Published in final edited form as:

JACC Cardiovasc Imaging. 2018 March ; 11(3): 437–446. doi:10.1016/j.jcmg.2017.02.020.

The External Validity of Prediction Models for the Diagnosis of Obstructive Coronary Artery Disease in Patients with Stable Chest Pain: Insights from the PROMISE trial

Tessa S.S. Genders, MD, PhD¹, Adrian Coles, PhD¹, Udo Hoffmann, MD, MPH², Manesh R. Patel, MD¹, Daniel B. Mark, MD, MPH¹, Kerry L. Lee, PhD¹, Ewout W. Steyerberg, PhD³, M.G. Myriam Hunink, MD, PhD^{4,5}, Pamela S. Douglas, MD¹ CAD Consortium and the PROMISE Investigators

¹Duke Clinical Research Institute, Duke University School of Medicine, Durham, NC, USA

²Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA ³Department of Public Health, Erasmus University Medical Center, Rotterdam, The Netherlands ⁴Department of Epidemiology and Radiology, Erasmus University Medical Center, Rotterdam, The Netherlands

⁵Department of Health Policy and Management, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, USA.

Abstract

Objective: To externally validate prediction models for the presence of obstructive coronary artery disease (CAD).

Background: A better assessment of the probability of CAD may improve the identification of patients who benefit from non-invasive testing.

Methods: PROMISE patients with stable chest pain who underwent either computed tomography angiography (CTA), invasive coronary angiography (ICA), or both, were included. Obstructive CAD was defined as ≥ 1 stenosis of $\geq 50\%$ on ICA. We assumed that patients with CTA showing 0% stenosis and a coronary artery calcium (CAC) score of 0 were free of obstructive CAD, and we multiply imputed the missing ICA results based on clinical variables and the CTA results for 2106 patients (61%). Predicted probabilities of CAD were calculated using published coefficients for 3 models: a basic model (age, sex, type of chest pain), a clinical model (basic model + diabetes, hypertension, dyslipidemia, and smoking), and a clinical + CAC score model. We assessed

Correspondence to: Dr. Pamela S. Douglas, Duke Clinical Research Institute, 7022 North Pavilion DUMC, PO Box 17969, Durham, NC 27715, USA; phone 919.681.2690; fax 919.668.7059; pamela.douglas@duke.edu.

Author Contributions: Dr. Douglas and Dr. Genders had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Disclosures: TSSG has nothing to disclose; AC has nothing to disclose; UH has received research grants (significant) from American College of Radiology Imaging Network, HeartFlow, and Siemens Healthcare; MRP has received research Grants (significant) from AstraZeneca, Jansen, and HeartFlow, and served on an advisory board for AstraZeneca, Jansen, Bayer, and Genzyme; DBM has consulted for Medtronic, CardioDx, and St. Jude Medical, and received research grants (significant) from Eli Lilly and Company, Medtronic, Bristol Myers Squibb, AstraZeneca, Merck & Company, Oxygen Therapeutics, and Gilead; KLL has nothing to disclose; EWS has nothing to disclose; MGMH has received royalties from Cambridge University Press, grants and non-financial support from European Society of Radiology, and non-financial support from the European Institute for Biomedical Imaging Research; PSD has received research grants (significant) from GE and HeartFlow.

discrimination and calibration, and compared published predictor effects to observed predictor effects.

Results: We included 3468 patients (1805 women; mean age 60y; 779 [23%] with obstructive CAD on CTA). The models demonstrated moderate to good discrimination, with c-statistics (95%CI) of 0.69 (0.67-0.72), 0.72 (0.69-0.74), and 0.86 (0.85-0.88), for the basic, clinical, and clinical + CAC score models, respectively. Calibration was satisfactory although the observed predictive effects of typical chest pain and diabetes were weaker, and the observed predictive effect of the CAC score was stronger than suggested by the prediction models. Among the 31% of patients for whom the clinical model predicted a low (10%) probability of CAD, actual prevalence was 7%; among the 48% for whom the clinical + CAC score model predicted a low probability, the observed prevalence was 2%.

Conclusions: Existing clinical prediction models can identify patients with a low probability of obstructive CAD. Obstructive CAD on ICA was imputed for 61% of patients; hence, further validation is necessary.

Introduction

Every year, millions of patients in the United States with stable chest pain undergo non-invasive diagnostic testing to investigate the presence of obstructive coronary artery disease (CAD) (1). The decision to proceed to invasive coronary angiography (ICA), is often based on the results of such non-invasive tests. However, 59% of stable symptomatic patients referred for elective ICA in the United States are free of obstructive CAD (2). A better strategy to select patients who might benefit from *invasive* testing is needed, which should begin by better risk stratifying patients who should undergo *non-invasive* testing (3).

