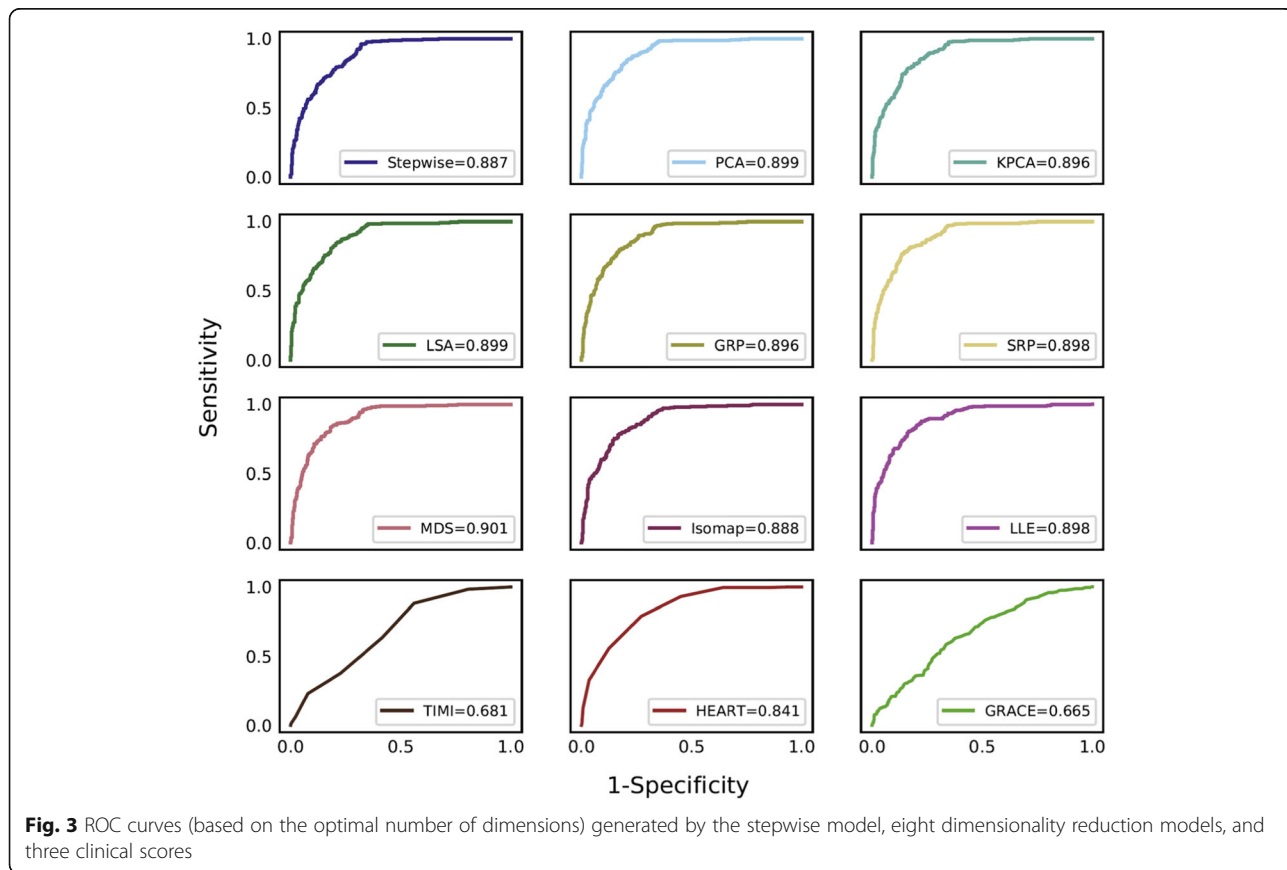
**Fig. 1** Variable preselection using *p*-value in univariable analysis for dimensionality reduction: **a** prediction area under the curve versus the *p*-value, and **(b)** the number of preselected variables versus the *p*-value



dimensionality. This may account for the relatively limited benefit of dimensionality reduction in our analysis.

Additionally, with comparable performance to the traditional stepwise model, transparency and interpretability of machine learning dimensionality reduction

models are constrained by complex algorithmic transformations of variables, leading to obstacles in the adoption of such models in real-world clinical settings. In contrast, traditional biostatistical approaches like logistic regression with stepwise variable selection deliver a simple





**Table 4** Comparison of performance of the HRnV models (based on 5-fold cross-validation), HEART, TIMI, and GRACE scores in predicting 30-day major adverse cardiac events (MACE). The cut-off values were defined as the points nearest to the upper-left corner on the ROC curves

