BMJ Open

Discrimination capability of pretest probability of stable coronary artery disease: a systematic review and meta-analysis suggesting how to improve validation procedures

Pierpaolo Mincarone ,[1] Antonella Bodini,[2] Maria Rosaria Tumolo,[1] Federico Vozzi,[3] Silvia Rocchiccioli,[3] Gualtiero Pelosi,[3] Chiara Caselli,[3] Saverio Sabina,[4] Carlo Giacomo Leo [4]

To cite: Mincarone P, Bodini A, Tumolo MR, *et al*. Discrimination capability of pretest probability of stable coronary artery disease: a systematic review and meta-analysis suggesting how to improve validation procedures. *BMJ Open* 2021;**11**:e047677. doi:10.1136/bmjopen-2020-047677

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2020-047677).

PM and AB contributed equally.

Received 04 December 2020
Accepted 01 June 2021

For numbered affiliations see end of article.

Correspondence to
Dr Carlo Giacomo Leo;
leo@ifc.cnr.it

ABSTRACT

Objective Externally validated pretest probability models for risk stratification of subjects with chest pain and suspected stable coronary artery disease (CAD), determined through invasive coronary angiography or coronary CT angiography, are analysed to characterise the best validation procedures in terms of discriminatory ability, predictive variables and method completeness.

Design Systematic review and meta-analysis.

Data sources Global Health (Ovid), Healthstar (Ovid) and MEDLINE (Ovid) searched on 22 April 2020.

Eligibility criteria We included studies validating pretest models for the first-line assessment of patients with chest pain and suspected stable CAD. Reasons for exclusion: acute coronary syndrome, unstable chest pain, a history of myocardial infarction or previous revascularisation; models referring to diagnostic procedures different from the usual practices of the first-line assessment; univariable models; lack of quantitative discrimination capability.

Methods Eligibility screening and review were performed independently by all the authors. Disagreements were resolved by consensus among all the authors. The quality assessment of studies conforms to the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2). A random effects meta-analysis of area under the receiver operating characteristic curve (AUC) values for each validated model was performed.

Results 27 studies were included for a total of 15 models. Besides age, sex and symptom typicality, other risk factors are smoking, hypertension, diabetes mellitus and dyslipidaemia. Only one model considers genetic profile. AUC values range from 0.51 to 0.81. Significant heterogeneity ($p<0.003$) was found in all but two cases ($p>0.12$). Values of I^2 >90% for most analyses and not significant meta-regression results undermined relevant interpretations. A detailed discussion of individual results was then carried out.

Conclusions We recommend a clearer statement of endpoints, their consistent measurement both in the derivation and validation phases, more comprehensive validation analyses and the enhancement of threshold validations to assess the effects of pretest models on clinical management.

Strengths and limitations of this study

- This is the first meta-analysis summarising the most up-to-date data on the discrimination capability of pretest probability models of stable coronary artery disease.
- The systematic review pays careful attention to the whole validation procedures.
- The majority of included studies were considered to be of high methodological quality.
- We considered pretest models developed in cohorts of patients referred for an anatomical test.
- The meta-analyses have a low reliability due to the small number of included studies and the very high heterogeneity.

PROSPERO registration number CRD42019139388.

INTRODUCTION

The leading cause of mortality and morbidity worldwide in 2019 was represented by cardiovascular disease with 523 million prevalent cases and 18.6 million deaths.[1] Among these, coronary artery disease (CAD) was reported in 197 million subjects and caused 9.14 million deaths. Stable CAD is typically caused by the build-up of plaques that limit blood flow and is characterised by reversible myocardial demand/supply mismatch usually inducible by exercise, emotion or other stress, and commonly associated with transient chest pain (stable angina pectoris).[2 3]

Stable CAD diagnosis is supported by non-invasive functional and/or anatomical testing,[2 3] and invasive coronary angiography (ICA).[2] To limit the risk of inappropriate examinations and their consequences on patients' and healthcare professionals' safety, and economic sustainability of healthcare

systems,[4–7] eligibility to diagnostic testing is established through models that provide a risk stratification of subjects based on a pretest probability (PTP) of CAD. Since the introduction of the Diamond-Forrester model (DFM)[8] and the Duke Clinical Score (DCS),[9] several alternative PTP models have been developed in cohorts of patients referred for ICA or coronary CT angiography (CCTA). Indeed, due to its very high sensitivity and negative predictive values, CCTA can substantially contribute to ruling out CAD.[10] The DFM and its more recent updates have been recommended in guidelines for stable symptomatic subjects.[3 11] Recent debates within scientific societies broach the question of the overestimation flaw of such models. The UK National Institute for Health and Care Excellence (NICE) has preferred no longer to resort to a probabilistic risk-stratification approach and adopt a simpler identification of anginal chest pain to decide for further testing.[12] The European Society of Cardiology (ESC) updated guideline that determines PTPs from the stratified prevalence of CAD in a contemporary cohort, instead of recurring to a prediction model as in the past. These new estimated risks are noticeably lower compared with the previous ones and then underestimation of the disease prevalence can be obtained in different populations.[13] US experts are debating on whether adopting the NICE diagnostic approach or keeping on using PTP.[14 15] To face the flaws on widely recognised PTP models highlighted by NICE and ESC, these organisations clearly underline the need for more information on the various risk factors acting as modifier of the PTP, especially in the low probability range,[11] and for the development and validation of new scores addressing outstanding uncertainties in the estimation of the PTP of CAD.[12]

This review provides several new contributions to the actual debate on how to ameliorate the PTP models developed for anatomically defined outcomes. It mainly focuses on external validation,[16] carries out a meta-analysis to identify the best results and characterises the best procedures in terms of discriminatory ability, significant predictive variables and method completeness. By highlighting some key issues that could be further improved on the development and validation phases, this work aims at stimulating more rigorous procedures for the comparison of different pretest models.

METHODS

This systematic review conforms to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement.[17 18]

Study inclusion and exclusion criteria

We identified studies that validated pretest models intended for the first-line assessment of patients with chest pain and a suspect of stable CAD. The disease was considered as a binary outcome determined through either ICA or CCTA. Reasons for exclusion were: (1) acute coronary syndrome, unstable chest pain, a history of myocardial infarction or previous revascularisation; (2) models that included a diagnostic procedure that does not reflect the usual practices of the first-line assessment[3 11]; (3) models based on a single predictive variable; and (4) lack of clearly stated discrimination capability. Unlike previous works,[19] external validation was primarily considered. We also included internal validation but limited it to k-fold cross-validation as a technique inspired by the same purposes of external validation. Moreover, papers referring to machine learning (ML)-based PTP models have been excluded as considered in a recent review focusing on CAD diagnosis by ML with aims close to ours.[20]

Only full papers were retained because other publications, for example, letters to editors, conference proceedings, etc, are usually not assessed for study quality. Only articles published in English and Italian were considered.

Searches

The databases Global Health (Ovid), Healthstar (Ovid) and MEDLINE (Ovid) were systematically searched (CGL, PM) on 22 April 2020 using several keywords including: angina pectoris, chest pain, coronary artery disease, coronary heart disease, coronary stenosis, stratification score, likelihood function, predictive model, pre-test probability, coronary angiography, cardiac catheterisation and computed tomography angiography. The same full electronic search strategy was applied to the three databases (no filter was used), and is reported in online supplemental file 1c. Citation searches were also performed on reference lists of definitively included studies.

Study selection

Eligibility screening was performed independently by all the authors. Preliminary screening was performed using Abstrackr[21] based on title and abstract with each paper assessed by two randomly assigned reviewers among the authors. Selected papers were assessed based on full text. Disagreements were resolved by consensus among all the authors.

Data extraction strategy

A data collection form was developed by three authors (AB, CGL, PM) and filled in by reviewers independently. Each selected paper was assigned for data extraction to the statistician (AB) and two randomly selected reviewers. Correspondence with the authors of the included studies was initiated if necessary. The reviewers worked independently and in plenary session meetings. Disagreements were resolved by consensus among all the authors. AB, CGL and PM reviewed the final form for internal consistency.

Study quality assessment

The quality assessment of included studies conforms to QUADAS-2 and was performed by four reviewers (AB, CGL, PM, MRT).[22] Due to the previously described features (1–4), we considered that the eligible works did not raise applicability concerns.

Data synthesis and statistical analysis

The discriminative performances of prediction models can be summarised using several methods and indices, and the area under receiver operating characteristic (ROC) curve (AUC) or c-statistics is certainly the best known and more suitable.[23] Then, it has been chosen as the main index for the purposes of this review. Sensitivity and specificity also describe the discrimination capability of the model for a given cut-off and thus provide an indication of clinical usefulness. However, the bivariate nature of this index is not suitable for direct comparisons and then we resorted to the associated AUC.

For the purposes of generalisation of a PTP model to populations that differ from the development population study, the computation of performance indices is not sufficient because a lower performance is usually expected.[16 24] Therefore, we also noted whether more extended validation procedures were performed in order to properly apply a model to new populations.

A random-effects meta-analysis of AUC values from validations of each identified model was performed using R Statistical Software (R Project for Statistical Computing, RRID:SCR_001905)[25] by meta[26] and auctestr[27] packages. Meta-regression was planned to explore the possible sources of unexplained heterogeneity by considering the following factors: (1) sample size, (2) prevalence and (3) anatomical test for outcome assessment.

Patient and public involvement

Patients and the public were not involved in this review.

RESULTS

Study selection

A total of 5711 studies were identified (three through reference lists of included studies) and 2685 different abstracts were screened. Out of the 71 relevant full-texts assessed for eligibility, 27 were finally included (figure 1).

Study characteristics

Table 1 summarises the selected studies in terms of model name, geographical location and population recruitment criteria. Sometimes the same model is referenced with different names across the papers, then table 1 indicates the original name and the one we adopted here.

Studies are mainly conducted in North America[28–37] or Europe.[38–46]

The updated DFM (uDFM),[28 38–40 42–50] and the CAD Consortium Clinical model (CADC-Clin)[28 31 34 39–42 46 50 51] are the most assessed models.

The quality of included studies is generally high due to the specific review question and adopted eligible criteria. Nevertheless, a risk of bias arises from a few specific issues. A few validation studies[29 33 37 43 51] do not declare that they enrolled only consecutive or random samples of patients. With respect to the index test, only one work adopted an optimal discriminating threshold in addition to prespecified ones.[37] Application of CCTA as a reference test yields a risk of bias in many studies[30 31 36 43 45 47 48 51 52] that do not report measures against misclassification of the test results. Finally, in four works,[31 35 38 51] patients did not receive the same reference test for the diagnosis of stable CAD. A graphical summary of the risk of bias is reported in online supplemental file 3.

Predictive variables

As shown in table 2, the identified models can be classified into two broad classes: basic models, including the DFM (based on age, sex and chest pain) and its updates, and clinical models, including the DCS and the models that extend the DFM by adding a few, mainly traditional,[53] risk factors. Within this quite classic framework, the Corus CAD model is distinguished by relating CAD to patients without diabetes to the expression levels of a set of genes. All the models were derived by logistic regression. Exceptions are: DFM, derived by a conditional probability analysis in the late 1970s; Corus CAD, obtained through Ridge regression; CONFIRM score, developed to predict adverse clinical events by fitting a Cox proportional hazards model and subsequently validated for diagnosis of CAD.

Cross-validation[51] and split sample[30 33] have been used in a few cases only.

