Prediction of cardiac death after adenosine myocardial perfusion SPECT based on machine learning

David Haro Alonso, MSc,[a] Miles N. Wernick, PhD,[a] Yongyi Yang, PhD,[a] Guido Germano, PhD,[b] Daniel S. Berman, MD,[b] and Piotr Slomka, PhD[b]

[a] Medical Imaging Research Center, Illinois Institute of Technology, Chicago, IL
[b] Departments of Imaging and Medicine, Cedars-Sinai Medical Center, Los Angeles, CA

Received Sep 22, 2017; accepted Feb 19, 2018
doi:10.1007/s12350-018-1250-7

Background. **We developed machine-learning (ML) models to estimate a patient's risk of cardiac death based on adenosine myocardial perfusion SPECT (MPS) and associated clinical data, and compared their performance to baseline logistic regression (LR). We demonstrated an approach to visually convey the reasoning behind a patient's risk to provide insight to clinicians beyond that of a "black box."**

Methods. **We trained multiple models using 122 potential clinical predictors (features) for 8321 patients, including 551 cases of subsequent cardiac death. Accuracy was measured by area under the ROC curve (AUC), computed within a cross-validation framework. We developed a method to display the model's rationale to facilitate clinical interpretation.**

Results. **The baseline LR (AUC = 0.76; 14 features) was outperformed by all other methods. A least absolute shrinkage and selection operator (LASSO) model (AUC = 0.77; *p* = .045; 6 features) required the fewest features. A support vector machine (SVM) model (AUC = 0.83; *p* < .0001; 49 features) provided the highest accuracy.**

Conclusions. **LASSO outperformed LR in both accuracy and simplicity (number of features), with SVM yielding best AUC for prediction of cardiac death in patients undergoing MPS. Combined with presenting the reasoning behind the risk scores, our results suggest that ML can be more effective than LR for this application. (J Nucl Cardiol 2019;26:1746–54.)**

Key Words: Cardiac death • risk model • machine learning • feature selection • data visualization

Abbreviations

SPECT	Single-photon emission computed tomography
MPS	Myocardial perfusion SPECT
SSS	Summed stress score
SRS	Summed rest score
LV	Left ventricular
ROC	Receiver operating characteristic
AUC	Area under the curve
SVM	Support vector machine
LASSO	Least absolute shrinkage and selection operator
LR	Logistic regression

See related editorial, pp. 1755–1758

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12350-018-1250-7) contains supplementary material, which is available to authorized users.

The authors of this article have provided a PowerPoint file, available for download at SpringerLink, which summarises thecontents of the paper and is free for re-use at meetings andpresentations. Search for the article DOI on SpringerLink.com

Reprint requests: Miles N. Wernick, PhD, Medical Imaging Research Center, Illinois Institute of Technology, 3440 S. Dearborn St., Suite 100, Chicago, IL 60616; wernick@iit.edu

1071-3581/$34.00

INTRODUCTION

It is well established that myocardial perfusion single-photon emission computed tomography (MPS)[1] can provide prognostic information for patients with coronary artery disease.[2–7] However, assessment of prognosis and consequently selection of appropriate therapies may be improved through consideration of a diverse set of additional clinical data that, although available for a given patient, may be challenging for the clinician to aggregate effectively. Therefore, it may be beneficial to provide the clinician with an accurate quantitative assessment of the risk of adverse outcomes, such as cardiac death, based on the combination of clinical information and MPS data. In this paper, we explore the extent to which the burgeoning field of machine learning (ML)[8–10] can provide risk assessments that are more accurate and useful than those produced by logistic regression (LR),[11] which is a traditional, commonly used model in this context.[12–14]

In this work, we compare several ML algorithms to a baseline LR model according to two criteria: (1) prediction accuracy, measured by using the area under the receiver operating characteristic curve (AUC); and (2) simplicity of the model, described by the number of predictors (known in ML as *features*).

Importantly, a clinician may not fully benefit from a computer-generated risk score without some explanation as to the reasoning behind it. Without the rationale behind the risk score, the prediction model may be perceived as an uninterpretable ''black box.'' To address this, we demonstrate a method of displaying a model's reasoning to the clinician. To facilitate interpretability, we sought to develop models that use a small number of features, so that the clinician can readily gain insight by viewing the model's findings using a simple display. Although, in general, the ML models are sophisticated in their theoretical underpinnings, the ML linear models that we studied are simply a weighted sum of the features; therefore, the results are easy to interpret.

MATERIALS AND METHODS

Patient Data

Our study was based on a set of clinical data for 8321 patients with a mean age of 71 ± 12 years and all clinical variables recorded, who had undergone dual-isotope MPS with adenosine stress and adenosine stress with walking as an adjunct[15] at Cedars-Sinai Medical Center between January 1991 and December 1999. Among these were 551 subsequent cases of cardiac death during a follow-up period of 3.15 ± 1.99 years. Cardiac death was confirmed by review of the death certificate, hospital chart, or physician's records.