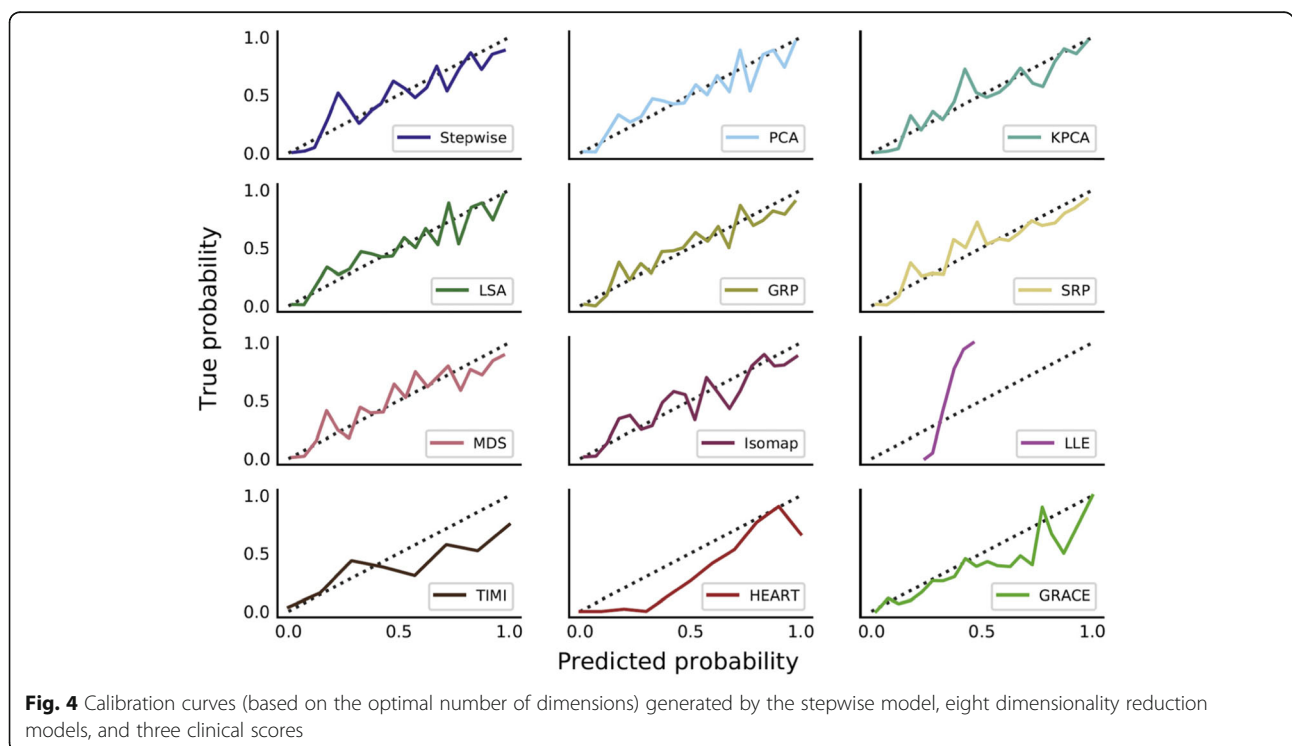
	AUC (95% CI)	Cut-off	Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)
<b>Stepwise</b>	0.887 (0.859–0.916)	0.3140	79.4 (74.3–84.4)	78.8 (75.4–82.3)	62.8 (57.5–68.2)	89.4 (86.7–92.2)
<b>PCA</b>	0.899 (0.872–0.926)	0.2881	85.4 (81.0–89.8)	78.5 (75.0–81.9)	64.1 (59.0–69.3)	92.3 (89.9–94.7)
<b>KPCA</b>	0.896 (0.869–0.923)	0.3489	81.8 (77.0–86.6)	82.1 (78.9–85.3)	67.3 (62.0–72.6)	90.9 (88.4–93.4)
<b>LSA</b>	0.899 (0.872–0.926)	0.2884	85.4 (81.0–89.8)	78.6 (75.2–82.1)	64.3 (59.1–69.5)	92.3 (89.9–94.7)
<b>GRP</b>	0.896 (0.868–0.923)	0.2965	85.0 (80.6–89.5)	78.5 (75.0–81.9)	64.0 (58.8–69.2)	92.1 (89.6–94.5)
<b>SRP</b>	0.898 (0.871–0.925)	0.2940	84.6 (80.1–89.1)	79.6 (76.2–82.9)	65.1 (59.9–70.3)	92.0 (89.5–94.4)
<b>MDS</b>	0.901 (0.874–0.928)	0.3095	83.4 (78.8–88.0)	81.6 (78.3–84.8)	67.1 (61.8–72.4)	91.6 (89.1–94.1)
<b>Isomap</b>	0.888 (0.860–0.917)	0.3468	78.5 (73.4–83.7)	82.7 (79.5–85.8)	67.1 (61.7–72.5)	89.5 (86.9–92.2)
<b>LLE</b>	0.898 (0.870–0.925)	0.3140	85.0 (80.6–89.5)	79.4 (76.0–82.8)	65.0 (59.8–70.2)	92.2 (89.7–94.6)
<b>HEART</b>	0.841 (0.808–0.874)	5	78.9 (73.9–84.0)	72.8 (69.1–76.5)	56.7 (51.4–61.9)	88.5 (85.5–91.4)
<b>TIMI</b>	0.681 (0.639–0.723)	2	63.6 (57.6–69.6)	58.4 (54.3–62.5)	40.8 (35.9–45.7)	78.0 (74.0–82.1)
<b>GRACE</b>	0.665 (0.623–0.707)	107	64.0 (58.0–70.0)	60.8 (56.7–64.9)	42.4 (37.3–47.4)	78.9 (75.0–82.8)

