International Journal of Cardiology 335 (2021) 130–136

Contents lists available at ScienceDirect

International Journal of Cardiology

journal homepage: www.elsevier.com/locate/ijcard

Improving patient identification for advanced cardiac imaging through machine learning-integration of clinical and coronary CT angiography data

Jan Walter Benjamins [a], Ming Wai Yeung [a], Teemu Maaniitty [b], Antti Saraste [b], Riku Klén [b], Pim van der Harst [a,c], Juhani Knuuti [b], Luis Eduardo Juarez-Orozco [a,b,c,*]

[a] University of Groningen, University Medical Center Groningen, Department of Cardiology, Hanzeplein 1, 9700RB Groningen, the Netherlands
[b] Turku PET Centre, University of Turku and Turku University Hospital, Kiinamyllynkatu 4-8, 20520 Turku, Finland
[c] Department of Cardiology, Hart and Lung Division, University Medical Centre Utrecht, Heidelberglaan 100, 3508, GA, Utrecht, the Netherlands

ARTICLE INFO

Article history:
Received 10 January 2021
Received in revised form 4 March 2021
Accepted 2 April 2021

Available online 6 April 2021

Keywords:
Computed tomography angiography
Coronary artery disease
Myocardial ischemia
Machine learning
Feature integration
Positron emission tomography

ABSTRACT

Background: Standard computed tomography angiography (CTA) outputs a myriad of interrelated variables in the evaluation of suspected coronary artery disease (CAD). But an important proportion of obstructive lesions does not cause significant myocardial ischemia. Nowadays, machine learning (ML) allows integration of numerous variables through complex interdependencies that optimize classification and prediction at the individual level. We evaluated ML performance in integrating CTA and clinical variables to identify patients that demonstrate myocardial ischemia through PET and those who ultimately underwent early revascularization.

Methods and results: 830 patients with CTA and selective PET were analyzed. Nine clinical and 58 CTA variables were integrated through ensemble-boosting ML to identify patients with ischemia and those who underwent early revascularization. ML performance was compared against expert CTA interpretation, calcium score and clinical variables.

While ML using all CTA variables achieved an AUC = 0.85, it was outperformed by expert CTA interpretation (AUC = 0.87, $p < 0.01$ for comparison), comparable to ML integration of CTA variables with clinical variables. However, the best performance was achieved by ML integration of expert CTA interpretation and clinical variables for both dependent variables (AUCs = 0.91 and 0.90, $p < 0.001$).

Conclusions: Machine learning integration of diagnostic CTA and clinical data may improve identification of patients with myocardial ischemia and those requiring early revascularization at the individual level. This could potentially aid in sparing the need for subsequent advanced imaging and better identifying patients in ultimate need for revascularization. While ML integrating all CTA variables did not outperform expert CTA interpretation, ML data integration from different sources consistently improves diagnostic performance.

1. Introduction

Coronary artery disease (CAD) represents one of the main causes of mortality worldwide. Non-invasive evaluation through advanced imaging has notably improved our capacity to identify patients with anatomically and functionally obstructive CAD. Recent data has demonstrated that the overall pre-test probability of obstructive CAD both in patients with stable angina and the general population has decreased [1,2]. Therefore, ruling-in obstructive CAD will increasingly require complementary sequential testing that can objectify an ischemic target for invasive treatment.

A significant proportion of anatomically obstructive coronary artery lesions are not functionally significant [3]. It has been proposed that a sequential approach considering initial anatomical evaluation through coronary computed tomography angiography (CTA) and further selective functional assessment through positron emission tomography (PET) myocardial perfusion imaging can improve diagnostic performance by identifying ischemia-causing coronary lesions (plaques) [4,5]. But not every healthcare facility using coronary CTA has effective access to advanced nuclear imaging.

Standard CTA interpretation is currently performed according to SCCT recommendations through systematic evaluation of 17 coronary artery segments [6]. In clinical routine, these segments are evaluated for the presence of atherosclerotic plaques, their degree of coronary

* Corresponding author at: Department of Cardiology, Heart and Lung Division, University Medical Centre Utrecht, Heidelberglaan 100, Utrecht 3584 CX, E03.511, P.O. box 85500, 3508, GA, Utrecht, the Netherlands.
E-mail address: l.e.juarezorozco@umcutrecht.nl (L.E. Juarez-Orozco).

https://doi.org/10.1016/j.ijcard.2021.04.009

luminal narrowing and the presence or absence of calcification. Such evaluation of the entire coronary artery tree is summarized in the resulting clinical report with emphasis on plaques deemed to be significantly obstructive (causing a > 50% diameter luminal narrowing). Nevertheless, formal integration of all resulting variables is only intuitively performed and their integration with clinical variables and risk factors if left to the clinical consideration of the attending physician.