Each patient was characterized by a broad array of features, including MPS and ECG findings, and many other clinical variables. The clinical variables considered numerous attributes of the patient, such as demographic risk factors, medications, hemodynamics, and the clinical history of the patient (see Table 1). The MPS findings included the clinically recorded visual summed stress score (SSS), summed rest score (SRS), and summed difference score, obtained from the 20-segment model and the corresponding converted % scores.[16] The images were not reprocessed in any other way.

Continuous features were scaled to the range 0 to 1. For consistency, discrete features having multiple categories were re-coded as a set of binary variables using one-hot encoding,[17] i.e., we created a binary feature for each different category of the discrete variable, and for each patient only one of the binary features takes the value 1. We ignored the features that had missing values, since more than half of their entries were missing and imputation methods would not lead to accurate

Table 1. Patient characteristics

Demographic risk factors	
Patient's age (years)	71 ± 12
Gender (male)	52%
Hypertension	60%
Diabetes	27%
Family history of CAD	23%
Hypercholesterolemia	44%
Smoking	13%
Angina	49%
Shortness of breath	12%
Medications	
Beta-blocker	15%
Calcium channel blocker	12%
Nitrate (any)	10%
Digoxin	10%
Clinical history	
History of MI	28%
History of PCI	16%
History of CABG	18%
Hemodynamics	
Rest systolic BP (mm Hg)	150 ± 28
Rest diastolic BP (mm Hg)	80 ± 12
Stress systolic BP (mm Hg)	135 ± 26
Stress diastolic BP (mm Hg)	72 ± 13
Rest HR (per min)	72 ± 14
Stress HR (per min)	92 ± 20

CAD, coronary artery disease; *MI*, myocardial infarction; *PCI*, percutaneous coronary intervention; *CABG*, coronary artery bypass grafting; *BP*, blood pressure; *HR*, heart rate results. After data cleaning and normalization, the final data set consisted of 122 features in the range from 0 to 1.

Score Validation and Statistical Analysis

We characterized the distribution of each continuous feature by way of its mean and standard deviation, while the binary categorical features are described by their prevalence within the data set. We fit and test the models in different sets of data to ensure unbiased performance evaluation. We used nested, stratified, fivefold cross-validation in our experiments. The data are divided in five parts (known as folds) of similar size, and then each fold is used only once for validating the performance of the model; the final performance is the average of the five iterations. We ''stratified'' the folds to preserve the percentage of patients of each class in the divisions, and we used a ''nested'' framework to optimize the search of the model's hyperparameters and avoid overfitting the model—an inner fivefold cross-validation is used within the training data to select the best values of the hyperparameters, while the outer cross-validation evaluates the performance of the ML models. We compared the performance of various methods for prediction of cardiac death using receiver operating characteristic (ROC) analysis, specifically the area under the curve (AUC), which summarizes ROC performance. We conducted a statistical analysis to compare the AUC values obtained from several ML and LR models. The difference in performance between a test model and a baseline LR model (described in the ''Results'' section) was considered to be statistically significant when the p value was less than 0.05, using the test described in the Reference [18].

Prediction Models

We implemented several variations of LR, and the following four categories of ML models[19–23]: (1) the least absolute shrinkage and selection operator (LASSO); (2) the support vector machine (SVM); (3) random forests; and (4) AdaBoost. The LASSO automatically performs *shrinkage*, in which an internal mechanism is used to automatically select a minimal number of predictive features. SVM is a technique that places an emphasis on distinguishing the example cases that are the most difficult to predict. The random forests method forms a consensus decision among multiple decision trees. AdaBoost iteratively refines a model by favoring example cases that are misclassified by previous iterations of the model.

LR. The simplicity and good performance of LR have led to its adoption as a widely used method to predict adverse clinical outcomes. To understand the characteristics of LR when applied to our data, and provide a basis for comparison to the ML models, we implemented three LR variations[24,25]: (1) LR with prognostically important features identified based on a Cox proportional hazards analysis; (2) LR applied to all 122 features (LR-All); and (3) stepwise LR, in which features are successively added or removed based on significance of the regression coefficients (LR-Stepwise). In the LR-Stepwise model, we set $p < .05$ as the threshold for entry of variables into the model, and $p > .10$ for their removal. We defined the first of these three LR variations as our baseline model (hence, we refer to it as LR-Baseline), because it is widely accepted; and, as we show in the Results section, it yielded the simplest (most interpretable) of the LR models, while still exhibiting good prediction accuracy.