AUC area under the curve, CI confidence interval, PPV positive predictive value, NPV negative predictive value, HEART History, ECG, Age, Risk factors and Troponin, TIMI Thrombolysis in Myocardial Infarction, GRACE Global Registry of Acute Coronary Events

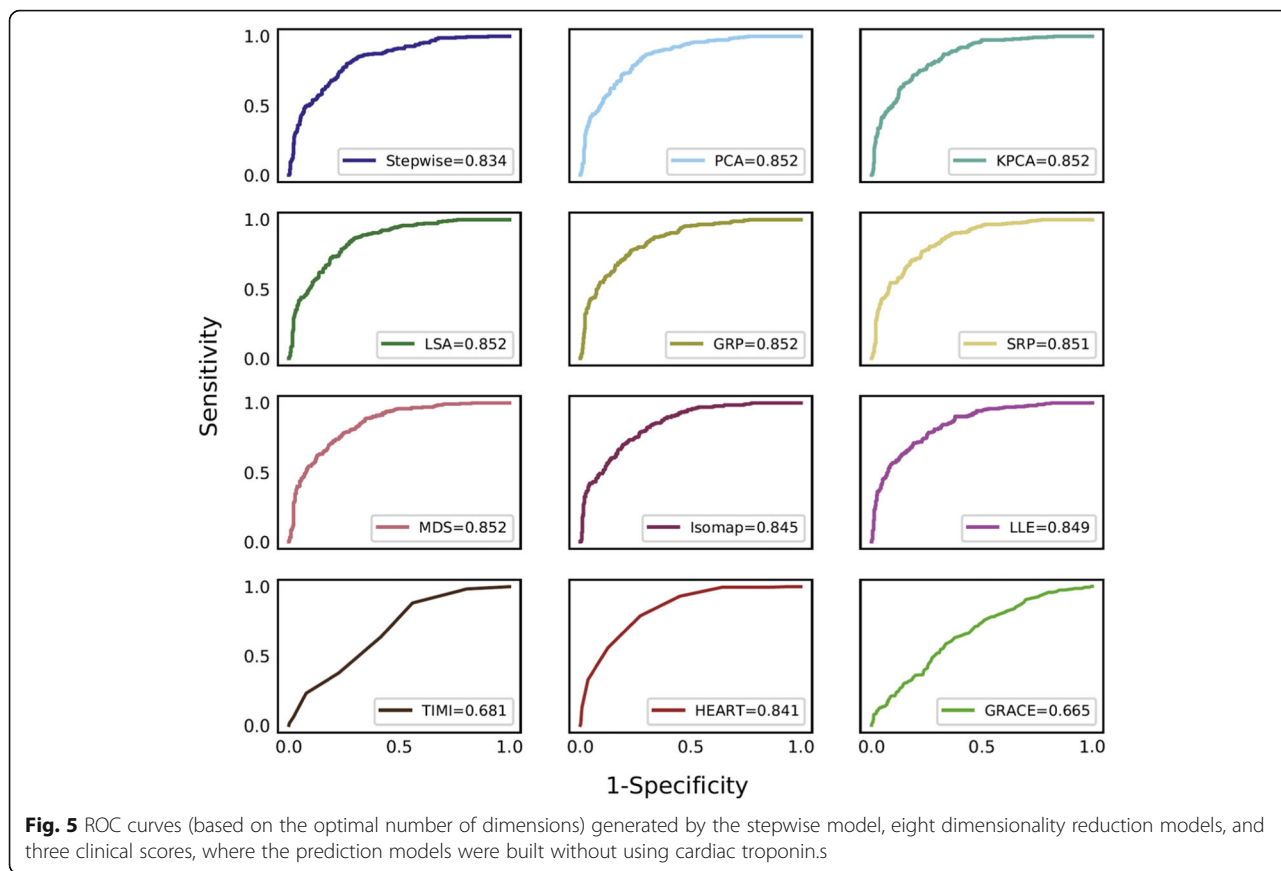
and transparent model, in which the absolute and relative importance of each variable can be easily interpreted and explained from the odds ratio. Marginal performance improvements should be weighed against these limitations in interpretability, which is an important consideration in clinical predictive modeling.

