Machine learning in the integration of simple variables for identifying patients with myocardial ischemia

Luis Eduardo Juarez-Orozco, MD, PhD,[a] Remco J.J. Knol, MD, PhD,[b] Carlos A. Sanchez-Catasus, PhD,[c] Octavio Martinez-Manzanera, PhD,[d] Friso M. van der Zant, MD, PhD,[b] and Juhani Knuuti, MD, PhD[a]

[a] Turku PET Centre, University of Turku and Turku University Hospital, Turku, Finland
[b] Cardiac Imaging Division Alkmaar, Department of Nuclear Medicine, Northwest Clinics, Alkmaar, The Netherlands
[c] Nuclear Medicine and Molecular Imaging, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands
[d] Department of Neurology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

Received Feb 26, 2018; accepted May 7, 2018
doi:10.1007/s12350-018-1304-x

Background. **A significant number of variables are obtained when characterizing patients suspected with myocardial ischemia or at risk of MACE. Guidelines typically use a handful of them to support further workup or therapeutic decisions. However, it is likely that the numerous available predictors maintain intrinsic complex interrelations. Machine learning (ML) offers the possibility to elucidate complex patterns within data to optimize individual patient classification. We evaluated the feasibility and performance of ML in utilizing simple accessible clinical and functional variables for the identification of patients with ischemia or an elevated risk of MACE as determined through quantitative PET myocardial perfusion reserve (MPR).**

Methods. **1,234 patients referred to Nitrogen-13 ammonia PET were analyzed. Demographic (4), clinical (8), and functional variables (9) were retrieved and input into a cross-validated ML workflow consisting of feature selection and modeling. Two PET-defined outcome variables were operationalized: (1) any myocardial ischemia (regional MPR < 2.0) and (2) an elevated risk of MACE (global MPR < 2.0). ROC curves were used to evaluate ML performance.**

Results. **16 features were included for boosted ensemble ML. ML achieved an AUC of 0.72 and 0.71 in identifying patients with myocardial ischemia and with an elevated risk of MACE, respectively. ML performance was superior to logistic regression when the latter used the ESC guidelines risk models variables for both PET-defined labels ($P < .001$ and $P = .01$, respectively).**

Conclusions. **ML is feasible and applicable in the evaluation and utilization of simple and accessible predictors for the identification of patients who will present myocardial ischemia and an elevated risk of MACE in quantitative PET imaging. (J Nucl Cardiol 2020;27:147–55.)**

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12350-018-1304-x) 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 the contents of the paper and is free for re-use at meetings and presentations. Search for the article DOI on SpringerLink.com.

Reprint requests: Luis Eduardo Juarez-Orozco, MD, PhD, Turku PET Centre, University of Turku and Turku University Hospital, Kiinamyllynkatu 4-8, 20520 Turku, Finland; *l.e.juarez.orozco@gmail.com*

1071-3581/$34.00

Key Words: Machine learning • myocardial ischemia • risk of MACE • PET

Abbreviations

BP	Blood pressure
CAD	Coronary artery disease
CV	Cardiovascular
HR	Heart rate
LVEF	Left ventricular ejection fraction
MACE	Major adverse cardiovascular events
MBF	Myocardial blood flow
ML	Machine learning
MPR	Myocardial perfusion reserve
PET	Positron emission tomography

See related editorial, pp. 156–158

INTRODUCTION

Currently, a vast number of variables are routinely obtained during the characterization of patients with suspected myocardial ischemia due to coronary artery disease (CAD). In daily clinical practice, a simultaneous process of thought occurs within the mind of clinicians during the evaluation of the available data in order to integrate both a diagnostic and prognostic probability statement for any given patient. To help this process, some clinical variables are considered (in scores implemented, for instance, in the current ESC guidelines[1–3]) in order to provide an estimate of the probability of having significant CAD or a certain risk of major cardiovascular (CV) events (MACE). As such, demographic and clinical characteristics can guide the selection of a non-invasive diagnostic technique[2] on the basis of their performance and availability. Ultimately, several techniques are commonly performed through the diagnostic ''path'' (e.g., a stress ECG followed by echocardiography, coronary computed tomography angiography [CCTA], or nuclear imaging), which delivers additional variables that complement the patients' pathologic profile.

Within this growing body of information, it is clear that many variables are interrelated to a lesser or greater extent. For instance, the presence or absence of type 2 diabetes mellitus in a female patient will correlate with a higher left ventricular ejection fraction (LVEF) as well as with body mass index (BMI), the presence of dyslipidemia, and other components of the metabolic syndrome.[4] These intrinsic correlations, both between predictors and between predictors and outcomes, can be complex (especially as the number of variables increases) and are usually overlooked by the traditional assumptions of linearity and predictor independence in linear models.

