

Predicting Long-term Mortality in Patients with Stable Angina Across the Spectrum of Dysglycemia: A Machine Learning Approach

Yu-Hsuan Li

Taichung Veterans General Hospital

Wayne Huey-Heng Sheu

Taichung Veterans General Hospital

Wen-Chao Yeh

National Tsing Hua University

Yung-Chun Chang

Taipei Medical University

I-Te Lee (✉ itlee@vghtc.gov.tw)

Taichung Veterans General Hospital <https://orcid.org/0000-0003-2665-3635>

Original investigation

Keywords: angiography, Brier score, Harrell's C-index, least absolute shrinkage and selection operator, machine learning, oral glucose tolerance test

Posted Date: November 5th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-101450/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: We aimed to develop and validate a model to predict mortality in patients with stable angina across the spectrum of dysglycemia.

Methods: A total of 1479 patients admitted for coronary angiography due to angina were enrolled. All-cause mortality was followed up and served as the primary endpoint. We compared the performance of different machine learning models for survival analysis and used least absolute shrinkage and selection operator (LASSO) to select important features. Performance was evaluated using Harrell's C-index and the Brier score. The models were validated with five-fold cross validation to predict long-term mortality.

Results: The features selected by LASSO were age, heart rate, plasma glucose levels at 30 min and 120 min during an oral glucose tolerance test (OGTT), use of angiotensin II receptor blockers, use of diuretics, and smoking history. The best performing model was built using a random survival forest with selected features. It had a good discriminative ability (Harrell's C-index: 0.829) and acceptable calibration (Brier score: 0.08) for predicting long-term mortality. Among patients with obstructive coronary artery disease confirmed by angiography, our model outperformed the Global Registry of Acute Coronary Events discharge score for mortality prediction (Harrell's C-index: 0.829 vs. 0.739, $P < 0.001$).

Conclusions: We developed a machine learning model to predict long-term mortality among patients with stable angina. With the integration of OGTT, the model could help to identify patients with stable angina at high risk of mortality across the spectrum of dysglycemia.

Background

Angina is a common symptom of ischemic heart diseases [1]. Based on coronary angiography, ischemic heart disease can be stratified as obstructive coronary artery disease (CAD) and non-obstructive CAD [1, 2]. Obstructive CAD has an established association with increased cardiovascular mortality and is the target of treatment in contemporary practice [3, 4]. However, a growing body of evidence has revealed that almost 50%~70% of patients with angina have no significant obstruction on coronary angiography [5, 6], and they are associated with a high risk of all-cause mortality and cardiovascular events [7, 8]. Since there is heterogeneity in pathogenesis among patients with stable angina, the European Society of Cardiology published associated guidelines to address differences in diagnosis and prognosis [9]. However, the optimal management strategy for patients with stable angina is still under debate, and a risk stratification model to identify patients at high risk and to tailor personal management is warranted [10–12].

Current risk stratification models for patients with ischemic heart disease are mainly based on patients with obstructive CAD, ranging from stable CAD to acute coronary syndrome [13]; however, a predictive model focused on patients with stable angina is still lacking. In addition, glucose perturbation represents a well-documented risk factor for atherosclerosis and is commonly seen in patients with ischemic heart disease [14]. A recent survey showed that dysglycemia detected by the oral glucose tolerance test (OGTT)

is prevalent in patients with CAD [15]. However, contemporary predictive models for patients with ischemic heart disease seldom integrate glucose indices into their parameters. For example, the Global Registry of Acute Coronary Events (GRACE) discharge score [16], which is a widely used risk score for patients with acute coronary syndromes and has recently had its predictive ability validated for patients with stable CAD [17], has long been criticized for not including glucose indices [18]. Herein, we aimed to use data collected from patients with stable angina undergoing coronary angiography and OGTT to build a machine learning model for predicting long-term mortality.

Methods

Setting and participants

Data were obtained from a prospective, observational study conducted at Taichung Veterans General Hospital. The study enrolled adult patients admitted for coronary angiography between April 2009 and December 2018 due to symptoms of typical angina and under suspicion of ischemic heart disease by cardiologists. Patients were excluded if they had undergone coronary artery bypass graft or they had been diagnosed with diabetes. All patients underwent OGTT after overnight fasting at an outpatient visit after discharge, and glucose levels were tested at fasting, at 30 minutes, and at 120 minutes during the OGTT. Normal glucose regulation was determined as the fasting plasma glucose (FPG) < 100 mg/dL and the glucose level at 120 minutes (OGTT 120 min) < 140 mg/dL. Newly diagnosed diabetes was defined as the FPG \geq 126 mg/dL or the OGTT 120 min \geq 200 mg/dL. Prediabetes was defined in patients with a glucose regulation between that of normal glucose regulation and diabetes. Obstructive CAD was defined as at least a lesion with \geq 50% stenosis and non-obstructive CAD was defined as no any lesion with \geq 50% stenosis in coronary angiography reports. Baseline data collected upon admission, coronary angiography reports, and OGTT results and the medications at the outpatient visit were considered as candidate variables for analysis. Variables with > 20% missing values were excluded, and multivariable imputation using the chained equation was applied for the remaining variables [19]. Mortality data up to December 2019 were retrieved from the Collaboration Center of Health Information Application, Department of Health, Executive Yuan, Taiwan, and served as the outcome of interest.

Model development and evaluation

Model development consisted of two parts. In the first part, we used all available variables to develop prediction models. We chose different machine learning methods, including random survival forest (RSF), gradient boosting machine learning algorithm (XGBoost) for survival analysis, and discrete-time survival model for neural networks [20, 21]. RSF is an ensemble tree method for analysis of right censored survival data. It can handle complex interactions among variables, including mixed data types and nonlinear relationships between variables. XGBoost is a novel boosting tree-based ensemble algorithm, whose performance is iteratively improved through optimization of a customized objective function. To handle time-to-event data, the objective function of XGBoost was set as Cox regression. The discrete-time survival model for neural networks was implemented in the Kera deep learning framework and was

trained with the maximum likelihood method using minibatch stochastic gradient descent. The likelihood function was used as the loss function, and it naturally incorporated non-proportional hazards. The prediction of these machine learning methods was the hazard ratio of each individual, and their results were compared with those of Cox proportional hazards models.