Recently, machine learning (ML) analytics have rapidly permeated into medical imaging research [7,8], offering the possibility to exploit complex dependencies within data to optimize classification and prediction tasks [9]. This capacity to handle numerous interrelated variables [10] serving as weak and strong classifiers renders ML suitable to integrate clinical data from a patient's risk profile and extensive CTA findings at the individual level. The great capability of ML analytics to process data efficiently warrants its implementation to complement and augment standard clinical practice.

Given the additional (although discrete [11]) radiation burden of PET and its exclusive access in advanced imaging centers, it is desirable to refine the identification of patients who may actually benefit from quantitative PET myocardial perfusion scanning, i.e. those who will present findings compatible with myocardial ischemia and that may trigger elective invasive revascularization.

Hence, the aim of this study was to explore the implementation and evaluate the performance of ML, with focus on the integration of various data-sources, namely clinical and extensive CTA characterization data, for the identification of patients that will further demonstrate myocardial ischemia through PET myocardial perfusion imaging and those who ultimately undergo early revascularization, at the individual level (see Graphical Abstract).

2. Methods

2.1. Patient population

We analyzed adult patients at low and intermediate risk of CAD from the Turku hybrid sequential PET/CTA registry from the PET Centre of the Turku University Hospital in Finland. Patients with prior cardiovascular disease, myocardial infarction, patients who underwent revascularization, as well as those suspected with cardiomyopathy were excluded. In total, data from 951 consecutive patients were available for analysis. CTA was performed initially in all patients, and thereon, after a prompt initial evaluation of the scan, the attending clinician indicated if a subsequent PET scan with ^{15}O-H_2O as the perfusion radiotracer was necessary due to a suspected anatomically obstructive CAD. Alternatively, PET imaging was not performed in patients deemed without obstructive CAD and therefore no regional ischemia, or alternatively, with clear high-grade stenosis (>90% of luminal narrowing). A total of 121 patients who had non-interpretable imaging or failure to adhere to the sequential scanning protocol were excluded. Fig. 1 shows the proportions of patients progressing through the imaging protocol.

The present study complies with the Declaration of Helsinki. The Ethics Committee of the Hospital District of Southwest Finland waived the need for written informed consent for retrospective collection of patient data.

2.2. Clinical variables

Available clinical variables extracted from the electronic medical records were sex, age, smoking status, type 2 diabetes mellitus, arterial hypertension, dyslipidemia, family history of CAD, type of chest pain and dyspnea.

2.3. CTA and PET acquisition

The sequential hybrid PET/CT acquisition protocols have been previously described [11]. CTA scans were performed in a 64-row PET/CT scanner (GE Discovery VCT or GE D690, General Electric Medical Systems, Waukesha, Wisconsin). Prior acquisition, 0 to 30 mg of metoprolol was administered intravenously to achieve a target heart rate of less than 60 bpm, and 1.25 mg of isosorbide dinitrate aerosol or alternatively 800 mg of sublingual nitrate (800 mg) were also administered. CTA utilized an intravenously administered low-osmolal iodine contrast agent (48–155 ml at 320–400 mg iodine/ml). Prospectively triggered acquisition was applied as feasible. Global calcium score (CaSc) was quantified and stored as a unique summary variable for utilization as a comparator in the statistical analysis (*see below*).

Selective dynamic quantitative PET myocardial perfusion imaging during stress was performed in the same imaging session, as previously