Machine learning (ML) constitutes a rapidly growing field of computer science which offers the possibility to elucidate and ''learn'' from patterns in large sets of data. ML approaches are able to define models from existing data by exploring and combining predictors in a non-linear and ''highly interactive'' fashion.[5] The resulting models can be applied then to new instances of unseen data in order to perform and optimize classification and prediction.[6,7]

On the other hand, from the available diagnostic techniques for cardiovascular disease, myocardial perfusion quantification through positron emission tomography imaging (PET) provides the best diagnostic performance in the detection of significant CAD as assessed by contemporary invasive coronary angiography standards.[8] Furthermore, PET has proven very good performance in the prediction of MACEs.[9,10] As such, quantitative PET perfusion imaging currently represents the strongest proxy for the presence of ischemia-causing CAD and the consequent risk of MACEs with the inherent advantage of being non-invasive.

Recently, it has been proposed that ML methods may improve detection of significant CAD and early revascularization through integration of clinical and SPECT data.[11,12] Likewise, it has demonstrated to improve all-cause mortality risk evaluation based on a wide range of CCTA predictors.[7] However, no evaluation has yet attempted to apply ML to numerous readily available (and interrelated) predictor variables for the identification of patients who will demonstrate myocardial ischemia and its associated CV risk as determined through quantitative PET myocardial perfusion.

Hence, we aimed to evaluate the feasibility and performance of ML in evaluating and utilizing numerous simple and readily accessible clinical and functional variables in order to identify patients who will demonstrate any significant regional myocardial ischemia or an elevated risk of MACEs (as demonstrated through quantitative Nitrogen-13 ammonia PET). Additionally, we explored how ML performance compared to that of a traditional logistic regression approach. For such purpose, we utilized the predictor variables stated in two widely used clinical models implemented in the current ESC Guidelines[2,3] namely, the Genders' model for obstructive CAD[1,2] and the Systematic Coronary Risk Evaluation (SCORE) model for risk of fatal CV events.[3]

METHODS

Study Population

One thousand two hundred and thirty-four patients were included in the analysis. Patients were retrospectively selected from the population referred to PET myocardial perfusion

imaging due to suspected myocardial ischemia at the department of nuclear medicine of the Northwest Clinics in the Alkmaar Medical Center, The Netherlands, up until December 2016. Patients with documented CAD, either as prior myocardial infarction (MI) or revascularization (PCI or CABG), were excluded from the present study.

The study was performed in accordance with the Declaration of Helsinki and all patients provided written informed consent for usage of their anonymous data.

Clinical and Diagnostic Data

Demographic (sex, age, BMI, and family history of CAD), clinical (type of chest complaints [typical angina, atypical angina or non-anginal], type 2 diabetes mellitus, smoking, dyslipidemia, hypertension, rest heart rate [HR], and resting systolic and diastolic blood pressure [BP]), complementary diagnostic data (abnormal rest ECG, abnormal stress ECG, stress HR, % of max HR, stress systolic and diastolic BP, rest LVEF, and stress LVEF), and the Duke clinical score were available and retrieved from the patient electronic file system.

Patent ECG-abnormalities were considered present when descriptive values reported a departure from normal voltage, duration and/or pattern ranges (and their changes) for waves, segments, and intervals.

PET Data Acquisition and Quantitative Perfusion Analysis

Every patient underwent a two-phase (rest and adenosine stress) PET scan with the use of Nitrogen-13 ammonia as the perfusion radiotracer. All image data were acquired in list mode on a Siemens Biograph-16 TruePoint PET/CT (Siemens Healthcare, Knoxville, USA) with the TrueV option (the axial field of view of 21.6 cm). This 3D system consists of a 16-slice CT and a PET scanner with four rings of lutetium oxyorthosilicate (LSO) detectors. Patients were instructed to fast overnight and to avoid the consumption of methylxanthines, caffeine-containing beverages, or medications for 24 hours before the study. The details of the acquisition-reconstruction protocol have been previously described.[13]

Based on the dynamic subsets, left ventricular contours were assigned automatically using the Syngo MBF software (Siemens Medical Solutions, Berlin, Germany) with minimum observer intervention when appropriate. With a previously described 2-compartment kinetic model for Nitrogen-13 ammonia, stress MBF and rest MBF values in mL/g/minute were computed for each sample on the polar map through the resulting time-activity curves for quantification.[14]

Myocardial perfusion reserve (MPR) was calculated as the ratio between stress and rest MBF. MPR was calculated regionally according to the main vascular territories (left anterior descending [LAD], left circumflex [LCx], and right coronary artery [RCA]) and within the whole left ventricular region. These measurements were utilized to label patients for ML as (1) having myocardial ischemia when at least one vascular territory demonstrated a regional MPR < 2.0[15] and (2) having an elevated risk of MACE when the global MPR resulted to be < 2.0.[9]

Probability of Ischemia-Causing CAD

The model implemented in the 2013 ESC Guidelines on the management of stable CAD[1,2] considers age, sex, and type of chest pain in order to estimate the probability of obstructive CAD. Based on the relevance of these three variables, we utilized them to estimate how they would perform in our patient sample using traditional logistic regression. We then adopted this estimation as reference to evaluate the ML workflow's performance in identifying patients with any regional myocardial ischemia (PET MPR < 2.0 in at least 1 main vascular territory) when utilizing all previously described readily available variables (see ''Clinical and Diagnostic Data'').

