A machine learning–based clinical decision support algorithm for reducing unnecessary coronary angiograms

J.D. Schwalm, MD, MSc,*† Shuang Di, MEd, MSc,‡§ Tej Sheth, MD,*†
Madhu K. Natarajan, MD,*† Erin O'Brien, BA,* Tara McCready, PhD,*
Jeremy Petch, PhD*†‡‖

*From the *Population Health Research Institute, McMaster University and Hamilton Health Sciences, Hamilton, Canada, †Department of Medicine, Division of Cardiology, McMaster University, Hamilton, Canada, ‡Centre for Data Science and Digital Health, Hamilton Health Sciences, Hamilton, Canada, §Dalla Lana School of Public Health, University of Toronto, Toronto, Canada, and ‖Institute for Health Policy, Management and Evaluation, University of Toronto, Toronto, Canada.*

BACKGROUND Conventional clinical risk scores and diagnostic algorithms are proving to be suboptimal in the prediction of obstructive coronary artery disease, contributing to the low diagnostic yield of invasive angiography. Machine learning could help better predict which patients would benefit from invasive angiography vs other noninvasive diagnostic modalities.

OBJECTIVE To reduce patient risk and cost to the healthcare system by improving the diagnostic yield of invasive coronary angiography through optimized outpatient selection.

METHODS Retrospective analysis of 12 years of referral data from a provincial cardiac registry, including all patients referred for invasive angiography of more than 1.4 million individuals in Ontario, Canada. Stable outpatients undergoing coronary angiography during the study period were included in the analysis. The training set (80% random sample, n = 23,750) was used to develop 8 prediction models in Python using grid-search cross-validation. The test set (20% random sample, n = 5938), evaluated the discrimination performance of each model.

RESULTS The machine-learning model achieved a substantially better performance (area under the receiver operating characteristics curve: 0.81) than existing models for predicting obstructive coronary artery disease in patients referred for invasive angiography. It significantly outperformed both the reference model and current clinical practice with a net reclassification index of 27.8% (95% confidence interval [CI]: [24.9%–30.8%], P value <.01) and 44.7% (95% CI: [42.4%–47.0%], P value <.01), respectively.

CONCLUSION This prediction model, when coupled with a point-of-care, online decision support tool to be used by referring physicians, could improve the diagnostic yield of invasive coronary angiography in stable, elective outpatients, thus improving patient safety and reducing healthcare costs.

KEYWORDS Coronary artery disease; Machine learning; Coronary angiography; Coronary computed tomographic angiography; Prediction model

(Cardiovascular Digital Health Journal 2022;3:21–30) © 2021 Heart Rhythm Society. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Background

Coronary artery disease (CAD) is the leading cause of death worldwide. The gold-standard test used to diagnose CAD is invasive coronary angiography. However, the low diagnostic yield of invasive angiography performed in elective patients has gained attention. Patel and associates[1] performed an analysis of the American College of Cardiology National Cardiovascular Data Registry that included nearly 400,000 patients undergoing elective angiography at 663 hospitals, without prior known CAD. This study showed that nearly 60% of angiograms had results showing <50% stenosis or normal findings.[1] Despite lower population rates of angiography in Canada, rates of normal (0% stenosis in major epicardial arteries) approaches 42%.[2] Current rates of nonobstructive CAD in Ontario, Canada (defined as <70% stenosis in major epicardial vessels or <50% stenosis of the left main artery) in patients undergoing elective angiography are approximately 53% (CorHealth QPMM Quarterly Reports, unpublished data, 2020). Wide variation in the frequency of nonobstructive CAD in patients undergoing invasive angiography suggests that there are opportunities to optimize referral practices.[2]

Obtaining the results of an angiogram with normal or nonsignificant findings may offer some important benefits

KEY FINDINGS

- A machine-learning model used to analyze multicenter administrative and clinical databases can accurately predict obstructive coronary artery disease in outpatients referred for coronary angiography compared to current strategies.

- This machine-based model has the potential to improve patient safety and decrease healthcare costs by optimizing referral for outpatient invasive coronary angiographies.

- While this machine learning–enabled predictive model achieves better performance than existing clinical risk scores, it does require external validation and local adaptation before being applied in other health regions.

to patients by definitively excluding obstructive CAD as a cause of symptoms and providing valuable prognostic insight. A retrospective analysis of patients referred for an invasive angiography demonstrated that <11% of referrals were deemed "inappropriate" for this test by current practice standards.[3] However, there are some drawbacks to having a high proportion of patients with negative invasive angiography results, despite being "indicated" to undergo this invasive procedure. First, these patients undergo the risk of the procedure (stroke death, myocardial infarction, vascular complications, and bleeding) without any potential to benefit from revascularization. Second, they consume resources better allocated to alternate patients, particularly in a value-based healthcare delivery model or a resource-constrained universal healthcare system. It is preferable for such patients to undergo a noninvasive assessment if similar diagnostic certainty can be achieved.

Conventional clinical risk scores and diagnostic algorithms are proving to be suboptimal in the prediction of obstructive CAD, contributing to a lack of efficiencies in referral practices and the identified low diagnostic yield of invasive angiography.[4] Novel research strategies that harness the powers of artificial intelligence and machine learning could help better predict patients that would benefit from invasive angiography vs other noninvasive diagnostic modalities like coronary computed tomographic angiography (CCTA).