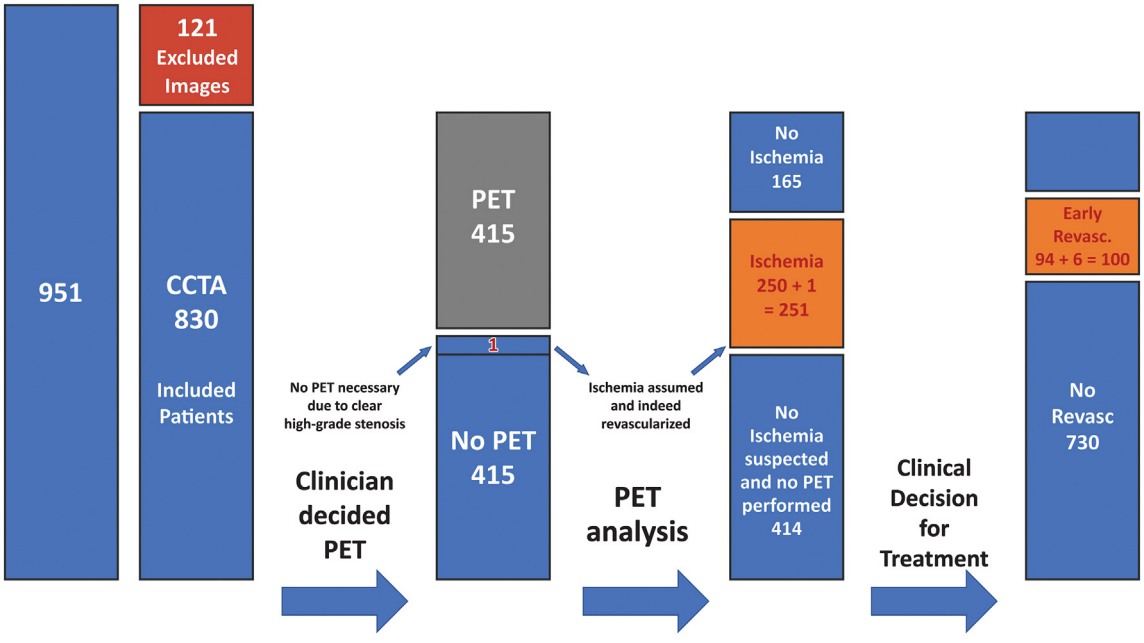

Fig. 1. Sequential hybrid imaging protocol and patient distribution.

described. A mean injected activity of 1042 ± 117 MBq of ^{15}O-H_2O was utilized, while an adenosine infusion at 140 μg/kg/min was used to induce vasodilator stress. All patients were previously instructed to refrain from methyl-xanthines (e.g. coffee, chocolate, tea) for 24 h before the PET study. In a few patients, perfusion imaging was delayed to the following days or weeks due to logistics or xanthine use.

2.4. Image interpretation

CTA data was analyzed according to the 17-segment system [6]. In particular, the coronary artery tree was described in terms of: 1) system dominance [right, left or co-dominance], 2) presence or absence of each coronary segment, 2) presence or absence of atherosclerotic plaque per segment, 3) visually estimated percentage of luminal narrowing [normal: 0%, non-obstructive: <50%, borderline obstructive: 50–69%, obstructive: 70–99% or 100%, prevalence are described in the supplement (table S1)] and 4) complete, partial or absent calcification of the plaque when present. In total, this extensive segmental description of the coronary artery tree delivered 58 variables.

CTA findings were also summarized in the clinical report as demonstrating: normal coronary arteries, non-obstructive CAD, borderline obstructive CAD or anatomically obstructive CAD. This was operationalized the *expert CTA interpretation variable*.

PET data was quantitatively analyzed using Carimas software (developed at the Turku PET Centre, Finland) in the standardized 17-segment AHA model [12]. Using a one-compartment kinetic model, absolute estimates of stress myocardial blood flow (MBF) were obtained per myocardial segment. The finding of at least one segment with a stress MBF of ≤2.4 ml/g/min was considered abnormal and indicative of myocardial ischemia [13].

All image interpretations were performed by experienced physicians and recorded in a standardized reporting system.

2.5. Machine learning

The present workflow was generated considering current state-of-the-art recommendations and following the Proposed Requirements for Cardiovascular Imaging-Related Machine Learning Evaluation (PRIME) Checklist [14]. Information on the aspects covered in this standardized approach can be consulted in the supplement (Table S2).

2.5.1. Data pre-processing and re-sampling

Supervised ML analytics were implemented in the Waikato Environment for Knowledge Analysis (WEKA open-source software, v.3.8.3, Hamilton, New Zealand). Input variables were standardized as per custom (to a value range from 0 to 1). A 10-fold cross-validation policy was applied to both the automated feature selection and later modeling. For the prediction of a positive PET result for myocardial ischemia the entire patient sample was utilized. Conversely, for the prediction of ulterior early revascularization a stratified sub-sampling (due to major class unbalance) was performed to achieve a 3:1 ratio of non-cases vs. cases.

2.5.2. Feature selection

In total, 58 CTA variables and 9 clinical parameters were available for the analyses. The information gain-based method known as Information Gain (Entropy) Ranking (IGR) was utilized for attribute (feature) selection using a criterion of >0.001 in order to consider the variable as potentially contributing to posterior modeling [9,10].

2.5.3. Modeling

We constructed comparative models using different combinations of variables as inputs, including the extensive CTA features, the expert (clinical) CTA interpretation variable, CaSc, and clinical variables. All input variables were integrated through an ensemble boosting algorithm (*LogitBoost*) using decision stumps with a training batch size of 100 and a shrinkage of 1.0. This approach has demonstrated to be robust and has outperformed other approaches in the integration of imaging variables in prediction tasks [9,10,15]. Briefly, LogitBoost utilizes weak base classifiers to iteratively adjust their relevance depending on misclassifications during training to create a compound strong classifier that conveys a pseudo-probability score binarized for the outcome variable of. In this case, the operationalized dependent variables were: 1) myocardial ischemia demonstrated by PET (i.e. patients that benefit from the performance of advanced nuclear imaging with PET), and 2) early (within 6 months) invasive coronary evaluation with revascularization (i.e. patients that warrant invasive evaluation/treatment even in the absence of advanced nuclear imaging).