Risk of MACE

The SCORE risk model, implemented in the 2012 European Guidelines on CV disease prevention in clinical practice,[3] considers age, sex, smoking, systolic blood pressure, and cholesterol in the estimation of the 10-year risk of fatal CV disease events. Because a reduced PET-measured MPR has shown to be a strong and independent predictor for MACE,[16] we utilized the aforementioned MPR threshold to label the outcome variable. Once again, we utilized the five SCORE variables to explore their basal performance on our sample using traditional logistic regression as reference. Thereon, we compared the performance of the ML workflow in identifying patients at an elevated risk of MACE (i.e., a global PET MPR < 2.0) utilizing all previously described readily available variables (see ''Clinical and Diagnostic Data'').

Machine Learning

In total, twenty-one readily available variables were considered and an ML workflow was created consisting of feature selection and posterior modeling, using a k(10)-fold stratified cross-validation (see ''Re-sampling'').[7] A visual summary is presented in Figure 1.

Feature selection. The selection of attributes (variables) was conducted through *information gain attribute ranking*, which gages the efficacy of the given variables in case classification during training iterations.[17] Variables that documented an information gain $\neq 0$ were subsequently used for the ML model.[18]

Model construction. The selected attributes for the identification of patients with myocardial ischemia and then, of patients at an elevated risk of MACE, were analyzed using ensemble boosting with *LogitBoost* (using decision stumps and random forests).[19] Briefly, *LogitBoost* combines weak base classifiers and iteratively adjusts their relevance based on misclassification in order to create a single strong classifier. The resulting classifier scores are then expressed as a pseudo-probability of the outcome category and binarized for

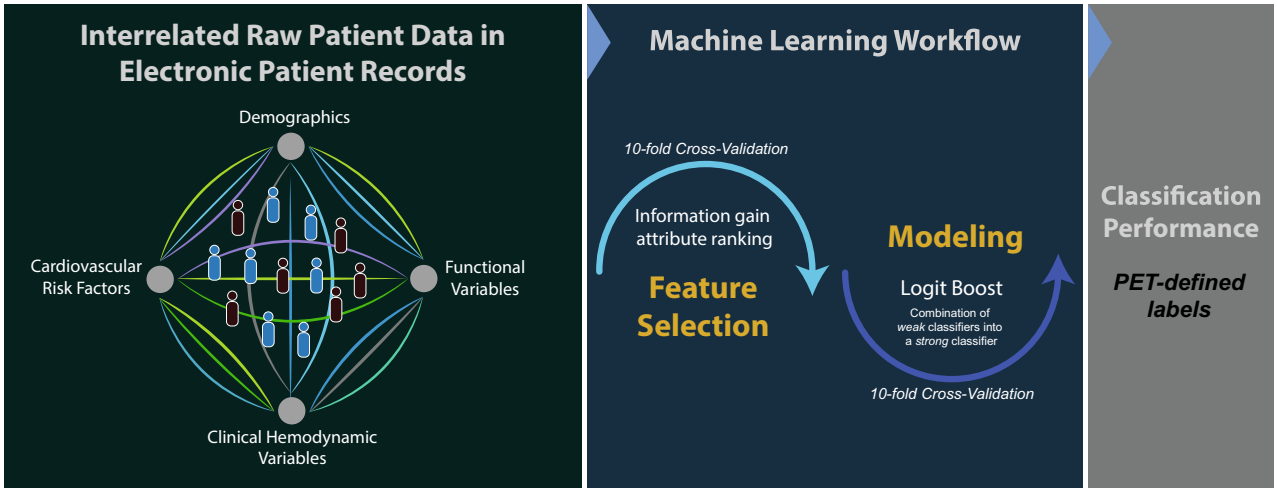

Figure 1. Schematic depiction of the study progression, ML workflow, and its components.

practical classification. Additionally, the performance of previous commonly reported algorithms was explored for comparison, namely, Naïve Bayes,[20] Random Tree[21], and Support Vector Machine (SVM).[22]

Re-sampling. Performance and error were evaluated through stratified 10-fold cross-validation to obtain stable and accurate estimates of overall performance.[23] Instead of a previously utilized method of randomly splitting data into training and testing sets, cross-validation utilizes all data ''*k-times*'' for both purposes as depicted in the Online Supplementary Resource e-Figure 1.

Exploration of the ML workflow Generalization. In order to explore the practical ML generalization adequacy, an unseen (hold-out) validation set of the latest 60 patients referred for PET imaging for suspected CAD in our center was evaluated with the resulting ML models for both outcome variables.

Statistical Analysis

Categorical variables are presented as frequencies, and continuous variables are expressed as mean ± SD. Areas under the curve (AUC) were used to report the ML algorithms' performance and were compared to that of logistic regression using the predictor variables contemplated by the models implemented in the ESC Guidelines through ROC curves analysis and pairwise comparisons according to the method reported by Delong.[24] ML performance was further compared with that of traditional logistic regression with the use of the ML-selected features. The supplementary AUCs, corresponding standard errors, and pairwise comparisons were documented for alternative methods.

