Minimal Patients' Clinical Variables to Accurately Predict Stress Echocardiography Outcome: Validation Study Using Machine Learning Techniques (Preprint)

Journal Item

Minimal Patients' Clinical Variables to Accurately Predict Stress Echocardiography Outcome: Validation Study Using Machine Learning Techniques

Mohamed Bennasar, Duncan Banks, Blaine Price, Attila Kardos

Submitted to: JMIR Cardio
on: November 08, 2019

Table of Contents

Minimal Patients' Clinical Variables to Accurately Predict Stress Echocardiography Outcome: Validation Study Using Machine Learning Techniques

Mohamed BennasarBSc, MSc, PhD, ; Duncan BanksBSc, PhD, ; Blaine PriceMSc, ; Attila KardosMD, PhD, FRCP, FESC,

Corresponding Author:
Duncan BanksBSc, PhD,
Phone: +441908659198
Email: duncan.banks@open.ac.uk

Abstract

Background: Stress echocardiography (SE) is a well-established diagnostic tool in assessing patients with suspected coronary artery disease (CAD). Cardiovascular risk factors are used in the assessment of the probability of CAD. The link between the outcome of SE and patients' variables including cardiovascular risk factors, current medication and anthropometric variables has not been widely investigated.

Objective: This study aims to use Machine Learning (ML) to predict significant CAD defined by positive SE results in patients with chest pain based on patients' anthropometrics, cardiovascular risk factors and medication as variables.

Methods: A ML framework is proposed to automate the prediction of SE results. The proposed framework consists of four stages; feature extraction, pre-processing, feature selection and classification stage. A mutual information-based feature selection method was used to investigate the amount of information that each feature carries to define the positive outcome of SE. Two classification algorithms, Support Vector Machine (SVM) with Radial Basis Function (RBF) kernel, and Random Forest classifiers have been deployed. Data from 529 patients have been used to train and validate the proposed framework. Their mean age was 61 (±12 SD). The data consists of the anthropological data and cardiovascular risk factors such as gender, age, weight, family history, diabetes, smoking history, hypertension, hypercholesterolaemia, prior diagnosis of CAD and prescribed medications at the time of the test. The results of the SE were defined as outcome. A total of 82 patients had positive (abnormal) and 447 negative (normal) results, respectively. The proposed framework has been evaluated using the whole dataset including the cases with prior diagnosis of CAD. Five folds cross validation was used to validate the performance of the proposed framework. We also investigated the model in the subset of patients with no prior CAD.

Results: The feature selection methods showed that prior diagnosis of CAD, sex, and prescribed medications such as angiotensin convertase enzyme inhibitor or angiotensin receptor blocker were the features that shared the most information about the outcome of SE. SVM classifiers showed the best trade-off between sensitivity and specificity and was achieved with three features. The best trade-off between sensitivity and specificity for the whole dataset accuracy was 66.63% with sensitivity and specificity 72.87%, and 67.67% respectively. However, for patients with no prior diagnosis of CAD only two features (sex and angiotensin convertase enzyme inhibitor or angiotensin receptor blocker use) were needed to achieve accuracy of 70.32% with sensitivity and specificity at 70.24%.

Conclusions: This pilot study shows that ML can predict the outcome of SE in detecting significant CAD based on only a few features: patient prior cardiac history, gender, and prescribed medication. Further research recruiting higher number of patients who underwent SE could further improve the performance of the proposed algorithm with the potential of facilitating patient's selection for early treatment / intervention with avoiding un-necessary downstream testing.

(JMIR Preprints 08/11/2019:16975)
DOI: https://doi.org/10.2196/preprints.16975

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
Please make my preprint PDF available to anyone at any time (recommended).

✓ **Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all u**
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Revised Paper

Minimal Patients' Clinical Variables to Accurately Predict Stress Echocardiography Outcome: Validation Study Using Machine Learning Techniques

Mohamed Bennasar [1], Duncan Banks [2], Blaine A. Price [1], Attila Kardos [3]

[1] School of Computing and Communication, STEM, The Open University, UK
[2] School of Life, Health and Chemical Sciences, STEM, The Open University, UK
[3] Department of Cardiology, Milton Keynes University Hospital NHS Foundation Trust, School of Medicine, University of Buckingham, UK

Corresponding Author:
Dr Duncan Banks, School of Life, Health and Chemical Sciences, STEM, The Open University, UK. duncan.banks@open.ac.uk; +44 (0)1908 659198

Abstract

Background: Stress echocardiography (SE) is a well-established diagnostic tool for suspected coronary artery disease (CAD). Cardiovascular risk factors are used in the assessment of the probability of CAD. The link between the outcome of SE and patients' variables including risk factors, current medication and anthropometric variables has not been widely investigated.

Objective: This study aims to use Machine Learning (ML) to predict significant CAD defined by positive SE results in patients with chest pain based on anthropometrics, cardiovascular risk factors and medication as variables. This could allow clinical prioritisation of patients with likely prediction of CAD, thus saving clinician time and improving outcomes.

Method: A ML framework is proposed to automate the prediction of SE results. The framework consists of four stages; feature extraction, pre-processing, feature selection and classification stage. A mutual information-based feature selection method was used to investigate the amount of information that each feature carries to define the positive outcome of SE. Two classification algorithms, Support Vector Machine (SVM) with Radial Basis Function (RBF) kernel, and Random Forest classifiers have been deployed. Data from 529 patients was used to train and validate the framework. Their mean age was 61 (±12 SD). The data consists of anthropological data and cardiovascular risk factors such as gender, age, weight, family history, diabetes, smoking history, hypertension, hypercholesterolaemia, prior diagnosis of CAD and prescribed medications at the time of the test. There were 82 positive (abnormal) and 447 negative (normal) SEbresults, respectively. The framework was evaluated using the whole dataset including the cases with prior diagnosis of CAD. Five folds cross validation was used to validate the performance of the framework. We also investigated the model in the subset of patients with no prior CAD.