AdaBoost and random forests. We implemented AdaBoost and random forests in only the standard fashion.[26,27]

SVM. SVM may be implemented as a linear model (i.e., the computed risk score being merely a weighted sum of the features), or as a more-complicated nonlinear version.[21] We compared two variations of linear SVM: SVM using all 122 features (SVM-All), and SVM trained within a sequential forward-selection procedure (SVM-SFS) to choose the best features.[28] We also compared two variations of nonlinear SVM: SVM with radial basis function kernel (SVM-RBF), and SVM with quadratic kernel (SVM-Q).[29]

LASSO. We implemented two prediction models based on the LASSO method. LASSO internally selects the features that it uses through an optimization mechanism that involves a so-called *shrinkage parameter*, which controls the extent to which the LASSO reduces the number of features.[30] In the standard method (LASSO-Standard) that we implemented, LASSO chose the shrinkage parameter automatically. In a modified approach (LASSO-Modified), we sought to determine the extent to which the LASSO could further shrink the number of required features, without substantial loss of prediction accuracy. Toward this end, LASSO-Modified was programmed to increase its shrinkage parameter as much as possible without causing LASSO's performance to drop below that of LR-Baseline.

Accuracy vs. Simplicity

The objective of our study was to study both the accuracy and complexity of various models, where we define complexity as the number of features required by the model to obtain its predictions. For purposes of interpretability, it would be desirable to identify a model that is effective with as few features as possible. Thus, we examine the number of required features, as well as the accuracy, when comparing the models.

Converting the Risk Score to a Probability

The LR and ML prediction models each seek to predict a normalized risk score that lies in the range from 0 to 1, with 1 representing the highest level of risk. We convert the score in an individual patient to a true probability of cardiac death by relating it to the outcomes of the population of patients in the data set.[31,32] Thus, the final output is a probability of cardiac death.

''Rationale'' Display

ML is sometimes perceived as yielding a model that is merely a ''black box,'' producing a decision without any explanation as to the reasoning behind it. To address this

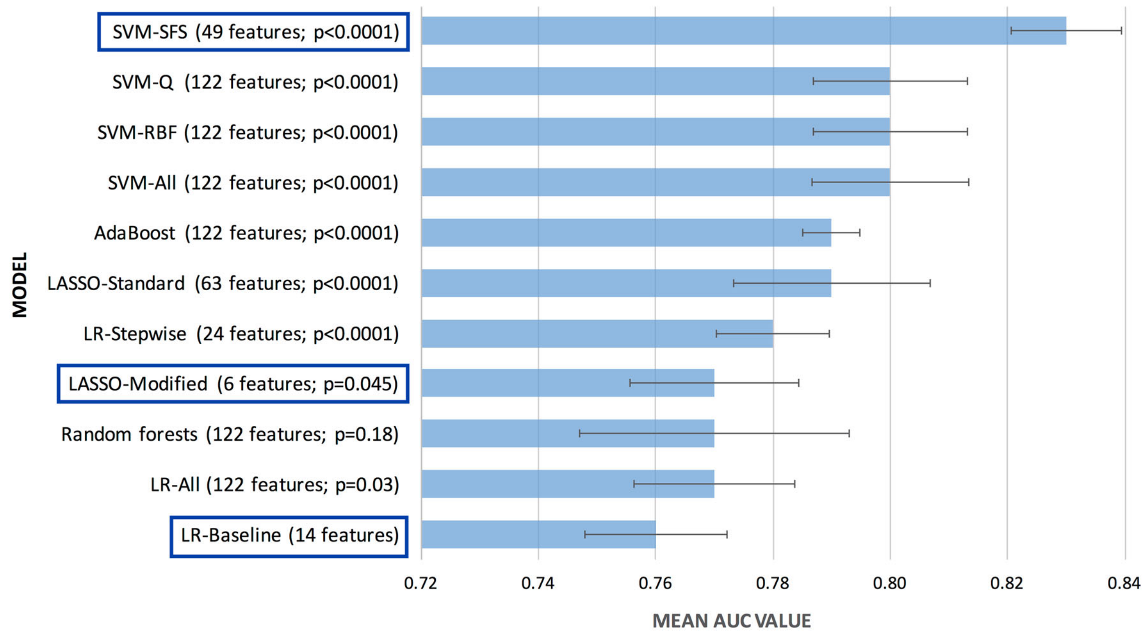

Figure 1. Performance comparison of all models tested. Mean AUC, with error bars representing one standard deviation, is shown for each model. For each model, we show how many features it required and the p value when comparing that model vs. LR-Baseline. We draw the reader's attention to three models in particular, indicated by a box drawn around each one's label. The results show that all of the methods we tested are able to surpass the baseline logistic regression (LR-Baseline) in terms of AUC performance. The LASSO-Modified method provides both better AUC performance and requires fewer features, so it surpasses LR-Baseline in both respects. If sheer AUC performance is the goal, then one would choose SVM-SFS, which produces substantially better performance than any of the other methods, the disadvantage being that it requires 49 features. In summary, LASSO-Modified produces the most parsimonious model, while SVM-SFS yields the best AUC performance.