A P-value $< .05$ was considered significant. Statistical analyses were performed with SPSS (IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 23.0. Armonk, NY) and MedCalc Statistical Software version 18.2.1 (MedCalc Software bvba, Ostend, Belgium). For the purposes of ML, we utilized the Waikato Environment for Knowledge Analysis (WEKA) open software, version 3.8.1.[25]

RESULTS

Study Population

The baseline characteristics of the population are shown in Table 1. There were 519 patients with myocardial ischemia (i.e., in at least one main coronary territory), while 387 patients had an increased risk of MACE (i.e., a global PET MPR < 2.0).

Machine Learning Workflow Performance

Feature selection. ML extracted 16 out of the 21 readily available variables for further modeling. Rest HR, systolic blood pressure, stress LVEF, and age showed the highest rank within the attributes for modeling the identification of both myocardial ischemia and an elevated risk of MACE. The full ML ranking and selection of variables for each considered outcome variable is depicted in Figure 2.

Identification of patients with myocardial ischemia. The ML algorithm (*LogitBoost*) yielded the best performance achieving an AUC of 0.72, which was significantly higher ($P < .001$) in comparison with the traditional logistic regression model utilizing the predictor variables of the model implemented in the ESC Guidelines[1,2] (AUC = 0.61), in the identification of patients with PET-documented myocardial ischemia. The ROC comparison is shown in Figure 3, Panel A. The analysis considering the sixteen ML-selected attributes was repeated with traditional logistic regression and documented a significantly lower AUC = 0.67

Table 1. Baseline characteristics across all attributes (predictor variables), *n* = 1,234

Variable	Summary statistic
Age (years)	67 ± 10
Female sex	694 (56%)
BMI (kg/m^2)	28 ± 5
Smoking	194 (16%)
Diabetes	157 (13%)
Dyslipidemia	420 (34%)
Hypertension	681 (55%)
Family history of CAD	405 (33%)
Duke clinical score (0–100)	43 ± 29
Rest HR (bpm)	68 ± 14
Stress HR (bpm)	97 ± 18
% of max HR	64 ± 11
Rest Systolic BP (mmHg)	136 ± 22
Rest diastolic BP (mmHg)	72 ± 11
Stress systolic BP (mmHg)	131 ± 20
Stress diastolic BP (mmHg)	68 ± 12
Abnormal rest ECG	444 (36%)
Abnormal stress ECG	
Inconclusive	632 (51%)
Abnormal	19 (2%)
Type of chest complaints[a]	
No angina	331 (27%)
Typical	241 (20%)
Atypical	662 (54%)
Rest LVEF (%)	67 ± 11
Stress LVEF (%)	68 ± 11

BMI, body mass index; *BP*, blood pressure; *CAD*, coronary artery disease; *HR*, heart rate; *LVEF*, left ventricular ejection fraction
[a]Complaints were also evaluated related to vasodilatory stimulus

as compared to ML ($P < .01$). The ESC Guidelines' model predictors (3 variables) were also analyzed using the *LogitBoost* algorithm and documented a lower predictive performance (AUC = 0.58).

Identification of patients at risk of MACE. ML modeling (in the form of *LogitBoost*) demonstrated an AUC = 0.71 in the identification of patients at an elevated risk of MACE, as evaluated by PET. Its performance was superior ($P < .01$) to that of the logistic regression model that utilized the predictors of the SCORE risk model (five variables) (AUC = 0.64). The analysis utilizing the sixteen selected attributes was repeated with traditional logistic regression and documented an AUC = 0.70, comparable to ML ($P = .08$). ROC curves are shown in Figure 3, panel B. Finally, the SCORE risk model variables were also analyzed using *LogitBoost* and also documented a decrease in predictive performance (AUC = 0.62).

The AUC-analysis results for the rest of the ML algorithms explored and traditional logistic regression can be found in the Online Supplementary Resource e-Table 1 and e-Figure 2. Overall, the performance of *LogitBoost* was found to be better overall and quantitatively comparable to that of *SVM* for the second outcome variable. Other methods were alternatively suboptimal for the proposed task.

Exploratory ML generalization. The results of the small exploratory evaluation of the ML models generalization performance tracked those of the modeling results for both the identification of patients with myocardial ischemia and an elevated risk of MACE (AUC = 0.75 for both outcome variables).

DISCUSSION

The present study has demonstrated that ML is feasible and can be useful in the evaluation and selection of numerous predictors, and in the identification of patients with any regional myocardial ischemia or an elevated risk of MACE (based on a regionally or globally hampered PET MPR, respectively). Furthermore, when comparing performance to that of a traditional regression approach employing the variables of the model implemented in the ESC Guidelines on the management of stable CAD[1,2] and the SCORE risk model variables, ML documented an improvement in the identification of patients with myocardial ischemia (functionally significant CAD) and at an elevated risk of MACE.

Currently, refining the identification of patients with ischemia-causing CAD as well as those at risk of potentially fatal CV outcomes is a central need in the practice of clinical cardiology. State-of-the-art algorithms, such as those implicated in ML, offer a novel approach to refine and boost our estimations both in the diagnostic and the prognostic clinical process through deep integration of numerous (and often times complexly interrelated) predictors,[6] both weak and strong.