Predictors were classified into four macro-areas: demography, medical history, clinical presentation/physical examination and biochemistry. The demographic macro-area is present in all models with the variables age and sex, while race is only included in the Expanded clinical model and PROMISE Minimal Risk model. The most used variables in the medical history macro-area are diabetes mellitus and hypertension. The clinical presentation/physical examination macro-area is present in all but the Corus CAD models. Only the Corus CAD and PROMISE Minimal Risk models do not include chest pain. The most used variable in the biochemistry macro-area is dyslipidaemia. The other risk factors are model specific: gene expression (Corus CAD), oestrogen status (Morise score), high-density lipoprotein cholesterol (PROMISE Minimal Risk model) and the high-sensitivity cardiac troponin (uDFM-cTn).

Discrimination capability

All the papers presented ROC curves and/or AUC values. In Adamson *et al*,[47] fixed thresholds only were analysed and the c-statistics associated with sensitivity and specificity reported. Table 3 reports the AUC values and their 95% CIs, while the summary of the meta-analyses conducted for the models with more than one validation is shown in figure 2, where models with a single validation are also considered for the sake of completeness. To carry out meta-analyses as complete as possible, the missing information about the SE of estimated AUC values was filled in by the 'se_auc' command of the auctestr package. Then, the (Gaussian) 95% CIs are reported in table 3. This computation only requires to know the study sample size and the prevalence, and is as better as the size of the

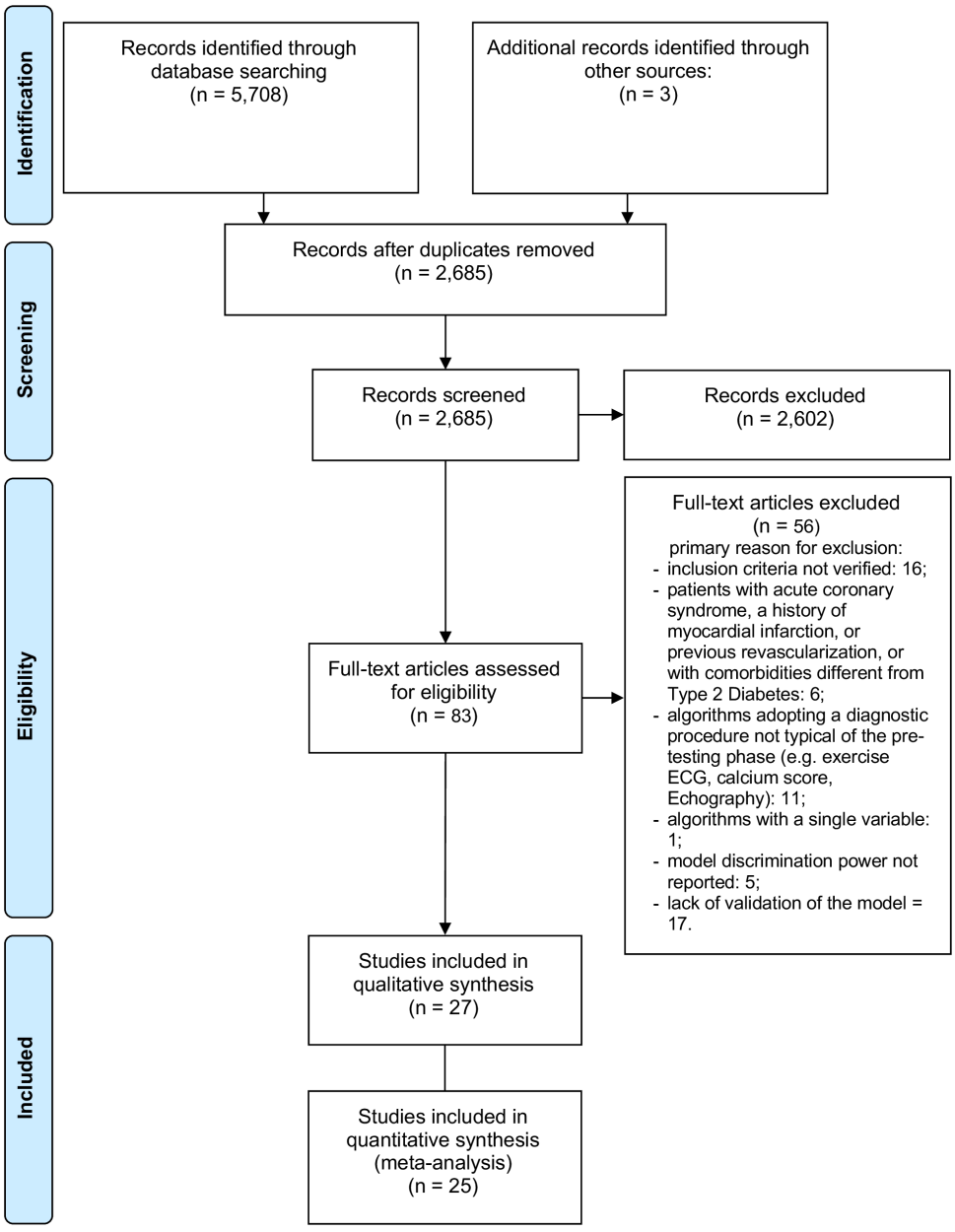

Figure 1 Flow diagram of the study selection process.

study is larger. For a small sample size, the computed SE is generally larger than the exact one and then CIs are more conservative. For only two papers, the conditions for inclusion in the meta-analyses are not met.[29 30]

AUC values range from 0.51[47] (almost failing) to approximately 0.81[51] (almost excellent). The statistical heterogeneity of the AUC values among the studies validating each PTP model was assessed by using the Cochran Q test and the I^2 statistic.[54] In all but two cases (CONFIRM score and Morise score), a statistically significant heterogeneity has been obtained, as expected ($p<0.003$). On the one hand, the lack of heterogeneity is unreliable, due to the low number (≤5) of included studies and the low power of the Cochran Q test. On the other hand, significant heterogeneity exceeds 0.90 for most analyses and even 0.95 undermining significant interpretations ([55] and references therein). Then, in the following the discussion of the pooled values is complemented by a detailed discussion of the individual results.

From the meta-analyses, uDFM-cTn and CONFIRM show the best performances (AUC=0.757and pooled AUC=0.7554, respectively). In slightly more detail, the extension of uDFM with the use of high-sensitivity cardiac troponin I (uDFM-cTn) has been validated in only one population where it showed a significantly higher AUC than uDFM alone (0.757 vs 0.738, p=0.025) and better calibration (Hosmer-Lemeshow (HL) p=0.0001 vs HL p=0.1123).[38] The substantially steady results of the CONFIRM score on several data sets are also confirmed on a validation data set consisting of subjects at the low extreme of traditional cardiovascular risk factor burden.[56]

Table 1 Characteristics of the studies on PTP for CAD

Study	Models/scores	Study centres	Population	
			Inclusion criteria	Exclusion criteria
Adamson *et al*[47]	DFM/CASS uDFM	1. Multicentre PROMISE trial, USA and Canada 2. Multicentre SCOT-HEART trial, Scotland (UK)	See PROMISE. Randomised to receive CCTA as non-initial non-invasive test. See SCOT-HEART. Randomised to the CCTA intervention arm.	See PROMISE and SCOT-HEART. Known CAD.
Adamson *et al*[38]	uDFM (baseline CADC model, in text) uDFM-cTn (baseline CADC model with the addition of troponin, in text)	Odense University Hospital, Denmark	Clinical stable prospectively enrolled patients with suspected angina pectoris scheduled for either ICA or CCTA.[70]	Suspected acute coronary syndrome. To avoid potential confounding effects on the biomarkers measured, patients with established atherosclerotic manifestations, including an abnormal 12-lead rest ECG, were excluded: known ischaemic heart disease, prior ischaemic stroke or transitory ischaemic attack, known peripheral artery disease (n=10), and p-creatinine >200 mmol/L. CCTA not performed or of poor technical quality, lack of informed consent, missing hs-cTnI measure or personal history.[70]
Almeida *et al*[39]	CADC-Clin (CAD Consortium 2, in text) DCS uDFM (CAD Consortium 1, in text)	Single centre in southwestern Europe	Patients with chest pain and suspected CAD referred to ICA.	Patients with a history of CAD, acute coronary syndrome or coronary revascularisation.
Baskaran *et al*[40]	CADC-Clin CONFIRM score uDFM	Multicentre SCOT-HEART trial, Scotland (UK)	See SCOT-HEART. Randomised to the CCTA intervention arm and with information on all variables needed for the analysis.	See SCOT-HEART. Known CAD.
Bittencourt *et al*[28]	CADC-Basic CADC-Clin uDFM (Diamond-Forrester score, in text)	Massachusetts General Hospital; Brigham and Women's Hospital (Massachusetts, USA)	Subjects ≥18 years who underwent CCTA for suspect of CAD.	Patients who were missing any of the clinical information needed to calculate the PTP, who had non-diagnostic CCTA images, who had incomplete follow-up information; with congenital heart disease, heart transplantation, or prior CAD, defined as prior percutaneous coronary interventions, coronary artery bypass graft surgery or myocardial infarction.
Daniels *et al*[29]	Corus CAD (gene expression score—GES, in text)	Multicentre PREDICT trial, USA	See PREDICT.	See PREDICT. Patients with diabetes.
Edlinger *et al*[41]	CADC-Clin	University Clinic of Cardiology at Innsbruck (Austria)	Patients were 18 years of age or older with chest pain or symptoms suggestive of CAD (predominantly dyspnoea) and/or non-invasive evidence of CAD referred for elective ICA.	(1) An elective ICA before or after heart transplantation, (2) an elective ICA prior to solid organ transplantation, (3) an elective ICA before heart valve repair or replacement, or with valvular heart disease as leading clinical diagnosis, (4) an isolated right heart catheterisation, (5) an electrophysiological procedure (pacemaker implantation or catheter ablation) as leading clinical indication, (6) an elective ICA because of a known or suspected congenital heart disease as leading clinical diagnosis (eg, atrial septal defect, ventricular septal defect or patent foramen ovale), or (7) when referred for other reasons (like myocardial biopsy, aortic aneurysms, myxoma, endocarditis or prior failed angiography). History of myocardial infarction.