The overarching goal of this project is to reduce patient risk and cost to the healthcare system by improving the diagnostic yield of invasive coronary angiography through optimized patient selection based on a Learning Health System environment. Specifically, we have developed a contemporary prediction tool using machine learning, trained on data from an established multicenter, regional registry, that could be integrated into diagnostic algorithms in order to improve the diagnostic yield of invasive coronary angiography. Our intention is that this clinical model will provide physicians with clinical decision support to optimize referrals to angiography vs noninvasive modalities.

Methods

Study design, data source, and study cohort

We retrospectively analyzed referral data from the CorHealth (Ontario, Canada) cardiac registry of 2 regional referral hospitals, spanning years 2008–2019. The CorHealth database is a prospectively collected registry that includes all patients referred for invasive angiography in the province of Ontario, Canada. This registry has been previously used in other observational and interventional studies.[5–7] This study was approved by Hamilton Health Sciences Research Ethics Board 4697.

Stable outpatients undergoing coronary angiography at the 2 hospitals during the study period were included in the analysis. Outpatients who met any of the following criteria were excluded: (1) with previous percutaneous coronary intervention, (2) with previous coronary artery bypass grafting, (3) with cardiogenic shock, (4) on glycoprotein IIb/IIIa inhibitors, and (5) with prior acute coronary syndromes.

The outcome of interest was obstructive CAD, defined as ≥70% stenosis in major epicardial vessels of >2 mm in diameter or ≥50% stenosis of the left main artery (as defined in CorHealth, Data Dictionary, unpublished). The predictors for model development were all the routinely collected variables in the provincial registry. Specifically, the predictors included demographic characteristics, patient referral information, anthropometric measures, clinical symptoms and risk factors, medical history, and tobacco habits. We were also able to include socioeconomic variables as a predictor, by linking to the 2016 Ontario Marginalization Index using postal codes of patient residence. The Ontario Marginalization Index is a validated data tool that combines a wide range of demographic indicators to characterize 4 dimensions of marginalization: residential instability, material deprivation, ethnic concentration, and dependency.[8] Lower scores on each dimension correspond to areas that are the least marginalized; higher scores on each dimension correspond to areas that are the most marginalized.

Statistical analysis

On the training set (80% random sample, n = 23,750), we developed 8 models to predict the probability of significant CAD. First, as the reference model, we fit a logistic regression model with 7 predictors derived from the dataset, including high-risk exercise stress testing/imaging, age >65 years, diabetes, previous myocardial infarction (MI), left ventricular ejection fraction <35%, Canadian Cardiovascular Society (CCS)/New York Heart Association (NYHA) class = 3 or 4, and creatinine >180 μmol/L. The reference model was a clinical model with predictors and was originally developed to predict the probability of high-risk CAD (eg, 3-vessel disease).[9] In our dataset, >2 of beta blockers / calcium channel blockers /nitrates was not available.
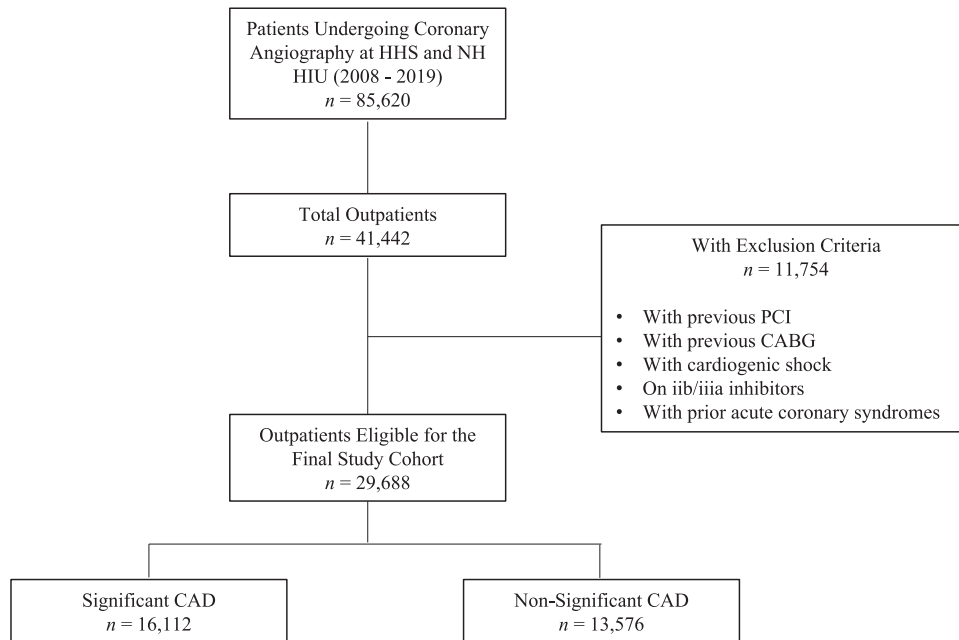
Figure 1   Study population profile.

Next, using all variables available in our dataset, we developed 7 machine-learning models, including logistic regression, logistic regression with Lasso regularization,[10] random forests,[11] gradient boosted decision trees,[12,13] Light Gradient Boosting Machine (LightGBM),[14] eXtreme Gradient Boosting (XGBoost),[15] and a multilayer perception deep neural network.[16] Fivefold cross-validation was used to find a set of optimal hyperparameters for each machine-learning algorithm. Python (version 2.7.17) was used for model development.