A handful of cardiovascular risk factors have been identified and utilized in the creation of models in order to predict the presence of disease and its clinical prognosis (e.g., the ones implemented in the 2013 ESC Guidelines on the management of stable CAD[1,2] and the 2012 ESC Guidelines on CV disease prevention[3]). Such models make use of a discrete number of variables (3 and 5, respectively) because they aim to offer simplicity and ease-of-use in general clinical guidance and because of the reductionist nature of the process of evaluation and their comparative importance against other predictors. Therefore, it is reasonable to explore alternative

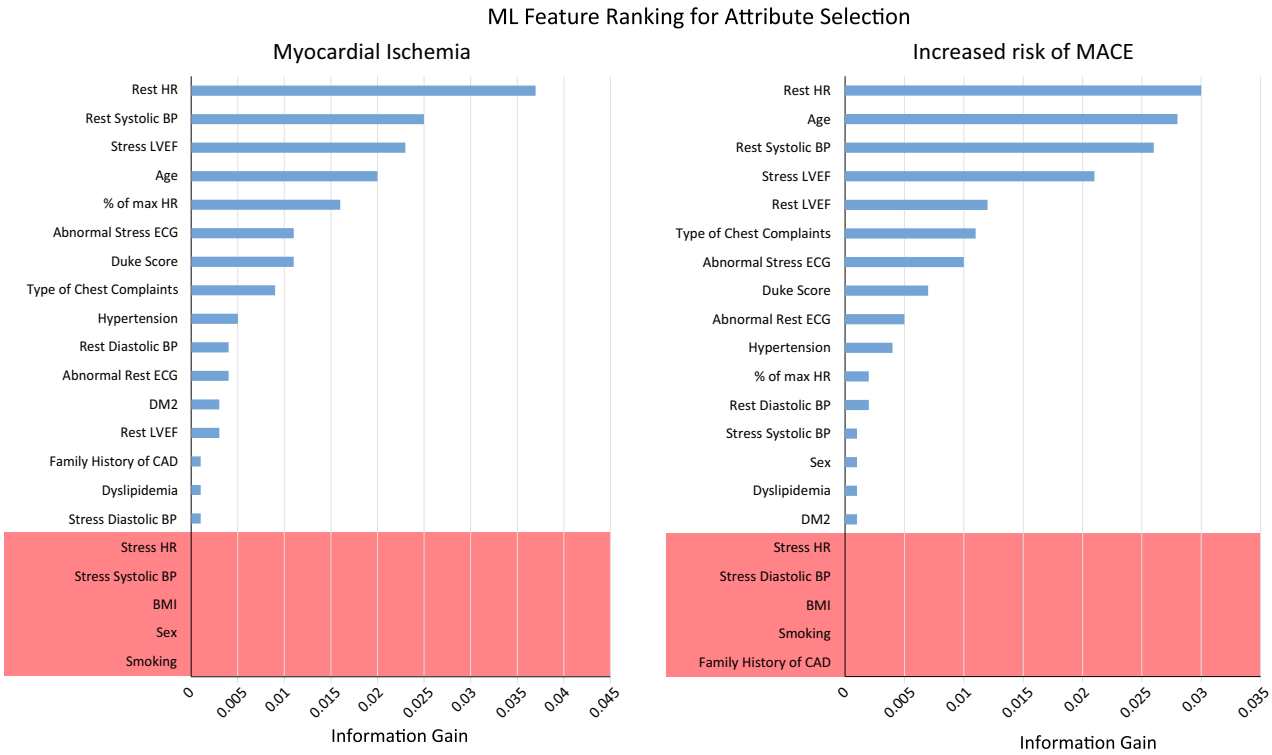

Figure 2. Information gain by attribute (predictor variable) in descending order for the considered outcome variables: myocardial ischemia (left) and an elevated risk of MACE (right). The variables that documented no information gain were excluded from the ML modeling (contained in the red area).

methods when encountering a large number of potentially interrelated variables that may further characterize the (patho) physiological status of individual patients. It is within this setting that ML can provide a useful alternative in predictor selection and patient classification.[7] Our results show that an ML workflow, in the form of feature selection and posterior application of the *LogitBoost* algorithm specifically, was feasible for the identification of patients with either of the two selected outcome variables. More discrete results yielded by alternative ML algorithms are reported in the supplementary material conveying that method exploration should be encouraged along with method fine-tuning in order to harness the potential benefits of ML in clinical research.

In order to evaluate our ML results contextually, we referred to the variables utilized in the models implemented in current ESC Guidelines, which are widely used in the standard evaluation of patients with suspected CAD. Even though the dependent variables we defined from quantitative cardiac PET measurements are similar to those used to validate the aforementioned models (i.e., any PET-documented regional myocardial ischemia is highly related to any anatomically obstructive CAD [$\geq$ 50% luminal narrowing], while an elevated risk of MACE according to PET will very likely relate to the 10-year risk of fatal CV outcomes), we underline that we did not equate these outcome variables to ours. Rather, we utilized the predictors employed by these validated models to test and compare their performance in our sample to that of our ML workflow because, by definition, such variables should offer the strongest one-dimensional surrogates of significant CAD and MACE risk available.