limitation and provide the clinician with helpful insight, we developed a data-visualization approach that shows the user two ways of understanding the model's risk assessment. First, we show the relative contributions of the risk factors (features)—using clinical and MPS variables—to a patient's risk assessment, expressed as a percentage. A risk factor's contribution to the score is the product of the risk factor with the corresponding model coefficient, expressed as a percentage of the total. Thus, the output can be interpreted to mean that ''X% of the patient's risk score is due to feature x, Y% is due to feature y, etc.'' Using unity-based normalization (scaling to the range [0,1]), as we have done in this study, guaranteed that these contributions are non-negative if the model coefficients and class labels are non-negative. That was the case in our study (the labels were 1 for death, 0 otherwise; and all the variables are positive predictors). In situations where one or more variables are negative predictors (which would lead to negative model coefficients), one can invert the scale of such variables, thus leading to non-negative contributions from the model coefficients.

As a second aid for ML decision understanding, we show the percentile ranking of the patient with respect to the patient population used for the ML development. This can be viewed as a ''severity'' rank, showing how severe is a given variable in relation to other patients. We illustrate these ''at-a-glance'' display approaches in the Results section.

RESULTS

Comparisons of Model Performance

The accuracy and complexity of the models are shown in Figure 1, in which each bar represents mean AUC performance of a model, and each error bar represents one standard deviation. In the label below each bar in the graph, we indicate the number of features used by that model, as well as the p value in comparison of the model's AUC performance to that of the baseline LR model (LR-Baseline).

The results obtained by the three LR variations were as follows: (1) LR-Baseline: mean AUC = 0.76, 14 features; (2) LR-All: AUC = 0.77, 122 features, p = .03; (3) LR-Stepwise: mean AUC = 0.78, 24 features, $p < .0001$. The features selected by the LR-Baseline method are shown in Table 2. Note that, of the LR methods, the LR-Baseline uses fewer features while maintaining good prediction accuracy.

Both of the linear SVMs outperformed LR-Baseline very significantly: SVM-SFS achieved mean AUC = 0.83 (49 features, $p < .0001$), and SVM-All yielded mean AUC = 0.80 (122 features, $p < .0001$). The two nonlinear variations SVM performed about the same as SVM-All, but they increase the complexity of the classifier; therefore, they offer no benefit in this dataset.

Random forests and AdaBoost achieved mean AUC = 0.77 (122 features, $p = .18$) and AUC = 0.79 (122 features, $p < .0001$), respectively.

In the standard method (LASSO-Standard) that we implemented, LASSO chose the shrinkage parameter automatically; LASSO-Standard yielded AUC = 0.79 ($p < .0001$) using 63 features.

Next, we sought to determine the extent to which the LASSO could further shrink the number of required features, without substantial loss of prediction accuracy. Toward this end, we developed a modified LASSO (LASSO-Modified), in which we programmed LASSO to increase its shrinkage parameter as much as possible without causing LASSO's performance to drop below that of LR-Baseline. By this approach, LASSO-Modified reduced the number of features to just six, while still producing better accuracy than LR-Baseline (AUC = 0.77; $p = .045$).

Table 2. Variables used in baseline logistic regression model (LR-Baseline)

Variables	Coefficients	p value
SSS	4.3480	< .0001
PRSSS	− 4.0694	< .0001
SRS	2.1991	< .0001
PRSDS	− 2.1159	< .0001
Patient's age	0.5459	.016
Abnormal rest ECG	0.3021	.028
Symptom	0.2191	.045
LV enlargement	0.1643	.001
Uninterpretable ECG	0.1400	.044
Perfusion defects	0.1302	< .001
Perfusion findings	0.0816	.002
Bypass surgery	0.0755	.045
Likelihood of ischemia	− 0.0260	.041
History of MI	0.0204	.025

SSS, summed stress score; *PRSSS*, percentage of LV abnormal; *SRS*, summed rest score; *PRSDS*, percentage of LV reversible; *ECG*, electrocardiogram; *LV*, left ventricular; *MI*, myocardial infarction
Logistic regression model using the variables that the univariable Cox proportional hazards model finds statistically significant (the *p* values are shown in the table)

Table 3. Variables used by LASSO-modified model

Variables	Coefficients
SSS	0.3089
Patient's age	0.2962
SRS	0.2325
LV enlargement	0.1916
Lung uptake	0.1747
Digoxin	0.0693

SSS, summed stress score; *SRS*, summed rest score; *LV*, left ventricular

Table 4. Features for two example patients in the data set

Variables	Patient A	Patient B
SSS	29	16
Patient's age	68 years	60 years
SRS	26	16
LV enlargement	Yes	Yes
Lung uptake	Severe	None
Digoxin	Yes	Yes
Suffered cardiac death	Yes	No

SSS, summed stress score; *SRS*, summed rest score; *LV*, left ventricular

The six features identified by LASSO-Modified, in order of importance, are as follows: summed stress score (SSS), age, summed rest score (SRS), left ventricular (LV) enlargement, lung uptake, and use of digoxin. The model coefficients are shown in Table 3.