The clinical value of a diagnostic test for CAD depends on the pre-test probability of CAD (4-7). Current guidelines uniformly recognize this and recommend to consider the pre-test probability before deciding whether to test. However, due to a lack of evidence on comparative effectiveness of imaging strategies, for a given pre-test probability and other factors, the test of choice may vary across countries (8-10). The traditional Diamond & Forrester model (11), which in combination with the model based on Coronary Artery Surgery Study (CASS) data (12) is currently recommended by the American College of Cardiology Foundation/American Heart Association stable ischemic heart disease guidelines (8), significantly overestimates the pre-test probability of obstructive CAD (13-15). Improved estimates of the pre-test probability may be obtained using updated prediction models that were developed by the CAD Consortium (13), and they can potentially help clinicians make better decisions as to which patients should undergo non-invasive testing. The current study aims to externally validate the CAD consortium prediction models for the presence of obstructive CAD in chest pain patients from the United States.

Methods

Study population

Our study population consisted of patients enrolled in the Prospective Multicenter Imaging Study for Evaluation of Chest Pain (PROMISE) trial, which has been described in detail

previously (16,17). In brief, the PROMISE trial was a pragmatic, multi-center randomized trial that compared outcomes of initial anatomic testing with the use of coronary computed tomography angiography (CTA) vs. initial functional testing (exercise electrocardiography, nuclear stress testing, or stress echocardiography) for patients with suspected CAD. Enrollment began on 27 July 2010 and was completed on 19 September 2013. Patients were symptomatic outpatients; men were aged ≥ 55 years, or 45-54 years with ≥ 1 cardiac risk factor (diabetes, peripheral arterial disease, cerebrovascular disease, current or past smoking, hypertension, or dyslipidemia), and women were aged ≥ 65 years, or 50-64 years with ≥ 1 cardiac risk factor. Patients with a history of acute myocardial infarction, known CAD, or revascularization were excluded. For the current study we selected PROMISE patients who were assigned to the anatomic testing strategy, presented with chest pain, and underwent either CTA, ICA, or both.

The Duke University Health System Institutional Review Board approved this study. A waiver of informed consent was granted for the current analysis. Patients previously consented for enrollment in PROMISE.

Risk factor definitions in PROMISE

Chest pain symptoms were defined as typical, atypical, or non-cardiac. Typical chest pain was defined as 1) substernal chest pain or discomfort that was 2) provoked by exertion or emotional stress and 3) relieved by rest and/or nitroglycerine. Atypical chest pain was defined as two of the before mentioned criteria. If one or none of the criteria was present, chest pain symptoms were categorized as non-cardiac (18). Hypertension was defined as a blood pressure $>140/90$ mmHg on at least two occasions ($>130/80$ mmHg for patients with diabetes or chronic kidney disease) and/or requiring anti-hypertensive treatment. Diabetes was defined as a history of diabetes, an elevated fasting serum glucose greater than 126 mg/dL (7 mmol/L), or the need for anti-diabetic agents. Dyslipidemia was defined as an elevated cholesterol level (total cholesterol >200 mg/dL [5.18 mmol/L], low-density lipoprotein >130 mg/dL [3.37 mmol/L], or high-density lipoprotein <40 mg/dL [1.04 mmol/L] in men and <50 mg/dL [1.30 mmol/L] in women) or treatment with cholesterol-lowering medications. Smoking was defined as current or past smoking. The coronary artery calcium (CAC) score was determined using the Agatston method (19). Risk factor definitions were essentially the same as those used during the development of the prediction models (13) (Table 1).

Prediction models

We validated three logistic regression models as previously developed and reported by the CAD Consortium (13). The basic model included age, sex, and type of chest pain. The clinical model included the variables of the basic model and additionally diabetes, hypertension, dyslipidemia, and smoking. The clinical + CAC score model included the CAC score in addition to the variables of the clinical model. For reference we also considered the validity of the updated Diamond & Forrester model (14).

Outcome

Obstructive CAD was defined as ≥ 1 stenosis of $\geq 50\%$ in one or more vessels (≥ 2.0 mm diameter) on ICA (within 90 days of the index test). Assessment of ICA results was not blinded for clinical information nor blinded for the CTA result. CTA and ICA interpretation was based on site-specific reads. ICA is often considered to be the reference standard for diagnosing CAD. CTA is an excellent test with a sensitivity of 96-100% and specificity of 85-92% compared with ICA (20,21). Given its high sensitivity for diagnosing obstructive CAD, a negative CTA ($<50\%$ stenosis) virtually excludes the presence of obstructive CAD on ICA, especially in low-risk populations (20). We assumed that patients with a completely normal CTA (0% stenosis and CAC score = 0) would be free of obstructive CAD on ICA, which is supported by the very low risk of adverse events in this subgroup (22).