From the 21 variables included, sixteen where selected through ML feature selection. Within these, we imported variables derived from common diagnostics undertaken previous to PET imaging. For instance, patient records included the evaluation of rest and stress ECG and their results were operationalized according to their interpretation (positive, negative, or inconclusive for abnormalities). Similarly, measurements of LVEF can generally be obtained from echocardiography although inter-observer variability should always be considered. As such, ML was feasible within the continuum of the characterization process in our institution in which clinical variables are systematically registered along with simple diagnostics in order to conform the body of information available for each individual patient at the moment of making crucial

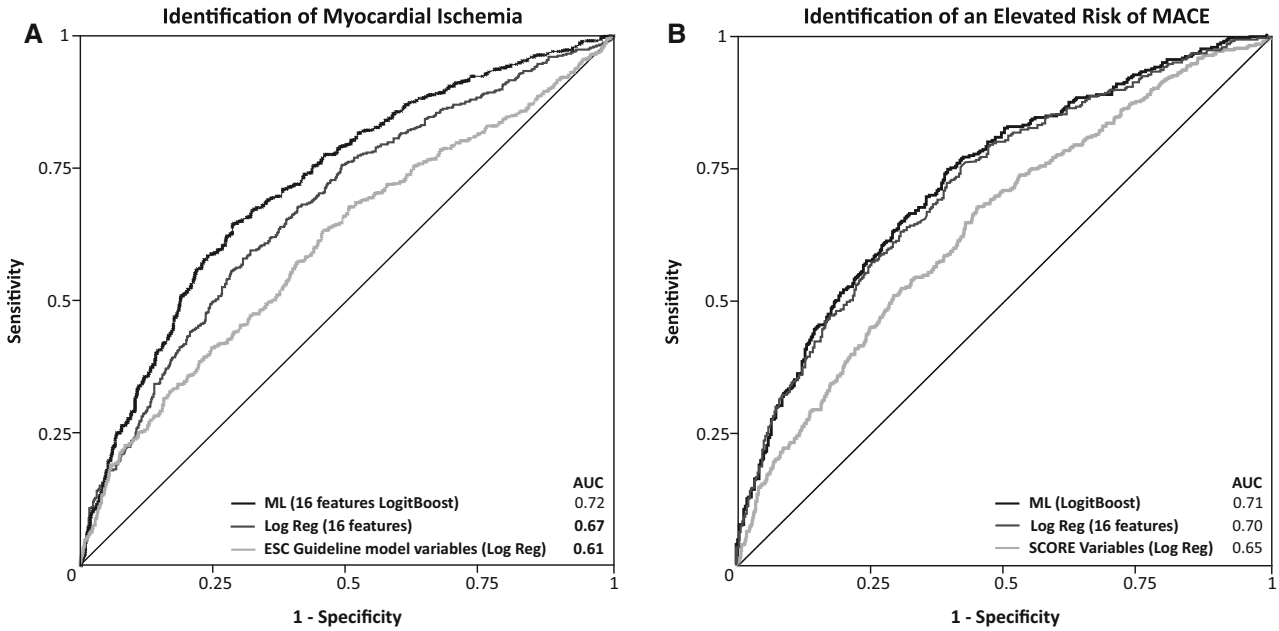

Figure 3. Receiver-operating characteristic curves (**A**) for the identification of patients with myocardial ischemia (regional PET MPR < 2.0 in at least one of the main vascular territories [LAD, LCx, and RCA]) with ML (*LogitBoost*) as compared to logistic regression using ML-selected features and logistic regression using the variables of model implemented in the ESC Guidelines on the management of stable CAD, and (**B**) for the identification of patients with an elevated risk of MACE (global PET MPR < 2.0) as compared to logistic regression using ML-selected features and logistic regression using the variables of the SCORE risk model implemented in the European Guidelines on CV disease prevention in clinical practice.

therapeutic decisions. Therefore, we support the notion that ML can be applied in any context where systematic recording of patient-pertinent information is conducted.

Our results show that the overall performance of ML was moderate (with AUCs in the range from 0.71 to 0.75) and therefore, it probably cannot yet replace diagnostic or risk estimations that further workup is able to provide. Nevertheless, when results were compared to those of utilizing the sets of variables considered in the ESC Guidelines models, ML performed better than the former and comparably to the latter. Remarkably, when ML was further applied to analyze only the constrained variables described in those same models (3 and 5, respectively), classification performance decreased. Furthermore, when the ML-selected features (16 in both models) were analyzed through a traditional logistic regression approach, performance improved but did not reach that of *LogitBoost* for the identification of patients with myocardial ischemia, while it also improved (to a level that was comparable to ML) in detecting patients at elevated risk of MACE. The implication of such results is two-fold: first, this supports the notion that ML may outperform traditional approaches when enough data and numerous interrelated covariates are to be analyzed for a complex particular problem; secondly, this suggests that ML-based methods can also aid through *feature selection* in the identification of the most relevant predictors (evaluated at a higher dimensional level) that can even improve the performance of traditional approaches (such as logistic regression) when simpler models are sufficient.