In the second part, we used features selected by the least absolute shrinkage and selection operator (LASSO)-derived Cox proportional hazards model to construct the predictive model. LASSO regularization has the ability to shrink the estimations and force certain coefficients to zero, thereby keeping only the important features [22]. Significant predictors with P values < 0.05 in the LASSO-derived Cox proportional hazards model were selected as input variables for the machine learning methods mentioned above.

The prediction models were internally validated using five-fold cross validation. Discrimination ability was assessed using the Harrell's C-index [23]. Calibration was evaluated using the Brier score [5]. The Brier score measures the mean squared difference between the predicted probability and the actual outcome, with a lower score indicating better calibrated predictions. To explain the contribution of each variable to the best performing model, Shapley values were utilized [24, 25]. Based on game theory, Shapley values can explain a model's prediction by computing the importance of each feature to the prediction.

Comparison with GRACE discharge score

For patients with obstructive CAD confirmed by coronary angiography, we calculated their GRACE discharge score and compared its performance with that of the best performing model. Net reclassification improvement (NRI) and integrated discrimination improvement (IDI) were used to evaluate the improvement in predictive power of our final model compared with the GRACE discharge score [26]. Analyses were performed using R 3.4 software (The R Project for Statistical Computing, Vienna, Austria) and Python (version 3.6).

Results

Characteristics of enrolled patients

A total of 1479 patients were included in the analyses, and 157 patients (10.6%) died during a median follow up of 6 years (interquartile range, 3.1–9.1 years). Table 1 lists the cohort's baseline characteristics. Briefly, patients who died tended to be older, with higher heart rates, higher urine albumin-to-creatinine ratios, higher uric acid levels, higher high-density lipoprotein cholesterol levels, higher FPG, and higher OGTT 120 min, whereas their body mass index, diastolic blood pressure, triglycerides, glutamic pyruvic transaminase, hemoglobin, and estimated glomerular filtration rate were lower compared with patients who survived. Patients who died during follow up had a lower prevalence of beta-blocker use and a higher prevalence of new-diagnosed diabetes, CAD history before admission, smoking history, and uses of angiotensin-converting enzyme inhibitors, angiotensin II receptor blockers (ARB), alpha blockers, and diuretics compared with patients who survived.

Table 1
Baseline characteristics of the study population

	Survival (N = 1322)	Death (N = 157)	P
Age (years)	58 ± 10	72 ± 12	< 0.001
Male (n, %)	238 (18%)	20 (13%)	0.122
Body height (cm)	165 ± 8	162 ± 8	< 0.001
Body weight (kg)	71.3 ± 11.7	66.4 ± 12.2	< 0.001
BMI (kg/m ²)	26.1 ± 3.4	25.2 ± 3.9	0.004
Waist (cm)	90.6 ± 8	90.5 ± 9.4	0.923
SBP (mmHg)	126 ± 17	128 ± 20	0.226
DBP (mmHg)	74 ± 10	69 ± 11	< 0.001
Heart rate (beat/min)	71 ± 11	75 ± 13	< 0.001
UACR (mg/g)	36 ± 176	111 ± 274	< 0.001
Uric acid (mg/dL)	6.6 ± 1.6	7.3 ± 2.0	< 0.001
Triglycerides (mg/dL)	147 ± 105	122 ± 78	0.004
Total cholesterol (mg/dL)	173 ± 39	169 ± 35	0.253
HDL cholesterol (mg/dL)	46 ± 12	49 ± 15	0.007
LDL cholesterol (mg/dL)	106 ± 34	103 ± 31	0.501
eGFR (mL/min/1.73 m ²)	82 ± 22	63 ± 21	< 0.001
GPT (U/L)	32 ± 32	26 ± 19	0.028
Hemoglobin (g/dL)	13.9 ± 1.5	13.2 ± 1.9	0.001
WBC (/μL)	6898 ± 2772	6923 ± 2176	0.914

*CAD history means patients had undergone percutaneous coronary intervention for obstructive CAD before admission

†Significant stenosis defined as stenosis ≥ 50%

ACE = angiotensin-converting enzyme, ARB = angiotensin II receptor blocker, BMI = body mass index, BP = blood pressure, eGFR = estimated glomerular filtration rate, FPG = fasting plasma glucose, HDL = high-density lipoprotein, LDL = low-density lipoprotein, OGTT = oral glucose tolerance test, SBP = systolic blood pressure, DBP = diastolic blood pressure, CCB = calcium channel blocker, CAD = coronary artery disease, WBC = white blood cell count, CK = creatine kinase, GPT = glutamic pyruvic transaminase, HbA1c = glycated hemoglobin, UACR = albumin-to-creatinine ratio.

	Survival (N = 1322)	Death (N = 157)	P
CK (U/L)	151 ± 314	180 ± 370	0.373
CKMB (U/L)	9 ± 16	11 ± 16	0.241
Troponin-T (ng/L)	3.9 ± 19.5	1.4 ± 3.6	0.209
OGTT (mg/dL)			
Glucose 0 min	95 ± 14	100 ± 19	< 0.001
Glucose 30 min	169 ± 32	169 ± 36	0.962
Glucose 120 min	145 ± 50	166 ± 59	< 0.001
HbA1c (%)	5.8 ± 0.6	6.1 ± 0.8	< 0.001
HbA1c (mmol/mol)	40.8 ± 6.4	43.4 ± 8.5	
Glucose status (n, %)			< 0.001
Normal glucose regulation	428 (32.4%)	43 (27.4%)	
Prediabetes	637 (48.2%)	51 (32.5%)	
Diabetes	257 (19.4%)	63 (40.1%)	
Smoking status (n, %)			< 0.001
Non-smoker	590 (44.6%)	52 (33.1%)	
Smoker	318 (24.1%)	18 (11.5%)	
Ex-smoker	414 (31.3%)	87 (55.4%)	
Medication (n, %)			
Antiplatelet	1226 (92.8%)	147 (93.6%)	0.830
ACE inhibitor	278 (21.0%)	46 (29.3%)	0.023
ARB	415 (31.4%)	69 (43.9%)	0.002