Comparing the eight dimensionality reduction algorithms, PCA and LSA use common linear algebra

techniques to learn to create principal components in a compressed data space, while MDS, Isomap, and LLE are nonlinear, manifold learning-based dimensionality reduction methods. As observed from our results, complex nonlinear algorithms did not show an obvious advantage over simple PCA and LSA methods in enhancing the predictive performance. Yet, nonlinear algorithms are more computationally complex and require



**Fig. 4** Calibration curves (based on the optimal number of dimensions) generated by the stepwise model, eight dimensionality reduction models, and three clinical scores



more computing memory. For example, KPCA and Iso-map have computational complexity of  $O(n^3)$  and memory complexity of  $O(n^2)$ , while PCA has computational complexity of  $O(\tilde{D}^3)$  and memory complexity of  $O(\tilde{D}^2)$  [39]. In applications of clinical predictive modeling,  $n$  — the number of patients — is usually larger than  $\tilde{D}$  — the number of variables; in our study,  $n$  is 795 and  $\tilde{D}$  is 29 or 30, depending on the inclusion of troponin. This suggests that linear algorithms may be preferred due to reduced computational complexity and memory while retaining comparable performance. Another observation in this study was that the impact of preselection (as shown in Fig. 1) on predictive performance was more substantial than that of dimensionality reduction, indicating the importance of choosing statistically significant candidate variables.

Our study also reiterates the value of HRnV-based prediction models for chest pain risk stratification. Among chest pain risk stratification tools in the ED, clinical scores like HEART, TIMI, and GRACE are currently the most widely adopted and validated [55, 56]. However, a common barrier to quick risk prediction using these traditional clinical scores is the requirement of cardiac troponin, which can take hours to obtain. To

address these difficulties, machine learning-based predictive models that integrate HRV measures and clinical parameters have been proposed [17, 22, 25, 26], including our development of HRnV, a novel alternative measure to HRV that has shown promising results in predicting 30-day MACE [28], which was the stepwise model in this paper. Both the dimensionality reduction-based predictive models and the stepwise model with troponin presented superior performance than HEART, TIMI, and GRACE scores. When troponin was not used, several dimensionality reduction-based models such as PCA, KPCA, and MDS still yielded marginally better performance than the original HEART score, while benefiting from generating the predictive scores in merely 5 to 6 min.

Additionally, Table 4 shows that all HRnV-based predictive models had higher specificities than the HEART score while all HRnV-based models except Isomap also improved on the already high sensitivity of the HEART score [21, 57]. The specificities of KPCA, Isomap, and MDS were significantly higher by an absolute value of almost 10%. Substantial improvements to the specificity of MACE predictive models may reduce unnecessary admission and thus minimize costs and resource usage [5]. This is particularly relevant in low-resource settings, for

example, the overburdened EDs in the current coronavirus disease 2019 (COVID-19) pandemic, where novel methods in resource allocation and risk stratification could alleviate the strain on healthcare resources [58].

There remains a need for further investigation into methods that utilize information from the full set of HRV and HRnV variables. From 174 variables in the initial data set, dimensionality reduction performed the best with a preselection of 30 variables, of which 19 were HRV and HRnV parameters. That is, the majority of the newly constructed HRnV parameters were removed based on the strict significance threshold of  $p < 0.02$  on univariable analysis. Therefore, novel HRnV measures were not fully used in prediction models of 30-day MACE, leaving room for further investigation of alternative ways of using them. Moving forward, it may be valuable to develop and evaluate deep learning frameworks [59] to synthesize novel low-dimensional representations of multidimensional information. Alternatively, building point-based, interpretable risk scores [60] can also be beneficial to implementation and adoption in real-world clinical settings, since designing inherently interpretable models is more favorable than explaining black box models [61].

We acknowledge the following limitations of this study. First, the clinical application (i.e., risk stratification of ED chest pain patients) was only one example of clinical predictive modeling, thus our conclusion on the effectiveness of machine learning dimensionality reduction algorithms may not be generalizable to other applications, particularly those with a larger number of variables. Second, only eight dimensionality reduction algorithms were investigated, while many other methods are available. Third, given the small sample size, we were unable to determine the threshold  $\tilde{P}$  and build predictive models with a separate training set; this also limited the stability check [62] for both logistic regression and machine learning models. Last, we did not build a workable predictive model for risk stratification of ED chest pain patients, although several models built in this study showed promising results compared to existing clinical scores. We aim to conduct further investigations.

## Conclusions

In this study we found that machine learning dimensionality reduction models showed marginal value in improving the prediction of 30-day MACE for ED chest pain patients. Being black box models, they are further constrained in clinical practice due to low interpretability. Whereas traditional stepwise prediction model showed simplicity and transparency, making it feasible for clinical use. To fully utilize the available information in building high-performing predictive models, we

suggest additional investigations such as exploring deep representations of the input variables and creating interpretable machine learning models to facilitate real-world clinical implementation.

