The performance of non-invasive tests to rule-in and rule-out significant coronary artery stenosis in patients with stable angina: a meta-analysis focused on post-test disease probability

Juhani Knuuti[1]*, Haitham Ballo[1†], Luis Eduardo Juarez-Orozco[1†], Antti Saraste[1], Philippe Kolh[2], Anne Wilhelmina Saskia Rutjes[3], Peter Jüni[4], Stephan Windecker[5], Jeroen J. Bax[6], and William Wijns[7]

[1]Turku PET Centre, Turku University Hospital and University of Turku, Kiinamyllynkatu 4-8, 20520, Turku, Finland; [2]Department of Biomedical and Preclinical Sciences, University of Liège, Sart Tilman B 35, 4000, Liège, Belgium; [3]Institute of Social and Preventive Medicine (ISPM), University of Bern, Mittelstrasse 43, 3012, Bern, Switzerland and Institute of Primary Health Care (BIHAM), University of Bern, Gesellschaftsstrasse 49, 3012, Bern, Switzerland; [4]Department of Medicine, Applied Health Research Centre (AHRC), Li Ka Shing Knowledge Institute of St. Michael's Hospital, Institute of Health Policy, Management and Evaluation, University of Toronto, 30 Bond St, ON M5B 1W8, Toronto, Canada; [5]Department of Cardiology, University Hospital Bern, Freiburgstrasse 4, 3010, Bern, Switzerland; [6]Department of Cardiology, Leiden University Medical Center, Albinusdreef 2, 2333ZA, Leiden, The Netherlands; and [7]The Lambe Institute for Translational Medicine and Curam, National University of Ireland, Galway and Saolta University Healthcare Group, University College Hospital Galway, Newcastle Rd, Galway, Ireland

Aims	To determine the ranges of pre-test probability (PTP) of coronary artery disease (CAD) in which stress electrocardiogram (ECG), stress echocardiography, coronary computed tomography angiography (CCTA), single-photon emission computed tomography (SPECT), positron emission tomography (PET), and cardiac magnetic resonance (CMR) can reclassify patients into a post-test probability that defines (>85%) or excludes (<15%) anatomically (defined by visual evaluation of invasive coronary angiography [ICA]) and functionally (defined by a fractional flow reserve [FFR] ≤0.8) significant CAD.
Methods and results	A broad search in electronic databases until August 2017 was performed. Studies on the aforementioned techniques in >100 patients with stable CAD that utilized either ICA or ICA with FFR measurement as reference, were included. Study-level data was pooled using a hierarchical bivariate random-effects model and likelihood ratios were obtained for each technique. The PTP ranges for each technique to rule-in or rule-out significant CAD were defined. A total of 28 664 patients from 132 studies that used ICA as reference and 4131 from 23 studies using FFR, were analysed. Stress ECG can rule-in and rule-out anatomically significant CAD only when PTP is ≥80% (76–83) and ≤19% (15–25), respectively. Coronary computed tomography angiography is able to rule-in anatomic CAD at a PTP ≥58% (45–70) and rule-out at a PTP ≤80% (65–94). The corresponding PTP values for functionally significant CAD were ≥75% (67–83) and ≤57% (40–72) for CCTA, and ≥71% (59–81) and ≤27 (24–31) for ICA, demonstrating poorer performance of anatomic imaging against FFR. In contrast, functional imaging techniques (PET, stress CMR, and SPECT) are able to rule-in functionally significant CAD when PTP is ≥46–59% and rule-out when PTP is ≤34–57%.
Conclusion	The various diagnostic modalities have different optimal performance ranges for the detection of anatomically and functionally significant CAD. Stress ECG appears to have very limited diagnostic power. The selection of a diagnostic technique for any given patient to rule-in or rule-out CAD should be based on the optimal PTP range for each test and on the assumed reference standard.
Keywords	Stable coronary artery disease • Non-invasive imaging • Pre-test likelihood • Post-test likelihood • Likelihood ratio

* Corresponding author. Tel: +358 2 313 2842, Fax: +358 2 231 8191, Email: Juhani.Knuuti@utu.fi

† These authors contributed equally to the study.

Introduction

Accurate detection of coronary artery disease (CAD) remains paramount in the practice of cardiology. Traditionally, the characterization of 'significant' CAD has relied upon visual evaluation of coronary artery stenosis during invasive coronary angiography (ICA). However, the severity of angiographic stenosis does not unequivocally reflect its functional significance.[1] Recently, the invasive assessment of fractional flow reserve (FFR) has been adopted to identify functionally significant coronary artery stenoses.[2] Yet, FFR evaluation is not without limitations as diffuse CAD and haemodynamic conditions have shown an influence on its estimation, it is inherently invasive and costly, and it still does not represent the most common practice in invasive evaluation of CAD.[3]

Stable CAD is understood as the condition characterized by episodes of inducible and reversible ischaemia commonly associated with transient chest discomfort. The current European and American guidelines on the management of stable CAD[2,4] recommend that patients with an intermediate pre-test probability (PTP) (ranging from 15 to 85%) of significant CAD should undergo non-invasive evaluation.[5,6] In subjects whose probability of a significant coronary artery narrowing is low (<15%), routine testing is not recommended. On the other hand, patients with a high probability (>85%) of the disease calls for direct therapeutic interventions.