Results: The feature selection methods showed that prior diagnosis of CAD, sex, and prescribed medications such as angiotensin convertase enzyme inhibitor or angiotensin receptor blocker were the features that shared the most information about the outcome of SE. SVM classifiers showed the best trade-off between sensitivity and specificity and was achieved with three features. Using only these three features we achieved an accuracy of 67.63% with sensitivity and specificity 72.87%, and 66.67% respectively. However, for patients with no prior diagnosis of CAD only two features (sex and angiotensin convertase enzyme inhibitor or angiotensin receptor blocker use) were needed to achieve accuracy of 70.32% with sensitivity and specificity at 70.24%.

Conclusion

This study shows that ML can predict the outcome of SE based on only a few features: patient prior cardiac history, gender, and prescribed medication. Further research recruiting higher number of patients who underwent SE could further improve the performance of the proposed algorithm with the potential of facilitating patient's selection for early treatment / intervention with avoiding un-necessary downstream testing.

Keywords

stress echocardiography, coronary heart disease, risk factors, machine learning, feature selection, risk prediction.

Introduction

Cardiovascular disease (CVD) is the leading cause of death in Western societies [1]. In the UK, 7.4 million people are living with CVD, which is more than twice the number of people who suffer from cancer and Alzheimer's disease. More than 43,000 people under the age of 75 die each year due to CVD costing national health services in the UK about £9 billion [2]. Coronary artery disease (CAD) is the most common form of CVD, that may lead to sudden death [3].

Diagnosing CAD early can save lives and reduce risk of myocardial infarction and stroke. Diagnostic procedures are typically performed in experienced cardiac centres to diagnose CAD and risk stratify patients using tests such as a stress echocardiogram (SE). Stress echocardiography is a diagnostic tool to assess the functionality of the heart and blood delivery under stress, such as treadmill or bicycle exercise test or following administration of a drug such as dobutamine. Dobutamine is a pharmacological agent that is administered intravenously to increases the heart rate in a similar way that would occur during physical exercise. During Dubutamine SE, incremental doses of dobutamine in three minutes stages is administered until the termination of the test criteria is achieved. The principle of SE is to increase the myocardial oxygen uptake/demand and in case the supply is insufficient due to blocked heart arteries than echocardiographic features of this mismatch can be detected by identifying regional wall motion abnormalities in the under-perfused heart muscle region during the test. Echocardiographic images are acquired at rest, during the intermediate stage, peak stress and in the recovery. The classical criteria were used as a termination of the test i.e. target heart rate achieved, development of typical chest pain symptoms with or without regional wall motion abnormalities, haemodynamically significant arrhythmias or development of symptomatic hypotension. Positive/abnormal SE is defined as developments of new regional wall motion abnormalities. The definition of wall motion abnormalities was: hypokinesia if the wall thickness was maintained and the endocardial excursion was between 5 and 2mm; akinesia if the wall thickness was reduced and the endocardial excursion was less than 2mm; dyskinesia if the wall thickness was reduced and the endocardial excursion was outward moving in systole. Dobutamine SE has a sensitivity and specificity of 83% and 86% respectively [4]. A computer-based algorhythms in image analysis and interpretation can play a significant role in the early diagnosis of CAD. Many machine learning-based methods have been devolved for image analysis to aid diagnosis and prognostic monitoring of CAD [5].

Machine Learning (ML) is a term used to define computer algorithms that can be trained to learn the patterns in training data. These algorithms are then effectively able to make predictions on unseen data. The ability of ML techniques to learn from experience without any explicit guidelines for the program or following any predefined rules is making these techniques increasingly popular in many domains [6]. ML in healthcare has enormous potential in supporting healthcare practitioners in decision making, enhancing diagnostic accuracy and in reducing healthcare cost [7]. Machine learning can be used as part of computer-aided clinician decision support system to assimilate patterns and act as an appropriate source of knowledge.

Several frameworks that employ ML for CAD prediction have been proposed [8]. These techniques are used either for predicting the outcome of observations or discovering the hidden pattern and structure in the data not readily recognisable to humans. The data often used for this kind of research includes: patient anthropometric data, blood test results and data obtained from various investigation modalities used in the diagnosis of CAD such as Electrocardiography (ECG), Computed Tomography Angiography, and Trans Thoracic Echocardiography (TTE) [8].

Clinical data have been used to predict coronary events: Voss et al. [9] used 10 years follow up data for 5159 middle-age men with a 6.3% incidence of coronary events during that period of time. Multi-Layer Perceptron (MLP) was used to build their model. The study involved 57 clinical and laboratory variables to train the MLP. The reported results showed that the area under the curve was 0.89. Gharehchopogh and Khalifelu [10] employed deep learning as a learning algorithm for building a prediction model, the learning algorithm was trained using data from 40 participants, including age, sex, hypertension, and smoking. The reported classification accuracy was 0.85 for heart failure cases.

Another study [11] employed ML on clinical and laboratory data of 378,256 patients to predict the first CVD event. The data used consisted of 30 attributes which included risk factors, laboratory data, medications, and information about history of CVD and other chronic diseases such as poor mental health, chronic obstructive pulmonary disease, kidney disease, and rheumatoid arthritis. The authors applied four ML algorithms namely: random forest, logistic regression, gradient boosting, and neural networks. The reported results showed that the best performance has been achieved by the neural network algorithm with sensitivity of 67.5%, and specificity 70.7%.