2.6. Statistical analysis

Continuous variables were presented as mean (standard deviation, SD) if normally distributed or median (interquartile range, IQR) if the distribution was skewed. Categorical variables were expressed as count (percentage). ML performance was evaluated through the area under the curve (AUC) obtained from receiver operating characteristic (ROC) analyses, as well as sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy and F1-score (a metric proper of ML evaluation) for the following models: 1) All CTA variables only, 2) all CTA variables + clinical variables, 3) expert CTA interpretation + clinical variables, 4) global CaSc + clinical variables, and 5) clinical variables only. Further, ML discrimination performance was compared to that of traditional regression analyses considering: 6) expert CTA interpretation only, and 7) global CaSc only. All AUC comparisons were performed with the method proposed by Delong [16]. A two-tailed p-value <0.05 was considered statistically significant. These statistical analyses were performed using MedCalc Statistical Software version 18.2.1 (MedCalc Software bvba, Ostend, Belgium) and SPSS 21.0 (SPSS Inc., Chicago, IL, USA).

3. Results

Out the eight-hundred and thirty patients were analyzed in this study, 415 underwent PET myocardial perfusion imaging as indicated by the attending physician during the hybrid imaging session. From the other 415 patients who did not undergo PET imaging, only one was considered to show clear high-grade stenosis conveying myocardial ischemia and PET was spared. Baseline characteristics of the whole sample through the hybrid imaging acquisitions are shown in Table 1.

3.1. Feature selection

Input variable characterization demonstrated a maximum rate of missing values among all CTA variables of 0.36% and 6% for CaSc, while

Table 1
Baseline characteristics by ischemia.

Variable	All patients (n = 830)	Patients with no ischemia (n = 579)	Patients with ischemia (n = 251)	p-value
Age (years)	61.2(9.5)	60.3(9.9)	63.3(8.3)	**<0.001**
Females (n)	454(54.7%)	370(63.9%)	84(33.5%)	**<0.001**
Smoking (n)	164(23.4%)	98(20.7%)	66(28.9%)	**0.016**
Diabetes status				**0.002**
Pre-diabetes (n)	138(21.2%)	89(20.7%)	49(22.2%)	
Diabetes mellitus (n)	117(18.0%)	64(14.9%)	53(24.0%)	
Arterial hypertension (n)	445(64.2%)	275(59.4%)	170(73.9%)	**<0.001**
Dyslipidemia (n)	515(74.3%)	337(71.5%)	178(80.2%)	**0.015**
Family history of CAD (n)	374(64.2%)	260(66.3%)	114(59.7%)	0.117
Type of chest pain				**<0.001**
Typical	192(24.9%)	106(20.0%)	86(35.7%)	
Atypical/Non-anginal	426(55.2%)	312(58.8%)	114(47.3%)	
Dyspnea	274(57.1%)	185(57.1%)	89(57.1%)	0.992
Global CaSc (HU)	28(0,260)	1(0,71.5)	358(103,850)	**<0.001**

the maximum among registered clinical variables was 30% for family history of CAD and 22% for diabetes mellitus.

The IGR evaluation for feature selection considering the two outcome variables (myocardial ischemia and early revascularization) is graphically depicted in supplementary Fig. S1.

In the case of identification of patients who demonstrated myocardial ischemia, all CTA variables (58 in total) significantly contributed with the ML modeling, while 8 of the 9 available clinical variables also were considered as relevant model components. Notably, the excluded variable was the presence of dyspnea. Conversely, the highest-ranking features related to the presence of plaque and the luminal narrowing in the proximal segments of the left anterior descending, circumflex and right coronary arteries, calcification in the left main, and sex.

Feature selection considering the identification of patients who underwent early revascularization (re-sampled $n = 400$) showed once again that all CTA variables significantly contributed with ML modeling, while also 8 of the 9 available clinical variables also were considered as relevant model components. The excluded variable in this case was age, while the most prominent features mostly related to the presence of plaque and the luminal narrowing in the proximal segments of the left anterior descending, circumflex and right coronary arteries.

3.2. Outcomes, ML classification performance and comparative discrimination

From the study sample (830 patients), 251 patients (30.2%) demonstrated myocardial ischemia. These represented 60.3% of the patients who underwent PET imaging (415 patients) as indicated on-site by the attending physician. On the other hand, 100 patients underwent early revascularization, which represented 12.1% of the total patient sample and 24.0% of those who underwent PET imaging.

Regarding the identification of patients who demonstrated myocardial ischemia through PET imaging, ML integration of all CTA variables (58) demonstrated an AUC = 0.85, while ML integration of all CTA variables + clinical variables significantly ($p < 0.01$) improved the AUC to 0.87. Comparatively, the logistic regression considering the summary CTA variable (expert interpretation) demonstrated an equivalent AUC = 0.87, while ML integration of the expert CTA interpretation variable + clinical variables significantly outperformed these models with an AUC = 0.91 ($p < 0.01$). Further, the ML integration of only the clinical variables achieved an AUC = 0.74, while the logistic regression considering global CaSc alone demonstrated an AUC = 0.82. The ML integration of CaSc + clinical variables performed significantly better (AUC = 0.84) ($p < 0.01$) than either component alone. The rest of the performance parameters are shown in Table 2 and the ROC curve analysis is depicted in Fig. 2.