In the group of patients with intermediate PTP of significant CAD, the current recommendations for the selection of the optimal non-invasive technique are broad and do not assign preference of one modality over another. Certain techniques are broadly available because of their relative low technical and personnel demands [such as stress electrocardiogram (ECG)] or good availability [stress echocardiography, coronary computed tomography angiography (CCTA), and single-photon emission computed tomography (SPECT)], while others, like positron emission tomography (PET) and stress cardiac magnetic resonance (CMR), although powerful, are much less available and their applicability is still limited by infrastructural and capacity requirements.[7]

It is expected that each technique has a particular range of PTP of significant CAD where the usefulness of its application is maximized. The performance of non-invasive techniques is generally reported in terms of sensitivity and specificity. Nevertheless, these numbers cannot be readily utilized in the clinical decision-making process. They can however be used to derive positive and negative likelihood ratios (LR+ and LR-), which constitute readily useful parameters of a test's accuracy that facilitate the selection of a diagnostic test for individual patients.[8] Given a PTP of significant CAD and the performance of a particular test by means of its LR's, one can assess the post-test probability of significant CAD after performing such test. Using this approach, one can estimate the range of PTP when a positive or negative test result can confidently rule-in (if the post-test probability goes beyond 85%) or rule-out (if the post-test probability drops below 15%) the disease.

As currently both anatomical (ICA) and functional (FFR) reference standards are utilized, it is rational to consider evidence using both standards.[9] The anatomical standard has been used in most of the studies available today and there is a massive amount of evidence, although functional information has gained increasing interest. It can be expected that some tests demonstrate better agreement with ICA while others with FFR. Therefore, integration of all available data may provide important clinical information for conscious selection of the tests.

The aim of the present systematic review and meta-analysis was to evaluate the diagnostic performance of stress ECG, stress echocardiography, CCTA, SPECT, PET, stress CMR, and ICA in the detection of anatomically and functionally significant CAD in order to determine the optimal range of PTP in the diagnostic application of each technique for ruling-in or ruling-out significant CAD.

Methods

The present systematic review was conducted in accordance to the Preferred Reporting items for Systematic Reviews and Meta-analysis (PRISMA)[10] recommendations and the MOOSE checklist (see Results Section and Supplementary material online, *Table S1*).[11]

Data sources

We performed a systematic search for original studies published until August 2017 that reported on the diagnostic performance of stress ECG, stress echocardiography, CCTA, SPECT, PET, stress CMR, and ICA for the detection of significant CAD.

The search was performed in electronic databases (Medline, Embase, PubMed, Scopus, The Cochrane Library, Web of Science, and ProQuest) using a broad strategy with a combination of MeSH terms and free text words sensitive to: identify studies concerning (i) the aforementioned diagnostic techniques, (ii) diagnostic performance, (iii) patients with intermediate PTP of the condition, and (iv) significant CAD. The search results were limited to the English language and to studies performed in humans. The full search string is reported in Supplementary material online, *Table S2*. Reference lists from relevant studies were scanned and cross-checked to identify potentially overlooked publications.

Study selection and quality assessment

Studies were included according to the following eligibility criteria: (i) the study aimed to investigate stable CAD (not acute coronary syndromes), (ii) either catheter-based X-ray angiography (ICA) or ICA with FFR evaluation were used as the reference standard for the diagnosis of stable CAD, (iii) the reported data was explicit or sufficient to extract numbers for true and false positive and negative results, and (iv) the study included a sample of at least 100 patients (for robustness). Selected studies were further divided according to the reference standard considered (ICA or FFR evaluation).

For each included study, the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) criteria were determined by two authors (L.E.J.-O. and H.B.). The QUADAS-2 tool assesses the study quality in different domains including patient selection, index test, reference standard, and flow of patients through the study considering the timing of the index test and reference standard. For each article, quality and applicability were assessed in the aforementioned domains as follows: 'yes' if concern existed based on enough description in the report, 'no' if there was no concern based on enough description in the report or 'unclear' if there was inadequate or insufficient information reported in the article to make a judgement.

Data extraction

Data were recorded according to the technique and reference standard utilized. The number of subjects, male to female patient proportion, age, type of stressor, tracer utilized (if any), stable CAD definition, and

prevalence were extracted. The number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN), as well as derived diagnostic performance variables were recorded.

Study review, quality evaluation, and data extraction were performed in parallel by two authors (A.S. and H.B.). Any specific discrepancies were resolved by consensus. If necessary, a third reviewer (J.K.) was considered to reach convergence.

Reference standard

Catheter-based ICA alone and ICA with FFR measurement were considered as the reference standards for the determination of anatomically significant and functionally significant CAD, respectively. Anatomic coronary narrowing >50% was considered as determinant of significant CAD and an FFR ≤0.80 was considered as functionally significant CAD.