## Abbreviations

ACS: Acute coronary syndrome; AMI: Acute myocardial infarction; AUC: Area under the curve; ApEn: Approximate entropy; CI: Confidence intervals; CABG: Coronary artery bypass graft; COVID-19: Coronavirus disease 2019; DFA: Detrended fluctuation analysis; ECG: Electrocardiogram; EHR: Electronic health records; ED: Emergency department; GRP: Gaussian random projection; GRACE: Global registry of acute coronary events; HRnV: Heart rate n-variability; HRV: Heart rate variability; HF: High frequency; HEART: History, ECG, Age, Risk factors, and initial Troponin; IQR: Interquartile range; KPCA: Kernel principal component analysis; LSA: Latent semantic analysis; LLE: Locally linear embedding; LF: Low frequency; MACE: Major adverse cardiac events; Mean NN: Average of R-R intervals; MDS: Multidimensional scaling; NPV: Negative predictive value; NN50: The number of times that the absolute difference between 2 successive R-R intervals exceeds 50 ms; NN50n: The number of times that the absolute difference between 2 successive  $RR_n/RR_{n,m}$  sequences exceeds  $50 \times n$  ms; PACS: Patient acuity category scale; PCI: Percutaneous coronary intervention; pNN50: NN50 divided by the total number of R-R intervals; pNN50n: NN50n divided by the total number of  $RR_n/RR_{n,m}$  sequences; PPV: Positive predictive value; PCA: Principal component analysis; ROC: Receiver operating characteristic; RMSSD: Square root of the mean squared differences between R-R intervals; RRI: R-R interval; SampEn: Sample entropy; SD: Standard deviation; SDNN: Standard deviation of R-R intervals; SRP: Sparse random projection; STEMI: ST-elevation myocardial infarction; TIMI: Thrombolysis in myocardial infarction; VLF: Very low frequency

## Acknowledgments

We would like to thank and acknowledge the contributions of doctors, nurses, and clinical research coordinators from the Department of Emergency Medicine, Singapore General Hospital.

## Authors' contributions

NL conceived the study and supervised the project. NL and MLC performed the analyses and drafted the manuscript. NL, MLC, ZXX, SLL, AFWH, DG, and MEHO made substantial contributions to results interpretation and critical revision of the manuscript. All authors read and approved the final manuscript.

## Funding

This work was supported by the Duke-NUS Signature Research Programme funded by the Ministry of Health, Singapore. The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

## Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Declarations

### Ethics approval and consent to participate

The ethical approval was obtained from the Centralized Institutional Review Board (CIRB, Ref: 2014/584/C) of SingHealth, in which patient consent was waived.

### Consent for publication

Not applicable.

### Competing interests

NL and MEHO hold patents related to using heart rate variability and artificial intelligence for medical monitoring. NL, ZXX, DG, and MEHO are advisers to TIIM SG. The other authors report no conflicts.

**Author details**

<sup>1</sup>Duke-NUS Medical School, National University of Singapore, Singapore, Singapore. <sup>2</sup>Health Services Research Centre, Singapore Health Services, Singapore, Singapore. <sup>3</sup>Institute of Data Science, National University of Singapore, Singapore, Singapore. <sup>4</sup>Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, Australia. <sup>5</sup>Department of Emergency Medicine, Singapore General Hospital, Singapore, Singapore. <sup>6</sup>SingHealth Duke-NUS Emergency Medicine Academic Clinical Programme, Singapore, Singapore.