*CAD history means patients had undergone percutaneous coronary intervention for obstructive CAD before admission

†Significant stenosis defined as stenosis ≥ 50%

ACE = angiotensin-converting enzyme, ARB = angiotensin II receptor blocker, BMI = body mass index, BP = blood pressure, eGFR = estimated glomerular filtration rate, FPG = fasting plasma glucose, HDL = high-density lipoprotein, LDL = low-density lipoprotein, OGTT = oral glucose tolerance test, SBP = systolic blood pressure, DBP = diastolic blood pressure, CCB = calcium channel blocker, CAD = coronary artery disease, WBC = white blood cell count, CK = creatine kinase, GPT = glutamic pyruvic transaminase, HbA1c = glycated hemoglobin, UACR = albumin-to-creatinine ratio.

	Survival (N = 1322)	Death (N = 157)	P
Alpha blocker	54 (4.1%)	16 (10.2%)	0.001
Beta blocker	368 (27.8%)	23 (14.6%)	0.001
CCB	689 (52.1%)	78 (49.7%)	0.622
Diuretics	153 (11.6%)	44 (28.0%)	< 0.001
CAD history* (n, %)	140 (10.6%)	26 (16.6%)	0.035
Number of coronary arteries with significant stenosis [†] (n, %)			0.007
0	611 (46.2%)	51 (32.5%)	
1	356 (26.9%)	48 (30.6%)	
2	245 (18.5%)	42 (26.8%)	
3	110 (8.3%)	16 (10.2%)	
Involved coronary arteries (n, %)			
Left artery	549 (41.5%)	81 (51.6%)	0.020
Right artery	271 (20.5%)	37 (23.6%)	0.429
Left circumflex	356 (26.9%)	62 (39.5%)	0.001
*CAD history means patients had undergone percutaneous coronary intervention for obstructive CAD before admission			
[†] Significant stenosis defined as stenosis \geq 50%			
ACE = angiotensin-converting enzyme, ARB = angiotensin II receptor blocker, BMI = body mass index, BP = blood pressure, eGFR = estimated glomerular filtration rate, FPG = fasting plasma glucose, HDL = high-density lipoprotein, LDL = low-density lipoprotein, OGTT = oral glucose tolerance test, SBP = systolic blood pressure, DBP = diastolic blood pressure, CCB = calcium channel blocker, CAD = coronary artery disease, WBC = white blood cell count, CK = creatine kinase, GPT = glutamic pyruvic transaminase, HbA1c = glycated hemoglobin, UACR = albumin-to-creatinine ratio.			

Feature importance and model performance

The significant predictors selected by the LASSO-derived Cox proportional hazards model included age, heart rate, glucose level at 30 minutes (OGTT 30 min), OGTT 120 min, CAD history, smoking history, use of angiotensin II receptor blockers, and use of diuretics (Table 2). The performance of the predictive models is shown in Table 3. The RSF model after feature selection had the highest Harrell's C-index (Harrell's C-index: 0.829) and acceptable calibration (Brier score: 0.08). Figure 1 shows the area under the

operating curve and the calibration plot of predicted risks at 10 years. Shapley values of the variables in the best performing model are shown in Fig. 2.

Table 2
Features selected by the LASSO-derived Cox proportional hazards model.

	HR	95%CI	P
Age	1.06	(1.04–0.08)	< 0.001
Heart rate	1.02	(1.01–1.04)	< 0.001
OGTT 30 min	0.98	(0.97–0.99)	< 0.001
OGTT 120 min	1.01	(1.01–1.02)	< 0.001
CAD history	1.66	(1.03–2.66)	0.040
Smoking history	1.31	(1.06–1.65)	0.013
ARB use	1.74	(1.18–2.55)	0.008
Diuretic use	1.57	(0.01–0.87)	0.048
ARB = angiotensin II receptor blocker, CAD = coronary artery disease, CI = confidence interval, HR = hazard ratio, LASSO = least absolute shrinkage and selection operator; OGTT = oral glucose tolerance test			

Table 3
Performance of different machine learning methods

	Harrell's C-index	Brier score
With all variables		
Cox regression	0.774	0.069
RSF	0.804	0.082
XGBoost	0.788	0.033
DNN	0.750	0.102
With selected features		
Cox regression	0.741	0.076
RSF	0.829	0.080
XGBoost	0.794	0.056
DNN	0.796	0.106
RSF = random survival forest, XGBoost = extreme gradient boosting, DNN = dense neural network		

Comparison with GRACE discharge score

Among patients with obstructive CAD confirmed by angiography during admission, Harrell's C-index of the best performing model was significantly greater than the GRACE discharge score (0.829 vs. 0.739, respectively; $P < 0.001$). The NRI (0.328, 95% confidence interval [CI]: 0.096–0.583, $P = 0.027$) and IDI (0.135, 95% CI: 0.068–0.203, $P = 0.007$) indices also showed improvement in predictive ability compared with the GRACE discharge score (Table 4).

Table 4
Comparison between GRACE discharge score and the best-performing model.

Model	Harrell's C-index (95%CI)	P	Absolute IDI (95% CI)	P	NRI (95% CI)	P
GRACE score	0.739 (0.683,0.796)					
RSF	0.829 (0.790,0.869)	< 0.001	0.135 (0.068, 0.203)	0.007	0.328 (0.096, 0.583)	0.027

CI = confidence interval; GRACE = Global Registry of Acute Coronary Events; IDI = Integrated discrimination improvement; NRI = Net reclassification improvement; RSF = Random survival forest with selected features

Discussion

In this study, we built a machine learning based model to predict all-cause mortality among patients with stable angina across the spectrum of dysglycemia. With glucose indices obtained from OGTT and other available clinical data, this model showed good discrimination and accuracy in predicting long-term mortality after coronary angiography. To the best of our knowledge, this study is among the first to compare state-of-the-art machine learning methods to predict survival in patients with stable angina, with an emphasis on OGTT results as important parameters.