Continued

Table 1 Continued

Study	Models/scores	Study centres	Population	
			Inclusion criteria	Exclusion criteria
Ferreira *et al*[42]	uDFM (modified DF, in text) CADC-Clin (CAD Consortium 2, in text) CONFIRM score	Unspecified, Portugal	Patients undergoing CCTA for the evaluation of CAD	Age <30 years; known CAD; suspected acute coronary syndrome; preoperative assessment; known left ventricular systolic dysfunction; asymptomatic patients (typically referred after a positive screening exercise test); symptoms other than chest pain. Patients with suspected CAD who were scheduled to undergo CCTA but had the procedure halted due to a high coronary artery calcium Agatston score. A threshold of 400 was used as a general guideline for withholding CCTA in these circumstances, but the decision was ultimately left to the performing physician, taking into consideration the clinical context and the distribution of calcium in the coronary tree.
Fordyce *et al*[30]	PROMISE minimal risk model (the originally published version has been subsequently corrected online, see Fordyce *et al*[58])	Multicentre PROMISE trial, USA and Canada	See PROMISE. Patients assigned to anatomical testing.	See PROMISE.
Fujimoto *et al*[67]	DCS K-score	Multicentre, Japan	Suspected CAD.	Patients with known CAD, showing poor image quality and patients with unassessable segments due to severe calcification.
Genders *et al*[43]	DFM	14 European centres	Patients aged 30–69 years with stable chest pain (typical, atypical or non-specific chest pain) and if ICA performed.	Patients meeting the following criteria: (1) acute coronary syndrome or unstable chest pain, (2) history of myocardial infarction or previous revascularisation (percutaneous coronary intervention or coronary artery bypass graft surgery), and (3) no informed consent.
	uDFM	Erasmus Medical Center, Rotterdam, the Netherlands[71]	Patients with stable chest pain and no history of CAD.[71]	Not undergoing CCTA or ICA.
Genders *et al*[51]	DCS	Multicentre EU and USA	Stable chest pain, referred for catheter-based or CT-based coronary angiography.	Acute coronary syndrome, unstable chest pain, history of myocardial infarction or previous revascularisation or no informed consent.
	CADC-Basic CADC-Clin	Multicentre EU and USA	Stable chest pain, referred for catheter-based or CT-based coronary angiography.	Acute coronary syndrome, unstable chest pain, history of myocardial infarction or previous revascularisation or no informed consent.
Genders *et al*[31]	CADC-Basic CADC-Clin	Multicentre PROMISE trial, USA and Canada	See PROMISE trial for the main criteria. Patients assigned to anatomical testing.	See PROMISE trial for the main criteria.
Jensen *et al*[44]	CORSCORE DCS DFM Morise score uDFM	Lillebælt Hospital, Vejle, Denmark	Patients with chest pain indicative of CAD referred for ICA.	Unstable angina or previous coronary intervention.
Min *et al*[52]	CONFIRM score (integer-based risk model, in text)	USA, Canada, South Korea and Austria (4 out of 5 sites of the phase II of CONFIRM trial)[72]	Patients ≥18 years old referred to CCTA for suspected stable CAD (CONFIRM trial[72]).	Patients with prior coronary revascularisation or myocardial infarction, asymptomatic, missing data.
Pickett *et al*[32]	DFM/CASS Morise score	Walter Reed Army Medical Center, Washington, USA	Patients referred for CCTA.	Known CAD.

Continued

Table 1 Continued

Study	Models/scores	Study centres	Population	
			Inclusion criteria	Exclusion criteria
Rademaker *et al*[45]	DCS DFM Morise score (new score, in text) uDFM	VU University Medical Center, Amsterdam, the Netherlands	Symptomatic women undergoing evaluation for CAD and referred for CCTA.	History of CAD (percutaneous coronary intervention, coronary artery bypass graft surgery or previous myocardial infarction), or absolute or relative contraindications for CCTA such as (1) significant severe arrhythmia; (2) pregnancy; (3) renal insufficiency (glomerular filtration rate <45 mL/min); (4) known allergy to iodinated contrast material.
Rosenberg *et al*[33]	Corus CAD (gene expression test, in text) Expanded clinical model score DFM/CASS	Multicentre PREDICT trial, USA	See PREDICT.	See PREDICT. Diabetes.
Teressa *et al*[34]	CADC-Basic CADC-Clin	1 centre in the USA	>18 years old evaluated in the emergency department of a major academic tertiary university hospital for chest pain, using CCTA as a primary diagnostic modality.	Known CAD, defined as history of acute myocardial infarction, percutaneous intervention, coronary artery bypass graft, or evidence of CAD by either anatomical (CCTA or cardiac catheterisation) or functional tests (positive stress test). Haemodynamically or clinically unstable patients, patients with ST segment changes or positive cardiac troponin (>0.04 ng/mL), impaired renal function (estimated glomerular filtration rate <50 mL/min/1.73 m^2), tachycardia, or contraindication to nitroglycerin or iodinated contrast. Inadequate documentation on chest pain characteristics, repeat CCTAs, unavailable calcium score and non-diagnostic examination.
Thomas *et al*[35]	Corus CAD (GES, in text) DFM Morise score	Multicentre COMPASS trial, USA	See COMPASS.	See COMPASS.
Voora *et al*[36]	Corus CAD	Multicentre PROMISE trial, USA and Canada	See PROMISE. Patients assigned to anatomical testing.	See PROMISE. Diabetes. RNA sample not passing quality control.
Voros *et al*[37]	Corus CAD (GES, in text) DFM	Multicentre PREDICT, USA and COMPASS US trials	See PREDICT and COMPASS.	See PREDICT and COMPASS. Diabetes excluded from PREDICT cohort.
Wang *et al*[56]	CONFIRM score	Not specified, China	Patients who underwent CCTA for stable chest pain and with 0 or 1 risk factors among smoking, hypertension, diabetes and hyperlipidaemia.	Acute coronary syndrome, previous CAD or coronary revascularisation, unassessable segments due to motion artefact, atrial fibrillation, aortic disease, New York Heart Association class III or IV heart failure, age >90 years old, pacemaker leads or missing data.
Winther *et al*[46]	uDFM CADC-Basic CADC-Clin	Multicentre Dan-NICAD trial, Denmark	Patients without known CAD referred to CCTA due to a history of symptoms suggestive of CAD.	Age <40 years; previous coronary revascularisation or myocardial infarction; unstable angina pectoris; estimated glomerular filtration rate <40 mL/min; pregnancy and contraindication for iodine-containing contrast medium, MRI, or adenosine (severe asthma, advanced atrioventricular block or critical aortic stenosis).

Continued

Table 1 Continued

Study	Models/scores	Study centres	Population	
			Inclusion criteria	Exclusion criteria
Yang *et al*[48]	High Risk Anatomy score	Multicentre CONFIRM trial, North America, Europe and Asia University of Ottawa Heart Institute Cardiac CT registry[72]	Patients ≥18 years old referred to CCTA for suspected stable CAD (CONFIRM trial).[72]	Documented CAD, history of myocardial infarction, coronary revascularisation, cardiac transplantation, congenital heart disease.
	uDFM	Multicentre CONFIRM trial, North America, Europe and Asia[72]	Patients ≥18 years old referred to CCTA for suspected stable CAD (CONFIRM trial).[72]	Documented CAD, history of myocardial infarction, coronary revascularisation, cardiac transplantation, congenital heart disease.
Zhang *et al*[49]	DCS uDFM	Tianjin Chest Hospital, Tianjin, China	Patients with stable chest pain and referred for CCTA.	Acute coronary syndrome, previous CAD or coronary revascularisation (percutaneous coronary intervention or coronary artery bypass grafting), impaired renal function (serum creatinine >120 µmol/L), New York Heart Association class III or IV heart failure, atrial fibrillation, aortic disease, age more than 90 years or patients with unassessable segments because of artefact.
Zho *et al*[50]	CADC-Clin (Genders clinical model, in text) DCS uDFM	Not specified, China	Patients who underwent CCTA for stable chest pain.	Acute coronary syndromes, previous CAD or coronary revascularisation (percutaneous coronary intervention or coronary artery bypass grafting), patients with unassessable segments due to motion artefact, atrial fibrillation, aortic disease, New York Heart Association class III or IV heart failure, age >90 years, presence of pacemaker leads or missing data.

The trials COMPASS, CONFIRM, PREDICT, PROMISE and SCOT-HEART were considered in several studies, and thus their main characteristics are fully reported in online supplemental file 2.
CAD, coronary artery disease; CADC, CAD Consortium; CADC-Basic, CADC Basic model; CADC-Clin, CADC Clinical model; CASS, Coronary Artery Surgery Study; CCTA, coronary CT angiography; DCS, Duke Clinical Score; DFM, Diamond-Forrester (DF) model; EU, European Union; hs-cTnI, high-sensitive cardiac troponin I; ICA, invasive coronary angiography; PTP, pretest probability; uDFM, updated DFM.

Table 2 PTP models' variables

Macro model/ score categories	Predicting variables	Model/score														
		CADC-Basic	CADC-Clin	CONFIRM score	CORSCORE	Corus CAD	DCS	DFM	DFM/CASS	Expanded clinical model score	K-score	HRA score	Morise score	PROMISE Minimal Risk model	uDFM	uDFM-cTn
		28 31 34 39–42 46 50 51	28 29 31 34 39–42 46 50–52 56	28 29 40–42 44 52 56	44	29 33 35–37	39 44 45 49–51 67	35 37 43–45 52	32 33 47	33	67	48	32 35 44 45	30	28 38–40 42–50	38
Demography	Age	√	√	√	√	√	√	√	√	√	√	√	√	√	√	
	Sex	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
	Race									√				√		
Medical history	Diabetes mellitus		√	√			√				√	√	√	√		
	Hypertension		√	√	√					√	√	√	√	√		
	Previous MI				√		√									
	Cerebral infarction										√					
	Peripheral vascular disease											√				
Clinical presentation/ physical examination	Chest pain	√	√	√	√		√	√	√	√	√	√	√		√	√
	Abnormal ECG						√									
	Obesity												√			
	Smoking		√	√	√		√				√	√	√	√		
	Family history of CAD			√								√	√	√		
	Other (specify)				Medically treated hyperchole sterolaemia					Medically treated hyperchole sterolaemia				Symptoms related to physical or mental stress		
Biochemistry	HDL cholesterol													√		
	Dyslipidaemia		√				√			√	√	√	√	√		
	Oestrogen status												√			
	Gene expression					√										
	Troponin															√
Others										Aspirin, antiplatelet, ACE inhibitor use, systolic blood pressure						

Continued

Table 2 Continued

Macro model/ score categories	Predicting variables	Model/score													
		CADC-Basic	CADC-Clin	CONFIRM score	CORSCORE	Corus CAD	DCS	DFM	DFM/CASS	Expanded clinical model score	K-score	HRA score	Morise score	PROMISE Minimal Risk model	uDFM	uDFM-cTn
Derivation method		Log	Log	Cox proportional hazards models	Log	Score derived by a Ridge regression	Log	Conditional probability analysis*	Log	Log	Log	Score derived by a multi variable log	Score derived by a log	Log	Log	Log

*In Genders *et al*,[43] to unravel the implicit coefficients of the predictors in this model, the authors performed a weighted linear regression on the log odds of the DF predictions per subgroup
CAD, coronary artery disease; CADC-Basic, CAD Consortium Basic model; CADC-Clin, CAD Consortium Clinical model; CASS, Coronary Artery Surgery Study; DCS, Duke Clinical Score; DFM, Diamond-Forrester (DF) model; HDL, high-density lipoprotein; HRA, High Risk Anatomy; Log, logistic regression; MI, myocardial infarction; uDFM, updated DFM; uDFM-cTn, updated Diamond-Forrester model - high-sensitivity cardiac troponin.