We interpret the results of the study as follows. (1) When provided with a large number of features, the ML methods produce the best overall prediction performance, in some cases by a large margin. (2) LR underperforms the ML methods, even when provided with large numbers of features. (3) If one's goal is to obtain good performance with a very simple model, the LASSO-Modified is the best: it provided slightly better performance than LR-Baseline, while requiring only six features. (4) On the other hand, if one's goal is strictly to achieve maximum performance, the SVM-SFS is the best overall by a substantial margin. Alternatively, both models can be used together to provide a fuller picture, with the LASSO-Modified helping with interpretation.

Displaying the Model's Rationale

To help the clinician to interpret the features driving the risk score for a given patient, we provide a graph

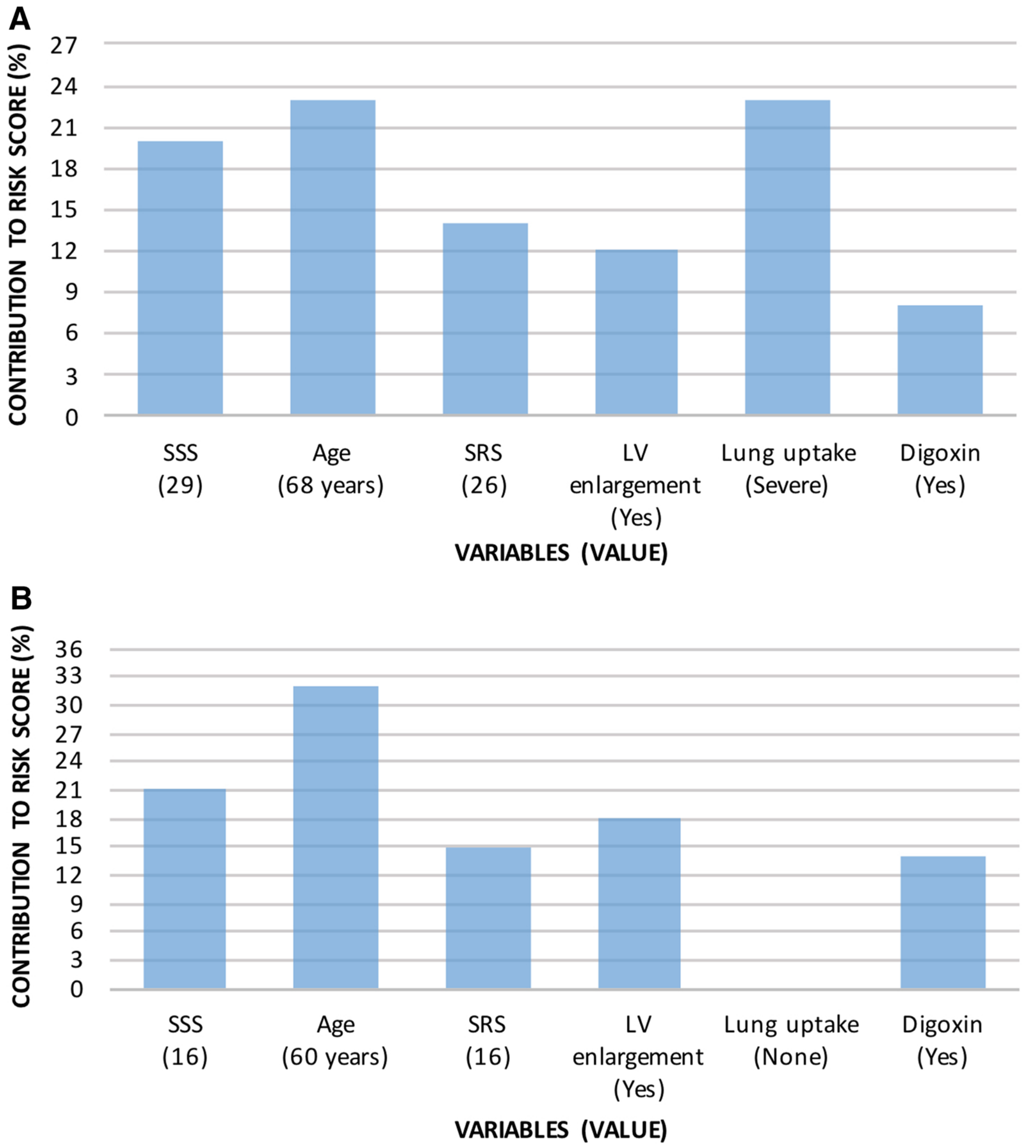

Figure 2. Rationale behind two patients' risk score in the case study. Proportion of **A** patient A and **B** patient B's total score explained by each variable. We show the value of the corresponding patient's variables in parenthesis. Patient A suffered cardiac death, patient B experienced no event.

showing the contribution of each feature, expressed as a percentage. Here, we use two patients (A and B) from the data set to illustrate the visualization approach.