On the test set (20% random sample, n = 5938), we evaluated the discrimination performance of each model using area under the receiver operating characteristics curve (AUROC) as the primary evaluation metric and area under the precision-recall curve, accuracy, sensitivity, specificity, positive predictive value, and negative predictive value as secondary evaluation metrics. The model with highest AUROC was selected as the final model. We then compared the reclassification performance of the final model against the reference model and current clinical practice as observed in our data using net reclassification improvement (NRI).[17]

For model interpretation purposes, we used permutation importance to assess the global importance of each predictor in predicting the probability of significant CAD.[18] Since machine-learning analyses include higher-order interactions between multiple predictors and capture nonlinear relationships between predictors and outcome, we used partial dependence plots to uncover and visualize such relationships.[19]

At the patient level, we used Shapley values to understand the magnitude and direction of each predictor's effect on predictions for individual patients.[20,21] To improve clinical utility of the model, we applied the local interpretable model-agnostic explanations framework to the final model and generated an explanation for each prediction result,[22] including the predicted probabilities of nonsignificant and significant CAD, as well as the contribution of each predictor to the prediction result. Since some categorical predictors have high cardinality, which makes the explanations difficult to interpret, we collapsed levels within a categorical predictor by summing the contribution of each level and used the sum score to approximate the contribution of the categorical predictor to the prediction result.

Results
Study population
During 2008–2019, the registry recorded 85,620 patients undergoing coronary angiography at the 2 hospitals. Of these, a total of 29,688 outpatients who met the inclusion criteria were included in the analysis, including 13,576 (45.72%) with nonsignificant CAD and 16,112 (54.27%) with significant CAD (Figure 1). Baseline characteristics for the whole study cohort and for patients according to different outcomes are presented in Table 1.

Discrimination performance
Table 2 presents the discrimination performance of different models. While all machine-learning models achieved better performance than the reference model, the LightGBM model showed the best discrimination performance (AUROC = 0.81) and was selected as the final model.

Reclassification performance
Reclassification performance of the selected LightGBM model against the reference model is shown in Table 3.

Table 1   Baseline characteristics of patients in the study cohort

Characteristics	Overall (N = 29,688)		Nonsignificant CAD (N = 13,576)		Significant CAD (N = 16,112)		P value	Missing
Demographic characteristics								
Sex, n (%)							<.001	1
Female	11,678	(39.3)	6610	(48.7)	5068	(31.5)		
Age (years), mean (SD)	65.6	(11.4)	64.2	(11.5)	66.8	(11.1)	<.001	0
Ethnicity, n (%)							.290	26,002
White	3478	(94.4)	1875	(94.7)	1603	(94.0)		
South Asian	68	(1.8)	29	(1.5)	39	(2.3)		
Asian	38	(1.0)	17	(0.9)	21	(1.2)		
Aboriginal	34	(0.9)	19	(1.0)	15	(0.9)		
Black	21	(0.6)	14	(0.7)	7	(0.4)		
Socioeconomic status								
Residential instability, mean (SD)	0.1	(1.0)	0.1	(1.0)	0.1	(1.0)	.026	530
Material deprivation, mean (SD)	0.1	(1.0)	0.1	(1.0)	0.1	(1.0)	.419	530
Ethnic concentration, mean (SD)	-0.4	(0.6)	-0.4	(0.6)	-0.4	(0.6)	.622	530
Dependency, mean (SD)	0.4	(1.2)	0.4	(1.2)	0.4	(1.2)	.002	530
Patient referral information								
Primary reason, n (%)							<.001	0
Coronary disease	25,856	(87.1)	11,739	(86.5)	14,117	(87.6)		
Other	3832	(12.9)	1837	(13.5)	1995	(12.4)		
Primary reason type, n (%)							<.001	3052
Elective, stable coronary disease	13779	(51.7)	5082	(42.2)	8697	(59.6)		
Rule out CAD	9150	(34.4)	5560	(46.2)	3590	(24.6)		
Other	3707	(13.9)	1396	(11.6)	2311	(15.8)		
Translator required, n (%)							.554	7780
Yes	161	(0.7)	71	(0.7)	90	(0.8)		
Dye allergy, n (%)							.002	1353
Yes	315	(1.1)	170	(1.3)	145	(0.9)		
Anthropometric measures								
Height (cm), mean (SD)	169.5	(11.2)	169.0	(11.4)	170.0	(10.9)	<.001	289
Weight (kg), mean (SD)	86.2	(21.3)	86.6	(21.5)	85.9	(21.1)	.005	212
Clinical symptoms and risk factors								
Ischemic change, n (%)							<.001	1864
Persistent	2841	(10.2)	1012	(8.0)	1829	(12.0)		
Transient with pain	336	(1.2)	105	(0.8)	231	(1.5)		
Transient without pain	127	(0.5)	48	(0.4)	79	(0.5)		
Exercise ECG risk, n (%)							<.001	0
High	6294	(21.2)	2178	(16.0)	4116	(25.5)		
Low	5531	(18.6)	2886	(21.3)	2645	(16.4)		
Functional imaging risk, n (%)							<.001	384
High	8358	(28.5)	3147	(23.6)	5211	(32.7)		
Low	5878	(20.1)	3137	(23.5)	2741	(17.2)		
LV method, n (%)							<.001	1378
Echo	22,416	(79.2)	10,792	(80.6)	11,624	(77.9)		
Other	3029	(10.7)	1185	(8.9)	1844	(12.4)		
Not done	2865	(10.1)	1407	(10.5)	1458	(9.8)		
LV function, n (%)							<.001	3054
≥50%	19,773	(74.2)	9577	(79.4)	10,196	(70.0)		
35%–49%	3472	(13.0)	1374	(11.4)	2098	(14.4)		
≤34%	2013	(7.6)	915	(7.6)	1098	(7.5)		
LV function value, mean (SD)	53.7	(12.9)	53.9	(13.2)	53.5	(12.5)	.219	23,904
Creatinine, n (%)							.985	279
Known	29,360	(99.8)	13,442	(99.8)	15,918	(99.8)		
Creatinine (mol/L), mean (SD)	100.2	(117.2)	96.9	(112.8)	103.1	(120.7)	<.001	458
CCS class, n (%)							<.001	0
0	8990	(30.3)	4736	(34.9)	4254	(26.4)		
1	4568	(15.4)	2274	(16.8)	2294	(14.2)		
2	8805	(29.7)	3873	(28.5)	4932	(30.6)		
3	6189	(20.8)	2261	(16.7)	3928	(24.4)		
4	1136	(3.8)	432	(3.2)	704	(4.4)		
NYHA class, n (%)							<.001	16
1	19,008	(64.1)	8375	(61.7)	10,633	(66.0)		
2	6433	(21.7)	3073	(22.7)	3360	(20.9)		