DFM, its DFM/Coronary Artery Surgery Study (CASS) version, uDFM and Morise score show the lowest pooled AUC values <0.70. In slightly more detail, DFM/CASS has the lowest pooled AUC value (0.61) due to the two threshold-based validations reported in.[47] By excluding these values from the meta-analysis, the pooled AUC value becomes closer to 0.70 (0.6861, 95% CI: 0.6312 to 0.7409) and heterogeneity decreases to a non-significant level (I^2=41.9%, p=0.19). With regard to the DFM and its DFM/CASS version, overestimation is usually reported, especially in women.[45] However, the DFM's inferior result is also due to the fact that usually it was not carefully validated but only used as a usual reference model[32 44 45] or as a basis to establish the performances of the Corus CAD model.[33 35 37] The only deep validation is presented in [43]. The Morise score and the Corus CAD are the only two models explicitly considering a female-specific factor (the oestrogen status and a sex-specific score, respectively): when directly compared with the same validating population, the Corus CAD had significantly higher AUC than the Morise score (0.79 vs 0.65, p<0.001).[35]

The uDFM and the CADC-Clin are the two most validated models with completely different performances (pooled AUC values: 0.6866 vs 0.7406). The uDFM updated and extended the traditional DFM to a contemporary cohort that included subjects 70 years and older. The CAD Consortium Basic model (CADC-Basic) can be considered as a further update on a different contemporary population (see table 2). The most complete validation of the uDFM, considering calibration-in-the-large, recalibration and eventually re-estimation, has been performed by the developers themselves[43] who obtained a valid overall effect of predictors. The other validating procedures limit themselves to AUC computation and to a rough assessment of under/overestimation, mainly by the HL goodness-of-fit test and related calibration plots (calibration-in-the-large is applied in one study[42]).

The CADC-Clin model shows good performances on validating populations by reaching estimated AUC values even >0.80, and this high performance level is generally confirmed in other validations by taking into account estimation uncertainty (95% CIs including 0.80).[28 34 40] Moreover, its performances significantly improve with respect to the related CADC-Basic.[28 31 34 51] The pooled AUC value (0.7406) is only slightly lower than the highest ones. It could even have been the best one if three highly performing validations[51] had presented all the data (ie, SE) for their inclusion in the meta-analysis. The generalisability of the CADC-Clin model to external populations was analysed by deep validation procedures.[31 34 41 46] Results on miscalibration analysis could be considered quite consistent across papers. This finding indicates smaller than expected effects of the diagnostic characteristics, chest pain typicality in particular.[31 34 41] Model calibration can be worse in women compared with men, a situation that also arises from the validation of other models (eg, DFM[43]). Despite different pooled AUC values, direct comparisons of either uDFM or CADC-Clin

Table 3 AUC values of PTP models

Model	Study	Outcome definition	Reference test	Sample size	Prevalence (%)	AUC (95% CI)
CADC-Basic	Bittencourt *et al*[28]	At least 1 segment (with a >2 mm diameter) with a lesion with ≥50% diameter stenosis	CCTA	2274	22	0.7517 (0.729 to 0.775)
	Genders *et al*[51]	≥1 diameter stenosis of ≥50% in ≥1 vessel	CCTA, ICA	Min: 471 Max: 1241	NA	Mean: 0.77
	Genders *et al*[31]	≥1 diameter stenosis of ≥50% in ≥1 vessel (≥2.0 mm diameter) by ICA. Patients with a completely normal CCTA (0% stenosis and coronary artery calcium score of 0) are considered as free of obstructive CAD on ICA.	CCTA, ICA	3468	23	0.69 (0.67 to 0.72)
	Teressa *et al*[34]	1 vessel with stenosis of 50%	CCTA	1981	10.4	0.77 (*0.731 to 0.809*)
	Winther *et al*[46]	Coronary diameter stenosis reduction ≥50% in all segments with a reference vessel diameter >2 mm	CCTA	1653	23.7	0.66 (0.63 to 0.69)
CADC-Clin	Almeida *et al*[39]	Stenosis of >50% in at least one major epicardial vessel	ICA	2234	58.5	0.683 (0661 to 0.706)
	Baskaran *et al*[40]	A stenosis causing ≥50% diameter stenosis	CCTA	1738	37.7	0.790 (0.768 to 0.811)
	Bittencourt *et al*[28]	At least 1 segment (with a >2 mm diameter) with a lesion with ≥50% diameter stenosis	CCTA	2274	22	0.791 (0.770 to 0.812)
	Edlinger *et al*[41]	Stenosis ≥50% diameter in at least one of the main coronary arteries	ICA	4888	44	0.69 (0.67 to 0.70)
	Ferreira *et al*[42]	Coronary diameter stenosis ≥50%	CCTA	1069	13.8	0.73 (0.71 to 0.76)
	Genders *et al*[51]	≥1 diameter stenosis of ≥50% in ≥1 vessel	CCTA, ICA	Min: 471 Mean: NA Max: 1241	NA	0.78 0.79 0.81
	Genders *et al*[31]	≥1 diameter stenosis of ≥50% in ≥1 vessel (≥2.0 mm diameter) by ICA. Patients with a completely normal CCTA (0% stenosis and coronary artery calcium score of 0) are considered as free of obstructive CAD on ICA.	CCTA, ICA	3468	23	0.72 (0.69 to 0.74)
	Teressa *et al*[34]	1 vessel with stenosis of 50%	CCTA	1981	10.4	0.80 (0.763 to 0.837)
	Winther *et al*[46]	Coronary diameter stenosis reduction ≥50% in all segments with a reference vessel diameter >2 mm	CCTA	1653	23.7	0.69 (0.66 to 0.72)
	Zhou *et al*[50]	≥1 lesion with ≥50% diameter stenosis or any non-assessable segments due to severe calcification	CCTA	5743	32.6	0.774 (0.761 to 0.788)
CONFIRM score	Baskaran *et al*[40]	A stenosis causing ≥50% diameter stenosis	CCTA	1738	37.7	0.749 (0.726 to 0.771)
	Ferreira *et al*[42]	Coronary diameter stenosis ≥50%	CCTA	1069	13.8	0.71 (0.66 to 0.75)
	Min *et al*[52]	≥50% luminal diameter stenosis in any coronary artery ≥1.5 mm in diameter	CCTA	2132	NA	0.76 (0.746 to 0.771)
	Wang *et al*[56]	≥1 lesion with ≥50% diameter stenosis or any non-assessable segments due to severe calcification	CCTA	0 risk factors (RF): 1201 1 RF: 2415	30.2 27.1	0.756 (0.731 to 0.781) 0.762 (0.742 to 0.783)

Continued

Table 3 Continued

Model	Study	Outcome definition	Reference test	Sample size	Prevalence (%)	AUC (95% CI)
CORSCORE	Jensen *et al*[44]	Lumen area diameter reduction ≥50% in ≥1 coronary artery	ICA	633	34.1	0.727 (*0.684 to 0.770*)
Corus CAD	Daniels *et al*[29]	At least one lesion in a major coronary artery (≥1.5 mm lumen diameter) ≥70% diameter stenosis by clinical read or ≥50% diameter stenosis by invasive QCA	ICA	Several subsets from a total of 1502	NA	Min: 0.64 Max: 0.72
	Rosenberg *et al*[33]	≥1 atherosclerotic plaque in a major coronary artery (≥1.5 mm lumen diameter) causing ≥50% luminal diameter stenosis by QCA	ICA	526	36.5	0.70 (0.68 to 0.72)
	Thomas *et al*[35]	≥1 diameter stenosis ≥50% in a major vessel on ICA by QCA (≥1.5 mm) or CCTA (≥2.0 mm)	CCTA, ICA	431	14.6	0.79 (0.72 to 0.84)
	Voora *et al*[36]	≥70% stenosis in major coronary artery or ≥50% left main stenosis	CCTA	1137	10.1	0.625 (0.573 to 0.678)
	Voros *et al*[37]	Outcome 50: ≥50% maximum diameter stenosis	CCTA	610	14	0.75 (0.70 to 0.80)
		Outcome 70: ≥70% maximum diameter stenosis	CCTA		NA	0.75 (0.67 to 0.83)
DCS	Almeida *et al*[39]	Stenosis of >50% in at least one major epicardial vessel	ICA	2234	58.5	0.685 (0.663 to 0.708)
	Fujimoto *et al*[67]	Lesions with diameter stenosis of ≥75% were defined to be obstructive stenotic lesions. As for left main trunk lesion, lesions with diameter stenosis ≥50% were defined to be obstructive stenotic lesions.	CCTA	361	34.1	0.688 (0.626 to 0.750)
	Genders *et al*[51]	Severe CAD defined as ≥70% diameter stenosis or ≥50% left main stenosis	CCTA, ICA	4426	NA	0.78 (0.76 to 0.81)
	Jensen *et al*[44]	Lumen area diameter reduction ≥50% in ≥1 coronary artery	ICA	633	34.1	0.718 (*0.674 to 0.762*)
	Rademaker *et al*[45]	>50% luminal diameter stenosis	CCTA	178	23.6	0.59 (0.51 to 0.66)
	Zhang *et al*[49]	≥1 lesion with ≥50% diameter stenosis	CCTA	Men: 3001	39	0.785 (*0.767 to 0.803*)
				Women: 2776	25	0.684 (*0.660 to 0.708*)
	Zhou *et al*[50]	≥1 lesion with ≥50% diameter stenosis or any non-assessable segments due to severe calcification	CCTA	5743	32.6	0.772 (0.759 to 0.786)

Continued

Table 3 Continued

Model	Study	Outcome definition	Reference test	Sample size	Prevalence (%)	AUC (95% CI)
DFM	Genders *et al*[43]	≥50% diameter stenosis in ≥1 vessel	ICA	1683	55.7	0.78 (0.76 to 0.79)
	Jensen *et al*[44]	Lumen area diameter reduction ≥50% in ≥1 coronary artery	ICA	633	34.1	0.642 (*0.596 to 0.688*)
	Min *et al*[52]	≥50% luminal diameter stenosis in any coronary artery ≥1.5 mm in diameter	CCTA	2132	NA	0.64 (0.628 to 0.659)
	Rademaker *et al*[45]	>50% luminal diameter stenosis	CCTA	178	23.6	0.56 (0.49 to 0.64)
	Thomas *et al*[35]	≥1 diameter stenosis ≥50% in a major vessel on ICA by QCA (≥1.5 mm) or CCTA (≥2.0 mm)	CCTA, ICA	431	14.6	0.69 (0.62 to 0.75)
	Voros *et al*[37]	Outcome 50: ≥50% maximum diameter stenosis	CCTA	610	14	0.65 (0.59 to 0.71)
		Outcome 70: ≥70% maximum diameter stenosis	CCTA		NA	0.63 (0.53 to 0.73)
DFM/CASS	Adamson *et al*[47]	≥70% area stenosis in any major epicardial vessel or ≥50% stenosis in the left main stem	CCTA	4541 (PROMISE)	11.8	0.510 (0.506 to 0.514)
			CCTA	1619 (SCOT-HEART)	22.2	0.560 (0.548 to 0.573)
	Pickett *et al*[32]			1027	6.82	0.72 (0.66 to 0.78)
	Rosenberg *et al*[33]	≥1 atherosclerotic plaque in a major coronary artery (≥1.5 mm lumen diameter) causing ≥50% luminal diameter stenosis by QCA	ICA	526	36.5	0.663 (0.638 to 0.688)
Expanded clinical model	Rosenberg *et al*[33]	≥1 atherosclerotic plaque in a major coronary artery (≥1.5 mm lumen diameter) causing ≥50% luminal diameter stenosis by QCA	ICA	526	36.5	0.732 (*0.686 to 0.778*)
HRA score	Yang *et al*[48]	High-risk CAD: left main coronary artery diameter stenosis ≥50%, 3-vessel disease (≥70%) or 2-vessel disease involving the pLAD artery	CCTA	7333	4.8	0.71 (0.69 to 0.74)
K-score	Fujimoto *et al*[67]	Lesions with diameter stenosis of ≥75% were defined to be obstructive stenotic lesions. As for left main trunk lesion, lesions with diameter stenosis ≥50% were defined to be obstructive stenotic lesions.	CCTA	361	34.1	0.712 (*0.656 to 0.770*)
Morise score	Jensen *et al*[44]	Lumen area diameter reduction ≥50% in ≥1 coronary artery	ICA	633	34.1	0.681 (*0.636 to 0.726*)
	Pickett *et al*[32]	≥50% visual luminal diameter stenosis in ≥1 epicardial coronary artery segment ≥1.5 mm in diameter	CCTA	1027	6.82	0.68 (0.63 to 0.74)
	Rademaker *et al*[45]	>50% luminal diameter stenosis	CCTA	178	23.6	0.67 (0.60 to 0.74)
	Thomas *et al*[35]	≥1 diameter stenosis ≥50% in a major vessel on ICA by QCA (≥1.5 mm) or CCTA (≥2.0 mm)	CCTA, ICA	431	14.6	0.65 (0.59 to 0.74)
PROMISE Minimal Risk model	Fordyce *et al*[30]	Minimal risk: normal CCTA and further conditions*	CCTA	1528	25.0	0.713 (0.684 to 0.742)
uDFM	Adamson *et al*[47]	≥70% area stenosis in any major epicardial vessel or ≥50% stenosis in the left main stem	CCTA	4541 (PROMISE)	11.8	0.510 (0.506 to 0.514)
			CCTA	1619 (SCOT-HEART)	22.2	0.594 (0.579 to 0.610)