Patient A was a case of cardiac death, while Patient B experienced no adverse cardiac event. Table 4 summarizes the features of Patients A and B that were used by the LASSO-Modified model, which estimated the probabilities of cardiac death for Patients A and B as 63% and 10%, respectively. Figure 2 shows the risk probability for each patient, above a graph that explains the rationale behind the probability assessment. As we saw in Table 4, patient A exhibited a high value for lung uptake, which highly influenced the estimated probability of cardiac mortality. This condition, along with the patient's age and the elevated value of the summed stress score (SSS), explains 66% of the total score. For Patient B, a younger person, lung uptake did not enter significantly into the risk score, and his SSS and SRS values were lower than those found in patient A, resulting in a reduction of the risk score. Consequently, the variables with the highest contribution to the patient's risk are his age and SSS, which explain 53% of the total score.

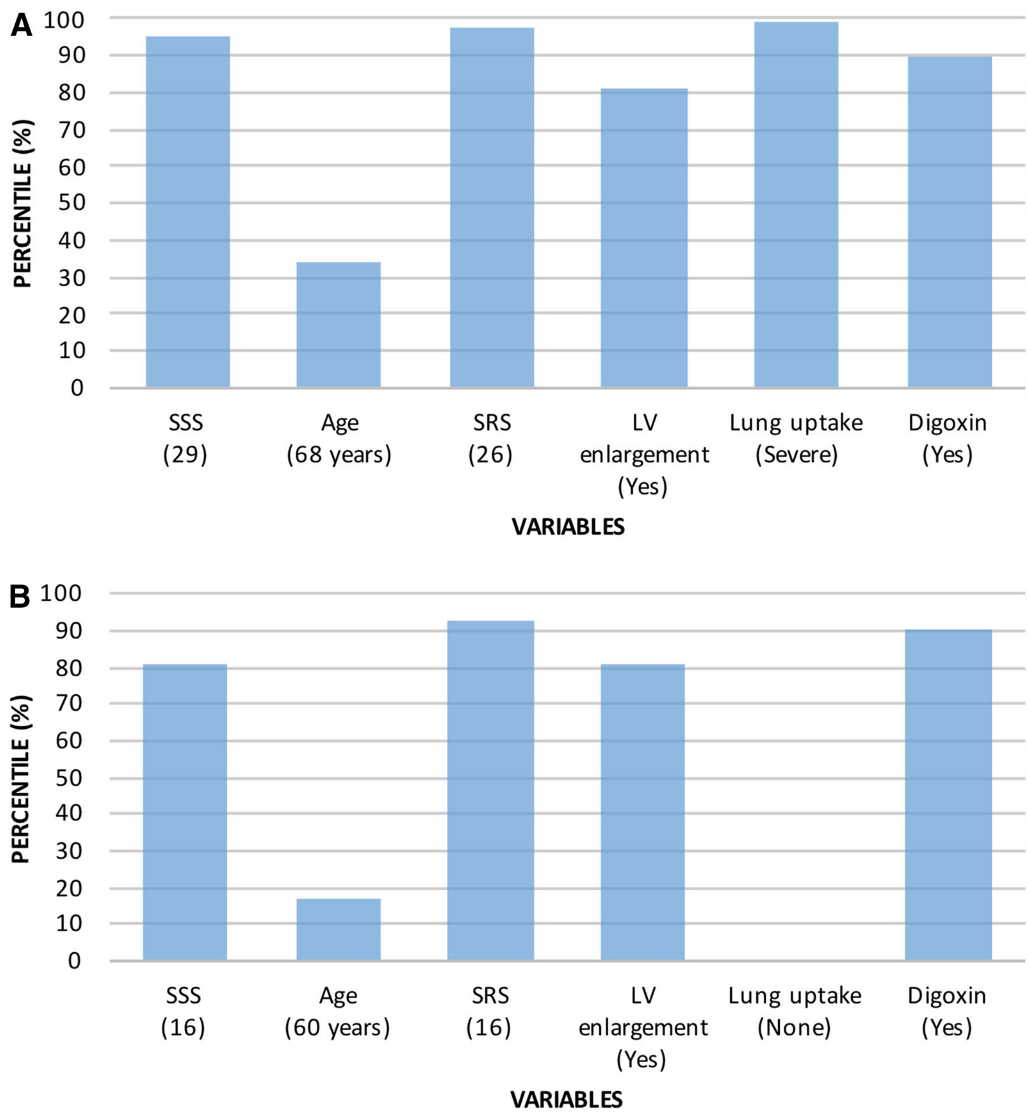

Figure 3. Percentile rank of two patient's characteristics in the case study. The graph compares the value of **A** patient A and **B** patient B's variables with those in the entire database. We show the value of the corresponding patient's variables in parenthesis.

Figure 3 shows the percentile rank of the features for patients A and B. This shows that several of the variables for Patient A are generally worse than those for Patient B, which explains the higher estimated risk for Patient A.