Received: 14 July 2020 Accepted: 5 April 2021

Published online: 17 April 2021

**References**

- Long B, Koyfman A. Best clinical practice: current controversies in evaluation of low-risk chest pain-part 1. *J Emerg Med*. 2016;51(6):668–76. <https://doi.org/10.1016/j.jemermed.2016.07.103>.
- Long B, Koyfman A. Best clinical practice: current controversies in the evaluation of low-risk chest pain with risk stratification aids. Part 2. *J Emerg Med*. 2017;52(1):43–51. <https://doi.org/10.1016/j.jemermed.2016.07.004>.
- Januzzi JL Jr, McCarthy CP. Evaluating chest pain in the emergency department: searching for the optimal gatekeeper. *J Am Coll Cardiol*. 2018; 71(6):617–9. <https://doi.org/10.1016/j.jacc.2017.11.065>.
- Pope JH, Aufderheide TP, Ruthazer R, Woolard RH, Feldman JA, Beshansky JR, et al. Missed diagnoses of acute cardiac ischemia in the emergency department. *N Engl J Med*. 2000;342(16):1163–70. <https://doi.org/10.1056/NEJM200004203421603>.
- Hollander JE, Than M, Mueller C. State-of-the-art evaluation of emergency department patients presenting with potential acute coronary syndromes. *Circulation*. 2016;134(7):547–64. <https://doi.org/10.1161/CIRCULATIONAHA.116.021886>.
- Antman EM, Cohen M, Bernink PJ, McCabe CH, Horacek T, Papuchis G, et al. The TIMI risk score for unstable angina/non-ST elevation MI: a method for prognostication and therapeutic decision making. *JAMA*. 2000;284(7):835–42. <https://doi.org/10.1001/jama.284.7.835>.
- Morrow DA, Antman EM, Charlesworth A, Cairns R, Murphy SA, de Lemos JA, et al. TIMI risk score for ST-elevation myocardial infarction: a convenient, bedside, clinical score for risk assessment at presentation: an intravenous nPA for treatment of infarcting myocardium early II trial substudy. *Circulation*. 2000;102(17):2031–7. <https://doi.org/10.1161/01.CIR.102.17.2031>.
- Fox KA, Dabbous OH, Goldberg RJ, Pieper KS, Eagle KA, Van de Werf F, et al. Prediction of risk of death and myocardial infarction in the six months after presentation with acute coronary syndrome: prospective multinational observational study (GRACE). *BMJ (Clinical research ed)*. 2006;333(7578):1091.
- Six AJ, Backus BE, Kelder JC. Chest pain in the emergency room: value of the HEART score. *Neth Hear J*. 2008;16(6):191–6. <https://doi.org/10.1007/BF03086144>.
- Backus BE, Six AJ, Kelder JC, Bosschaert MA, Mast EG, Mosterd A, et al. A prospective validation of the HEART score for chest pain patients at the emergency department. *Int J Cardiol*. 2013;168(3):2153–8. <https://doi.org/10.1016/j.ijcard.2013.01.255>.
- Six AJ, Cullen L, Backus BE, Greenslade J, Parsonage W, Aldous S, et al. The HEART score for the assessment of patients with chest pain in the emergency department: a multinational validation study. *Crit Path Cardiol*. 2013;12(3):121–6. <https://doi.org/10.1097/HPC.0b013e31828b327e>.
- Chen X-H, Jiang H-L, Li Y-M, Chan CPY, Mo J-R, Tian C-W, et al. Prognostic values of 4 risk scores in Chinese patients with chest pain: prospective 2-Centre cohort study. *Medicine*. 2016;95(52):e4778. <https://doi.org/10.1097/MD.0000000000004778>.
- Jain T, Nowak R, Hudson M, Frisoli T, Jacobsen G, McCord J. Short- and Long-term prognostic utility of the HEART score in patients evaluated in the emergency Department for Possible Acute Coronary Syndrome. *Crit Path Cardiol*. 2016;15(2):40–5. <https://doi.org/10.1097/HPC.000000000000070>.
- Sakamoto JT, Liu N, Koh ZX, Fung NX, Heldeweg ML, Ng JC, et al. Comparing HEART, TIMI, and GRACE scores for prediction of 30-day major adverse cardiac events in high acuity chest pain patients in the emergency department. *Int J Cardiol*. 2016;221:759–64. <https://doi.org/10.1016/j.ijcard.2016.07.147>.
- Sun BC, Laurie A, Fu R, Ferencik M, Shapiro M, Lindsell CJ, et al. Comparison of the HEART and TIMI risk scores for suspected acute coronary syndrome in the emergency department. *Crit Path Cardiol*. 2016;15(1):1–5. <https://doi.org/10.1097/HPC.000000000000066>.
- Poldervaart JM, Langedijk M, Backus BE, Dekker IMC, Six AJ, Doevendans PA, et al. Comparison of the GRACE, HEART and TIMI score to predict major adverse cardiac events in chest pain patients at the emergency department. *Int J Cardiol*. 2017;227:656–61. <https://doi.org/10.1016/j.ijcard.2016.10.080>.