For patients with obstructive CAD, several predictive models have been developed to predict major cardiovascular events and mortality. For example, the GRACE discharge score has been recently validated its accuracy to predict mortality 2 years after coronary angiography with an area under curve of 0.61 for patients with stable CAD [17]. The ABC-CHD score [27] with risk factors identified by the Cox proportional hazards model, including age, biomarkers (N-terminal prohormone of brain natriuretic peptide and troponin-T), and clinical histories (smoking, diabetes, and presence of peripheral artery disease), has good discriminatory ability (Harrell's C-index: 0.71) and calibration for three-year mortality. However, previous models seldom include patients with non-obstructive CAD. It has been reported that more than half of patients with angina have no obstructive CAD during coronary angiography [28–30]. Since a

sizable proportion of patients have angina has non-obstructive CAD, our model, which was derived from a cohort in which almost 50% of patients had non-obstructive CAD, is more representative of real-world patients with stable angina. Even for patients with obstructive CAD, our model outperformed the GRACE discharge score to predict long-term mortality.

The predictors detected by our LASSO-derived Cox proportional hazards models were age, diuretic use, ARB use, heart rate at admission, OGTT 120 min, and OGTT 30 min. Although some are well-established risk factors for mortality and have been included in previous predictive models for patients with CAD, using OGTT results as parameters for risk stratification has not been investigated before. Our study showed high accuracy in mortality prediction with the integration of OGTT results. According to the EUROASPIRE study, OGTT 120 min is a predictor of major cardiovascular events and mortality for patients without diabetes [31]. Chattopadhyay et al. [32] showed that with the adjustment of OGTT 120 min, GRACE discharge score has an improved prognostic ability for patients with acute coronary syndrome. Similar to above studies, our model containing OGTT 120 min outperformed the GRACE discharge score for predicting mortality among patients with stable angina, supporting the importance of OGTT 120 min for mortality prediction among patients with ischemic heart disease. Our model also revealed potentially unidentified predictors, such as OGTT 30 min. OGTT 30 min has predictive value for developing type 2 diabetes and is associated with inflammatory markers [33, 34]; however, its role in mortality prediction has not been previously evaluated. Based on the Shapley value derived from our model, OGTT 30 min also contributed to mortality prediction in patients with stable angina. Further prospective studies are warranted to elucidate the prognostic value of OGTT 30 min.

There are several clinical applications of this model. Our model is derived from patients with stable angina. More than half of patients with stable angina have non-obstructive CAD, and approximately 20%~40% of patients with CAD still suffer from angina symptoms after revascularization [4]. There is heterogeneity in prognosis among patients with stable angina, and it is important to stratify their risk and tailor their management strategy. However, contemporary clinical practice mainly focuses on prevention and management of obstructive CAD [35], despite the fact that the risk of major cardiovascular events and mortality among patients with non-obstructive CAD is increased [36]. Our model could help to identify patients with stable angina at a high risk of mortality. In addition, our model emphasized the importance of OGTT. Screening for dysglycemia using OGTT in patients undergoing percutaneous coronary intervention has long been proposed and is also recommended in European Society of Cardiology guidelines [36, 37]. However, adherence to this recommendation is poor [38], partially because the prognostic role of OGTT is less clear. Our model, which adds prognostic value to the OGTT for patients with stable angina, could increase adherence to this recommendation.

The major strength of this study is that we used the LASSO-derived Cox proportional hazards model for feature selection and advanced machine learning methods for model development. Only six variables were needed in our model after utilization of LASSO regularization, and most of them, except OGTT, were available from electronic health records, making it a convenient tool to implement in clinical practice. The best performing machine method in our study was RSF. Previous predictive models for CAD were usually

built using Cox proportional hazards models; however, several assumptions must be met before applying the Cox proportional hazards model. Conversely, machine learning methods, such as RSF, can handle non-linear, complex relationships between features without assumptions, thus widening their clinical application. However, there are still some limitations which should be highlighted in the current study. First, this cohort has been conducted since 2009, and contemporary anti-diabetic medications, such as sodium glucose co-transporters 2 inhibitors and glucagon-like peptide-1 receptor agonists, which can reduce mortality risk in patients with type 2 diabetes, were seldom been prescribed. Second, our model has not been externally validated with other independent datasets, so its performance in other datasets is unknown.

Conclusions

We developed a machine learning model, containing OGTT results and other clinically available parameters to predict all-cause mortality among patients with stable angina. This model could help to identify patients at a high risk of mortality. With the integration of glucose indices from OGTT, patients with dysglycemia could be identified early, enabling their risk of mortality to be accurately evaluated.

Abbreviations

ARB: angiotensin II receptor blockers; CAD: coronary artery disease; FPG: fasting plasma glucose; GRACE: Global Registry of Acute Coronary Events; IDI: integrated discrimination improvement; LASSO: least absolute shrinkage and selection operator; NRI: net reclassification improvement; OGTT 120 min: glucose level at 120 minutes; OGTT 30 min: glucose level at 30 minutes; OGTT: oral glucose tolerance test; RSF: random survival forest; XGBoost: gradient boosting machine learning algorithm;

Declarations

Ethics approval and consent to participate:

The study complied with the Declaration of Helsinki and was approved by the Institutional Review Board of Taichung Veterans General Hospital. Written consent was obtained from each patient before the study procedures were performed.

Consent for publication:

Not applicable.

Availability of data and materials:

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests:

The authors declare that they have no competing interests.