Table 1   (*Continued*)

Characteristics	Overall (N = 29,688)		Nonsignificant CAD (N = 13,576)		Significant CAD (N = 16,112)		P value	Missing
3	2706	(9.1)	1326	(9.8)	1380	(8.6)		
4	371	(1.3)	181	(1.3)	190	(1.2)		
Medical history								
History of MI, n (%)							<.001	0
Yes	3282	(11.1)	848	(6.2)	2434	(15.1)		
Recent MI, n (%)							.694	0
Yes	144	(0.5)	63	(0.5)	81	(0.5)		
History of cerebrovascular disease, n (%)							<.001	590
Yes	2184	(7.5)	818	(6.2)	1366	(8.6)		
History of peripheral vascular disease, n (%)							<.001	13
Yes	1618	(5.5)	472	(3.5)	1146	(7.1)		
Possible intracardiac thrombus, n (%)							.94	557
Yes	50	(0.2)	22	(0.2)	28	(0.2)		
History of infective endocarditis, n (%)							.005	51
Yes	32	(0.1)	23	(0.2)	9	(0.1)		
Active endocarditis, n (%)							.076	29,656
Yes	7 (n = 29,688)	(21.9)	3 (n = 13,576)	(13.0)	4 (n = 16,112)	(44.4)		
Congenital heart disease, n (%)							.133	7787
Yes	112	(0.5)	70	(0.6)	42	(0.4)		
History of congestive heart failure, n (%)							.003	33
Yes	2759	(9.3)	1336	(9.8)	1423	(8.8)		
Anticoagulant, n (%)							<.001	0
None	26,092	(87.9)	11,645	(85.8)	14,447	(89.7)		
Dialysis, n (%)							.436	5
Yes	790	(2.7)	350	(2.6)	440	(2.7)		
Diabetes, n (%)							<.001	3
Yes	8671	(29.2)	3506	(25.8)	5165	(32.1)		
Diabetes control, n (%)							<.001	21,010
On oral hypoglycemics	4970	(57.3)	1950	(56.2)	3020	(58.0)		
Insulin treatment	2506	(28.9)	1036	(29.9)	1470	(28.2)		
Managed by diet only	906	(10.4)	409	(11.8)	497	(9.5)		
No treatment	296	(3.4)	72	(2.1)	224	(4.3)		
Hypertension, n (%)							<.001	2
Yes	20,165	(67.9)	8656	(63.8)	11,509	(71.4)		
Hyperlipidemia, n (%)							<.001	6
Yes	20,431	(68.8)	8470	(62.4)	11,961	(74.3)		
COPD, n (%)							<.001	14
Yes	1863	(6.3)	942	(6.9)	921	(5.7)		
Tobacco habits								
History of smoking, n (%)							<.001	1070
Never	13,363	(46.7)	6385	(49.4)	6978	(44.5)		
Former	9252	(32.3)	4084	(31.6)	5168	(32.9)		
Current	6003	(21.0)	2455	(19.0)	3548	(22.6)		

CAD = coronary artery disease; CCS = Canadian Cardiovascular Society; COPD = chronic obstructive pulmonary disease; ECG = electrocardiogram; LV = left ventricular; MI = myocardial infarction; NYHA = New York Heart Association.

The LightGBM model significantly outperformed the reference model with an NRI of 27.84% (95% confidence interval [CI]: [24.85%–30.83%], *P* value <.001). Of the 2715 patients with nonsignificant CAD, 738 were correctly reclassified to the nonsignificant CAD category and 287 were reclassified to the significant CAD category. Of the 3223 patients with significant CAD, 742 were correctly reclassified to the significant CAD category and 380 were reclassified to the nonsignificant CAD category. The LightGBM model also significantly outperformed current clinical practice with an NRI of 44.70% (95% CI: [42.41%–46.98%], *P* value <.001).