Statistical analysis

We used multiple imputation methods with chained equations (23) to account for missing clinical variables (24) and to impute missing ICA results (see eMethods) (25-28). We assumed missing clinical variables occurred at random, and missing variables were predicted using the CTA result, ICA result, CAC score, age, sex, type of chest pain, hypertension, dyslipidemia, diabetes, smoking, and body mass index. McFadden's R^2 for the imputation model ranged from 0.52 to 0.59 across iterations, which indicates excellent model fit. Twenty datasets were created that contained identical information except for the imputed missing variables, for which the variability across datasets reflects the uncertainty associated with imputations. We used previously published logistic regression model equations (see eTable 1 in the Supplement) to calculate predicted probabilities for each patient.

For each prediction model, we calculated the c-statistic (discrimination), constructed calibration plots (29), and performed a detailed assessment of model validity (see eMethods, eTable 2, eFigure 1) (28-31). The first validation step consisted of assessing calibration-in-the-large, which compares the mean predicted probability and the mean observed frequency of obstructive CAD in the validation data. The ideal value is zero difference. The second validation step was assessment of the overall predictive effect, which included graphical assessment in a calibration plot and estimation of a calibration slope. A slope different from 1 indicates that the overall combined predictive effect of the predictors in the model was different from the overall effect as observed in the validation data. The last validation step was re-estimating the predictor effects in the validation data and calculating the difference with the predictor effects in the prediction models. A p-value < 0.05 was considered statistically significant. All analyses were performed using STATA/SE 14.1 (StataCorp, TX, USA). Calibration plots were constructed using R version 3.3.0 (2016).

Results

Of the 4996 patients assigned to the anatomic testing (CTA) strategy in PROMISE, 3468 patients presented with chest pain and either underwent CTA, ICA, or both (1663 men, 1805 women; mean age 58 and 61, respectively). Nearly all patients (99.6%) underwent CTA, and $n=3335$ (96%) had interpretable CTA results, showing $\geq 50\%$ stenosis in 779 patients (23%). Of those who underwent CTA, 441 were referred for ICA which showed obstructive CAD in

75% (n=332). Of those patients with <50% on CTA (n=2556), 38 patients (1.5%) were referred for ICA which demonstrated obstructive disease in 13.

When compared with the patient characteristics of the derivation cohort (13), PROMISE patients were more likely to have cardiac risk factors such as diabetes, hypertension, dyslipidemia, and smoking (Table 1). The distribution of type of chest pain was different as well, with fewer patients with typical and non-cardiac chest pain in PROMISE. The proportion of patients with $\geq 50\%$ stenosis on CTA was 23% in PROMISE and 25% in the derivation cohort. Overall 452 (13%) patients in PROMISE were referred for ICA and 848 (19%) patients were referred for ICA in the derivation cohort. The proportion of ICA results showing absence of obstructive CAD was 25% in PROMISE and 48% in the derivation cohort.

Some patients had missing values for body mass index (n=34, 1%), CAC score (n=375, 11%), CTA (n=133, 4%), and ICA (n=3016, 87%). Of the patients without ICA results, 910 had a completely normal CTA (0% stenosis and CAC score = 0). We assumed that these patients would be free of obstructive CAD on ICA, leaving 2106 (61%) patients with missing ICA results. Of these, 1608 had <50% stenosis on CTA for whom almost all imputed ICA results were negative (range 93-98% across imputations), 289 had moderate CAD (50-70% stenosis) on CTA for whom most imputed ICA results were positive (range 59-72% across imputations), and 127 had severe CAD ($\geq 70\%$ stenosis or $\geq 50\%$ left main stenosis) for whom the majority of imputed ICA results were positive (range 73-89% across imputations).

The externally validated c-statistic was 0.69 (95%CI 0.67-0.72) for the basic model, 0.72 (95%CI 0.69-0.74) for the clinical model, and 0.86 (95%CI 0.85-0.88) for the clinical + CAC Model (Figure 1). Overall calibration was satisfactory for all three models. The clinical model overestimated for higher predicted probabilities, which was most pronounced in women. The basic and clinical model were less well calibrated in women versus men, however calibration of the clinical + CAC model was near perfect regardless of sex (Figure 1). The overall predictor effects were slightly smaller in PROMISE (calibration slope < 1). Specifically, the presence of typical chest pain or diabetes was less predictive (p=0.001 and p=0.04, respectively) in PROMISE as compared with the effect as suggested by the prediction models, and the predictive effect of the CAC score was stronger (p=0.02) (eTable 2). In a sensitivity analysis we used the CTA result as the outcome of interest (eFigure 2) and the subgroup with available ICA results (eFigure 3), eliminating the need to rely on imputed data. The prediction models underestimated the probability of obstructive CAD based on CTA which is expected as the prediction models were developed to predict CAD on ICA. Given the relatively less favorable positive predictive value of CTA, especially in low probability populations (20), the prevalence of CAD based on CTA is generally higher than based on ICA. In spite of this, the CAD consortium models demonstrated better performance compared to the guideline recommended probabilities which significantly overestimate the probability of obstructive CAD on CTA (15). The updated Diamond & Forrester model (14) significantly overestimated the probability of CAD (poor calibration), and its predictor effects were stronger than observed in the validation data (eFigure 4).