DISCUSSION

Modeling Results

The goal of this work was to develop a model that can accurately predict an individual patient's risk of cardiac mortality, and at the same time allow the clinician to understand the reasoning behind the model's assessment. It would be preferable to obtain a model that uses as few features as possible, so that the relative importance of the risk factors can be readily appreciated within the time constraints of clinical practice.

All the models tested outperformed LR-Baseline in terms of prediction accuracy, with SVM-SFS producing the best performance overall. The LASSO-Modified model was the simplest of the models tested, requiring only six features, while still outperforming LR-Baseline, which required more than twice as many[14] features.

The SVM-SFS model produced the highest overall prediction accuracy, reaching AUC = 0.83, as compared to AUC = 0.76 for LR-Baseline, and AUC = 0.77 for LASSO-Modified. However, SVM-SFS required 49 features, which we anticipate might be too many features to be readily reviewed and interpreted in the clinic. Nevertheless, we can imagine situations in which the clinician may wish to obtain the best accuracy possible (by using SVM-SFS), but may turn to the LASSO-Modified output to understand the roles of the underlying risk factors.

Limitations

Although cross-validation was used for unbiased performance evaluation, deriving the predictive model from historical data may not generalize in a similar way to future data due to changes in the patient's characteristics. Rozanski et al. demonstrated in Reference [33] a progressive decline in the prevalence of abnormal SPECT studies, from 40.9% in 1991 to 8.7% in 2009. They also found significant temporal trends in the average patient age, gender, clinical history, and hemodynamics parameters. These temporal trends may involve potential challenges in applying machine learning algorithms developed on historical data to new prospective data.

Our analysis ignores the survival time and simply predicts cardiac death at any time; therefore, we do not distinguish between near- and long-term cardiac death prediction. Nevertheless, the average follow-up period that we use is relatively short (3.15 years). There are emerging machine learning methods in which the time-to-event can be used[34]; however, the current models do not provide the patient-specific explanation capability that we have demonstrated in our study. No analysis adjusting for the time-to-event has been reported in cardiovascular imaging using machine learning. The exact prediction of time of cardiac death based on medical imaging in cardiology is less precise than in other fields. Further studies, especially in cohorts with long average follow-up time, may incorporate time-to-event information in the machine learning models.

NEW KNOWLEDGE GAINED

We have developed and compared several models for predicting risk of cardiac death for individual patients undergoing nuclear cardiology studies. Our study demonstrates the potential benefit of machine learning methods over logistic regression for prediction of cardiac death event. To aid in interpretation of the risk assessments provided by the models, we have demonstrated two clinically practical methods of visualizing the relevance and impact of the risk factors.

CONCLUSIONS

We used machine learning to predict a patient's risk of cardiac death based on the combination of MPS findings and other clinical variables, which maximize the prediction accuracy and minimize the number of input variables. We have shown that an SVM algorithm can be the most accurate but requires large number of features, and that a LASSO model can achieve good accuracy with only six variables. We develop a data-visualization tool based on the rationale behind the LASSO's risk score to allow the clinician to interpret the results. We also provide the percentile rank that offers an understanding of the relative severity of the patient's characteristics.

Acknowledgments

Research reported in this article was supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health under Award Numbers R01HL122484 and R01HL089765. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Disclosure

DH.A., Y.Y., and M.N.W. from the Illinois Institute of Technology have nothing to disclose. G.G., D.S.B., and P.S. from the Cedars-Sinai Medical Center have nothing to disclose.

References

1. Wernick MN, Aarsvold JN. Emission tomography: The engineering and physics of PET and SPECT. San Diego: Academic Press; 2004. p. 596.
2. Hachamovitch R, Berman DS, Kiat H, Cohen I, Lewin H, Amanullah A, et al. Incremental prognostic value of adenosine stress myocardial perfusion single-photon emission computed tomography and impact on subsequent management in patients with or suspected of having myocardial ischemia. Am J Cardiol 1997;80:426-33.
3. Hachamovitch R, Berman DS, Shaw LJ, Kiat H, Cohen I, Cabico JA, et al. Incremental prognostic value of myocardial perfusion single photon emission computed tomography for the prediction of cardiac death. Circulation 1998;97:535-43.
4. Klocke FJ, Baird MG, Lorell BH, Bateman TM, Messer JV, Berman DS, et al. ACC/AHA/ASNC guidelines for the clinical use of cardiac radionuclide imaging—executive summary. Circulation 2003;108:1404-18.
5. Sharir T, Ben-Haim S, Merzon K, Prochorov V, Dickman D, Ben-Haim S, et al. High-speed myocardial perfusion imaging: initial