Funding:

This work was supported by Taichung Veterans General Hospital, Taichung, Taiwan (grant number TCVGH-1093503C), the Ministry of Science and Technology, Taiwan (grant number MOST 109-2314-B-075A-004), and National Health Research Institute (grant number NHRI-EX109-10927HT). The funders had no role in the decision to submit the manuscript for publication.

Authors' contributions:

Y.L. contributed to the study design, data collection, data analysis, and drafting of the manuscript. W.H.S. contributed to the study design and data collection. W.Y. contributed to the data analysis. Y.C. contributed to data analysis, and drafting and editing of the manuscript. I.L. contributed to the study design, data collection, and editing of the manuscript. All authors performed a critical revision of the manuscript for important intellectual content. I.L. is the guarantor of this work and takes full responsibility for the work as a whole, including the study design, access to data, and the decision to submit and publish the manuscript.

Acknowledgements:

The authors acknowledge the Cardiovascular Center of Taichung Veterans General Hospital for their support.

References

1. Ford TJ, Berry C. Angina: contemporary diagnosis and management. *Heart* 2020;106:387-398.
2. Ferrari R, Camici PG, Crea F, Danchin N, Fox K, Maggioni AP, et al. Expert consensus document: A 'diamond' approach to personalized treatment of angina. *Nat Rev Cardiol* 2018;15:120-132.
3. Sundaram V, Bloom C, Zakeri R, Halcox J, Cohen A, Bowrin K, et al. Temporal trends in the incidence, treatment patterns, and outcomes of coronary artery disease and peripheral artery disease in the UK, 2006-2015. *Eur Heart J* 2020;41:1636-1649.
4. Marzilli M, Crea F, Morrone D, Bonow RO, Brown DL, Camici PG, et al. Myocardial ischemia: From disease to syndrome. *Int J Cardiol* 2020;314:32-35.
5. Brier GW. Verification of forecasts expressed in terms of probability. *Monthly Weather Review* 1950;78:1-3.
6. Pepine CJ, Anderson RD, Sharaf BL, Reis SE, Smith KM, Handberg EM, et al. Coronary microvascular reactivity to adenosine predicts adverse outcome in women evaluated for suspected ischemia results from the National Heart, Lung and Blood Institute WISE (Women's Ischemia Syndrome Evaluation) study. *J Am Coll Cardiol* 2010;55:2825-2832.

7. Jespersen L, Hvelplund A, Abildstrøm SZ, Pedersen F, Galatius S, Madsen JK, et al. Stable angina pectoris with no obstructive coronary artery disease is associated with increased risks of major adverse cardiovascular events. *Eur Heart J* 2012;33:734-744.
8. Schulman-Marcus J, Hartaigh B, Gransar H, Lin F, Valenti V, Cho I, et al. Sex-Specific Associations Between Coronary Artery Plaque Extent and Risk of Major Adverse Cardiovascular Events: The CONFIRM Long-Term Registry. *JACC Cardiovasc Imaging* 2016;9:364-372.
9. Knuuti J, Wijns W, Saraste A, Capodanno D, Barbato E, Funck-Brentano C, et al. 2019 ESC Guidelines for the diagnosis and management of chronic coronary syndromes: The Task Force for the diagnosis and management of chronic coronary syndromes of the European Society of Cardiology (ESC). *European Heart Journal* 2019;41:407-477.
10. Maron DJ, Hochman JS, Reynolds HR, Bangalore S, O'Brien SM, Boden WE, et al. Initial Invasive or Conservative Strategy for Stable Coronary Disease. *N Engl J Med* 2020;382:1395-1407.
11. Kunadian V, Chieffo A, Camici PG, Berry C, Escaned J, Maas A, et al. An EAPCI Expert Consensus Document on Ischaemia with Non-Obstructive Coronary Arteries in Collaboration with European Society of Cardiology Working Group on Coronary Pathophysiology & Microcirculation Endorsed by Coronary Vasomotor Disorders International Study Group. *Eur Heart J* 2020;ehaa503. (doi: 10.1093/eurheartj/ehaa503. Online ahead of print)
12. Dweck MR, Newby DE. Non-obstructive coronary artery disease can no longer be ignored. *Eur Heart J Cardiovasc Imaging* 2020;21:489-490.
13. Buccheri S, D'Arrigo P, Franchina G, Capodanno D. Risk Stratification in Patients with Coronary Artery Disease: A Practical Walkthrough in the Landscape of Prognostic Risk Models. *Interv Cardiol* 2018;13:112-120.
14. Kuhl J, Jorneskog G, Wemminger M, Bengtsson M, Lundman P, Kalani M. Long-term clinical outcome in patients with acute coronary syndrome and dysglycaemia. *Cardiovasc Diabetol* 2015;14:120.
15. Ferrannini G, De Bacquer D, De Backer G, Kotseva K, Mellbin L, Wood D, et al. Screening for Glucose Perturbations and Risk Factor Management in Dysglycemic Patients With Coronary Artery Disease-A Persistent Challenge in Need of Substantial Improvement: A Report From ESC EORP EUROASPIRE V. *Diabetes Care* 2020;43:726-733.
16. Shuvy M, Beerli G, Klein E, Cohen T, Shlomo N, Minha S, et al. Accuracy of the Global Registry of Acute Coronary Events (GRACE) Risk Score in Contemporary Treatment of Patients With Acute Coronary Syndrome. *Can J Cardiol* 2018;34:1613-1617.
17. Zhao X, Li J, Xian Y, Chen J, Gao Z, Qiao S, et al. Prognostic value of the GRACE discharge score for predicting the mortality of patients with stable coronary artery disease who underwent percutaneous coronary intervention. *Catheter Cardiovasc Interv* 2020;95 Suppl 1:550-557.
18. Liu XJ, Wan ZF, Zhao N, Zhang YP, Mi L, Wang XH, et al. Adjustment of the GRACE score by HemoglobinA1c enables a more accurate prediction of long-term major adverse cardiac events in acute coronary syndrome without diabetes undergoing percutaneous coronary intervention. *Cardiovasc Diabetol* 2015;14:110.

19. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? *Int J Methods Psychiatr Res* 2011;20:40-49.
20. Chen T, Guestrin C. Xgboost: A scalable tree boosting system. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, 2016*. p.785-794.
21. Gensheimer MF, Narasimhan B. A scalable discrete-time survival model for neural networks. *PeerJ*. 2019;7:e6257.
22. Teppola P, Taavitsainen VM. Parsimonious and robust multivariate calibration with rational function Least Absolute Shrinkage and Selection Operator and rational function Elastic Net. *Anal Chim Acta* 2013;768:57-68.
23. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Stat Med* 2011;30:1105-1117.
24. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence* 2020;2:56-67.
25. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: *31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA*, p. 4765-4774.
26. Pencina MJ, D'Agostino RB, Sr., Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med* 2011;30:11-21.
27. Lindholm D, Lindbäck J, Armstrong PW, Budaj A, Cannon CP, Granger CB, et al. Biomarker-Based Risk Model to Predict Cardiovascular Mortality in Patients With Stable Coronary Disease. *J Am Coll Cardiol* 2017;70:813-826.
28. Arnold SV, Grodzinsky A, Gosch KL, Kosiborod M, Jones PG, Breeding T, et al. Predictors of Physician Under-Recognition of Angina in Outpatients With Stable Coronary Artery Disease. *Circ Cardiovasc Qual Outcomes* 2016;9:554-559.
29. Lanza GA, Crea F. Primary coronary microvascular dysfunction: clinical presentation, pathophysiology, and management. *Circulation* 2010;121:2317-2325.
30. Pepine CJ, Ferdinand KC, Shaw LJ, Light-McGroary KA, Shah RU, Gulati M, et al. Emergence of Nonobstructive Coronary Artery Disease: A Woman's Problem and Need for Change in Definition on Angiography. *J Am Coll Cardiol* 2015;66:1918-1933.
31. Shahim B, De Bacquer D, De Backer G, Gyberg V, Kotseva K, Mellbin L, et al. The Prognostic Value of Fasting Plasma Glucose, Two-Hour Postload Glucose, and HbA(1c) in Patients With Coronary Artery Disease: A Report From EUROASPIRE IV: A Survey From the European Society of Cardiology. *Diabetes Care* 2017;40:1233-1240.
32. Chattopadhyay S, George A, John J, Sathyapalan T. Adjustment of the GRACE score by 2-hour post-load glucose improves prediction of long-term major adverse cardiac events in acute coronary syndrome in patients without known diabetes. *Eur Heart J* 2018;39:2740-2745.
33. Chen G, Zou X, Yao J, Jiang Q, Zhang Y, Tu M, et al. The correlation between the oral glucose tolerance test 30-minutes plasma glucose and risk factors for diabetes and cardiovascular diseases:

- a cross-sectional epidemiological study of diabetes in Fujian Province in the South-East of China. *J Endocrinol Invest* 2011;34:e115-120.
34. Tsuboi A, Minato S, Yano M, Takeuchi M, Kitaoka K, Kurata M, et al. Association of serum orosomucoid with 30-min plasma glucose and glucose excursion during oral glucose tolerance tests in non-obese young Japanese women. *BMJ Open Diabetes Res Care* 2018;6:e000508.
 35. Taqueti VR, Everett BM, Murthy VL, Gaber M, Foster CR, Hainer J, et al. Interaction of impaired coronary flow reserve and cardiomyocyte injury on adverse cardiovascular outcomes in patients without overt coronary artery disease. *Circulation* 2015;131:528-535.
 36. Shahim B, Gyberg V, De Bacquer D, Kotseva K, De Backer G, Schnell O, et al. Undetected dysglycaemia common in primary care patients treated for hypertension and/or dyslipidaemia: on the need for a screening strategy in clinical practice. A report from EUROASPIRE IV a registry from the EuroObservational Research Programme of the European Society of Cardiology. *Cardiovasc Diabetol* 2018;17:21.
 37. Grant PJ, Cosentino F. The 2019 ESC Guidelines on diabetes, pre-diabetes, and cardiovascular diseases developed in collaboration with the EASD: New features and the 'Ten Commandments' of the 2019 Guidelines are discussed by Professor Peter J. Grant and Professor Francesco Cosentino, the Task Force chairmen. *Eur Heart J* 2019;40:3215-3217.
 38. Hu DY, Pan CY, Yu JM, China Heart Survey G. The relationship between coronary artery disease and abnormal glucose regulation in China: the China Heart Survey. *Eur Heart J* 2006;27:2573-2579.