The potential clinical usefulness of the three prediction models is illustrated with a categorization of patients as low (<10%), intermediate (10-90%), and high probability (>90%) (Table 2, eTable 3-4) for the predicted probability of obstructive CAD. The basic and clinical model did not predict any probabilities >90%, whereas the clinical + CAC score model predicted a probability of >90% for <1% of patients. The clinical model was able to identify 31% of patients as having a low probability, of whom 7% had obstructive CAD, whereas the clinical + CAC model identified 48% as having a low probability, of whom only 2% had obstructive CAD (Table 2).

Discussion

Summary of key findings

The CAD consortium prediction models provided overall accurate estimates of the pre-test probability in the PROMISE trial, even though some predictors (i.e., typical chest pain and diabetes) appeared to be less predictive and the CAC score appeared to be more predictive than expected. The models can be used to estimate the pre-test probability in patients with chronic stable angina who are considered for non-invasive imaging tests. Including the CAC score significantly improved the prediction of obstructive CAD, suggesting that the CAC score could add important information to the diagnostic work-up. For example, patients with a low probability based on the clinical + CAC model (e.g. those with CAC score of zero) could potentially defer further diagnostic imaging. The CRESCENT trial compared a tiered CT strategy (CTA was deferred for patients with a low probability and CAC=0) with functional testing showing that this was safe and effective (32).

Traditional prediction tools such as the Diamond & Forrester model (11,14,15) and the Duke Clinical Score (13,33) have been shown to overestimate the probability of CAD. The European Society of Cardiology adopted the updated Diamond & Forrester model, because it significantly reduces the overestimation by the traditional Diamond & Forrester model for patients referred to ICA (14,34). However, contemporary patient populations who are considered for non-invasive testing have an even lower probability of disease (15,16). This overestimation was confirmed in the current analysis, where the updated Diamond & Forrester model substantially overestimated the probability of CAD in PROMISE trial patients (eFigure 3). Based on these findings, a revision of the 2012 American College of Cardiology Foundation/American Heart Association stable ischemic heart disease guideline (8) may review the current recommendation of using the traditional Diamond & Forrester model (11) and CASS model (12). We believe the CAD consortium prediction models should receive strong consideration as these models were developed and validated in contemporary patient populations that better represent the low-risk patients who are currently considered for non-invasive testing.

Interestingly, model calibration was better in men compared to women, which could be related to sex differences in symptom presentation, prevalence of risk factors (35,36), and overall prevalence of disease (obstructive CAD on CTA was 30% in men and 18% in women). We observed a significantly smaller predictive effect of the presence of typical chest pain than expected. Patients with typical chest pain are generally considered to have a very high likelihood of CAD and are oftentimes directly referred for ICA. Given the fact that

PROMISE included participants whose physicians believed that non-urgent, non-invasive cardiovascular testing was necessary, a substantial proportion of patients with typical chest pain may not have been included in PROMISE but instead have been referred directly to ICA (2). This may have attenuated the predictive effect of the presence of typical chest pain in the PROMISE data, although further research is needed to answer this question.

Limitations of the study

Several limitations of our study deserve mentioning. First of all, only 13% of this subset of PROMISE patients underwent ICA. We relied on multiple imputations, since an analysis restricted to patients who underwent ICA would be subject to verification bias (37). One could consider using the CTA result as the outcome, but the presence or absence of CAD can be misclassified by CTA as a result of false negative and false positive test results, respectively. In a population with a 20% prevalence of obstructive CAD, the positive predictive value of $\geq 50\%$ stenosis on CTA is estimated to be approximately 70% (20). This means that a non-negligible proportion of patients with $\geq 50\%$ stenosis on CTA will be free of obstructive CAD on ICA. Using the CTA result as the outcome will overestimate the true prevalence of obstructive CAD. We therefore opted to use an alternative method in which we assumed absence of obstructive CAD on ICA for those with a completely normal CTA (0% stenosis and CAC score = 0). We believe this is justified based on the well-established negative predictive value of a negative ($<50\%$ stenosis) CTA, which is $>99\%$ regardless of the CAC score (20) and the very low probability ($<1\%$) of $\geq 50\%$ stenosis for patients with a CAC score of zero (38,39). In our assumption we were as restrictive as possible and assumed a negative ICA result only for those patients with both 0% stenosis on CTA and a CAC score of zero. We subsequently multiply imputed the remaining missing ICA results based on the CTA result and clinical variables. By using this approach, we took advantage of the correlation between CTA and ICA, while taking into account the uncertainty of not having the reference standard result for some patients, which is reflected by the variability across imputed datasets (eTable 5). In a sensitivity analysis (eFigure 2 and 3) we used the CTA result as outcome of interest as well as the subgroup with available ICA results eliminating the need to rely on imputed data. Results showed worse calibration and discrimination, most likely related to the differences in disease prevalence. These findings should be interpreted in the context of the above mentioned limitations.