A notable feature of the present study is that the outcome variables were determined from myocardial perfusion quantification in absolute terms with Nitrogen-13 ammonia PET, which currently represents the reference standard for perfusion evaluation[26] as well as the best-performing non-invasive technique in the identification of functionally significant CAD.[8] As such, we believe that our operationalized outcome variables are of substantial clinical value even if they do not directly correspond to those used in standard diagnostic models (like luminal stenosis assessed with invasive coronary angiography). To the best of our knowledge, this constitutes the first large-scale (> 1,000 patients) report on the performance of ML using numerous readily available variables to predict and identify labels defined by the results of absolute quantitative PET myocardial perfusion data. As an additional value, we implemented the resulting ML models in a small sample of the most recent consecutive patients (without previous MI or

revascularization) referred to PET myocardial perfusion, and although it was not intended for extensive validation of the model, the performance obtained suggests adequate generalization of the ML modeling.

A valuable insight is that readily available variables (especially if numerous) can make use of novel ML approaches to identify probable results of performing full quantitative cardiac PET imaging. However, this benefit will probably maximize when further studies address more refined classifications. Our study only sets an initial pragmatic base by utilizing relatively coarse categories (any myocardial ischemia vs no ischemia and a binary risk category for MACE). In this sense as well, it was noticed that ML seemed to perform better in identifying patients without myocardial ischemia and without an elevated risk of MACE based PRC areas for negative cases (not shown in the tables). This falls in line with the current performance of available clinical scores and suggests that our ML approach may initially aid in ruling-out the need for PET imaging.

Cardiac PET has been restricted by its accessibility, cost, and mandatory infrastructure, and this has been a major challenge for its widespread implementation. Therefore, we could only speculate that, as ML performance is able to progressively improve through continuous data incorporation,[27] it may eventually help focalize the need of such a complex technique by selecting those patients in which performing myocardial perfusion with PET will effectively have incremental value by characterizing ischemic burden and location.

The present study naturally carries the limitations of any observational study. However, this kind of large-scale retrospective analysis is a main target of the data-driven approaches of ML. Also, some of the predictor variables were operationalized slightly differently from their operationalization in the original models taken as reference. This may not be an issue given that only the concepts expressed by the predictors in the reference models (contained in the ESC Guidelines[1–3]) were extracted for our ML analysis and we did not aim to evaluate these previous models per se. Finally, the simultaneous handling of numerous predictors can be challenging and overwhelming in the minds of busy physicians. Despite their potential, ML models can be difficult to interpret. This caveat should be considered during their implementation, especially in clinical evaluation applications. Nevertheless, ML could represent a great supplement in the optimization of diagnostic and prognostic estimations (i.e., precision medicine) and even clinical alerts, specially integrated in a background of constantly updated electronic patient records across all levels of care.[7] This continuous flow of information characterizes the situations where ML has exploded in its practical applications, ranging from Google's search engine, improving speech or image recognition, and all the way to self-driving cars,[28] which are constantly updated by a large integrated network. We believe, therefore, that a similar integration could be featured as a powerful added clinical tool as more efforts are currently directed to process automation[29] in medical data managing and storage.

NEW KNOWLEDGE GAINED

ML learning seems to be useful both in the identification of predictor variables and in the integration of such predictors through modeling for the identification of patients who ultimately will demonstrate abnormal perfusion and a higher risk of MACE. ML can outperform simpler models probably due to its capacity in exploration of higher dimensional interrelations between numerous predictors. ML could aid in improving patient selection for PET myocardial perfusion imaging.

CONCLUSIONS

ML is feasible and applicable in the identification and utilization of simple and readily accessible predictors for the identification of patients who will present myocardial ischemia and an elevated risk of MACE as determined through quantitative PET myocardial perfusion imaging. The implementation of ML in the optimization of diagnostic technique selection, cardiovascular diagnosis, and prognostic estimations warrants further research.

Disclosures

Dr. Juarez-Orozco has no relevant disclosures, Dr. Knol has no relevant disclosures, Dr. Sanchez-Catasus has no relevant disclosures, Dr. Martinez-Manzanera has no relevant disclosures, and Dr. van der Zant has no relevant disclosures. Dr. Knuuti reports personal fees from AstraZeneca (modest relationship), outside the submitted work. None of the other authors have relevant disclosures.

References

1. Genders TSS, Steyerberg EW, Alkadhi H, Leschka S, Desbiolles L, Nieman K, et al. A clinical prediction rule for the diagnosis of coronary artery disease: validation, updating, and extension. Eur Heart J 2011;32:1316-30.
2. Task Force Members, Montalescot G, Sechtem U, Achenbach S, Andreotti F, Arden C, et al. 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. Eur Heart J 2013;2013:2949-3003.