Continued

Table 3 Continued

Model	Study	Outcome definition	Reference test	Sample size	Prevalence (%)	AUC (95% CI)
	Adamson *et al*[38]	Luminal cross-sectional area stenosis of ≥70% (approximating to a 50% diameter stenosis) in at least 1 major epicardial vessel or ≥50% in the left main stem	CCTA, ICA	487	19.3	0.738 (0.687 to 0.788)
	Almeida *et al*[39]	Stenosis of >50% in at least one major epicardial vessel	ICA	2234	58.5	0.664 (0.641 to 0.687)
	Baskaran *et al*[40]	A stenosis causing ≥50% diameter stenosis	CCTA	1738	37.7	0.767 (0.744 to 0.790)
	Bittencourt *et al*[28]	At least 1 segment (with a >2 mm diameter) with a lesion with ≥50% diameter stenosis	CCTA	2274	22	0.714 (0.689 to 0.737)
	Ferreira *et al*[42]	Coronary diameter stenosis ≥50%	CCTA	1069	13.8	0.70 (0.67 to 0.72)
	Genders *et al*[43]	≥50% diameter stenosis in ≥1 vessel	ICA	471	NA	0.76 (0.71 to 0.81)
	Jensen *et al*[44]	Lumen area diameter reduction ≥50% in ≥1 coronary artery	ICA	633	34.1	0.714 (*0.670 to 0.758*)
	Rademaker *et al*[45]	>50% luminal diameter stenosis	CCTA	178	23.6	0.61 (0.53 to 0.68)
	Winther *et al*[46]	Coronary diameter stenosis reduction ≥50% in all segments with a reference vessel diameter >2 mm	CCTA	1653	23.7	0.65 (0.61 to 0.68)
	Yang *et al*[48]	High-risk CAD: left main coronary artery diameter stenosis ≥50%, 3-vessel disease (≥70%) or 2-vessel disease involving the pLAD artery	CCTA	24 251	3.6	0.64 (0.62 to 0.67)
	Zhang *et al*[49]	≥1 lesion with ≥50% diameter stenosis	CCTA	Men: 3001	39	0.782 (*0.764 to 0.800*)
				Women: 2776	25	0.678 (*0.654 to 0.702*)
	Zhou *et al*[50]	≥1 lesion with ≥50% diameter stenosis or any non-assessable segments due to severe calcification	CCTA	5743	32.6	0.765 (0.751 to 0.779)
uDFM-cTn	Adamson *et al*[38]	Luminal cross-sectional area stenosis of ≥70% (approximating to a 50% diameter stenosis) in at least 1 major epicardial vessel or ≥50% in the left main stem	CCTA, ICA	487	19.3	0.757 (0.706 to 0.808)

Values in Italic are derived by the statistician (AB).
*Further conditions are considered and should be all present, in addition to normal CCTA, for a subject to be at minimal risk: (1) coronary artery calcium score was 0 or was not obtained; (2) no evidence of atherosclerosis; (3) overall study quality was diagnostic (ie, sufficient data quality for interpretation); (4) left ventricular function was normal or not reported; (5) no wall motion abnormalities were present or not reported; and (6) no relevant cardiovascular incidental findings that could account for the patients' symptoms (ie, aortic dissection or pulmonary embolism) were noted. All patients with normal CCTA results were included in the minimal risk cohort in the absence of any of the following adjudicated clinical events during the median 25-month follow-up period: all-cause death, non-fatal MI, unstable angina hospitalisation or revascularisation during the entire follow-up period
AUC, area under receiver operating characteristic curve; CAD, coronary artery disease; CADC-Basic, CAD Consortium Basic model; CADC-Clin, CAD Consortium Clinical model; CASS, Coronary Artery Surgery Study; CCTA, coronary CT angiography; DCS, Duke Clinical Score; DFM, Diamond-Forrester model; HRA, High Risk Anatomy; ICA, invasive coronary angiography; MI, myocardial infarction; NA, not available; pLAD, proximal left anterior descending; PTP, pretest probability; QCA, quantitative coronary angiography; uDFM, updated DFM.

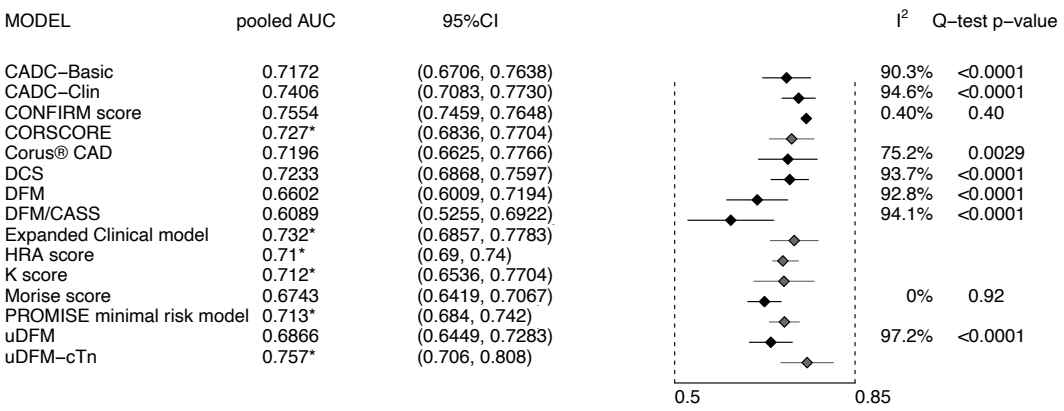

Figure 2 Summary of the meta-analyses. Models that were validated by one study only are denoted by area under receiver operating characteristic curve (AUC)* and a grey colour in the graphic. CAD, coronary artery disease; CADC-Basic, CAD Consortium Basic model; CADC-Clin, CAD Consortium Clinical model; CASS, Coronary Artery Surgery Study; DCS, Duke Clinical Score; DFM, Diamond-Forrester model; HRA, High Risk Anatomy; uDFM, updated DFM.

with the CONFIRM history-based score do not lead to a clear evaluation of the advantages of one over the other in terms of AUC,[40 42] while the CONFIRM score proves to be better than the DFM.[52] Figures 3 and 4 show the forest plot of the meta-analyses for uDFM and CADC-Clin model, the two most validated models. The heterogeneity for the uDFM model is not significantly reduced by removing the two threshold validations in Adamson *et al*[47] (I^2=95% vs I^2=97.4%). For the uDFM and CADC-Clin models, a meta-regression analysis was also conducted which did not lead to any significant result.

The traditional DCS generally overestimates prevalence and shows a lack of fit by the HL test. Moreover, miscalibration results from a reduced effect of sex and chest pain typicality and an increased effect of diabetes and dyslipidaemia.[51]

The Corus CAD model stands out from the other models because it defines an age-specific and sex-specific gene expression score. Validation is performed by AUC comparisons, HL test and additivity to DFM and other models. The validation procedures show significant AUC improvement when the score is added to other models (eg, 0.81 vs 0.65 when added to Morise score, with non-overlapping CIs[35]; 0.721 vs 0.663 when added to DFM, p=0.003[33]; not shown in the table). Testing the Corus CAD model on different data sets from an extension of the original validation population provides results very similar to the original ones.[29]

Finally, the Minimal Risk model upsets the usual point of view because it aims to directly identify patients with chest pain and normal coronary arteries. Unfortunately, the only other external validation published up to the date of our search[57] cannot be considered here because it was based on a former version of Fordyce *et al*[30] that included some computational errors.[58]

DISCUSSION

External validation is an indispensable tool for investigating the generalisability of a PTP model to populations

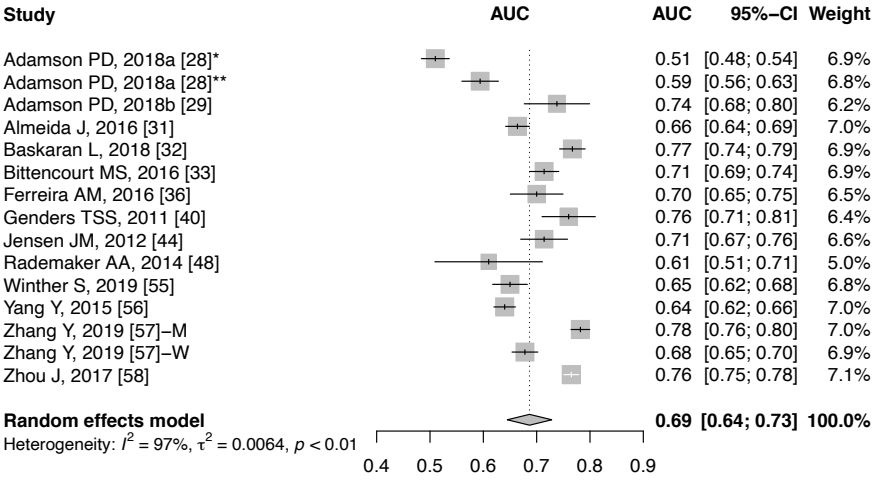

Figure 3 Forest plot of the meta-analysis for the updated Diamond-Forrester model. *PROMISE trial; **SCOT-HEART trial. AUC, area under receiver operating characteristic curve.