In this pilot study we aimed to investigate the performance of a ML algorithm in predicting the SE outcome in patients investigated for suspected CAD. Unlike previous research, we are testing a sophisticated feature selection method to investigate the significance of cardiovascular risk factors, current medication and anthropometric data in this prediction.

Methods

Population and Data Sources

The cohort of patients was derived from the Cardiology Department at Milton Keynes University Hospital (MKUH). Anonymised clinical data had been extracted from patients' electronic records, predominantly based of the very detailed SE reports introduced prospectively by one author (AK), a senior cardiologist, at the time of the development of SE services in the Hospital. We included all the patients (N=563) examined using Dobutamine Stress Echo between 2002 and 2004 with available data. However, we excluded 34 patients who had incomplete clinical data about their risk factors

leaving 529 for this study.

This study uses real patients' data which can raise some ethical concerns, such as the patient's permission to use their data and any confidential information that may exposed because of this research. This was resolved by having hospital staff, the direct clinical care provider, anonymise the records before they were sent for analysis. This study was registered by the Institutional Clinical Governance Department, MKUH - Clinical governance project reference number: 33.

Table 1 summarises patients' characteristics for the whole population and separately for the two groups with positive and negative SE results respectively. All of these patients had a complete data set for the anthropometric variables, risk factors such as gender, age, weight, family history (defined as having a first degree relative who had a myocardial infarction or died suddenly below the age of 60), diabetes, smoking status, hypertension, hypercholesterolemia, prior history of CAD, the prescribed medication related to CAD including: beta receptor blockers, calcium channel blockers, angiotensin convertase enzyme inhibitor or angiotensin receptor blocker (ACEI/ARB), antiplatelets, nitrates, statins, diuretics, as well as the SE results. The features in the table describe the number of patients who have that risk factor positive, for example; 306 of the total 529 participants had hypertension, and 313 of them had abnormal serum cholesterol level

Table 1. Characteristics of patients and their SE outcome

Group	Feature	Total (529)	SE-positive (82)	SE-negative (447)
Risk Factors	Sex	M= 249, F= 280	M=61, F=21	M=188, F=257
	Age	61.23 (±11.83 SD)	62.92 (±10.56 SD)	60.93 (±12.06 SD)
	Weight	80.82 (±17.25 SD)	83.06 (±15.88 SD)	80.45 (±17.49 SD)
	Hypertension (yes =1, no =0)	306	42	264
	Hypercholesterolaemia(yes=1,no=0)	313	50	263
	Ex-smoker = -1 Non-smoker = 1 Smoker= 0	107 330 92	24 40 18	83 290 74
	Diabetes Mellitus (yes =1, no =0)	99	19	80
	Family History (yes =1, no =0)	223	35	188
	Prior diagnosis of Coronary Artery Disease (yes =1, no =0)	123	40	83
Medications	Beta Receptor Blocker (yes =1, no =0)	281	58	223
	Calcium Channel Blocker (yes =1, no =0)	137	22	115
	(ACE-I/ARB) (yes =1, no =0)	258	59	199
	Antiplatelet therapy (yes =1, no =0)	344	64	280
	Nitrate (yes =1, no =0)	159	36	123
	Statin (yes =1, no =0)	314	59	255
	Diuretic (yes =1, no =0)	129	23	106

Proposed framework

The collected data were used to predict the outcome of the stress test based on the patient's clinical information. Figure 1 shows the architecture of the framework that was used to study the risk factors and the medication (referred to as features in this article), then we used these features to investigate the prediction power of this clinical data. Raw data was received as a mixture of text and numerical values. Therefore, the first stage in the proposed framework was the pre-processing stage where

natural language processing was used to extract and quantify the needed information from the text which included sex, age, weight, risk factors, medications, and the final outcome of the stress test (positive/negative). The criteria shown in Table 1 were used to convert the text into numerical values.

Feature normalisation stage: this is the second stage used for continuous features (age and weight) normalised using the formula:

$$(\ Normalised\ feature = \frac{raw\ feature - mean}{standrad\ deviation}\ ).$$

Discretisation stage: Two normalised features were then discretised using the Equal Width Discretisation (EWD) method [12]. In this method, the value of these features is allocated to one of the decimal numbers between 1 and 10. This method divides the range of the feature values into 10 bins with equal width.

$$Width = \frac{(Maximum - minimum)}{N}$$

Where *N* is the number bins. Each feature value is assigned to a bin based on the range it falls into. The reason of having discretisation stage is that most of the machine learning algorithms perform better with discretised data [13]. Due the bias of the feature selection stage to the continuous features [14], the discretisation stage is also needed to discretise these features before they are submitted to the feature selection stage.

Feature selection stage: This is the fourth stage in this framework, feature selection is a set of techniques that are used to measure the significance of each feature for predicting the class label (outcome of the stress test). In this study, the Joint Mutual Information Maximisation (JMIM) filter feature selection method [15] is used to rank the features according the amount of information that the feature adds to the selected subset. The method measures the amount of information that each feature shares with the class. At the end of this stage all the features (sex, age, weight, risk factors, and medications) will be ranked based on their significance in predicting the class label. This method has been developed based on information theory [16], the mechanism of the method is explained below.