Second, although the clinical + CAC score demonstrated the best model performance, the CAC score is not currently recommended in the diagnostic work-up for patients with stable chest pain. Therefore, the CAC score is often not available when the pre-test probability is calculated, which currently limits the clinical usefulness of the clinical + CAC model and suggests that the CAC score should be considered in future algorithms.

Third, the overall CAD prevalence in the PROMISE trial was low, which limited our ability to assess calibration at higher disease prevalence. However, the aim of the current analysis was to show validity of predictions for patients who are being considered for non-invasive testing, which, in the contemporary era, are generally low probability populations.

Finally, the current analysis focused on diagnosing obstructive CAD and did not consider prognostic information. Simultaneous prediction of cardiovascular adverse event rates potentially improves decision-making regarding non-invasive diagnostic testing.

Future research

Although it is widely recognized that the pre-test probability should be used to determine which patient should be considered for non-invasive testing, the test of choice varies across countries and across guidelines (8-10). Whereas the European (10) and US (8) guidelines recommend exercise electrocardiography for patients with an intermediate probability of CAD, the National Institute for Health Care Excellence guidelines in the United Kingdom (9) recommend against using exercise electrocardiography for diagnosing CAD. Generally, non-invasive imaging tests are recommended for patients with an intermediate probability; however, which test to use in what patient remains to be elucidated. Further, to date few guidelines have included the use of CAC as part of the initial evaluation, although our data suggests this may improve predictive models. Comparative effectiveness trials and cost-effectiveness analyses should be performed to determine the optimal imaging strategy, taking into account long-term outcomes and comparing multiple diagnostic strategies (40).

Clinical implications

Our results suggest that the basic model, clinical model, and the clinical + CAC score model can be used to reliably estimate the pre-test probability of CAD in low-risk patients who present with stable chest pain and who do not have a history of CAD. (The prediction models are publicly available at www.rcalc.com; click on 'Formulas,' then 'All Formulas,' and then, under 'Coronary Artery Disease,' a link to 'Prediction Model for the Presence of Obstructive Coronary Artery Disease' can be found. This website is free to use and provides a calculator for the three models simultaneously.) The estimated pre-test probability can be used to determine the need for further diagnostic work-up and help decide which test to use (5).

The prediction models were most useful for identifying patients with a low probability of CAD. Although the clinical + CAC model was able to predict probabilities > 90% (high probability), this occurred in <1% of patients. The number of patients in each probability category changes depending on the thresholds that are used to define a low vs. intermediate vs. high probability; however, there is no consensus which thresholds should be used. The European Society of Cardiology (10) and National Institute for Health and Care Excellence (9) guidelines recommend considering other causes of chest pain as opposed to further testing for CAD in patients with an estimated probability of <10-15% for CAD. Using the clinical model, 31% of patients are identified as having a probability of > 10% for CAD, in whom one may consider refraining from further testing for CAD, which potentially reduces unnecessary procedures and costs. Including the CAC score significantly improved model performance, which suggests that the CAC score may be useful to guide further management (9,32,41).

Clinical prediction models for the presence of CAD provide well-calibrated predictions and can help clinicians to identify patients with a low probability of CAD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgment

Funding: The PROMISE trial was supported by grants R01HL098237, R01HL098236, R01HL98305 and R01HL098235 from the National Heart, Lung, and Blood Institute (NHLBI). The authors are solely responsible for the design and conduct of this study, all study analyses, the drafting and editing of the paper and its final contents. This paper does not necessarily represent the official views of NHLBI.