Variable importance

To gain insights into the contribution of each predictor to the model, we estimated the importance of each predictor according to the selected final model using permutation importance (Figure 2). While traditional variables used in existing prediction models are ranked important (ie, sex, age, functional

Table 2 Discrimination performance of the reference model and machine learning models

Model	Primary metric	Secondary metrics					
	AUROC	AUPRC	Accuracy	Sensitivity	Specificity	PPV	NPV
LightGBM	0.81	0.84	0.73	0.75	0.71	0.75	0.70
XGBoost	0.79	0.82	0.72	0.75	0.69	0.74	0.70
Gradient boosted Decision trees	0.79	0.82	0.72	0.75	0.69	0.74	0.70
Random forests	0.79	0.81	0.71	0.76	0.66	0.73	0.70
Logistic regression (Lasso)	0.78	0.81	0.72	0.75	0.69	0.74	0.70
Logistic regression	0.78	0.80	0.71	0.74	0.68	0.73	0.69
Deep neural network	0.77	0.79	0.70	0.75	0.64	0.71	0.68
Reference model	0.62	0.65	0.58	0.63	0.53	0.61	0.55

AUPRC = area under the precision-recall curve; AUROC = area under the receiver operating characteristics curve; LightGBM = Light Gradient Boosting Machine; NPV = negative predictive value; PPV = positive predictive value; XGBoost = eXtreme Gradient Boosting.

imaging risk), other nontraditional variables also prove to be important, including congenital heart disease, need for a translator, referring physician, and the month in which the patient is referred for invasive coronary angiogram.

Variable effects and interactions

We visualized the relationship between predictors and significant CAD captured by the LightGBM model using a partial dependence plot (Supplemental Appendix 1), which shows the marginal effect 1 predictor has on the predicted probability of significant CAD according to the LightGBM model.

We observed complex relationships between continuous predictors and significant CAD in the partial dependence plots. For example, for the predictor age for patients under 40 years, the probability of significant CAD is low, and it decreases as age increases, whereas for patients aged over 40 years, as age increases, the probability of significant CAD increases. Similarly, for the predictor creatinine level under 75 μmol/L, the probability of significant CAD increases rapidly with creatinine level, whereas when creatinine level is over 75 μmol/L, the probability of significant CAD increases slowly with creatinine level. For socioeconomic variables, the relationships between the probability of significant CAD and marginalization dimensions (eg, dependency and residential instability) are complex and nonlinear. However, from the plots we can see an overall trend that patients who are more marginalized have higher probabilities of significant CAD.

We also used 2-way partial dependence plots to uncover and visualize feature interactions between predictors

(Supplemental Appendix 2). As shown on the left plot, when creatinine level is >120 μmol/L, the probability of significant CAD shows strong dependence on CCS class and is almost independent of creatinine level; however, when creatinine level is <120 μmol/L, the probability of significant CAD is dependent on the values of both predictors. Similarly, we observed an interaction between weight and NYHA class when weight is >70 kg.

At the patient level, we used a Shapley value summary plot (Supplemental Appendix 3) to understand the magnitude and direction of each predictor's effect on individual patients, according to the XGBoost model. In the summary plot, each point is a Shapley value for a predictor and a patient. The position on the y-axis is determined by the variable and on the x-axis by the Shapley value. For numerical variables, color represents the value of the variable from low (blue) to high (pink); for categorical variables, the blue color represents 0 and red represents 1.

In this plot, we observed that the directions of most predictors' effect are mostly the same for all patients, but the magnitudes of effect vary. For example, for categorical predictor history of MI, having previous MI is associated with higher risk for significant CAD for all patients, but the magnitudes of the effect are different for different patients.

Local explanations

We applied the local interpretable model-agnostic explanations framework to the LightGBM model and generated an explanation for each prediction result, including the predicted probability and the contribution of each predictor to the

Table 3 Reclassification performance of the LightGBM model against the reference model

Reference model	LightGBM model		Net correctly reclassified
	Nonsignificant CAD	Significant CAD	
Nonsignificant CAD			
Nonsignificant CAD	1168	287	16.61%
Significant CAD	738	522	
Significant CAD			
Nonsignificant CAD	442	742	11.23%
Significant CAD	380	1659	