The value of Mutual Information (MI) between any two of variable can be calculated using entropy. It is the amount of uncertainty about a random variable. Suppose $F = \{f_1, f_2, \ldots . f_N\}$ is a discrete variable, and $C = \{c, c, \ldots . c_N\}$ is a class label, the probability density function is

$$p(f_i) = \frac{number\ of\ instants\ with\ value\ f_i}{N}\ .$$

The MI between *F* and *C* is defined as:

$$I(F;C) = \sum_{i=1}^{N}\sum_{j=1}^{M} p(\ f_i, c_j)\ \log(\frac{p(\ f_i, c_j)}{p(\ f_i)\ p(c_j)})$$

The JMIM method employs the ‘maximum of the minimum’ criterion. The feature selected by the JMIM method is the one that maximise the following goal function:

$$f_{JMIM} = arg\,max_{f_i \in F-S}(¿min¿¿f_s \in S(I(f_i, f_s; C)))¿¿$$

Where $I(f_i, f_s; C)$ is the joint mutual information between the candidate feature and the features that already selected in the previous iteration. The method employs the following forward greedy search algorithm:

1. Set $F \leftarrow$ "initial set of n features;" $S \leftarrow$ "empty set".
2. (calculating the mutual information with the class label) for each feature $f_i \in F$ compute $I(C; f_i)$.
3. Find the feature f_i that maximizes $I(C; f_i)$; set $F \leftarrow F \setminus \{f_i\}$; set $S \leftarrow \{f_i\}$.
4. (Greedy selection) Repeat until $|S| = k$:
   a) (Computation of the joint mutual information between variables) for all pairs of variables $I(f_i, f_s; C)$ with $f_i \in F, f_s \in S,$ compute $I(f_i, f_s; C)$, if it is not already available.
   b) (Selection of the next feature) Choose feature.
   $f_{JMIM} = arg\,max_{f_i \in F-S}(¿min¿¿f_s \in S(I(f_i, f_s; C)))¿¿$
5. Output the set S containing the selected futures.

The method does not rank the features based on their individual discriminative power, it selects the features that provide the most information as a subgroup, the interaction information between the features affect is important in selecting the next significant feature. Therefore, if the list of submitted features is changed the rank order may be different. The whole data is submitted to the JMIM method to identify the significant subset of features (clinical variables). Smialowski et al. [17] reported that the feature selection stage should be included within folds-cross validation. However, that can cause instability to the results of the feature selection as submitting data with different instances may lead to different values of probability density function which consequently may lead to changes in the order of the significant features. This paper aims to define the clinical variables that can best precict the outcome of SE in the diagnosis of CAD. Therefore, the whole data is used at the feature selection stage to take the advantage of each valuable instance in the data.

Classification stage: The outcome of the stress test is predicted at this stage. Two alternative classifiers are tested at this stage: Support Vector Machine (SVM) [18], and Random Forest classifiers [19]. The performance of each classifier was evaluated using 5 folds cross-validation. The dataset is imbalanced, there are more than four times negative SE in the data than positive SE cases. To overcome this problem more weight was given to the minority class, a ratio 4:1 has been used with SVM, this means giving the minority class 4 times the weight that is given to the majority class. Due to this skewness in the number of classes, classification accuracy will not be a good measure for the performance, as it will be affected mainly by the ability of the classifier to recognise the majority of classes correctly. Therefore, sensitivity and specificity were used to provide a measure for the performance of the classifier in correctly classifying each class.

The data was randomly divided into five folds, four of them were used to train the classifier and one for testing, then this process was repeated four more times, at each time four folds were used for training and one of the folds that had never been used for testing before is used to test the classifier. At each time the accuracy, sensitivity, and specificity were calculated. The overall accuracy, sensitivity, and specificity are the average of the five.

To find the subset of features that produces the best prediction performance, the classifier is trained and tested after adding every feature according to its rank that identified at feature selection stage.

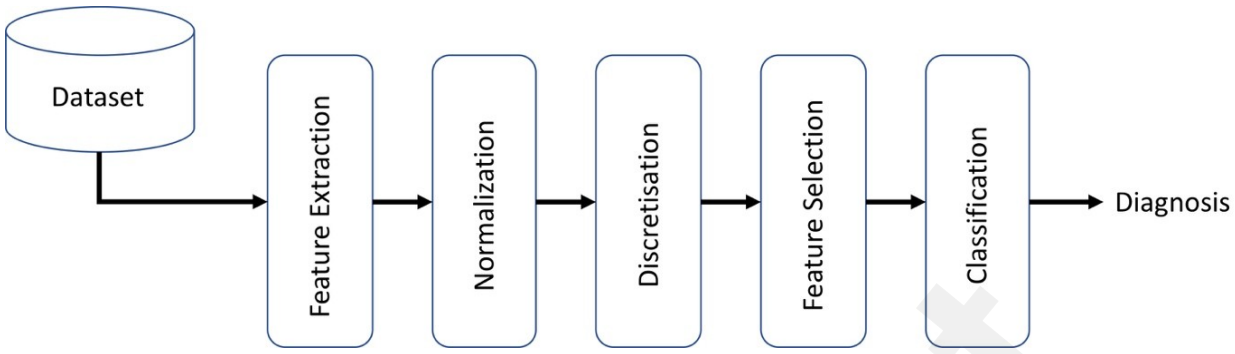

Figure 1. Experiment framework

Results

The proposed framework was used to study the whole dataset including the risk factors and medications, and also it was used to study a sub-set of the dataset that excluded the cases with prior CAD to investigate the influence of this variable on the performance of the model. Table 2 shows the characteristics of this subset of patients referred in the rest of this paper as the sub-dataset (Table 2).