References

1. Ladapo JA, Blecker S, Douglas PS. Physician decision making and trends in the use of cardiac stress testing in the United States: an analysis of repeated cross-sectional data. *Ann Intern Med* 2014;161:482–90. [PubMed: 25285541]
2. Patel MR, Dai D, Hernandez AF, et al. Prevalence and predictors of nonobstructive coronary artery disease identified with coronary angiography in contemporary clinical practice. *Am Heart J* 2014;167:846–52 e2. [PubMed: 24890534]
3. Diamond GA, Kaul S. Low diagnostic yield of elective coronary angiography. *N Engl J Med* 2010;363:93; author reply 94–5.
4. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. *N Engl J Med* 1980;302:1109–17. [PubMed: 7366635]
5. Genders TS, Petersen SE, Pugliese F, et al. The optimal imaging strategy for patients with stable chest pain: a cost-effectiveness analysis. *Ann Intern Med* 2015;162:474–84. [PubMed: 25844996]
6. Ladapo JA, Jaffer FA, Hoffmann U, et al. Clinical outcomes and cost-effectiveness of coronary computed tomography angiography in the evaluation of patients with chest pain. *J Am Coll Cardiol* 2009;54:2409–22. [PubMed: 20082932]
7. Min JK, Gilmore A, Budoff MJ, Berman DS, O'Day K. Cost-effectiveness of coronary CT angiography versus myocardial perfusion SPECT for evaluation of patients with chest pain and no known coronary artery disease. *Radiology* 2010;254:801–8. [PubMed: 20177094]
8. Fihn SD, Gardin JM, Abrams J, et al. 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS Guideline for the diagnosis and management of patients with stable ischemic heart disease: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, and the American College of Physicians, American Association for Thoracic Surgery, Preventive Cardiovascular Nurses Association, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. *J Am Coll Cardiol* 2012;60:e44–e164. [PubMed: 23182125]
9. Cooper A, Timmis A, Skinner J, Guideline Development G. Assessment of recent onset chest pain or discomfort of suspected cardiac origin: summary of NICE guidance. *BMJ* 2010;340:c1118. [PubMed: 20335334]
10. 2013 ESC guidelines on the management of stable coronary artery disease: the Task Force on the management of stable coronary artery disease of the European Society of Cardiology. *Eur Heart J* 2013;34:2949–3003. [PubMed: 23996286]
11. Diamond GA, Forrester JS. Analysis of probability as an aid in the clinical diagnosis of coronary-artery disease. *N Engl J Med* 1979;300:1350–8. [PubMed: 440357]
12. Chaitman BR, Bourassa MG, Davis K, et al. Angiographic prevalence of high-risk coronary artery disease in patient subsets (CASS). *Circulation* 1981;64:360–7. [PubMed: 7249303]
13. Genders TS, Steyerberg EW, Hunink MG, et al. Prediction model to estimate presence of coronary artery disease: retrospective pooled analysis of existing cohorts. *BMJ* 2012;344:e3485. [PubMed: 22692650]
14. Genders TS, Steyerberg EW, Alkadhi H, et al. A clinical prediction rule for the diagnosis of coronary artery disease: validation, updating, and extension. *Eur Heart J* 2011;32:1316–30. [PubMed: 21367834]

15. Cheng VY, Berman DS, Rozanski A, et al. Performance of the traditional age, sex, and angina typicality-based approach for estimating pretest probability of angiographically significant coronary artery disease in patients undergoing coronary computed tomographic angiography: results from the multinational coronary CT angiography evaluation for clinical outcomes: an international multicenter registry (CONFIRM). *Circulation* 2011;124:2423–32, 1-8. [PubMed: 22025600]
16. Douglas PS, Hoffmann U, Patel MR, et al. Outcomes of anatomical versus functional testing for coronary artery disease. *N Engl J Med* 2015;372:1291–300. [PubMed: 25773919]
17. Douglas PS, Hoffmann U, Lee KL, et al. PROspective Multicenter Imaging Study for Evaluation of chest pain: rationale and design of the PROMISE trial. *Am Heart J* 2014;167:796–803 e1. [PubMed: 24890527]
18. Diamond GA. A clinically relevant classification of chest discomfort. *J Am Coll Cardiol* 1983;1:574–5. [PubMed: 6826969]
19. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Jr., Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. *J Am Coll Cardiol* 1990;15:827–32. [PubMed: 2407762]
20. Schuetz GM, Zacharopoulou NM, Schlattmann P, Dewey M. Meta-analysis: noninvasive coronary angiography using computed tomography versus magnetic resonance imaging. *Ann Intern Med* 2010;152:167–77. [PubMed: 20124233]
21. von Ballmoos MW, Haring B, Juillerat P, Alkadhi H. Meta-analysis: diagnostic performance of low-radiation-dose coronary computed tomography angiography. *Ann Intern Med* 2011;154:413–20. [PubMed: 21403076]
22. Villines TC, Hulten EA, Shaw LJ, et al. Prevalence and severity of coronary artery disease and adverse events among symptomatic patients with coronary artery calcification scores of zero undergoing coronary computed tomography angiography: results from the CONFIRM (Coronary CT Angiography Evaluation for Clinical Outcomes: An International Multicenter) registry. *J Am Coll Cardiol* 2011;58:2533–40. [PubMed: 22079127]
23. van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. *Stat Med* 1999;18:681–94. [PubMed: 10204197]
24. Janssen KJ, Donders AR, Harrell FE Jr., et al. Missing covariate data in medical research: to impute is better than to ignore. *J Clin Epidemiol* 2010;63:721–7. [PubMed: 20338724]
25. de Groot JA, Janssen KJ, Zwinderman AH, Moons KG, Reitsma JB. Multiple imputation to correct for partial verification bias revisited. *Stat Med* 2008;27:5880–9. [PubMed: 18752256]
26. Harel O, Zhou XH. Multiple imputation for correcting verification bias. *Stat Med* 2006;25:3769–86. [PubMed: 16435337]
27. Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. *Stat Med* 1991;10:585–98. [PubMed: 2057657]
28. Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating: Springer, 2008.
29. Van Calster B, Nieboer D, Vergouwe Y, De Cock B, Pencina MJ, Steyerberg EW. A calibration hierarchy for risk models was defined: from utopia to empirical data. *J Clin Epidemiol* 2016.
30. Steyerberg EW, Borsboom GJ, van Houwelingen HC, Eijkemans MJ, Habbema JD. Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. *Stat Med* 2004;23:2567–86. [PubMed: 15287085]
31. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J* 2014;35:1925–31. [PubMed: 24898551]
32. Lubbers M, Dedic A, Coenen A, et al. Calcium imaging and selective computed tomography angiography in comparison to functional testing for suspected coronary artery disease: the multicentre, randomized CRESCENT trial. *Eur Heart J* 2016;37:1232–43. [PubMed: 26746631]
33. Pryor DB, Harrell FE Jr., Lee KL, Califf RM, Rosati RA. Estimating the likelihood of significant coronary artery disease. *Am J Med* 1983;75:771–80. [PubMed: 6638047]
34. Jensen JM, Voss M, Hansen VB, et al. Risk stratification of patients suspected of coronary artery disease: comparison of five different models. *Atherosclerosis* 2012;220:557–62. [PubMed: 22189201]