CAD = coronary artery disease; LightGBM = Light Gradient Boosting Machine.
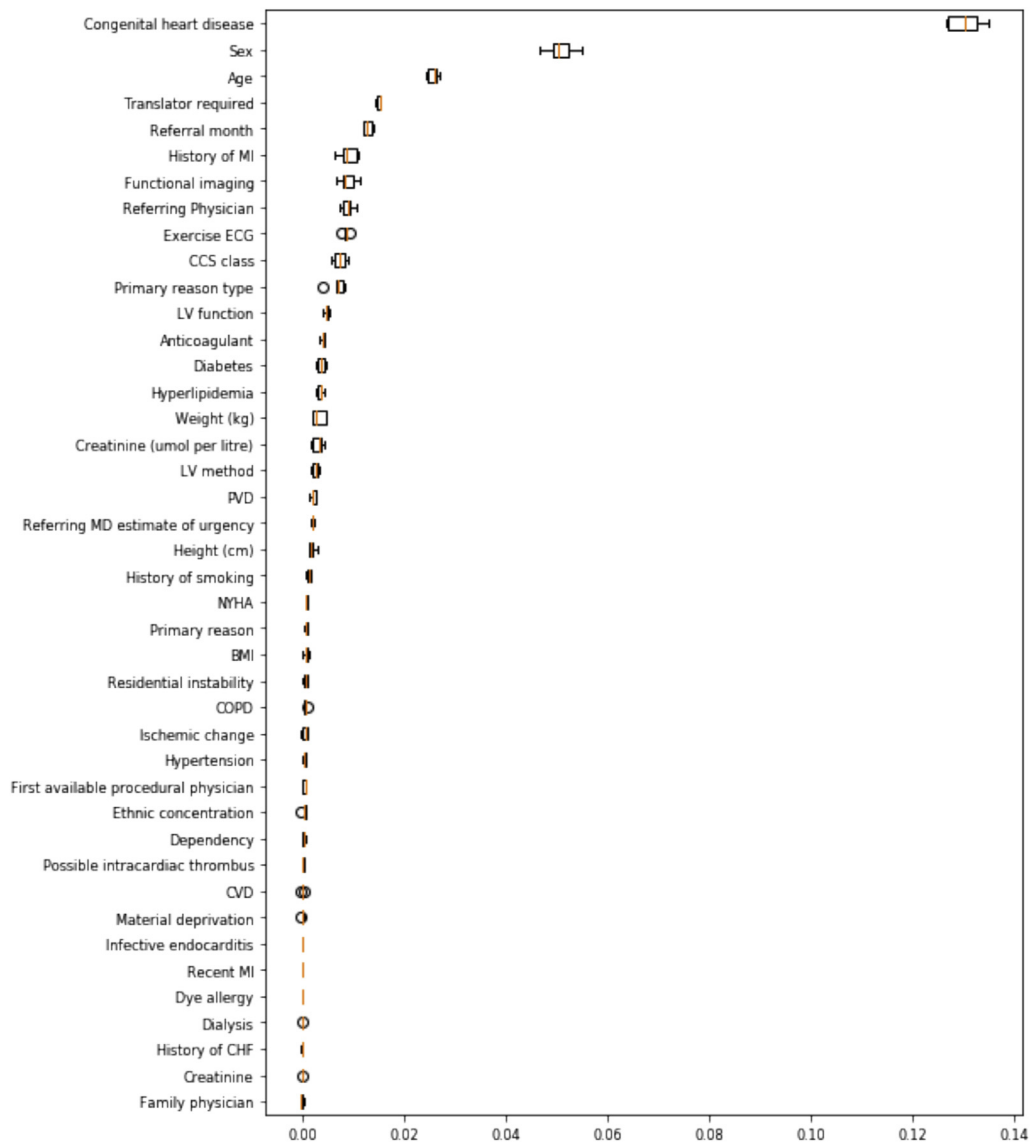
Figure 2   Estimated of the importance of each predictor according to the selected final model using permutation importance.

prediction result. For instance, in Figure 3 we present a true-negative case (ie, label and prediction result are both nonsignificant CAD). According to the LightGBM model, the predicted probability of nonsignificant CAD is 0.63. Congenital heart disease, sex, and referring physician are the predictors that contribute most to this prediction result. In this figure, green indicates a positive correlation and red indicates a negative correlation (analogous to positive and negative coefficients in a logistic regression).

Discussion

We present the first report of a machine learning–based model trained on existing administrative and clinical registries across 2 cardiac centers and 8 referring hospital corporations in Ontario, Canada. This model achieves substantially better performance than a baseline model for predicting obstructive CAD in patients referred for invasive angiography.

This model is unique for several reasons. First, this machine learning–based predictive model was generated based on a large, diverse cohort including all patients referred for invasive coronary angiography over 12 years from 8 hospital corporations covering a catchment population of more than 1.4 million. This large cohort lends significant power to this predictive model. The baseline demographics of all referred patients are cross-referenced against referring internist and cardiologist consultation letters. This registry of baseline demographics and coronary angiogram outcomes has been used in multiple observational and interventional studies.[5–7]

Second, the predictive model generated from this project can directly influence healthcare processes, potentially resulting in improved patient safety and reduced costs to the system. Implementing this model as part of a point-of-care decision
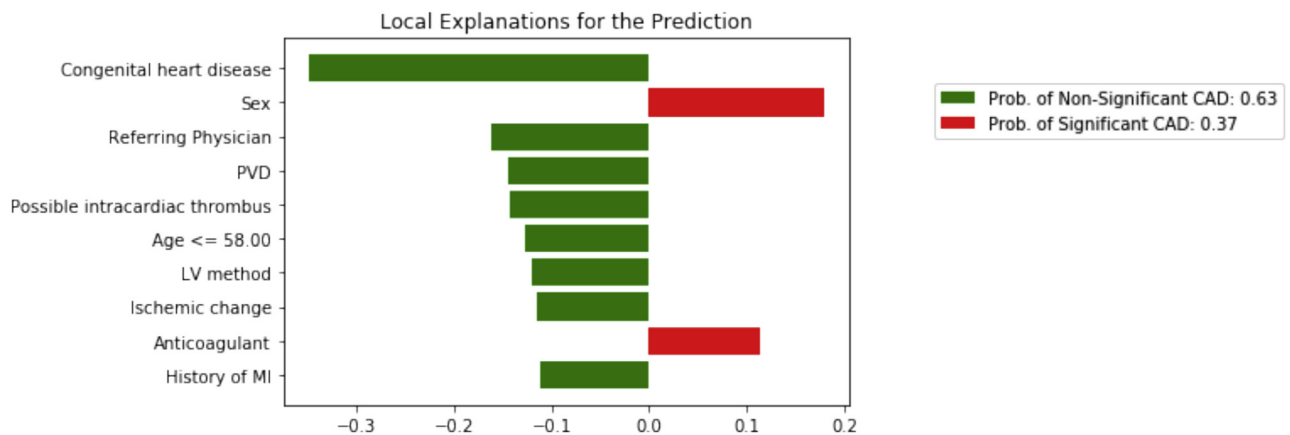
Figure 3 Example of a true negative case (i.e. label and prediction result are both non-significant for CAD) by applying the LIME framework to the LightGBM model to generate an explanation for each prediction result, including the predicted probability and the contribution of each predictor to the prediction result.