Considering the identification of patients who ultimately underwent early revascularization, ML integration of all CTA variables achieved an AUC = 0.85, while ML integration of all CTA variables + clinical variables significantly improved the AUC to 0.86 ($p < 0.01$). Comparatively, the logistic regression considering the expert CTA interpretation variable significantly outperformed the aforementioned models with an AUC = 0.88 ($p < 0.01$), while further ML integration of the expert CTA interpretation variable + clinical variables again showed the best performance with an AUC = 0.90 (all pairwise $p < 0.01$). Thereon, ML integration of only the clinical variables only achieved an AUC = 0.74, while the logistic regression considering global CaSc demonstrated a significantly higher AUC = 0.78 ($p < 0.01$). Once again, ML integration of CaSc + clinical variables outperformed either component alone with an AUC = 0.83 (all pairwise $p < 0.01$). Complementary performance metrics are also shown in Table 2.

Table 2
Performance metrics of all models for in the identification of patients who demonstrate myocardial ischemia (A) or underwent early revascularization (B).

Model Features A	Sens	Specif	PPV	NPV	Acc.	F1-score	AUC
Clinical variables	0.37	0.87	0.55	0.76	0.72	0.44	0.74
CaSc	0.34	0.96	0.79	0.77	0.77	0.48	0.82
Expert CTA interpretation variable	0.77	0.92	0.80	0.90	0.87	0.78	0.87
All CTA variables	0.59	0.92	0.75	0.84	0.82	0.66	0.85
CaSc + Clinical variables	0.59	0.88	0.68	0.83	0.79	0.63	0.84
Expert CTA interpretation variable + Clinical variables	0.73	0.93	0.81	0.89	0.87	0.77	0.91
All CTA variables + Clinical variables	0.63	0.91	0.76	0.85	0.83	0.69	0.87
B							
Clinical variables	0.39	0.91	0.60	0.82	0.78	0.47	0.74
CaSc	0.23	0.96	0.66	0.79	0.78	0.34	0.78
Expert CTA interpretation variable	0.96	0.85	0.68	0.99	0.88	0.80	0.88
All CTA variables	0.53	0.92	0.69	0.85	0.82	0.60	0.85
CaSc + Clinical variables	0.55	0.89	0.63	0.86	0.81	0.59	0.83
Expert CTA interpretation variable + Clinical variables	0.81	0.87	0.68	0.93	0.86	0.74	0.90
All CTA variables + Clinical variables	0.50	0.91	0.66	0.85	0.81	0.57	0.86

4. Discussion

The present study evaluated the performance of integrating conventional CTA and clinical data through ML in order to identifying patients with myocardial ischemia as demonstrated through PET imaging or requiring early revascularization (within 6 months). Comparisons were performed between modeling permutations involving: clinical data (9 variables), the extensive CTA features (58 variables), the clinical expert interpretation of the CTA findings as performed in daily practice (i.e. the *expert CTA interpretation variable*), and the global CaSc measurement (1 variable), while contemplating two outcome variables, namely: the finding of myocardial ischemia through PET myocardial perfusion and the ulterior occurrence of early revascularization after imaging.

Interestingly, results showed that ML integration of the multiple data sources consistently improved the identification of patients that could benefit from subsequent PET imaging (as they provide orthogonal information). Hence, this supports the compound use of ML and clinical interpretation to maximize analytical benefit. This became evident in the addition of clinical variables (demographics, cardiovascular risk factors and symptomatology) to the model with the CTA variables only, which improvied its performance (from 0.85 to 0.87 for classification of ischemia). A comparable improvement was observed when integrating the expert CTA interpretation variable with the clinical variables (from 0.87 to 0.91).

Seemingly, the expert CTA interpretation variable (operationalized according to the clinical interpretation of the expert reader and embodied in the imaging report) conveys comparable and in some instances, better discriminative performance than the entire body of variables (58) emerging from the segmental CTA description of the coronary tree. This rises three aspects increasingly recognized in this novel era of data analytics, namely: 1) that interrelated variable integration, although possible, may sometimes represent an overkill in situations where simpler modeling is warranted, 2) that ML can also be utilized for feature selection in order to optimize the attributes utilized in classical non-ML modeling (e.g. logistic regression), and 3) that expert interpretation in itself still represents the result of complex clinical thinking and data integration that is useful and robust.

The logistic regression analysis of global CaSc showed, as expected, that its value for the identification of patients with myocardial ischemia or early revascularization is lower than that of CTA interpretation data either in its extended or summarized form. In some circumstances, however, this difference may not hinder its use in light of the intrinsic advantages of CaSc (low radiation burden, cost, accessibility and speed of acquisition) at the individual-patient level. Moreover, the emergence of automatic CaSc evaluation by means of *deep learning* further enhances the attractiveness of this feature [17].