35. Hemal K, Pagidipati NJ, Coles A, et al. Sex Differences in Demographics, Risk Factors, Presentation, and Noninvasive Testing in Stable Outpatients With Suspected Coronary Artery Disease: Insights From the PROMISE Trial. *JACC Cardiovasc Imaging* 2016;9:337–46. [PubMed: 27017234]
36. Pagidipati NJ, Hemal K, Coles A, et al. Sex Differences in Functional and CT Angiography Testing in Patients With Suspected Coronary Artery Disease. *J Am Coll Cardiol* 2016;67:2607–16. [PubMed: 27058908]
37. Begg CB, Greenes RA. Assessment of diagnostic tests when disease verification is subject to selection bias. *Biometrics* 1983;39:207–15. [PubMed: 6871349]
38. Genders TS, Pugliese F, Mollet NR, et al. Incremental value of the CT coronary calcium score for the prediction of coronary artery disease. *Eur Radiol* 2010;20:2331–40. [PubMed: 20559838]
39. Haberl R, Becker A, Leber A, et al. Correlation of coronary calcification and angiographically documented stenoses in patients with suspected coronary artery disease: results of 1,764 patients. *J Am Coll Cardiol* 2001;37:451–7. [PubMed: 11216962]
40. Hunink MG, Weinstein MC, Wittenberg E, et al. *Decision making in health and medicine: Integrating evidence and values*: Cambridge University Press, 2014.
41. Mouden M, Timmer JR, Reiffers S, et al. Coronary artery calcium scoring to exclude flow-limiting coronary artery disease in symptomatic stable patients at low or intermediate risk. *Radiology* 2013;269:77–83. [PubMed: 23788718]

Figure 1. Calibration-in-the-Large, Calibration Slope, and C-Statistic for the Basic Model, Clinical Model, and Clinical + CAC Model

Spikes at the bottom of the graph indicate the probability distribution for those with and without CAD. Triangles indicate deciles of patients according to predicted probability, with 95% confidence intervals for the observed proportions of patients with obstructive CAD. (The observed proportion with obstructive CAD was based on the average outcome across imputations. The confidence intervals [grey area] in these graphs do not reflect the uncertainty associated with differences across imputations.) Calibration-in-the-large quantifies the difference in mean predicted probability and the mean observed proportion of obstructive CAD. The ideal value is zero difference. A recalibration slope different from 1 indicates that the overall combined predictive effect of the predictors in the model was different from the overall effect as observed in the validation data. See eMethods and eFigure 1 for more details.

Table 1.

Patient Characteristics in the Subset of PROMISE Trial Patients as Compared to the Patient Characteristics of the Low-Prevalence Populations in the Derivation Cohort^a