support program available to referring physicians and/or triage staff of cardiac catheterization laboratories could optimize the patient population undergoing expensive and invasive cardiac investigations. If a patient is deemed to be "low risk" as per the prediction model, then an invasive coronary angiogram may not be the ideal next step in the diagnostic algorithm. Triage staff and/or referring physicians could then be directed to a more appropriate test (ie, CCTA or functional imaging) depending on patient eligibility. Implementation of this predictive model into the routine triage of patients referred for invasive angiograms could significantly reduce costs to the healthcare system and improve patient safety. As an example, review of the registry data used to train our model demonstrates that there are about 2474 elective outpatient invasive angiograms undertaken annually that would be eligible for application of this predictive model. Within this population, the rate of nonobstructive disease is 45.72%. Based on the predictive model's NRI of 44.70% compared to current clinical practice, implementation of this model could result in an absolute reduction of up to 794 "potentially" unnecessary invasive angiograms in our region. This would result in 1 less death, stroke, or MI per year based on the cited risks of invasive angiography.[23] Furthermore, this strategy would offer an estimated cost savings of US$1,149,712 even if all 794 patients underwent a CCTA as an alternative test (direct cost difference is US$1448 less for CCTA compared to invasive angiogram) every year in this health region.[24] Given that all 794 patients may not be candidates for CCTA (ie, atrial fibrillation as a contraindication in approximately 10%) or may have an inadequate assessment (ie, high calcium score), the real-world benefits may be more modest but still significant. External validation of this predictive model, incorporation into a user-friendly decision support program, and evaluation of implementation are required next steps prior to scale and spread of this model.

Finally, no available variable was discounted. Traditional clinical risk scores or prediction models use a finite number of explanatory variables owing to cohort size and statistical limitations.[25] However, the strength of the machine-learning model is that it can incorporate a large number of variables, including highly correlated ones. We have been able to include variables that have been previously discounted regarding their importance as a predictor of obstructive CAD. Our analysis reveals that several nontraditional variables (including congenital heart disease, need for a translator, referring physician, and the month in which a patient is referred) are important in successfully predicting obstructive vs nonobstructive CAD. Clinically it makes sense that patients referred for invasive angiography with a suspicion or known congenital heart disease are less likely to have obstructive CAD. These patients are often younger and in need of structural interventions rather than coronary revascularization. The variable "need for a translator" indicates a person who experiences a language barrier when accessing health care, which in Canada frequently indicates the individual is of indigenous heritage, an immigrant, or a refugee, all 3 of which can face significant marginalization and barriers in accessing health care, which in turn have long been associated with negative health outcomes.[26] The predictive power of referring physician is a result of practice variation among physicians in this study's catchment area. Practice variation is well documented in medicine and a descriptive analysis of our data confirms that some physicians refer patients with nonobstructive CAD for an angiogram at a higher rate than their colleagues. This model was able to learn these referral patterns from the data. Finally, the month of referral is a predictor of obstructive CAD. This variable has likely been recognized by healthcare providers practicing in high-volume centers. As an example, it is more common for patients with obstructive CAD (ie, concerning symptom profile, whether reported or not) to present during months of the year associated with the holidays. Therefore, given reduced referrals and clinical "slowdowns" during the months of December, July, and August, the rate of significant disease identified increases.

Furthermore, this model can account for complex interactions between and within included variables, as outlined in Supplemental Appendix 2. As an example, when creatinine

level is beyond 120 µmol/L, the probability of obstructive CAD shows strong dependence on CCS class and is almost independent of creatinine level; however, when creatinine level is below 120 µmol/L, the probability of significant CAD is dependent on the values of both predictors. Such an interaction could not be determined using traditional statistical modeling.

While this machine learning–enabled predictive model achieves better performance than existing clinical risk scores, it does have a number of limitations. First, this model requires external validation before being implemented clinically. Further, the model could not necessarily be applied in other health regions without some local adaptation, since the model must learn physician referral patterns from local data (ie, the model would need to be retrained in each new jurisdiction to accurately model local practice variation). Moreover, although social deprivation variables are important predictors in our model, such variables may not be available in every region (though other indexes of social deprivation, such as the CDC's Social Vulnerability Index, may provide an adequate analogue). Second, given the number of variables included, this prediction model is not ideal as a "bedside" risk assessment, limiting its deployment to regions able to implement an online referral program for invasive testing. The outputs from this proposed model could help inform point-of-care, online decision support regarding invasive vs noninvasive testing in the work-up of stable outpatients with possible obstructive CAD. Third, the social deprivation index we relied upon provides area-level measure of deprivation and so is only a proxy for individual-level measures. However, we addressed this limitation, at least in part, by using full 6-digit postal codes to link to the smallest area unit possible (dissemination area), minimizing the risk that measurement error could result in an ecological fallacy. Finally, if the model were to be deployed clinically, it would require some level of updating over time. This is because the model might contribute to reducing factors like practice variation, which would alter the predictive power of the referring physician variable, potentially generating predictions influenced by a previous state that no longer holds. Addressing this limitation will require either manual retraining of the model once a change in practice variation has been detected or the implementation of an incremental learning approach,[27] though such approaches are not without their challenges.[28]

Conclusion

A novel prediction tool, using machine learning to analyze an established regional cardiac registry, is more accurate than traditional risk scores at predicting the presence of obstructive CAD in outpatients referred for invasive coronary angiography. This model, when coupled with point-of-care, online decision support, could improve the diagnostic yield of invasive coronary angiography, thus improving patient safety and reducing healthcare costs.