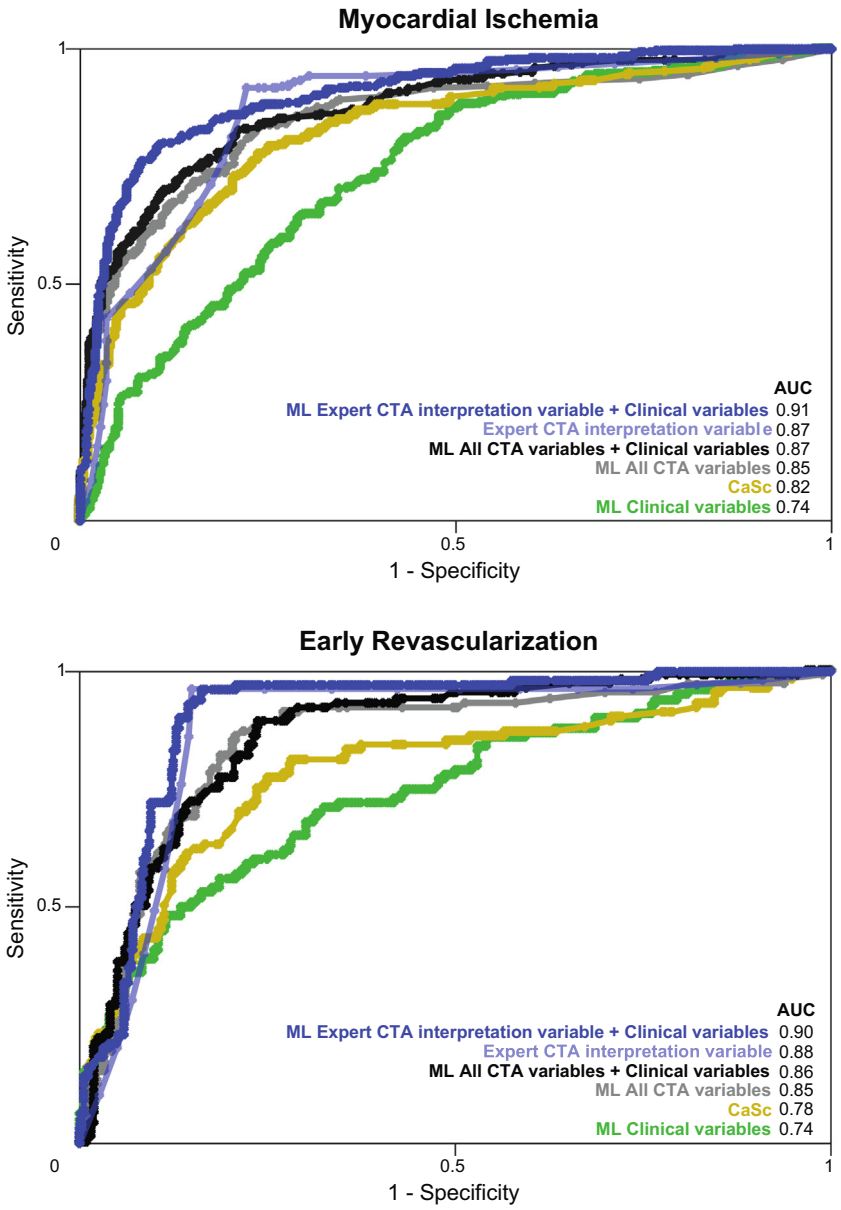

Fig. 2. ROC curve analysis of (10-fold cross validation) ML performance for the identification of myocardial ischemia on the top [$n = 830$] and early revascularization on the bottom [$n = 400$].

The established protocol for CTA and subsequent PET imaging indicated by the attending physician in our site has been developed to aid the identified need to optimize resource utilization and burden to the patient, it is supported in parallel by the recommendations in the NICE and ESC guidelines for the evaluation of patients with stable angina and chronic coronary syndromes. In this sample, roughly 60% of the patients who were selected to undergo PET imaging had findings compatible with myocardial ischemia, while successively a quarter of these patients ultimately underwent early revascularization. Therefore, current patient selection within the imaging protocol averages an accuracy of ~80% in identifying the need to query myocardial ischemia through perfusion imaging and of ~62% in predicting ulterior revascularization. Our results showed that ML integration of the expert CTA interpretation variable and all the considered clinical variables can achieve (cross-validated) accuracies equal to 87% and 86% in such tasks, respectively. This depicts a substantial gain to be harvested from the identification of patients that can benefit from PET imaging through ML. It should be noted that none of the patients who were not deemed to require PET imaging were revascularized, which inputs the assumption of the absence of myocardial ischemia in this subset. We acknowledge that this operational proposition neglects the possible presence of some degree of ischemia either due to CAD or microvascular dysfunction is some of these patients. Yet, we are studying this in parallel as this registry is longitudinal in nature and will report on such outcomes in due time.