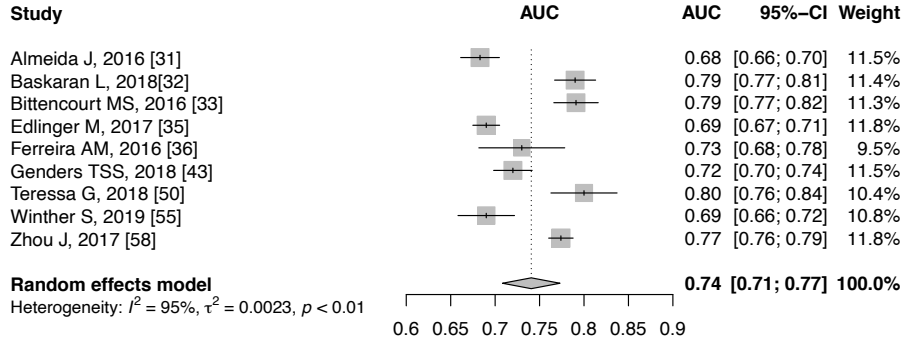

Figure 4 Forest plot of the meta-analysis for the CAD Consortium Clinical model. AUC, area under receiver operating characteristic curve; CAD, coronary artery disease.

that differ from the development population study. This process can use different approaches, from the computation of indices to more complex procedures that aim at understanding how the original model should adapt to the new population. The papers included in this review mainly relied on AUC. The advantage of this index lies in being suitable both for individual evaluations and for rigorous comparisons. However, the AUC is a summary: only the whole ROC curve will allow evaluation of the clinical usefulness of a test by showing the true positive and false positive fractions that will be obtained for any eventually chosen cut-off.

Most of the papers included in this review did not provide a careful assessment of the discriminative performances of the validated model with respect to a well-defined threshold, but limited to compute sensitivity and specificity with respect to the thresholds suggested by either European or American guidelines. Studies on the CAD Consortium models and the Corus CAD model are exceptions. As far as the CAD Consortium models are concerned, clinical usefulness is assessed at cut-offs that vary from 5% to 20%. A cut-off of 14.75 (15 in subsequent works) was identified for the Corus CAD model in the main work,[33] a value that corresponds to a disease likelihood of 20% on a validation data set (positivity for index ≤15). Notably, Corus CAD recently lost Medicare coverage in the USA.[59] The very low AUC values obtained by Adamson *et al*[47] at the cut-off of 15% in the comparison of the performance of major guidelines for the assessment of stable chest pain including risk-based strategies are representative of a general clinical protection approach leading clinicians to prefer a very high sensitivity, which of course implies low specificity.[60 61]

Despite the fact that all the models are obtained by regression techniques, which allow the interpretation of the effect of the predictor on the outcome of interest, very few papers[31 34 41 43] address a complete validation procedure without rejecting a model after obtaining a poor preliminary performance on the new population by some test. Rather, a different model is developed, without any further in-depth analysis of the failure reason. Regardless of the quality of the new developed model, the lack of adequate consideration of in-depth validation procedures involves the loss of the information captured by the initial study and hinders a deep understanding of how effect size of relevant risk factors can change in a different geographical or setting framework.[24] For instance, deep validation procedures like miscalibration analysis allow questioning the effect of chest pain typicality in different data sets.[31 34 41] This finding is consistent with what was recently noted by Di Carli and Gupta[62]: angina remains a common presenting symptom in a high proportion of patients with cardiac condition who do not show obstructive lesions in their coronary angiograms.

The diagnostic question is central in the determination of which diagnostic pathway and test is the most appropriate[62 63] and also affects statistical analysis. A carefully defined outcome should be required to provide a reliable basis for the evaluation of the effect of any predictive variable.[64] When referring to validation specifically, the application of a statistical model to predict an outcome different from the originally intended one raises some concerns and, eventually, should be explicitly noted. In data-driven models, the outcome definition in the population study also influences predictor selection. Thus, a small AUC value in the validation set does not necessarily indicate a lower performance of the original model on the new population. Instead, it suggests that the model may not be appropriate for the context.[57]

Despite meta-regression not being able to statistically assess the portion of heterogeneity explained by differences in sample size, prevalence and choice of the anatomical reference test, differences between studies in terms of the way the outcomes are defined and measured contribute to the methodological heterogeneity we narratively highlighted in this review.[65 66]

The main strengths of this review were the large number and high quality of included studies, the attention paid to validation procedures, as well as to AUC values alone and the careful consideration of different aspects yielding heterogeneity, as well as statistical heterogeneity alone.

The study had limitations. Most studies mainly refer to Western populations with a minority of studies referring to Asian subjects (Japan, South Korea and China).[48–50 52 56 67] Another limitation was that most of the studies did not investigate the use of any threshold. Pooled AUC values

from meta-analyses can provide only an approximate summary of the discrimination capacity of most of the models, due to the low number of validating studies. This also affects the analysis of heterogeneity due to the low power of the test, and the feasibility of meta-regression.[68] Although the focus of our meta-analysis was not a measure of an intervention effect, the meta-analysis was limited in the consideration of other possible sources of heterogeneity, mainly clinical like mean age or proportion of women. However, a multivariable analysis considering all the study-related variables together would have been unreliable, due to the low number of validations for most of the models.

Finally, in this review, we only considered pretest models developed in cohorts of patients referred for ICA or CCTA. Our choice was determined by main guidelines and traditional, well-established models. However, the need of models that are able to predict functionally significant CAD has been underlined,[69] for prognostic purposes as well. Nevertheless, how these alternative models could be used in a risk-stratification approach to guide further patient–clinician decision-making has not been assessed yet.

CONCLUSIONS

Several agencies and scientific organisations emphasise the need for increasing the knowledge on how the prediction of the disease can be modified according to the risk factors present in any specific study population or, possibly, in any particular patient. This would indeed improve the precision of the estimated clinical likelihood of CAD. However, the increasing availability of large data sets and the highly improved computational power seem to have directed large part of recent researches towards model development rather than model validation.[16] First of all, our review makes an important selection among the many developed models by mainly considering those externally validated. Then, it provides insights into the effects of traditional and emerging risk factors, biomarkers and comorbidities on the PTP of obstructive CAD. Finally, our findings lead to the following important recommendations. To achieve a more robust exploitation of PTP models in decision-making processes, significant endpoints should be more clearly stated and consistently measured both in the derivation and validation phases. In addition, more comprehensive validation analyses should be adopted to understand model weaknesses and variations. Finally, increased efforts are still needed to threshold validation and to analyse the effect of PTP on clinical management.

Author affiliations
[1]Institute for Research on Population and Social Policies, National Research Council, Brindisi, Italy
[2]Institute for Applied Mathematics and Information Technologies "Enrico Magenes", National Research Council, Milan, Italy
[3]Institute of Clinical Physiology, National Research Council, Pisa, Italy
[4]Institute of Clinical Physiology, National Research Council, Lecce, Italy

Acknowledgements The authors express their great appreciation to Dr Philip D Adamson for the valuable discussion on some methodological aspects. The authors would also like to thank Dr Tommaso Leo for the clarifications provided on some clinical aspects and Roberto Guarino (National Research Council of Italy, Institute of Clinical Physiology, Lecce) for his technical support (informatics tools for document management and title and abstract screening). Finally, the authors express their appreciation for the insightful comments of the two reviewers, Professor William Wijns and Dr Marc Dewey, that allowed the improvement of the manuscript.

Contributors AB, CGL, PM, GP and SS provided substantial contribution to the conception of the work. CGL and PM performed the literature search and retrieved selected publications. All the authors (PM, AB, MRT, FV, SR, GP, CC, SS and CGL) contributed to the extraction and interpretation of data. AB carried out the meta-analysis and the meta-regression analysis. AB, CGL, PM and MRT assessed the quality of included studies. All the authors contributed to drafting the work. AB, CGL and PM revised it critically. All the authors approved the version to be published and are accountable for all aspects of the work. CGL is responsible for the overall content as the guarantor.

Funding Part of this work was supported by the European Union Horizon 2020 research and innovation programme under grant agreement no 689068—Project 'Simulation Modelling of coronary ARTery disease:a tool for clinical decision support (SMARTool)'.

Disclaimer This publication reflects only the authors' view and the commission, which has no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript, is not responsible for any use that may be made of the information it contains. The funding source (European Commission) had no role in the study. All the authors are independent from funders, had full access to all of the data in the study, and can take responsibility for the integrity of the data and the accuracy of the data analysis.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval Our work of systematic review and meta-analyses of published literature does not require a research ethics approval.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplemental information.

ORCID iDs
Pierpaolo Mincarone http://orcid.org/0000-0002-8452-5642
Carlo Giacomo Leo http://orcid.org/0000-0002-1357-4578