Disclosures

The authors have no conflicts to disclose.

Authorship

All authors attest they meet the ICMJE criteria for authorship.

Patient Consent

Patient consent was waived for this research as all data were collected from retrospective databases and no identifying information was used.

Ethics Statement

The authors designed the study and gathered and analyzed the data according to the Helsinki Declaration guidelines on human research. The research protocol used in this study was reviewed and approved by the Hamilton Integrated Research Ethics Board #4697.

Appendix
Supplementary data

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.cvdhj.2021.12.001.

References

1. Patel MR, Peterson ED, Dai D, et al. Low diagnostic yield of elective coronary angiography. N Engl J Med 2010;362:886–895.
2. Levitt K, Guo H, Wijeysundera HC, et al. Predictors of normal coronary arteries at coronary angiography. Am Heart J 2013;166:694–700.
3. Anderson GM, Pinfold SP, Hux JE, Naylor CD. Case selection and appropriateness of coronary angiography and coronary artery bypass graft surgery in British Columbia and Ontario. Can J Cardiol 1997;13:246–252.
4. Natarajan MK, Gafni A, Yusuf S. Determining optimal population rates of cardiac catheterization: a phantom alternative? CMAJ 2005;173:49–52.
5. Ivers NM, Schwalm JD, Bouck Z, et al. Interventions supporting long term adherence and decreasing cardiovascular events after myocardial infarction (ISLAND): pragmatic randomised controlled trial. BMJ 2020;369:m1371.
6. Schwalm JD, Ivers Nm, Natarajan MK, et al. Cluster randomized controlled trial of Delayed Educational Reminders for Long-term Medication Adherence in ST-Elevation Myocardial Infarction (DERLA -STEMI). Am Heart J 2015; 170:903–913.
7. Ivers N, Schwalm JD, Jackevicius C, Guo H, Tu JV, Natarajan MK. Length of initial prescription at hospital discharge and long-term medication adherence for patients with coronary artery disease: a population-level retrospective cohort study. Can J Cardiol 2013;29:1408–1414.
8. Matheson FI, Dunn JR, Smith KLW, Moineddin R, Glazier RH. Ontario Marginalization Index-User Guide, version 1.0. St. Michael's Hospital: Ontario: Centre for Research on Inner City Health; 2016.
9. Natarajan MK, Xie C, Gafni A, et al. Development and validation of a risk score to manage patients referred for coronary angiography. Abstract. Can J Cardiol 2006;22(Suppl D):683.
10. Tibshirani R. 1996. Regression shrinkage and selection via the Lasso. J R Stat Soc Series B Stat Methodol 1996;58:267–288.
11. Breiman L. Random forests. J Mach Learn 2001;45:5–32.
12. Schapire RE. A brief introduction to boosting. IJCAI 1999;99:1401–1406.
13. Friedman JH. Greedy function approximation: a gradient boosting machine. Annals of Statistics 2001;29:1189–1232.

14. Ke G, Meng Q, Finley T, et al. LightGBM: a highly efficient gradient boosting decision tree. Oral presentation at the 31st Conference on Neural Information Processing Systems, NIPS 2017. CA: Long Beach; 2017.

15. Chen T, Guestrin C. 2016. XGBoost: a scalable tree boosting system, https://arxiv.org/pdf/1603.02754.pdf.

16. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature 1986;323:533–536.

17. Thomas LE, O'Brien EC, Piccini JP, D'Agostino RB, Pencina MJ. Application of net reclassification index to non-nested and point-based risk prediction models: a review. Eur Heart J 2019;40:1880–1887.

18. Altmann A, Toloşi L, Sander O, Lengauer T. Permutation importance: a corrected feature importance measure. Bioinformatics 2010;26:1340–1347.

19. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second edition. Springer Science & Business Media; 2009.

20. Lundberg SM, Lee S. A unified approach to interpreting model predictions. In: Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, Garnett R, eds. Advances in Neural Information Processing Systems 30. Curran Associates, Inc; 2017. p. 4765–4774.

21. Štrumbelj K, Kononenko I. Explaining prediction models and individual predictions with feature contributions. Knowl Inf Syst 2014;41: 647–665.

22. Ribeiro MT, Singh S, Guestrin C. 'Why should I trust you?' Explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016; p. 1135–1144.

23. Tavakol M, Ashraf S, Brener SJ. Risks and complications of coronary angiography: a comprehensive review. Glob J Health Sci 2012;4:65–93.

24. Rudziński PN, Kruk M, Kępka C, et al. Assessing the value of coronary artery computed tomography as the first-line anatomical test for stable patients with indications for invasive angiography due to suspected coronary artery disease. Initial cost analysis in the CAT-CAD randomized trial. J Cardiovasc Comput Tomogr 2020;14:75–79.

25. Goldstein BA, Navar AM, Carter RE. Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges. Eur Heart J 2017;38:1805–1814.

26. Bowen S. Language barriers in access to healthcare. Government of Canada. Published November 2001. https://www.canada.ca/en/health-canada/services/health-care-system/reports-publications/health-care-accessibility/language-barriers.html. Accessed December 2020.

27. Gepperth A, Hammer B. Incremental learning algorithms and applications. European Symposium on Artificial Neural Networks (ESANN). 2016. Bruges, Belgium. hal-01418129. https://hal.archives-ouvertes.fr/hal-01418129/file/article.pdf.

28. Lee CS, Lee AY. Clinical applications of continual learning machine learning. Lancet Digit Health 2020;2:279–281.