Characteristic	Available	Subset of PROMISE (N=3468)		Derivation cohort (low prevalence) ¹³ (N=4426)		
		Measure / category	Value	Available	Measure / category	Value
Age, y	3468 (100%)	Mean (SD)	59.5 (8)	4422 (99.9%)	Mean (SD)	57.2 (12)
		IQR	53-65		IQR	49-66
		Range	45-89		Range	18-92
Male sex	3468 (100%)		1663 (48%)	4426 (100%)		2406 (54%)
Chest pain ^b	3468 (100%)	Typical	453 (13%)	4424 (99.9%)	Typical	759 (17%)
		Atypical	2643 (76%)		Atypical	2699 (61%)
		Non-cardiac	372 (11%)		Non-cardiac	966 (22%)
Diabetes ^c	3468 (100%)		721 (21%)	4238 (96%)		622 (15%)
Hypertension ^d	3468 (100%)		2198 (63%)	4300 (97%)		2475 (58%)
Dyslipidaemia ^e	3468 (100%)		2309 (67%)	4255 (96%)		2194 (52%)
Smoking ^f	3468 (100%)		1760 (51%)	4273 (97%)		1231 (29%)
Body mass index ^g	3434 (99%)	Mean (median)	30 (29)	4117 (93%)	Mean (median)	28 (27)
CAC score ^h	3093 (89%)	Mean (SD) / median	191 (456) / 20	4009 (91%)	Mean (SD) / median	160 (399) / 4
		0	1127 (36%)		0	1777 (44%)
		0-<10	279 (9%)		0-<10	402 (10%)
		10-<100	711 (23%)		10-<100	749 (19%)
		100-<400	552 (18%)		100-<400	606 (15%)
		400	424 (14%)		400	475 (12%)
Coronary CT angiography	3335 (96%)	No obstructive CAD	2556 (77%)	4287 (97%)	No obstructive CAD	3232 (75%)
		Moderate CAD	378 (11%)		Moderate CAD	505 (12%)
		Severe CAD	401 (12%)		Severe CAD	550 (13%)
Invasive coronary angiography	452 (13%)	No obstructive CAD	113 (25%)	848 (19%)	No obstructive CAD	406 (48%)
		Moderate CAD	70 (15%)		Moderate CAD	177 (21%)
		Severe CAD	269 (60%)		Severe CAD	265 (31%)
Coronary CT angiography and	401 (12%)	No obstructive CAD	38 (9%)	742 (17%)	No obstructive CAD	230 (31%)
		Moderate CAD	89 (22%)		Moderate CAD	277 (37%)
		Severe CAD	274 (69%)		Severe CAD	235 (9%)
Invasive coronary angiography		No obstructive CAD	101 (25%)		No obstructive CAD	356 (48%)
		Moderate CAD	60 (15%)		Moderate CAD	172 (23%)
		Severe CAD	240 (60%)		Severe CAD	214 (29%)

Abbreviations: CAC = coronary artery calcium; CAD = coronary artery disease; CT = computed tomography; IQR = interquartile range; PROMISE = Prospective Multicenter Imaging Study for Evaluation of Chest Pain; SD = standard deviation.

^aDerivation cohort is adapted with permission from Genders et al., *BMJ* 2012 (13). Listed risk factor definitions apply to the derivation cohort, but definitions were essentially the same as in PROMISE (see Methods). All values are given as number (percentage of listed total) unless indicated otherwise. Moderate CAD is defined as 50-70% stenosis, severe CAD as 70% stenosis or 50% left main stenosis.

^bAccording to the traditional chest pain classification (18) (see Methods).

^cDefined as fasting glucose levels of ≥ 126 mg/dL (≥ 7 mmol/L) or treatment with either diet intervention, oral glucose lowering agent or insulin.

^dDefined as a blood pressure $\geq 140/90$ mmHg or the use of antihypertensive medication.

^eDefined as a total cholesterol ≥ 200 mg/dL (≥ 5.2 mmol/L) or treatment with lipid-lowering drugs.

^fCurrent or past smoking.

^gDefined as weight in kilograms divided by the squared height in meters (kg/m^2).

^hAgatston score as measured by computed tomography (19).

Table 2.

Clinical Usefulness for the 3 Prediction Models in Categorizing Patients as Low (< 10%), Intermediate (10-90%), and High Probability (≥ 90%)^a

Predicted probability	Basic model		Clinical model		Clinical + CAC model ^b	
	N (%)	CAD, ^c N (%) No CAD, N (%)	N (%)	CAD, ^c N (%) No CAD, N (%)	N (%)	CAD, ^c N (%) No CAD, N (%)
Low (< 10%)	1210 (35)	119 (10%) 1091 (90%)	1070 (31)	78 (7) 992 (93)	1474 (48)	29 (2) 1445 (98)
Intermediate (≥10-90%)	2258 (65)	595 (26%) 1663 (74%)	2398 (69)	636 (27) 1762 (73)	1605 (52)	594 (37) 1011 (63)
High (≥ 90%)	0 (0)	-	0 (0)	-	14 (0.5)	13 (93) 1 (7)
	3468 (100)		3468 (100)		3093 (100)	
10% vs. >10%						
Sensitivity (%)		83		89		95
Specificity (%)		40		36		59

Abbreviations: CAC = coronary artery calcium; CAD = coronary artery disease.

^aValues represent numbers of patients (percentage) unless indicated otherwise. See eTable 3 and eTable 4 for results with alternative definitions for low vs. intermediate vs. high probability.

^bIn patients with available CAC scores.

^cThe number of patients with obstructive CAD on invasive coronary angiography was based on the average outcome across imputations.