Figures

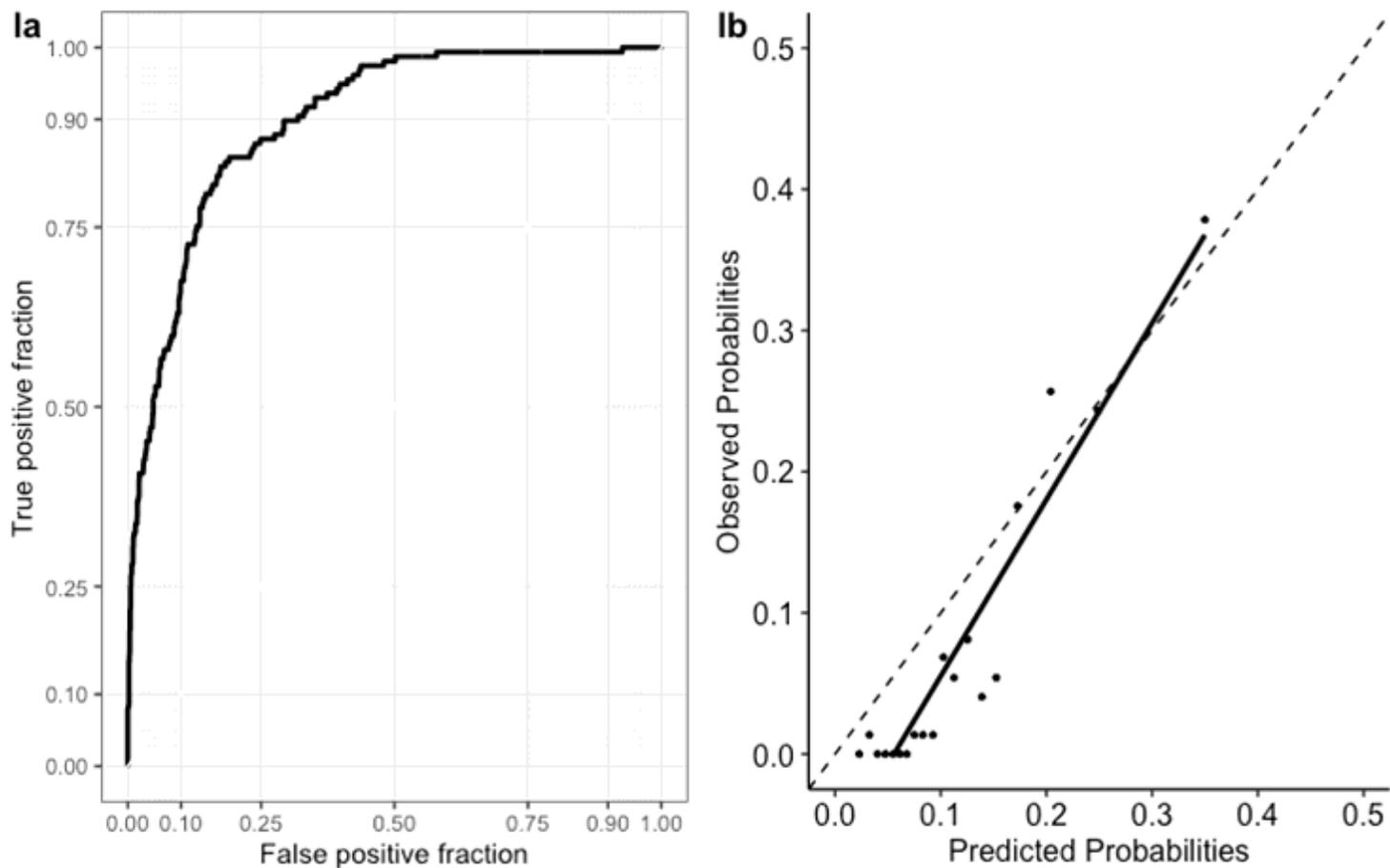


Figure 1

Area under the curve of the best performing model for prediction of all-cause mortality among patients with stable angina (A). The closer to the dotted line indicates the better the prediction in the calibration plot of the best performing model (B).

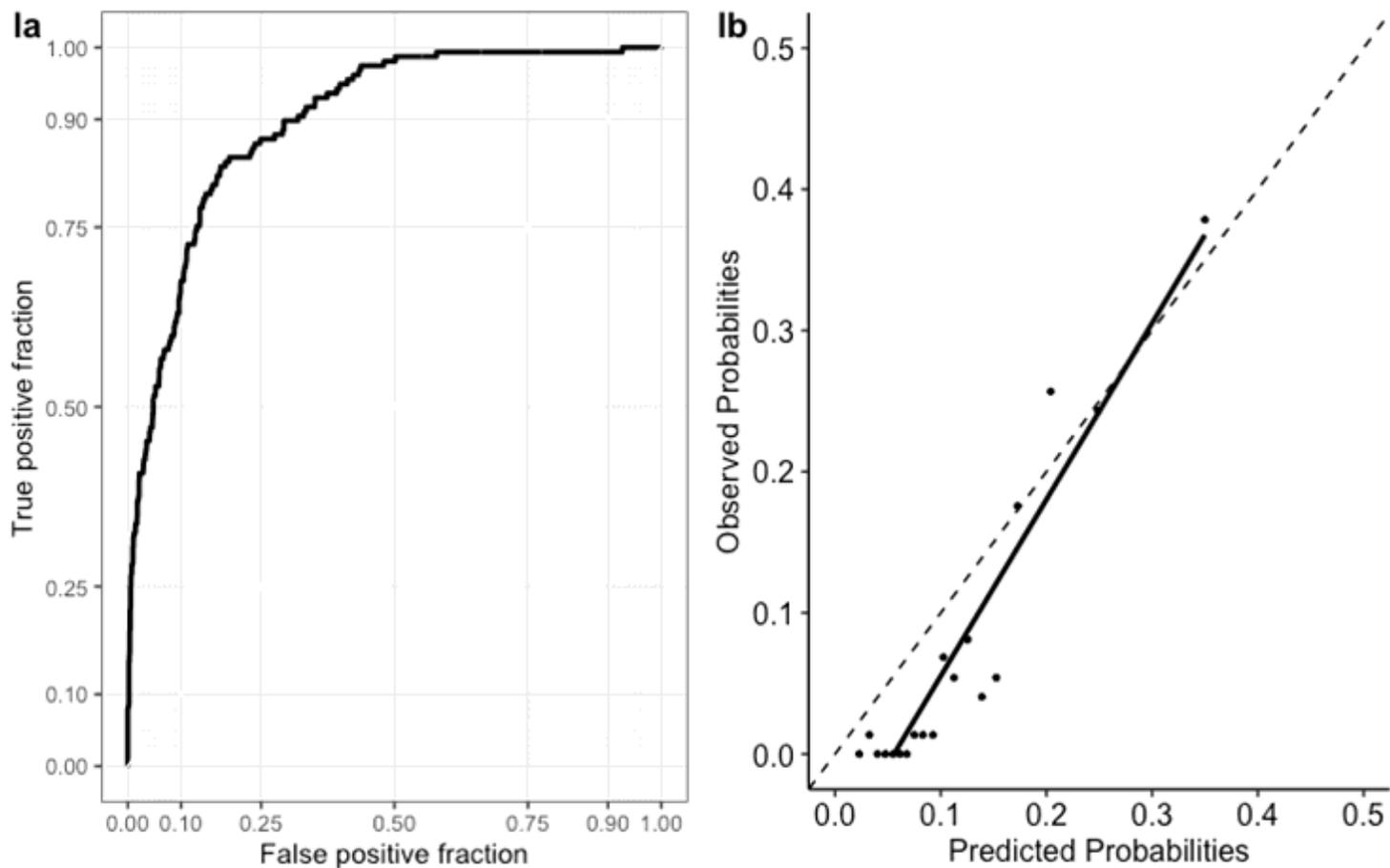


Figure 1

Area under the curve of the best performing model for prediction of all-cause mortality among patients with stable angina (A). The closer to the dotted line indicates the better the prediction in the calibration plot of the best performing model (B).