The prevalence of abnormal SE was 15.5%. Four hundred and forty-seven patients had negative SE results (84.5%). There were fewer women than men within the positive group, and the opposite in those with negative SE results. Mean age 62.92 (±10.56 SD) and 60.93 (±12 SD) in the positive and negative groups respectively. (Table1)

Feature selection: The feature selection stage was used to rank the features (clinical variables) in the whole dataset, the significant features for the whole dataset are depicted by Table 3. The table shows that for the whole cohort of patients' CAD is the most significant feature for predicting SE outcome, followed by sex, ACE-I/ARB usage, and smoking status.

Table 2. Characteristics of patients and their SE outcome in those with no prior ischaemic heart disease

Group	Feature	SE-positive (42)	SE-negative (364)
Risk Factors	Sex	M=33, F=9	M=147 F=217
	Age	64.28(±9.80 SD)	60.96(±12.20 SD)
	Weight	80.31(±13.18 SD)	80.46(±17.37 SD)
	Hypertension	25	216
	Hypercholesterolaemia	33	219
	Ex-smoker Non-smoker Smoker	13 17 12	13 227 65
	Diabetes Mellitus	6	67
	Family History	17	132
Me	Beta Receptor Blocker	28	166

dications	Calcium Channel Blocker	9	89
	ACE-I/ARB	32	148
	Antiplatelet therapy	34	216
	Nitrate	24	95
	Statin	29	201
	Diuretic	12	87

The results showed that prior CAD has the strongest power to distinguish between positive and negative SE results. Sex appeared second because most of the positive cases were male, and most of the negative were female. ACEI/ARB usage is the only applied medication among the 5 most significant features. On the other hand, age, family history, and diabetes appeared the least contributory features in this model.

Table 3. Feature rankings

No	Features
1	Prior diagnosis of coronary artery disease
2	Sex
3	ACE-I/ARB
4	Weight
5	Smoking status
6	Beta Receptor Blocker
7	Hypercholesterolaemia
8	Antiplatelet therapy
9	Statin
10	Nitrate
11	Hypertension
12	Calcium Channel Blocker
13	Diuretic
14	Diabetes Mellitus
15	Family History
16	Age

Feature selection has been also applied on the sub-dataset. The order of the features was slightly different as the prior CAD feature was excluded from the data. Table 4 depicts the order of the features, it shows that sex, ACE-I/ARB, cholesterol, nitrates, and smoking status are the five most significant features. The only difference from the previous results when the whole dataset was used is the swap between serum cholesterol and smoking status. Serum cholesterol status became the third most significant feature followed by nitrates medication which was not among the most important.

Table 4. Feature ranking in the model for patients with no prior ischaemic heart disease

No	Feature
1	Sex
2	ACE-I/ARB
3	Hypercholesterolaemia
4	Nitrate
5	Smoking status
6	Statin
7	Weight
8	Beta Receptor Blocker
9	Antiplatelet therapy
10	Hypertension
11	Diuretic
12	Calcium Channel Blocker
13	Diabetes Mellitus
14	Family History

15	Age

Classification: As mentioned earlier two classification algorithms were used in this study; namely: SVM, and RF. The results showed that the performance of the two classifiers were close to each other. However, the SVM slightly outperformed the RF classifier. In this paper only results produced by the SVM are presented.