- Sakamoto JT, Liu N, Koh ZX, Guo D, Heldeweg ML, Ng JC, et al. Integrating heart rate variability, vital signs, electrocardiogram, and troponin to triage chest pain patients in the ED. *Am J Emerg Med*. 2018;36(2):185–92.
- Engel J, Heeren MJ, van der Wulp I, de Bruijne MC, Wagner C. Understanding factors that influence the use of risk scoring instruments in the management of patients with unstable angina or non-ST-elevation myocardial infarction in the Netherlands: a qualitative study of health care practitioners' perceptions. *BMC Health Serv Res*. 2014;14(1):418. <https://doi.org/10.1186/1472-6963-14-418>.
- Wu WK, Yiadom MY, Collins SP, Self WH, Monahan K. Documentation of HEART score discordance between emergency physician and cardiologist evaluations of ED patients with chest pain. *Am J Emerg Med*. 2017;35(1): 132–5. <https://doi.org/10.1016/j.ajem.2016.09.058>.
- Ras M, Reitsma JB, Hoes AW, Six AJ, Poldervaart JM. Secondary analysis of frequency, circumstances and consequences of calculation errors of the HEART (history, ECG, age, risk factors and troponin) score at the emergency departments of nine hospitals in the Netherlands. *BMJ Open*. 2017;7(10): e017259. <https://doi.org/10.1136/bmjopen-2017-017259>.
- Laureano-Phillips J, Robinson RD, Aryal S, Blair S, Wilson D, Boyd K, et al. HEART score risk stratification of low-risk chest pain patients in the emergency department: a systematic review and meta-analysis. *Ann Emerg Med*. 2019;74(2):187–203. <https://doi.org/10.1016/j.annemergmed.2018.12.010>.
- Ong MEH, Goh K, Fook-Chong S, Haaland B, Wai KL, Koh ZX, et al. Heart rate variability risk score for prediction of acute cardiac complications in ED patients with chest pain. *Am J Emerg Med*. 2013;31(8):1201–7. <https://doi.org/10.1016/j.ajem.2013.05.005>.
- Rajendra Acharya U, Paul Joseph K, Kannathal N, Lim CM, Suri JS. Heart rate variability: a review. *Med Biol Eng Comput*. 2006;44(12):1031–51. <https://doi.org/10.1007/s11517-006-0119-0>.
- Liu N, Koh ZX, Chua ECP, Tan LML, Lin Z, Mirza B, et al. Risk scoring for prediction of acute cardiac complications from imbalanced clinical data. *IEEE J Biomed Health Inform*. 2014;18(6):1894–902. <https://doi.org/10.1109/JBHI.2014.2303481>.
- Liu N, Lin Z, Cao J, Koh ZX, Zhang T, Huang G-B, et al. An intelligent scoring system and its application to cardiac arrest prediction. *IEEE Trans Inf Technol Biomed*. 2012;16(6):1324–31. <https://doi.org/10.1109/TITB.2012.2212448>.
- Heldeweg ML, Liu N, Koh ZX, Fook-Chong S, Lye WK, Harms M, et al. A novel cardiovascular risk stratification model incorporating ECG and heart rate variability for patients presenting to the emergency department with chest pain. *Crit Care*. 2016;20(1):179. <https://doi.org/10.1186/s13054-016-1367-5>.
- Liu N, Koh ZX, Goh J, Lin Z, Haaland B, Ting BP, et al. Prediction of adverse cardiac events in emergency department patients with chest pain using machine learning for variable selection. *BMC Med Inform Decis Mak*. 2014; 14(1):75. <https://doi.org/10.1186/1472-6947-14-75>.
- Liu N, Guo D, Koh ZX, Ho AFW, Xie F, Tagami T, et al. Heart rate n-variability (HRnV) and its application to risk stratification of chest pain patients in the emergency department. *BMC Cardiovasc Disord*. 2020;20(1):168. <https://doi.org/10.1186/s12872-020-01455-8>.
- Meloun M, Militký J, Hill M, Brereton RG. Crucial problems in regression modelling and their solutions. *Analyst*. 2002;127(4):433–50. <https://doi.org/10.1039/b110779h>.
- Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*. 2013;36(1):27–46. <https://doi.org/10.1111/j.1600-0587.2012.07348.x>.
- Gui J, Andrew AS, Andrews P, Nelson HM, Kelsey KT, Karagas MR, et al. A robust multifactor dimensionality reduction method for detecting gene-gene interactions with application to the genetic analysis of bladder Cancer susceptibility. *Ann Hum Genet*. 2011;75(1):20–8. <https://doi.org/10.1111/j.1469-1809.2010.00624.x>.
- Ritchie MD, Hahn LW, Roodi N, Bailey LR, Dupont WD, Parl FF, et al. Multifactor-dimensionality reduction reveals high-order interactions among