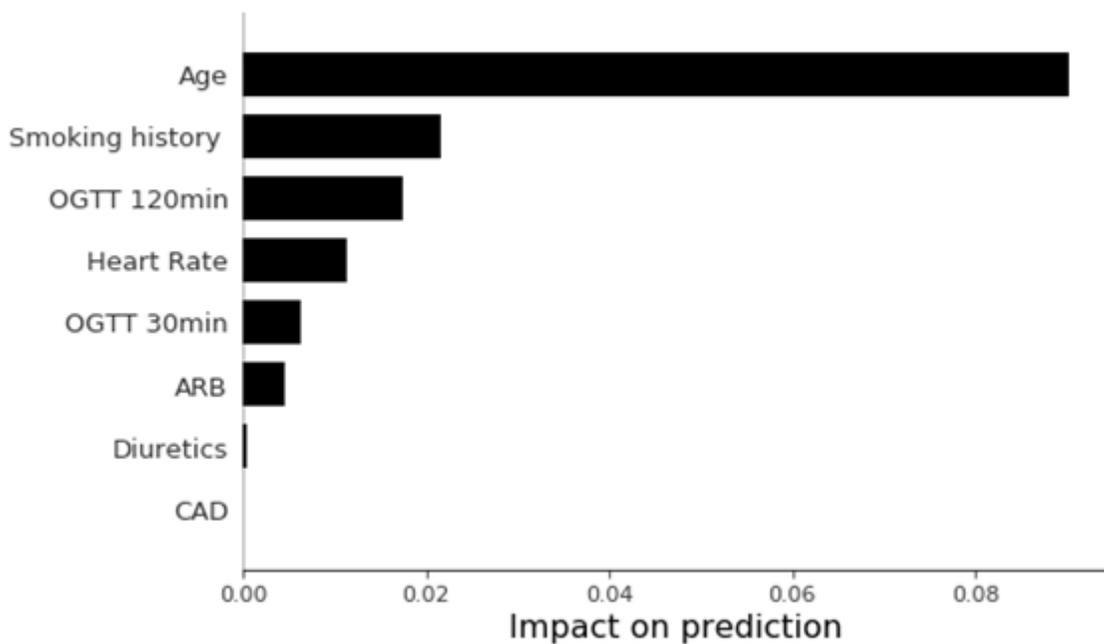


Figure 2

Impact of selected features on model prediction. The impact on prediction was based on the SHAP value, which was calculated from the best performing model. (ARB = angiotensin II receptor blocker, CAD= coronary artery disease, OGTT= oral glucose tolerance test, SHAP = Shapley Additive Explanation values)

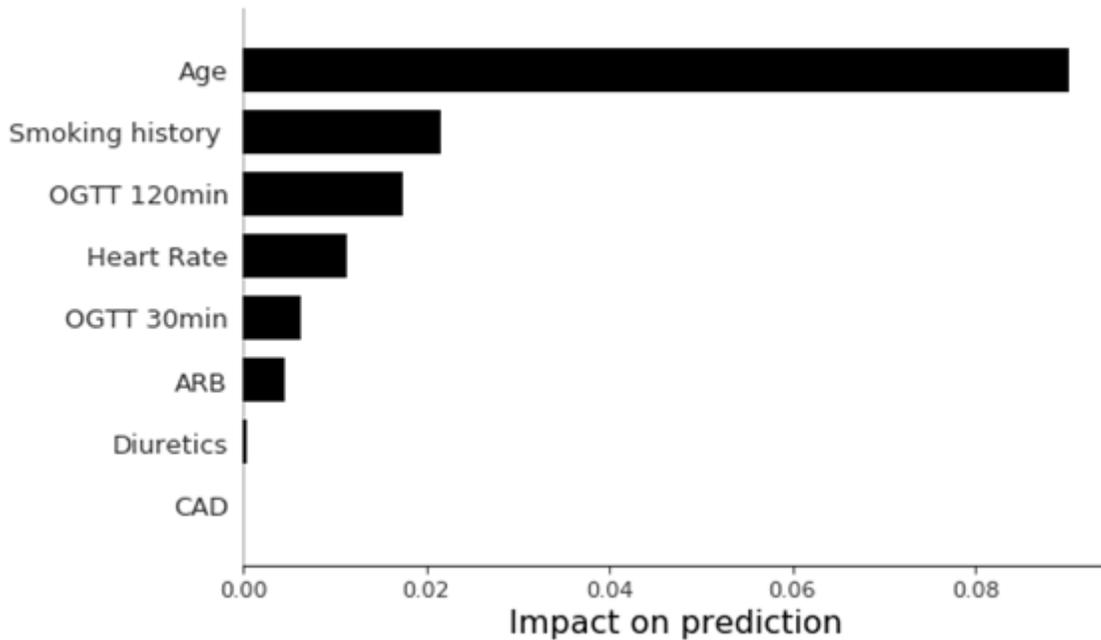


Figure 2

Impact of selected features on model prediction. The impact on prediction was based on the SHAP value, which was calculated from the best performing model. (ARB = angiotensin II receptor blocker, CAD= coronary artery disease, OGTT= oral glucose tolerance test, SHAP = Shapley Additive Explanation values)