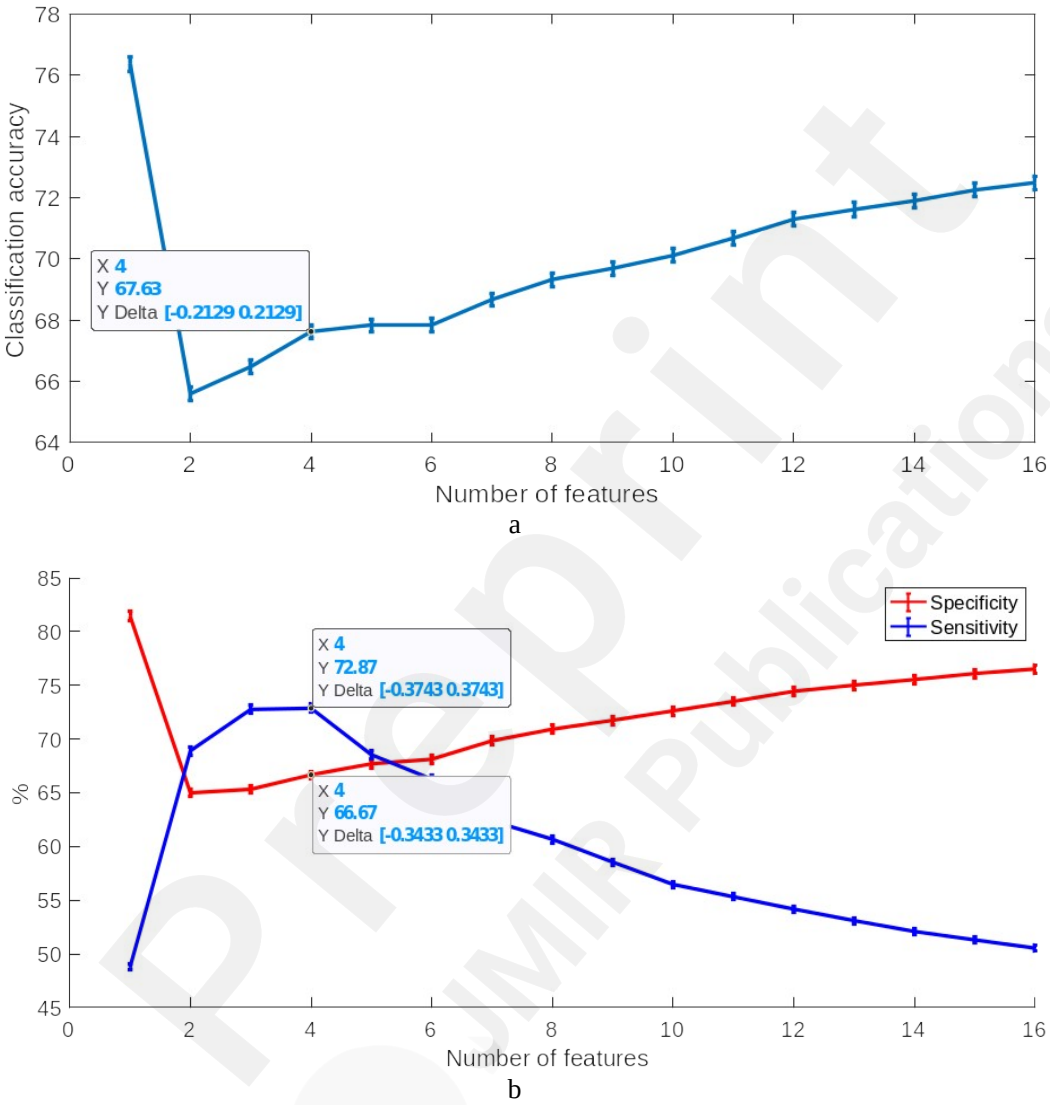

Figure 2. Performance of SVM classifier: (a) classification accuracy ± standard error, (b) sensitivity and specificity ± standard error.

Figure 2 shows that the best trade-off between sensitivity and specificity was achieved by the subset of the most significant four features (prior CAD, sex, weight, and ACE-I/ARB usage) which were 72.87%, and 66.67% respectively. The classification accuracy was 67.63%. The figure also showed that when more features were added the sensitivity started to decrease and the specificity started to increase therefore the accuracy is correlated more with the sensitivity due to this skewness in the number of classes. This drop in the sensitivity means that the rest of the features are either redundant or irrelevant for recognising the positive cases. When whole features were used only about 50% of the positive cases were classified correctly. On the other hand RF showed a slightly lower performance when the value of sensitivity, specificity, and classification accuracy were all the same (69.2%); this figure has been achieved with the most significant four features.

The experiment was repeated on patients' data with no known prior CAD. The performance of the classification stage is depicted in figure 3. The sensitivity was slightly affected by excluding patients with CAD as a feature from the data, however, the specificity increased. The classifier produced the best trade-off between the sensitivity and specificity with only using two features (sex, and ACE-I/ARB usage) which were 70.24% for both of them. The accuracy also increased to 70.32%, due to the increase of the specificity figure.

To test the robustness of the proposed framework, the experiment was repeated again, the model trained without sex features. The results showed that the best performance has been achieved with the best four features (prior CAD, ACE-I/ARB, Beta Receptor Blocker, and Smoking status) which were 72.87% and 60.23 respectively. The accuracy decreased to 62.19%.