3. Perk J, De Backer G, Gohlke H, Graham I, Reiner Z, Verschuren M, et al. European Guidelines on cardiovascular disease prevention in clinical practice (version 2012). The Fifth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of nine societies and by invited experts). Eur Heart J 2012;33:1635-701.
4. Mottillo S, Filion KB, Genest J, Joseph L, Pilote L, Poirier P, et al. The metabolic syndrome and cardiovascular risk a systematic review and meta-analysis. J Am Coll Cardiol 2010;56:1113-32.
5. Obermeyer Z, Emanuel EJ. Predicting the future—Big data, machine learning, and clinical medicine. N Engl J Med 2016; 375:1216-9.
6. Deo RC. Machine learning in medicine. Circulation 2015;132: 1920-30.
7. Motwani M, Dey D, Berman DS, Germano G, Achenbach S, Al-Mallah MH, et al. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: A 5-year multicentre prospective registry analysis. Eur Heart J 2016;52:468-76.
8. Danad I, Raijmakers PG, Driessen RS, Leipsic J, Raju R, Naoum C, et al. Comparison of coronary CT angiography, SPECT, PET, and hybrid imaging for diagnosis of ischemic heart disease determined by fractional flow reserve. JAMA Cardiol 2017;2:1100-7.
9. Herzog BA, Husmann L, Valenta I, Gaemperli O, Siegrist PT, Tay FM, et al. Long-term prognostic value of 13N-ammonia myocardial perfusion positron emission tomography added value of coronary flow reserve. J Am Coll Cardiol 2009;54:150-6.
10. Ziadi MC, DeKemp RA, Williams KA, Guo A, Chow BJW, Renaud JM, et al. Impaired myocardial flow reserve on Rubidium-82 positron emission tomography imaging predicts adverse outcomes in patients assessed for myocardial ischemia. J Am Coll Cardiol 2011;58:740-8.
11. Arsanjani R, Xu Y, Dey D, Vahistha V, Shalev A, Nakanishi R, et al. Improved accuracy of myocardial perfusion SPECT for detection of coronary artery disease by machine learning in a large population. J Nucl Cardiol 2013;20:553-62.
12. Arsanjani R, Dey D, Khachatryan T, Shalev A, Hayes SW, Fish M, et al. Prediction of revascularization after myocardial perfusion SPECT by machine learning in a large population. J Nucl Cardiol 2015;22:877-84.
13. Juárez-Orozco LE, van der Zant FM, Slart RHJA, Lazarenko SV, Alexanderson E, Tio RA, et al. Type 2 diabetes mellitus correlates with systolic function during myocardial stress perfusion scanning with Nitrogen-13 ammonia PET. J Nucl Cardiol 2017;24:1305-11.
14. Hutchins GD, Schwaiger M, Rosenspire KC, Krivokapich J, Schelbert H, Kuhl DE. Noninvasive quantification of regional blood flow in the human heart using N-13 ammonia and dynamic positron emission tomographic imaging. J Am Coll Cardiol 1990;15:1032-42.
15. Dey D, Diaz Zamudio M, Schuhbaeck A, Juarez Orozco LELE, Otaki Y, Gransar H, et al. Relationship between quantitative adverse plaque features from coronary computed tomography angiography and downstream impaired myocardial flow reserve by 13N-ammonia positron emission tomography: A pilot study. Circ Cardiovasc Imaging. 2015;8:e003255.
16. Juárez-Orozco LE, Tio RA, Alexanderson E, Dweck M, Vliegenthart R, El Moumni M, et al. Quantitative myocardial perfusion evaluation with positron emission tomography and the risk of cardiovascular events in patients with coronary artery disease: a systematic review of prognostic studies. Eur Heart J Cardiovasc Imaging. 2017. https://doi.org/10.1093/ehjci/jex331.
17. Hall MA, Holmes G. Benchmarking attribute selection techniques for discrete class data mining. IEEE Trans Knowl Data Eng 2003;15:1437-47.
18. Guyon I, Elisseeff A. An introduction to variable and feature selection. J Mach Learn Res 2003;3:1157-82.
19. Friedman J, Hastie T, Tibshirani R. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). Ann Stat 2000;28:337-407.
20. Miranda E, Irwansyah E, Amelga AY, Maribondang MM, Salim M. Detection of cardiovascular disease risk's level for adults using naive Bayes classifier. Healthc Inform Res 2016;22:196-205.
21. Breiman L. Random forests. Mach Learn 2001;45:5-32.
22. Vapnik V, Golowich S, Smola A. Support vector method for function approximation, regression estimation, and signal processing. In: Proceedings of the Advances in Neural Information Processing Systems 9. Cambridge: MIT Press; 1996. p. 281-87.
23. Molinaro AM, Simon R, Pfeiffer RM. Prediction error estimation: A comparison of resampling methods. Bioinformatics 2005;21:3301-7.
24. 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.
25. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The WEKA data mining software. SIGKDD Explor Newsl. 2009;11:10.
26. Qayyum AA, Kastrup J. Measuring myocardial perfusion: The role of PET, MRI and CT. Clin Radiol 2015;70:576-84.
27. Waljee AK, Higgins PDR. Machine learning in medicine: A primer for physicians. Am J Gastroenterol 2010;105:1224-6.
28. Hadsell R, Sermanet P, Ben J, Erkan A, Scoffier M, Kavukcuoglu K, et al. Learning long-range vision for autonomous off-road driving. J Field Robot 2009;26:120-44.
29. Slomka PJ, Dey D, Sitek A, Motwani M, Berman DS, Germano G. Cardiac imaging: working towards fully-automated machine analysis & interpretation. Expert Rev Med Devices 2017;14:197-212.