Modeling performances described for both operationalized outcome variables showed similar tendencies, which attests to their interrelatedness. However, a sizeable proportion of patients with proven ischemia were not revascularized in our sample, while also a few patients without ischemia were indeed revascularized. This is explained by the cases in which the attending physician and patient opted for optimal medical therapy instead, supported by a reduced ischemic burden, as well as those in which revascularization was expected to yield some symptomatic relief. We recognize that these patterns may rather input

noise in the modelled outcome that for the moment is unavoidable. Yet, these insights hint the type of refinements in collected data that should be pursued in the near future in order to refine ML estimations.

This study delivers information useful to advance the notion of utility of ML analytics in the integration of sensible diagnostic data and to navigate further through the optimization of patient selection for advanced imaging techniques such as PET. We envision that ML analytics for this purpose may be of value when drawbacks inherent to advanced imaging may be limiting for the comprehensive evaluation and tailored management of patients suspected with CAD.

5. Study limitations

The present study carries all the limitations concerning retrospective data analyses. Nevertheless, current standards in data analysis benefit from relatively large samples presently only available in existing datasets not originally intended for optimizing ML implementation. In the case of the present study, our analyses operated in a rather large sample for the standard pertaining advanced myocardial perfusion imaging with PET. Another limitation may be found in the apparent lack of "external" validation of the ML analyses, e.g. in data from alternative unseen centers, which may raise the concern of the accuracy of generalizability of the performances reported. Our results reflect the estimates emerging from the selected 10-fold cross-validation integrating performances from 10 models ultimately tested in 10% of actual unseen data per fold. And notably, this policy is currently the preferred technique for the appraisal of performance and its associated uncertainty (error) in ML workflows as it offers stable more stable results by removing the variability expected in other methods of prediction error quantification [18]. Additionally, the rapid emergence of novel ML implementations such as dedicated use of deep learning on cardiac imaging may suggest that these results could be further optimized in several ways. In this study, we opted for the practicality of applying non-deep learning machine learning to existing structured numerical data, which considerably cuts down on computational power needs and makes results more readily available for deployment in existing databases containing conventional CTA data available in every center that preforms such a study around the world without the need for special pre-processing for experimental imaging biomarker that may currently only be found in highly-specialized imaging centers. Finally, myocardial ischemia was defined by the finding of at least one segment with a stress MBF of ≤2.4 ml/g/min. Consequently, high prevalence of minor cases of ischemia could influence the results of this study. However, the mean number of segments having abnormal perfusion per patient was approximately seven, suggesting a minimum risk of gross overestimation of myocardial ischemia. Furthermore, there may be benefit in informing clinicians over the probability of myocardial ischemia even if minor since it may explain anginal symptoms and support the need for therapy optimization, and also the probability of need-to-revascularize, independently. Finally, it should be noted that the decision for sequential PET was dependent on the expert CTA interpretation, which might partially explain a relative superior performance of expert CTA interpretation.

6. Conclusions

Machine learning integration of diagnostic CTA and clinical data may improve the selection of patients that can benefit from advanced perfusion imaging with PET, i.e. those who demonstrate myocardial ischemia and those who ultimately undergo early revascularization, at the individual level. This could potentially spare the need for subsequent advanced imaging and better help identify patients in ultimate need for revascularization. The implementation of ML analytics in the full body of standard CTA imaging variables may not render unequivocally better results than its expert clinical interpretation, but it consistently improves performance when integrating data from other sources. Further research into the real-world implementation of ML-based support systems for the selection of patients for cardiac imaging, is warranted.

Funding

No relevant sources of funding for this particular project.

Declaration of Comepting Interest

Nothing to disclose.

Acknowledgements

The work of J.W. Benjamins and M.W. Yeung was supported by the Research Project CVON-AI (2018B017), which is financed by the PPP Allowance made available by Top Sector Life Sciences & Health to the Nederlandse Hartstichting to stimulate public-private partnerships. This work reflects only the author's view, not that of the funders. Stichting LSH-TKI or Hartstichting or the Ministry of Economic Affairs is not responsible for any use that may be made of the information it contains.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijcard.2021.04.009.