- estrogen-metabolism genes in sporadic breast Cancer. *Am J Hum Genet.* 2001;69(1):138–47. <https://doi.org/10.1086/321276>.
33. Akhbardeh A, Jacobs MA. Comparative analysis of nonlinear dimensionality reduction techniques for breast MRI segmentation. *Med Phys.* 2012;39(4):2275–89. <https://doi.org/10.1118/1.3682173>.
  34. Balvay D, Kachenoura N, Espinoza S, Thomassin-Naggara I, Fournier LS, Clement O, et al. Signal-to-noise ratio improvement in dynamic contrast-enhanced CT and MR imaging with automated principal component analysis filtering. *Radiology.* 2011;258(2):435–45. <https://doi.org/10.1148/radiol.10100231>.
  35. Tarvainen MP, Cornforth DJ, Jelinek HF. Principal component analysis of heart rate variability data in assessing cardiac autonomic neuropathy. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 2014. p. 6667–70.
  36. Vest AN, Da Poian G, Li Q, Liu C, Nemati S, Shah AJ, et al. An open source benchmarked toolbox for cardiovascular waveform and interval analysis. *Physiol Meas.* 2018;39(10):105004. <https://doi.org/10.1088/1361-6579/aae021>.
  37. Maimon O, Rokach L. Data mining and knowledge discovery handbook. Berlin: Springer Publishing Company, Incorporated; 2010. <https://doi.org/10.1007/978-0-387-09823-4>.
  38. Zhang Z. Variable selection with stepwise and best subset approaches. *Ann Transl Med.* 2016;4(7):136. <https://doi.org/10.21037/atm.2016.03.35>.
  39. van der Maaten LJP, Postma EO, van den Herik HJ. Dimensionality Reduction: A Comparative Review. In: Tilburg University Technical Report TiCC-TR 2009–005. Tilburg: Tilburg University; 2009.
  40. Nguyen LH, Holmes S. Ten quick tips for effective dimensionality reduction. *PLoS Comput Biol.* 2019;15(6):e1006907. <https://doi.org/10.1371/journal.pcbi.1006907>.
  41. Liu N, Wang H. Weighted principal component extraction with genetic algorithms. *Appl Soft Comput.* 2012;12(2):961–74. <https://doi.org/10.1016/j.asoc.2011.08.030>.
  42. Pan Y, Ge SS, Al Mamun A. Weighted locally linear embedding for dimension reduction. *Pattern Recogn.* 2009;42(5):798–811. <https://doi.org/10.1016/j.patcog.2008.08.024>.
  43. Artoni F, Delorme A, Makeig S. Applying dimension reduction to EEG data by principal component analysis reduces the quality of its subsequent independent component decomposition. *Neuroimage.* 2018;175:176–87. <https://doi.org/10.1016/j.neuroimage.2018.03.016>.
  44. Diamantaras KI, Kung SY. Principal component neural networks: theory and applications. New Jersey: Wiley; 1996.
  45. Schölkopf B, Smola AJ, Müller KR. Kernel principal component analysis. In: *Advances in kernel methods: support vector learning.* Cambridge: MIT Press; 1999. p. 327–52.
  46. Landauer TK, Foltz PW, Laham D. An introduction to latent semantic analysis. *Discourse Process.* 1998;25(2–3):259–84. <https://doi.org/10.1080/01638539809545028>.
  47. Dasgupta S. Experiments with random projection. In: *Proceedings of the sixteenth conference on uncertainty in artificial intelligence.* Stanford: Morgan Kaufmann Publishers Inc; 2000. p. 143–51.
  48. Li P, Hastie TJ, Church KW. Very sparse random projections. In: *Proceedings of the 12th ACM SIGKDD international conference on knowledge discovery and data mining.* Philadelphia: Association for Computing Machinery; 2006. p. 287–96.
  49. Mead A. Review of the development of multidimensional scaling methods. *J Royal Stat Soc Ser D (The Statistician).* 1992;41(1):27–39.
  50. Tenenbaum JB, de Silva V, Langford JC. A global geometric framework for nonlinear dimensionality reduction. *Science (New York, NY).* 2000;290(5500):2319–23.
  51. Roweis ST, Saul LK. Nonlinear dimensionality reduction by locally linear embedding. *Science (New York, NY).* 2000;290(5500):2323–6.
  52. Gravesteyn BY, Nieboer D, Ercole A, Lingsma HF, Nelson D, van Calster B, et al. Machine learning algorithms performed no better than regression models for prognosis in traumatic brain injury. *J Clin Epidemiol.* 2020; 122:95–107. <https://doi.org/10.1016/j.jclinepi.2020.03.005>.
  53. Nusinovici S, Tham YC, Chak Yan MY, Wei Ting DS, Li J, Sabanayagam C, et al. Logistic regression was as good as machine learning for predicting major chronic diseases. *J Clin Epidemiol.* 2020;122:56–69. <https://doi.org/10.1016/j.jclinepi.2020.03.002>.
  54. Lee JA, Verleysen M. Nonlinear dimensionality reduction. New York: Springer; 2007. <https://doi.org/10.1007/978-0-387-39351-3>.
  55. D'Ascenzo F, Biondi-Zoccai G, Moretti C, Bollati M, Omedè P, Sciuto F, et al. TIMI, GRACE and alternative risk scores in acute coronary syndromes: a meta-analysis of 40 derivation studies on 216,552 patients and of 42 validation studies on 31,625 patients. *Contemp Clin Trials.* 2012;33(3):507–14. <https://doi.org/10.1016/j.cct.2012.01.001>.
  56. Liu N, Ng JCY, Ting CE, Sakamoto JT, Ho AFW, Koh ZX, et al. Clinical scores for risk stratification of chest pain patients in the emergency department: an updated systematic review. *J Emerg Crit Care Med.* 2018;2:16.
  57. Byrne C, Toarta C, Backus B, Holt T. The HEART score in predicting major adverse cardiac events in patients presenting to the emergency department with possible acute coronary syndrome: protocol for a systematic review and meta-analysis. *Syst Rev.* 2018;7(1):148.
  58. Liu N, Chee ML, Niu C, Pek PP, Siddiqui FJ, Ansah JP, et al. Coronavirus disease 2019 (COVID-19): an evidence map of medical literature. *BMC Med Res Methodol.* 2020;20(1):177. <https://doi.org/10.1186/s12874-020-01059-y>.
  59. Xie J, Girshick R, Farhadi A. Unsupervised deep embedding for clustering analysis. In: *Proceedings of the 33rd international conference on international conference on machine learning - volume 48.* New York: [JMLR.org](https://doi.org/10.1007/978-0-387-09823-4); 2016. p. 478–87.
  60. Xie F, Chakraborty B, Ong MEH, Goldstein BA, Liu N. AutoScore: a machine learning-based automatic clinical score generator and its application to mortality prediction using electronic health records. *JMIR Med Inform.* 2020; 21798.
  61. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Machine Intelligence.* 2019;1(5):206–15. <https://doi.org/10.1038/s42256-019-0048-x>.
  62. Heinze G, Wallisch C, Dunkler D. Variable selection – a review and recommendations for the practicing statistician. *Biom J.* 2018;60(3):431–49. <https://doi.org/10.1002/bimj.201700067>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