REFERENCES

1 Roth GA, Mensah GA, Johnson CO, *et al*. Global burden of cardiovascular diseases and risk factors, 1990-2019: update from the GBD 2019 study. *J Am Coll Cardiol* 2020;76:2982–3021.
2 Task Force Members, Montalescot G, Sechtem U, *et al*. 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.
3 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: Executive summary: 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;60:2564–603.
4 Carpeggiani C, Landi P, Michelassi C, *et al*. Trends of increasing medical radiation exposure in a population hospitalized for cardiovascular disease (1970-2009). *PLoS One* 2012;7:e50168.
5 Alter DA, Stukel TA, Newman A. Proliferation of cardiac technology in Canada: a challenge to the sustainability of Medicare. *Circulation* 2006;113:380–7.
6 Lucas FL, DeLorenzo MA, Siewers AE, *et al*. Temporal trends in the utilization of diagnostic testing and treatments for cardiovascular disease in the United States, 1993-2001. *Circulation* 2006;113:374–9.
7 Leo CG, Carpeggiani C, Picano E. Cost and benefit in cardiovascular imaging: the quest for economic sustainability. *Int J Cardiovasc Imaging* 2010;26:613–6.
8 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.
9 Pryor DB, Harrell FE, Lee KL, *et al*. Estimating the likelihood of significant coronary artery disease. *Am J Med* 1983;75:771–80.
10 Gonzalez JA, Lipinski MJ, Flors L, *et al*. Meta-Analysis of diagnostic performance of coronary computed tomography angiography, computed tomography perfusion, and computed Tomography-Fractional flow reserve in functional myocardial ischemia assessment versus invasive fractional flow reserve. *Am J Cardiol* 2015;116:1469–78.
11 Knuuti J*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). *Russ J Cardiol*;25:119–80.
12 National Institute for Health and Clinical Excellence (NICE).. *Chest pain of recent onset: assessment and diagnosis. CG95*. London, UK, 2016. https://www.nice.org.uk/guidance/ cg95
13 Bing R, Singh T, Dweck MR, *et al*. Validation of European Society of cardiology pre-test probabilities for obstructive coronary artery disease in suspected stable angina. 2020;6:293–300.
14 Villines TC. Coronary CTA Should Be the Initial Test in Most Patients With Stable Chest Pain: PRO - American College of Cardiology, 2018. Available: https://www.acc.org/latest-in-cardiology/articles/2018/05/21/06/37/coronary-cta-pro [Accessed 6 Dec 2019].
15 Maini R, Moscona J, Yousuf T. Coronary CTA Should Be the Initial Test in Most Patients With Stable Chest Pain: CON - American College of Cardiology, 2018. Available: https://www.acc.org/latest-in-cardiology/articles/2018/05/21/06/37/coronary-cta-con [Accessed 6 Dec 2019].
16 Adibi A, Sadatsafavi M, Ioannidis JPA. Validation and utility testing of clinical prediction models: time to change the approach. *JAMA* 2020;324:235–6.
17 Liberati A, Altman DG, Tetzlaff J, *et al*. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009;339:b2700.
18 Booth A, Clarke M, Ghersi D, *et al*. An international registry of systematic-review protocols. *Lancet* 2011;377:108–9.
19 He T, Liu X, Xu N, *et al*. Diagnostic models of the pre-test probability of stable coronary artery disease: a systematic review. *Clinics* 2017;72:188–96.
20 Alizadehsani R, Abdar M, Roshanzamir M, *et al*. Machine learning-based coronary artery disease diagnosis: a comprehensive review. *Comput Biol Med* 2019;111:103346.
21 Wallace BC, Small K, Brodley CE. *Deploying an interactive machine learning system in an evidence-based practice center: abstrackr. In: ACM International Health Informatics Symposium (IHI)*: 819–24. http://www.cebm.brown.edu/software
22 Whiting PF, Rutjes AWS, Westwood ME, *et al*. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529.
23 Cook NR. Quantifying the added value of new biomarkers: how and how not. *Diagn Progn Res* 2018;2:14.
24 Moons KGM, Kengne AP, Grobbee DE, *et al*. Risk prediction models: II. external validation, model updating, and impact assessment. *Heart* 2012;98:691–8.
25 R Core Team. R: a language and environment for statistical computing, 2020. Available: https://www.r-project.org/
26 Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health* 2019;22:153–60.
27 Gardner J. Dates and times made easy with lubridate. model evaluation with auctestr, 2017. Available: https://github.com/jpgard/auctestr [Accessed 5 May 2021].
28 Bittencourt MS, Hulten E, Polonsky TS, *et al*. European Society of Cardiology-Recommended coronary artery disease Consortium pretest probability scores more accurately predict obstructive coronary disease and cardiovascular events than the diamond and Forrester score: the partners registry. *Circulation* 2016;134:201–11.
29 Daniels SE, Beineke P, Rhees B, *et al*. Biological and analytical stability of a peripheral blood gene expression score for obstructive coronary artery disease in the predict and COMPASS studies. *J Cardiovasc Transl Res* 2014;7:615–22.
30 Fordyce CB, Douglas PS, Roberts RS, *et al*. Identification of patients with stable chest pain deriving minimal value from noninvasive testing: the promise Minimal-Risk tool, a secondary analysis of a randomized clinical trial. *JAMA Cardiol* 2017;2:400–8.
31 Genders TSS, Coles A, Hoffmann U, *et al*. 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. *JACC Cardiovasc Imaging* 2018;11:437–46.
32 Pickett CA, Hulten EA, Goyal M, *et al*. Accuracy of traditional age, gender and symptom based pre-test estimation of angiographically significant coronary artery disease in patients referred for coronary computed tomographic angiography. *Am J Cardiol* 2013;112:208–11.
33 Rosenberg S, Elashoff MR, Beineke P, *et al*. Multicenter validation of the diagnostic accuracy of a blood-based gene expression test for assessing obstructive coronary artery disease in nondiabetic patients. *Ann Intern Med* 2010;153:425–34.
34 Teressa G, Zhang M, Lavenburg P, *et al*. Validity of Coronary Artery Disease Consortium Models for Predicting Obstructive Coronary Artery Disease & Cardiovascular Events in Patients with Acute Chest Pain Considered for Coronary Computed Tomographic Angiography. *Am J Cardiol* 2018;122:1310–21.
35 Thomas GS, Voros S, McPherson JA, *et al*. A blood-based gene expression test for obstructive coronary artery disease tested in symptomatic nondiabetic patients referred for myocardial perfusion imaging the COMPASS study. *Circ Cardiovasc Genet* 2013;6:154–62.
36 Voora D, Coles A, Lee KL, *et al*. An age- and sex-specific gene expression score is associated with revascularization and coronary artery disease: insights from the prospective multicenter imaging study for evaluation of chest pain (promise) trial. *Am Heart J* 2017;184:133–40.
37 Voros S, Elashoff MR, Wingrove JA, *et al*. A peripheral blood gene expression score is associated with atherosclerotic plaque burden and stenosis by cardiovascular CT-angiography: results from the predict and COMPASS studies. *Atherosclerosis* 2014;233:284–90.
38 Adamson PD, Hunter A, Madsen DM, *et al*. High-Sensitivity cardiac troponin I and the diagnosis of coronary artery disease in patients with suspected angina pectoris. *Circ Cardiovasc Qual Outcomes* 2018;11:e004227.
39 Almeida J, Fonseca P, Dias T, *et al*. Comparison of coronary artery disease Consortium 1 and 2 scores and Duke clinical score to predict obstructive coronary disease by invasive coronary angiography. *Clin Cardiol* 2016;39:223–8.
40 Baskaran L, Danad I, Gransar H, *et al*. A Comparison of the Updated Diamond-Forrester, CAD Consortium, and CONFIRM History-Based Risk Scores for Predicting Obstructive Coronary Artery Disease in Patients With Stable Chest Pain: The SCOT-HEART Coronary CTA Cohort. *JACC Cardiovasc Imaging* : 2019;12:1392-1400.
41 Edlinger M, Wanitschek M, Dörler J, *et al*. External validation and extension of a diaGnostic model for obstructive coronary artery disease: a cross-sectional predictive evaluation in 4888 patients of the Austrian coronary artery disease risk determination in Innsbruck by diaGnostic angiography (CARDIIGAN) cohort. *BMJ Open* 2017;7:e014467.
42 Ferreira AM, Marques H, Tralhão A, *et al*. Pre-test probability of obstructive coronary stenosis in patients undergoing coronary CT angiography: comparative performance of the modified diamond-Forrester algorithm versus methods incorporating cardiovascular risk factors. *Int J Cardiol* 2016;222:346–51.
43 Genders TSS, 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.
44 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.
45 Rademaker AAEM, Danad I, Groothuis JGJ, *et al*. Comparison of different cardiac risk scores for coronary artery disease in symptomatic women: do female-specific risk factors matter? *Eur J Prev Cardiol* 2014;21:1443–50.
46 Winther S, Nissen L, Westra J, *et al*. Pre-test probability prediction in patients with a low to intermediate probability of coronary artery disease: a prospective study with a fractional flow reserve endpoint. *Eur Heart J Cardiovasc Imaging* 2019;20:1208–18.

47 Adamson PD, Newby DE, Hill CL, *et al*. Comparison of International Guidelines for Assessment of Suspected Stable Angina: Insights From the PROMISE and SCOT-HEART. *JACC Cardiovasc Imaging* 2018;11:1301–10.
48 Yang Y, Chen L, Yam Y, *et al*. A clinical model to identify patients with high-risk coronary artery disease. *JACC Cardiovasc Imaging* 2015;8:427–34.
49 Zhang Y, Liu Y, Zhang H, *et al*. Impact of sex-specific differences in calculating the pretest probability of obstructive coronary artery disease in symptomatic patients: a coronary computed tomographic angiography study. *Coron Artery Dis* 2019;30:124–30.
50 Zhou J, Liu Y, Huang L, *et al*. Validation and comparison of four models to calculate pretest probability of obstructive coronary artery disease in a Chinese population: a coronary computed tomographic angiography study. *J Cardiovasc Comput Tomogr* 2017;11:317–23.
51 Genders TSS, Steyerberg EW, Hunink MGM, *et al*. Prediction model to estimate presence of coronary artery disease: retrospective pooled analysis of existing cohorts. *BMJ* 2012;344:e3485.
52 Min JK, Dunning A, Gransar H, *et al*. Medical history for prognostic risk assessment and diagnosis of stable patients with suspected coronary artery disease. *Am J Med* 2015;128:871–8.
53 Wilson PWF. Assessing coronary heart disease risk with traditional and novel risk factors. *Clin Cardiol* 2004;27:7–11.
54 Glaser A. *High - Yield Biostatistics Epidemiology, and Public Health*. 4th edn. Philadelphia, PA, 2013.
55 Imrey PB. Limitations of meta-analyses of studies with high heterogeneity. *JAMA Netw Open* 2020;3:e1919325.
56 Wang M, Liu Y, Zhou X, *et al*. Coronary calcium score improves the estimation for pretest probability of obstructive coronary artery disease and avoids unnecessary testing in individuals at low extreme of traditional risk factor burden: validation and comparison of confirm score and genders extended model. *BMC Cardiovasc Disord* 2018;18:176.
57 Adamson PD, Fordyce CB, McAllister DA, *et al*. Identification of patients with stable chest pain deriving minimal value from coronary computed tomography angiography: an external validation of the promise minimal-risk tool. *Int J Cardiol* 2018;252:31–4.
58 Fordyce CB, Douglas PS, Udelson JE. Errors in programming and coding affecting cohorts included in the study deriving and validating the promise Minimal-Risk tool. *JAMA Cardiol* 2018;3:1253.
59 CardioDx, maker of heart disease test, shutting down as Medicare rescinds coverage - SFChronicle.com. Available: https://www.sfchronicle.com/business/article/CardioDx-maker-of-heart-disease-test-shutting-13518778.php [Accessed 26 Jul 2019].
60 Richardson WS, Wilson MC, Guyatt GH, *et al*. Users' Guides to the Medical Literature. *JAMA* 1999;281:1214.
61 Knottnerus JA. *The evidence base of clinical diagnosis*. BMJ Books, 2002. https://www.libreriacortinamilano.it/scheda-libro/j-andre-knottnerus/the-evidence-base-of-clinical-diagnosis-9780727915719-19164.html
62 Di Carli MF, Gupta A. Estimating Pre-Test Probability of Coronary Artery Disease. *JACC Cardiovasc Imaging* 2019;12:1401–4.
63 Hecht HS, Shaw L, Chandrashekhar YS, *et al*. Should NICE guidelines be universally accepted for the evaluation of stable coronary disease? A debate. *Eur Heart J* 2019;40:1440–53.
64 Hlatky MA, Greenland P, Arnett DK, *et al*. Criteria for evaluation of novel markers of cardiovascular risk: a scientific statement from the American heart association. *Circulation* 2009;119:2408–16.
65 West SL, Gartlehner G, Mansfield AJ. *Discussion - Comparative Effectiveness Review Methods: Clinical Heterogeneity*. Rockville, MD: Agency for Healthcare Research and Quality (US), 2010. https://www.ncbi.nlm.nih.gov/books/NBK53308/
66 Deeks JJ, Higgins JP, Altman DG. Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JP, Thomas J, Chandler J, *et al*, eds. *Cochrane Handbook for systematic reviews of interventions. version 6.2. Cochrane*, 2021. https://training.cochrane.org/handbook/current/chapter-10#section-10-10
67 Fujimoto S, Kondo T, Yamamoto H, *et al*. Development of new risk score for pre-test probability of obstructive coronary artery disease based on coronary CT angiography. *Heart Vessels* 2015;30:563–71.
68 Thompson SG, Higgins JPT. How should meta-regression analyses be undertaken and interpreted? *Stat Med* 2002;21:1559–73.
69 Caselli C, Rovai D, Lorenzoni V, *et al*. A New Integrated Clinical-Biohumoral Model to Predict Functionally Significant Coronary Artery Disease in Patients With Chronic Chest Pain. *Can J Cardiol* 2015;31:709–16.
70 Madsen DM, Diederichsen ACP, Hosbond SE, *et al*. Diagnostic and prognostic value of a careful symptom evaluation and high sensitive troponin in patients with suspected stable angina pectoris without prior cardiovascular disease. *Atherosclerosis* 2017;258:131–7.
71 Nieman K, Galema T, Weustink A, *et al*. Computed tomography versus exercise electrocardiography in patients with stable chest complaints: real-world experiences from a fast-track chest pain clinic. *Heart* 2009;95:1669–75.
72 Min JK, Dunning A, Lin FY, *et al*. Rationale and design of the confirm (coronary CT angiography evaluation for clinical outcomes: an international multicenter) registry. *J Cardiovasc Comput Tomogr* 2011;5:84–92.