clinical comparison with conventional dual detector anger camera imaging. JACC: Cardiovasc Imaging 2008;1:156-63.
6. Heller GV, Herman SD, Travin MI, Baron JI, Santos-Ocampo C, Mcclellan JR. Independent prognostic value of intravenous dipyridamole with technetium-99 m sestamibi tomographic imaging in predicting cardiac events and cardiac-related hospital admissions. J Am Coll Cardiol 1995;26:1202-8.
7. Arsanjani R, Xu Y, Dey D, Fish M, Dorbala S, Hayes S, et al. Improved accuracy of myocardial perfusion SPECT for the detection of coronary artery disease using a support vector machine algorithm. J Nuclear Med 2013;54:549-55.
8. Bishop CM. Pattern recognition. Mach Learn 2006;128:1-58.
9. Duda RO, Hart PE, Stork DG. Pattern classification. New York: Wiley; 2012.
10. Wernick MN, Yang Y, Brankov JG, Yourganov G, Strother SC. Machine learning in medical imaging. Signal Process Mag IEEE 2010;27:25-38.
11. Hastie T, Tibshirani R, Friedman J, Hastie T, Friedman J, Tibshirani R. The elements of statistical learning. Berlin: Springer; 2009.
12. Hachamovitch R, Hayes SW, Friedman JD, Cohen I, Berman DS. Stress myocardial perfusion single-photon emission computed tomography is clinically effective and cost effective in risk stratification of patients with a high likelihood of coronary artery disease (CAD) but no known CAD. J Am Coll Cardiol 2004;43:200-8.
13. Nakajima K, Matsuo S, Okuyama C, Hatta T, Tsukamoto K, Nishimura S, et al. Cardiac event risk in Japanese subjects estimated using gated myocardial perfusion imaging, in conjunction with diabetes mellitus and chronic kidney disease. Circ J 2012;76:168-75.
14. Hachamovitch R, Hayes SW, Friedman JD, Cohen I, Berman DS. Comparison of the short-term survival benefit associated with revascularization compared with medical therapy in patients with no prior coronary artery disease undergoing stress myocardial perfusion single photon emission computed tomography. Circulation 2003;107:2900-7.
15. Berman DS, Kang X, Hayes SW, Friedman JD, Cohen I, Abidov A, et al. Adenosine myocardial perfusion single-photon emission computed tomography in women compared with men: Impact of diabetes mellitus on incremental prognostic value and effect on patient management. J Am Coll Cardiol 2003;41:1125-33.
16. Berman DS, Kiat H, Friedman JD, Wang FP, Van Train K, Matzer L, et al. Separate acquisition rest thallium-201/stress technetium-99 m sestamibi dual-isotope myocardial perfusion single-photon emission computed tomography: A clinical validation study. J Am Coll Cardiol 1993;22:1455-64.
17. Raschka S. Python machine learning. Birmingham: Packt Publishing Ltd; 2015.
18. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837-45.
19. Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc Ser B (Methodol) 1996;58:267-88.
20. Vapnik VN, Vapnik V. Statistical learning theory. New York: Wiley; 1998.
21. El-Naqa I, Yang Y, Wernick MN, Galatsanos NP, Nishikawa RM. A support vector machine approach for detection of microcalcifications. Med Imaging IEEE Trans 2002;21:1552-63.
22. Breiman L. Random forests. Mach Learn 2001;45:5-32.
23. Freund Y, Schapire RE. A desicion-theoretic generalization of on-line learning and an application to boosting. Eur Conf Comput Learn Theory. 1995. https://doi.org/10.1006/jcss.1997.1504.
24. Cox DR. Regression models and life-tables. Breakthr Stat 1992;527-541.
25. Greenland S. Modeling and variable selection in epidemiologic analysis. Am J Public Health 1989;79:340-9.
26. Ho TK. The random subspace method for constructing decision forests. IEEE Trans Pattern Anal Mach Intell 1998;20:832-44.
27. Schapire RE. The boosting approach to machine learning: An overview. Nonlinear Estim Classif 2003;149-71.
28. Pudil P, Novovičová J, Kittler J. Floating search methods in feature selection. Pattern Recognit Lett 1994;15:1119-25.
29. Scholkopf B, Smola AJ. Learning with kernels: Support vector machines, regularization, optimization, and beyond. Cambridge: MIT Press; 2001.
30. Park MY, Hastie T. L1-regularization path algorithm for generalized linear models. J R Stat Soc 2007;69:659-77.
31. Niculescu-Mizil A, Caruana R. Predicting good probabilities with supervised learning. In: Proceedings of the 22nd international conference on machine learning; 2005. p. 625–632.
32. Platt J. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. Adv Large Margin Classif 1999;10:61-74.
33. Rozanski A, Gransar H, Hayes SW, Min J, Friedman JD, Thomson LE, et al. Temporal trends in the frequency of inducible myocardial ischemia during cardiac stress testing: 1991 to 2009. J Am Coll Cardiol 2013;61:1054-65.
34. Mogensen UB, Ishwaran H, Gerds TA. Evaluating random forests for survival analysis using prediction error curves. J Stat Softw 2012;50:1.