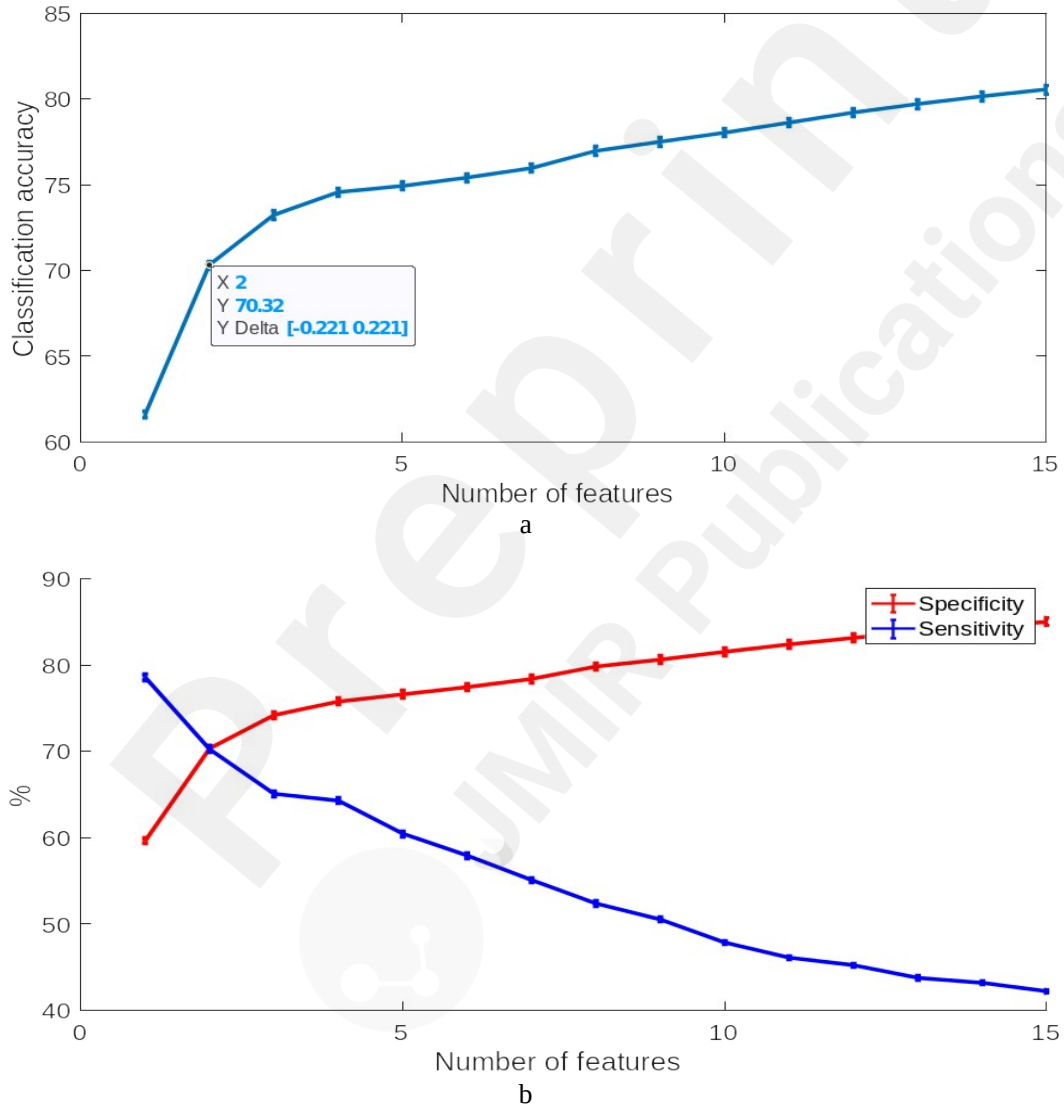

Figure 3. Performance of SVM classifier for sub-dataset: (a) classification accuracy ± standard error, (b) sensitivity and specificity ± standard error.

Discussion

Feature selection is used as part of the proposed framework; these techniques have the capability to

investigate the multidimensional relation between the features and the class label. In previous research [11] the classifiers were used to rank the features based on their importance for the classification algorithm, which are very specific to the classifier. This study employed the classifier *independent* feature selection method (JMIM) to investigate the relation between features of clinical data and class label (test outcome) in the context of the features that have been selected in the previous iterations. This means that once we have selected the features they can be used with any classifier. SVM and RF classifiers were tested in this study and SVM slightly outperformed RF in our dataset.

The results of the feature selection and classification stages showed that prior CAD feature is the most important risk factor for distinguishing between positive and negative cases. Sex, weight, and smoking status are among the group of most significant five features, and the only current medication that is within this group is ACE-I/ARB. Hypertension, diabetes, and positive family history are shown as the least significant features for the discrimination task. Only four features were needed to achieve the best performance (prior CAD, sex, weight, and ACE-I/ARB usage), which means by knowing only this information about the patients, the proposed framework is able to classify 72.87% of the positive cases, and 66.67% of the negative cases correctly which outperformed the previous study [11]. ACE-I/ARB is a used for several cardiovascular conditions and secondary prevention after an acute coronary event. This feature carries information about these conditions, that is why it is the most powerful predictor of SE outcome. For patients with no prior history of CAD, knowing the sex and whether the patient is taking ACE-I/ARB is sufficient to predict SE outcome in the majority of cases.

To study the robustness of the framework, the performance was tested without any information about the prior diagnosis of CAD. Once this was tested the order of the feature became different; cholesterol and nitrates medication became among the most significant of the five features. The feature ‘weight’ was less significant in this model. This change in the order of the features can be attributable to the information interaction between them.
Because prior diagnosis of CAD is such a power predictor of a positive SE, the other features contribute so little information by comparison, and it is hard to see their value. However, once these patients are removed from the dataset, we can see the predictive power of the other features for patients with no previous history of CAD.

Features like age, diabetes and family history are shown to be less significant for discriminating between positive and negative cases. It also showed that the information about the medications added significant value and could enhance the discrimination power of the clinical data. It also showed that interaction between features is important and can affect the order of the selected subset. Moreover, increasing the granularity of the value of the risk factors may improve their discriminative power by using continuous instead of categorial variables.

Strength

To our knowledge, this is the first study that has investigated applying ML techniques to a simple dataset of patient anthropometrics and cardiovascular risk factor profiles and cardioactive medications to predict positive/abnormal SE results. The study also investigated the performance of different ML techniques and employed a sophisticated feature selection method to study the significance of the clinical attributes. This method considers the interaction between clinical variables when analysing their significance and the class label. The proposed framework outperforms the other tools that have been proposed in the literature [11] in predicting CAD by more than 9%. The proposed framework can also be employed on data collected using other cardiovascular stress tests aimed to detect inducible ischaemia.

Limitations and future work

In this paper we report preliminary results using only 529 patients. The data includes only patients' anthropometric and clinical data that has been collected during the patient's hospital visit. Including more data from patient medical records could enhance the generic behaviour of any proposed model and improve the performance of the developed model. As we have a large dataset going back nearly 20 years, the model could be extended to predict mortality due to a cardiovascular event.

ML techniques can offer the very promising prospect of faster and more accurate diagnosis (especially for high risk groups), prioritising higher risk patients and increasing the capacity of clinicians. However, it is well known that most ML techniques are considered to be "black-boxes", where the model produces difficult to interpret results. Despite the black-box nature of various ML approaches, feature selection techniques can improve the understanding of the relationship between the diagnosis and clinical attributes. Data visualisation methods can improve the understanding of the produced model and interpretation of the output.

None of the clinical information detailing the results of the positive stress test such as wall motion score index is included with the clinical data. Inclusion may further differentiate between high and low risk patients.

Acknowledgement

This research was part funded by UK EPSRC grants EP/P01013X/1 (STRETCH) to BP and MB, EP/R013144/1 (SAUSE) to BP and an Open University Priority Research Area grant to DB.