Supplementary Material

Additional FILE 1 - Search strategy – It is the full search string adopted in OVID.

Additional File 2 – Study design and Eligibility Criteria of main studies – It provides details on the main studies cited in Table 1.

Additional File 3 – Proportion of studies with low, high or unclear risk of bias – It is a summary of the quality assessment according to QUADAS2.

Additional FILE 1 - Search strategy

The same search strategy has been applied to the three considered databases. The number reported in each single line refers to the sum of papers retrieved in each database as these are queried at the same time.

▼ **Search History** (28)

# ▲	Searches	Results
1	▶ Angina Pectoris/ or Angina Pectoris.af. or Angina, Stable/ or (Stable Angina* or Chronic Angina*).af.	98847
2	▶ Chest Pain/ or Chest Pain*.af.	72609
3	▶ Coronary Heart Disease/ or (CHD or Coronary Disease* or Coronary Heart Disease*).af.	412639
4	▶ Coronary Artery Disease/ or (CAD or Coronary Artery Disease* or Coronary Arteriosclerosis or Coronary Atheroscleros*).af.	288034
5	▶ Coronary Stenosis/ or (Coronary Stenos* or Artery Stenos*).af.	64518
6	▶ or/1-5	770062
7	▶ stratification score*.af.	492
8	▶ Likelihood Functions/ or Likelihood Function*.af. or (likelihood adj5 disease).af. or CAD likelihood.af. or "predict* CAD".af.	46652
9	▶ forecasting/ or (pre-test probabilit* or PTP or predictive model* or prediction or forecast*).af.	664543
10	▶ (probability adj5 disease).af.	8725
11	▶ or/7-10	716518
12	▶ Coronary Angiography/ or Angiograph*.af.	574196
13	▶ Angiocardiography/ or Angiocardiograph*.af.	21132
14	▶ Cardiac Catheterization/ or (Cardiac Catheterization* or Heart Catheterization*).af.	122622
15	▶ Computed Tomography Angiography/ or (Coronary Computed Tomography Angiograph* or CCTA).af.	21751
16	▶ or/12-15	677008
17	▶ 6 and 11 and 16	6069
18	▶ (ANIMALS not HUMANS).sh.	7244023
19	▶ 17 not 18	5985
20	▶ limit 19 to english language or limit 19 to italian	5708
21	▶ remove duplicates from 20	2682

Additional File 2 – Study design and Eligibility Criteria of main studies

Acronym	**COMPASS**
Name	Coronary Obstruction Detection by Molecular Personalized Gene Expression (Corus CAD or ASGES)
ClinicalTrials.gov Identifier	NCT01117506
Study design	Observational prospective study. The study enrolled a patient population that presented with stable chest pain syndrome or anginal equivalent and referred for stress myocardial profusion imaging
Study Center(s)	Multicenter trial: US (7 centers)
Inclusion Criteria:	• Ages 45-90 for women; 35-90 for men. • Stable chest pain syndrome (typical or atypical) or anginal equivalent in the judgment of the investigator (e.g., pain in the neck, jaw, arm or shoulder or dyspnea possibly due to cardiac ischemia). • Referred for a stress test using Myocardial Perfusion Imaging. • The patient has signed the appropriate Institutional Review Board approved Informed Consent Form
Exclusion Criteria	• History of known MI or significant CAD. • Current MI or acute coronary syndrome. • Current New York Heart Association (NYHA) class III or IV congestive heart failure symptoms. • Severe regurgitant or stenotic cardiac valvular lesion. • Severe left ventricular systolic dysfunction (LVEF ≤ 35 % documented in the last year); if no assessment was performed or documented in the year preceding enrolment, presume normal LVEF.

- Active systemic infection (diagnosed by a combination of clinical symptoms and laboratory testing, including but not limited to fever, leukocytosis, positive blood cultures, pneumonia, urinary tract infection, or abscess in the preceding 2 months) or chronic infection (e.g., HIV, Hepatitis B or C, Tuberculosis).
- Protocol-specified rheumatologic, autoimmune or hematologic conditions (e.g., rheumatoid arthritis, systemic lupus erythematosis, polymyalgia rheumatica, or systemic sarcoidosis).
- Known or suspected diabetes mellitus or documented Hemoglobin A1c (HbA1c) ≥ 6.5; presume normal HbA1c if none documented.
- Total WBC ≥ 11,000 cells/ul and platelet count ≤ 75,000 cells/ul from a CBC with differential drawn within 7 days prior to enrollment [WBC ≥ 11,000 cells/ul and platelet count ≤ 75,000 cells/ul from a CBC drawn > 7 days prior need to be re-drawn at enrollment].
- Recipient of any organ transplant.
- Immunosuppressive or immunomodulatory therapy including any dose of systemic corticosteroids in the preceding 2 months.
- Chemotherapy in the preceding year.
- Major surgery in the preceding 2 months.
- Blood or blood product transfusion in the preceding 2 months.
- Subjects for whom all forms (stress or pharmacologic) of MPI are contraindicated.
- Subjects for whom invasive coronary angiography or coronary CT angiography is contraindicated, including IV beta-blocker.
- Subjects who planned to decline research CCTA or invasive coronary

	angiography, regardless of MPI result. • Subjects with history of atrial fibrillation/flutter or frequent irregular or rapid heart rhythms. • Known history of renal insufficiency (serum creatinine ≥ 2.0 mg/dL), or severe allergy to iodinated contrast.
Acronym	**CONFIRM**
Name	COroNary CT Angiography Evaluation For Clinical Outcomes: An InteRnational Multicenter Registry
ClinicalTrials.gov Identifier	NCT01443637
Study design	Observational prospective study. Patients included in the CONFIRM Registry are those that have previously undergone clinically-indicated CCTA as part of their standard of care.
Study Center(s)	Multicenter trial: North America, Europe and Asia
Inclusion Criteria:	• Age > 18 years • Evaluation by CCTA with 64-detector rows or greater for CAD evaluation as part of standard of care • Interpretable CCTA • Prospective data collection for CAD risk factors.
Exclusion Criteria	• No explicit patient exclusion criteria are defined.
Acronym	**PREDICT**
Name	Personalized Risk Evaluation and Diagnosis (Using Corus CAD or ASGES) in

	the Coronary Tree
ClinicalTrials.gov Identifier	NCT00500617
Study design	Observational prospective study. The study enrolled patients undergoing clinically indicated invasive coronary artery angiogram or CT angiogram
Study Center(s)	Multicenter trial: US
Inclusion Criteria:	• Age 21 to 99 Years • Referral for a coronary angiogram (either invasive X-ray angiography or coronary CTA) • Any one of the following clinical syndromes: o chest pain syndrome, stable angina, or anginal equivalent suggesting myocardial ischemia o low-risk unstable angin, or o asymptomatic individuals with a high probability of CAD.
Exclusion Criteria	• History of myocardial infarction or known CAD • Current Myocardial Infarct (MI), acute coronary syndrome with high-risk features or unstable angina with high-risk features • New York Heart Association (NYHA) class III or IV congestive • Inability to give informed congestive heart failures • Severe left ventricular systolic dysfunction (LVEF<35%) • Severe regurgitant or stenotic cardiac valve lesion • Active or chronic systemic infection • Rheumatologic, autoimmune or hematologic conditions • Any organ transplant • Immunosuppressive therapy

	• Chemotherapy in the preceding year • Major blood or blood product transfusion in the preceding 2 months
Acronym	**PROMISE**
Name	PROspective Multicenter Imaging Study for Evaluation of Chest Pain
ClinicalTrials.gov Identifier	NCT01174550
Study design	Interventional (Clinical Trial) randomized study
Study Center(s)	Multicenter trial: North America
Inclusion Criteria:	• New or worsening chest pain suspicious for clinically significant coronary artery disease (CAD) • no prior evaluation for this episode of symptoms • planned non-invasive testing for diagnosis • men age ≥55 years • men age ≥45 years with increased probability of coronary artery disease (CAD) due to either (A. Diabetes Mellitus (DM) requiring medical treatment OR Peripheral Arterial Disease (PAD) defined as documented >50% peripheral arterial stenosis treated medically or invasively OR cerebrovascular disease (stroke, documented > 50% carotid stenosis treated medically or invasively) OR B. At least one of the following cardiovascular risk factors: 1-Ongoing tobacco use, 2-Hypertension, 3-Abnormal ankle brachial index (ABI) defined as less than <0.9, 4-Dyslipidemia • women age ≥65 years

	• women age ≥50 years with increased probability of coronary artery disease (CAD) due to either (A. Diabetes Mellitus (DM) requiring medical treatment OR Peripheral Arterial Disease (PAD) defined as documented >50% peripheral arterial stenosis treated medically or invasively OR cerebrovascular disease (stroke, documented > 50% carotid stenosis treated medically or invasively) OR B. At least one of the following cardiovascular risk factors: 1-Ongoing tobacco use, 2-Hypertension, 3-Abnormal ankle brachial index (ABI) defined as less than <0.9, 4-Dyslipidemia • Serum creatinine ≤ 1.5 mg/dL within the past 90 days • Negative urine/serum pregnancy test for female subjects of child-bearing potential
Exclusion Criteria	• Diagnosed or suspected acute coronary syndrome (ACS) requiring hospitalization or urgent or emergent testing; Elevated troponin or creatine kinase-myocardial band (CK-MB) • Hemodynamically or clinically unstable condition systolic blood pressure (BP) < 90 mmHg, atrial or ventricular arrhythmias, or persistent resting chest pain felt to be ischemic despite adequate therapy) • Known coronary artery disease (CAD) with prior Myocardial infarction (MI), percutaneous coronary intervention (PCI), coronary artery bypass graft (CABG) or any angiographic evidence of coronary artery disease (CAD) ≥50% lesion in a major epicardial vessel • Any invasive coronary angiography or non-invasive anatomic or functional cardiovascular (CV) test for detection of coronary artery

	disease (CAD), including coronary tomographic angiography (CTA) and exercise electrocardiogram (ECG), within the previous twelve (12) months • Known significant congenital, valvular (> moderate) or cardiomyopathic process (hypertrophic cardiomyopathy or reduced systolic left ventricular function (LVEF) ≤ 40%)) which could explain cardiac symptoms • Contraindication to undergoing a coronary tomographic angiography (CTA), including but not limited to: a. Allergy to iodinated contrast agent, b. Unable to receive beta blockers unless heart rate < 65 beats per minute, c. Pregnancy • Life expectancy < 2 years • Unable to provide written informed consent or participate in long-term follow-up
Acronym	**SCOT-HEART**
Name	Scottish COmputed Tomography of the HEART Trial
ClinicalTrials.gov Identifier	NCT01149590
Study design	Interventional (Clinical Trial) randomized study
Study Center(s)	Multicenter trial: Scotland (UK)
Inclusion Criteria:	• 18 and ≤75 years of age • Attendance at the Rapid Access Chest Pain Clinic
Exclusion Criteria	• Inability or unwilling to undergo computed tomography scanning, such as exceeding weight tolerance of scanner

- Severe renal failure (serum creatinine >200 µmol/L or estimated glomerular filtration rate <30 mL/min)
- Previous recruitment to the trial
- Major allergy to iodinated contrast agent
- Unable to give informed consent
- Known pregnancy
- Acute coronary syndrome within 3 months

Additional File 3 – Proportion of studies with low, high or unclear risk of bias