References

[1] J. Knuuti, W. Wijns, A. Saraste, D. Capodanno, E. Barbato, C. Funck-Brentano, et al., 2019 ESC guidelines for the diagnosis and management of chronic coronary syndromes, Eur. Heart J. (2019) 1–71.
[2] L.E. Juarez-Orozco, A. Saraste, D. Capodanno, E. Prescott, H. Ballo, J.J. Bax, et al., Impact of a decreasing pre-test probability on the performance of diagnostic tests for coronary artery disease, Eur. Heart J. Cardiovasc. Imaging 20 (11) (2019 Nov 1) 1198–1207.
[3] P.A.L. Tonino, W.F. Fearon, B. De Bruyne, K.G. Oldroyd, M.A. Leesar, P.N. Ver Lee, et al., Angiographic versus functional severity of coronary artery stenoses in the FAME study fractional flow reserve versus angiography in multivessel evaluation, J. Am. Coll. Cardiol. 55 (25) (2010 Jun 22) 2816–2821.
[4] S. Kajander, U. Ukkonen, H. Sipilä, M. Teräs, J. Knuuti, Low radiation dose imaging of myocardial perfusion and coronary angiography with a hybrid PET/CT scanner, Clin. Physiol. Funct. Imaging 29 (1) (2009 Jan) 81–88.
[5] A. Thomassen, H. Petersen, A.C.P. Diederichsen, H. Mickley, L.O. Jensen, A. Johansen, et al., Hybrid CT angiography and quantitative 15O-water PET for assessment of coronary artery disease: comparison with quantitative coronary angiography, Eur. J. Nucl. Med. Mol. Imaging 40 (12) (2013 Dec) 1894–1904.
[6] G.L. Raff, A. Abidov, S. Achenbach, D.S. Berman, L.M. Boxt, M.J. Budoff, et al., SCCT guidelines for the interpretation and reporting of coronary computed tomographic angiography, J. Cardiovasc. Comput. Tomogr. 3 (2) (2019) 122–136.
[7] A. Esteva, B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M. Blau, et al., Dermatologist-level classification of skin cancer with deep neural networks, Nature. 542 (7639) (2017 Jan 2) 115–118.
[8] L.E. Juarez-Orozco, O. Martinez-Manzanera, A.E. Storti, J. Knuuti, Machine learning in the evaluation of myocardial ischemia through nuclear cardiology, Curr. Cardiovasc. Imaging Rep. 12 (2) (2019 Feb 9) 5.
[9] M. Motwani, D. Dey, D.S.D. Berman, G. Germano, S. Achenbach, M.M.H. Al-Mallah, 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. 52 (4) (2016) 468–476.
[10] L.E. Juarez-Orozco, Knol RJJ, C.A. Sanchez-Catasus, O. Martinez-Manzanera, F.M. van der Zant, J. Knuuti, Machine learning in the integration of simple variables for identifying patients with myocardial ischemia, J. Nucl. Cardiol. 27 (2020) 147–155, https://doi.org/10.1007/s12350-018-1304-x.
[11] S. Kajander, E. Joutsiniemi, M. Saraste, M. Pietilä, H. Ukkonen, A. Saraste, et al., Cardiac positron emission tomography/computed tomography imaging accurately detects anatomically and functionally significant coronary artery disease, Circulation. 122 (6) (2010 Aug 10) 603–613.
[12] M.D. Cerqueira, N.J. Weissman, V. Dilsizian, A.K. Jacobs, S. Kaul, W.K. Laskey, et al., Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart. A statement for healthcare professionals from the cardiac imaging Committee of the Council on clinical cardiology of the American Heart Association, Circulation. 105 (4) (2002 Jan 29) 539–542.
[13] T. Maaniitty, I. Stenström, J.J. Bax, V. Uusitalo, H. Ukkonen, S. Kajander, et al., Prognostic value of coronary CT angiography with selective PET perfusion imaging in coronary artery disease, JACC Cardiovasc. Imaging 10 (11) (2017 Nov) 1361–1370.

[14] P.P. Sengupta, S. Shrestha, B. Berthon, E. Messas, E. Donal, G.H. Tison, et al., Proposed requirements for cardiovascular imaging-related machine learning evaluation (PRIME): a checklist: reviewed by the American College of Cardiology Healthcare Innovation Council, JACC Cardiovasc. Imaging 13 (9) (2020) 2017–2035.
[15] R. Arsanjani, Y. Xu, D. Dey, V. Vahistha, A. Shalev, R. Nakanishi, et al., Improved accuracy of myocardial perfusion SPECT for detection of coronary artery disease by machine learning in a large population, J. Nucl. Cardiol. 20 (4) (2013 Aug) 553–562.
[16] E.R. DeLong, D.M. DeLong, D.L. Clarke-Pearson, Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach, Biometrics. 44 (3) (1988 Sep) 837–845.
[17] S.G.M. van Velzen, N. Lessmann, B.K. Velthuis, Bank IEM, D.H.J.G. van den Bongard, T. Leiner, et al., Deep learning for automatic calcium scoring in CT: validation using multiple cardiac CT and chest CT protocols, Radiology. 295 (1) (2020) 66–79.
[18] A.M. Molinaro, R. Simon, R.M. Pfeiffer, Prediction error estimation: a comparison of resampling methods, Bioinformatics. 21 (15) (2005 Aug 1) 3301–3307.