References

1. Global, regional, and national age–sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. The Lancet 2015 Jan;385(9963):117–171. [doi: 10.1016/S0140-6736(14)61682-2]

2. British Heart Foundation. British Heart Foundation UK Factsheet [Internet]. 2019 Aug. Available from: https://www.bhf.org.uk/-/media/files/research/heart-statistics/bhf-cvd-statistics-uk-factsheet.pdf?la=en

3. Patidar S, Pachori RB, Rajendra Acharya U. Automated diagnosis of coronary artery disease using tunable-Q wavelet transform applied on heart rate signals. Knowl-Based Syst 2015 Jul;82:1–10. [doi: 10.1016/j.knosys.2015.02.011]

4. 2013 ESC guidelines on the management of stable coronary artery disease. The Task Force on the management of stable coronary artery disease of the European Society of Cardiology. Rev Esp Cardiol Engl Ed 2014 Feb;67(2):135. [doi: 10.1016/j.rec.2013.11.008]

5. Safdar S, Zafar S, Zafar N, Khan NF. Machine learning based decision support systems (DSS) for heart disease diagnosis: a review. Artif Intell Rev 2018 Dec;50(4):597–623. [doi: 10.1007/s10462-017-9552-8]

6. Singh G, Al'Aref SJ, Van Assen M, Kim TS, van Rosendael A, Kolli KK, Dwivedi A, Maliakal G, Pandey M, Wang J, Do V, Gummalla M, De Cecco CN, Min JK. Machine learning in cardiac CT: Basic concepts and contemporary data. J Cardiovasc Comput Tomogr 2018 May;12(3):192–201. [doi: 10.1016/j.jcct.2018.04.010]

7. Mastoi Q, Wah TY, Gopal Raj R, Iqbal U. Automated Diagnosis of Coronary Artery Disease: A Review and Workflow. Cardiol Res Pract 2018;2018:1–9. [doi: 10.1155/2018/2016282]

8. Johnson KW, Torres Soto J, Glicksberg BS, Shameer K, Miotto R, Ali M, Ashley E, Dudley JT. Artificial Intelligence in Cardiology. J Am Coll Cardiol 2018 Jun;71(23):2668–2679. [doi: 10.1016/j.jacc.2018.03.521]

9. Voss R, Cullen P, Schulte H, Assmann G. Prediction of risk of coronary events in middle-aged men in the Prospective Cardiovascular Münster Study (PROCAM) using neural networks. Int J Epidemiol 2002 Dec 1;31(6):1253–1262. [doi: 10.1093/ije/31.6.1253]

10. Gharehchopogh FS, Khalifelu ZA. Neural Network application in diagnosis of patient: A case study. Int Conf Comput Netw Inf Technol [Internet] Abbottabad, Pakistan: IEEE; 2011 [cited 2019 Sep 27]. p. 245–249. [doi: 10.1109/ICCNIT.2011.6020937]

11. Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? Liu B, editor. PLOS ONE 2017 Apr 4;12(4):e0174944. [doi: 10.1371/journal.pone.0174944]

12. Dougherty J, Kohavi R, Sahami M. Supervised and Unsupervised Discretization of Continuous Features. Mach Learn Proc 1995 [Internet] Elsevier; 1995 [cited 2019 Sep 27]. p. 194–202. [doi: 10.1016/B978-1-55860-377-6.50032-3]

13. Bennasar M, Setchi R, Hicks Y. Unsupervised Discretization Method based on Adjustable Intervals. Front Artif Intell Appl 2012;79–87. [doi: 10.3233/978-1-61499-105-2-79]

14. Bennasar M, Setchi R, Hicks Y. Feature Interaction Maximisation. Pattern Recognit Lett 2013 Oct;34(14):1630–1635. [doi: 10.1016/j.patrec.2013.04.002]

15. Bennasar M, Hicks Y, Setchi R. Feature selection using Joint Mutual Information Maximisation. Expert Syst Appl 2015 Dec;42(22):8520–8532. [doi: 10.1016/j.eswa.2015.07.007]

16. Cover TM, Thomas JA. Elements of information theory. 2nd ed. Hoboken, N.J: Wiley-Interscience; 2006. ISBN:978-0-471-24195-9

17. Smialowski P, Frishman D, Kramer S. Pitfalls of supervised feature selection. Bioinformatics 2010 Feb 1;26(3):440–443. [doi: 10.1093/bioinformatics/btp621]

18. Saunders C, Stitson MO, Weston J. Support Vector Machine Reference Manual [Internet]. Department of Computer Science, Royal Holloway; 1998 p. 26. Report No.: CSD-TR-98-03, 1998. Available from: https://eprints.soton.ac.uk/258959/1/SVM_Reference.pdf

19. Breiman L. Random Forests. Mach Learn 2001 Oct 1;45(1):5–32. [doi: 10.1023/A:1010933404324]

Abbreviations

1-D: One dimension
ACE-I/ARB: Angiotensin Converting Enzyme Inhibitor/Angiotensin Receptor Blocker
CAD: Coronary Artery Disease
CNN: Convolutional Neural Network
CVD: Cardiovascular Diseases
ECG: Electrocardiogram
EWD: Equal Width Discretisation
JMIM: Joint Mutual Information Maximisation
MI: Mutual Information
ML: Machine Learning
MLP: Multi-Layer Perceptron
RBF: Radial Basis Function
SD: Standard deviation
SE: Stress echocardiography
SVM: Support Vector Machine
TTE: Trans Thoracic Echocardiography