Data synthesis and statistical analysis

Hierarchical bivariate random-effects models were constructed to combine individual study-level data on the sensitivities and specificities across studies. This model takes the correlation between sensitivity and specificity into account, and is described in detail elsewhere.[12] The bivariate model used parametrization to render summary points for sensitivity and specificity with 95% confidence intervals (CIs) for each of the imaging techniques. We used an unstructured covariance matrix allowing all variances and covariances to be distinct. We then derived summary estimates of the LR+ and LR- with their CIs from the model estimates. For echocardiography and SPECT, more than one type of stressor was used. We compared if a model distinguishing by type of stressor had a better model fit than a model grouping all stressor techniques together. The analysis was performed separately for anatomically and functionally significant CAD (according to the reference standard used). We used the *P*-value from the likelihood ratio test to determine if the model with a covariate for the type of stressor fitted the data better than a model without such covariate. If the *P*-value was 0.05 or less, we depicted summary estimates for a specific type of stressor.

Utility of non-invasive approaches according to pre-test probability of stable coronary artery disease

Once the positive and negative LRs of each non-invasive diagnostic technique were obtained for both accepted reference standards, the ranges and in which every single technique allows to confidently rule-in CAD, rule-out CAD, or both were input into a colour-coded graph. Additionally, we created a supplemental colour-coded suggestion over the structure of the current ESC guidelines stable CAD PTP table to depict the suggested utility of each diagnostic technique at each level of risk based on age, sex, and type of symptoms.

Results

Study characteristics

The study selection flowchart is shown in *Figure 1*. Specific characteristics and the full reference for each selected study can be consulted in Supplementary material online, *Table S3*. After eligibility assessment and technique subgroup characterization, 13 studies on stress ECG, 12 studies on exercise stress echocardiography, 30 on dobutamine stress echocardiography, nine studies on CCTA, 28 studies on exercise and adenosine or dipyridamole stress SPECT, 13 on exercise stress SPECT, three studies on PET, and 11 on stress CMR were

considered for the pooled analysis on anatomically significant CAD. On the other hand, two studies in ICA, seven studies on CCTA, five on exercise stress SPECT, four on PET, and five on stress CMR were considered for the pooled analysis on functionally significant CAD.

Study heterogeneity and quality

Risk of bias in the included studies, as assessed with the QUADAS-2 score, showed important variation across diagnostic modalities. Overall, PET, CCTA, and stress CMR showed a low risk of bias and therefore, did not raise substantial concerns of applicability. However, these modalities conveyed the smallest number of studies included. Conversely, the proportions of unclear ratings for ECG and echocardiography studies related to the year when these were performed. For the oldest studies, insufficient data for this assessment is commonly reported. SPECT studies generally rated less well showing a balanced proportion of unclear and high risk of bias in all domains. Supplementary material online, *Figure S1* shows this assessment across techniques in an ascending order of risk. Overall quality per type of reference standard is shown in *Figure 2*.

Performance estimates

The pooled analysis considering anatomically significant CAD included a total of 2442 patients for stress ECG, 4302 for stress echo (with exercise or vasodilator), 2756 for CCTA, 4346 for exercise stress SPECT, 6551 for exercise and adenosine or dipyridamole stress SPECT, 418 for PET, and 3393 for stress CMR. Further, the pooled analysis considering functionally significant CAD included 954 for ICA, 1140 patients for CCTA, 740 for exercise stress SPECT, 709 for PET, and 588 for stress CMR. Some studies evaluated several techniques or technique subgroups simultaneously. Such studies were included as independent entries in more than one pooled analysis per technique.

Table 1 summarizes the performance estimates for every diagnostic technique according to each reference standard. Some techniques had various subcategories typically according to the type of stressor utilized. Some of these subcategories are less commonly used or did not yield adequate information for a summary estimate (e.g. stress echo with dobutamine stress *n* = 30, dobutamine stress SPECT *n* = 2, and dobutamine stress CMR *n* = 2) and were not included in these estimates.

Considering anatomically significant CAD, there were 11 vasodilatory stress echocardiography studies and analysis considering >50% as significant stenosis yielded a sensitivity of 0.75 (0.70–0.80) and specificity of 0.91 (0.86–0.94). These summary estimates were not statistically different from the summary estimates obtained for exercise stress echo (likelihood ratio test *P* = 0.386) and were consequently pooled together. The summary estimates obtained from 27 dobutamine stress echocardiography studies were 0.81 (0.77–0.85) for sensitivity and 0.84 (0.81–0.87) for specificity and given that these estimates were significantly different from exercise stress echocardiography (likelihood ratio test *P* = 0.012), they were not pooled together but their references can be consulted in the Supplementary Material online.

When anatomically significant CAD was used as reference standard, the LR– of different tests varied from 0.04 to 0.68. The best performance in ruling-out CAD was achieved using CCTA and poorest with stress ECG. The LR+ varied from 1.53 to 5.87. The best performance for ruling-in CAD was achieved using PET and the poorest
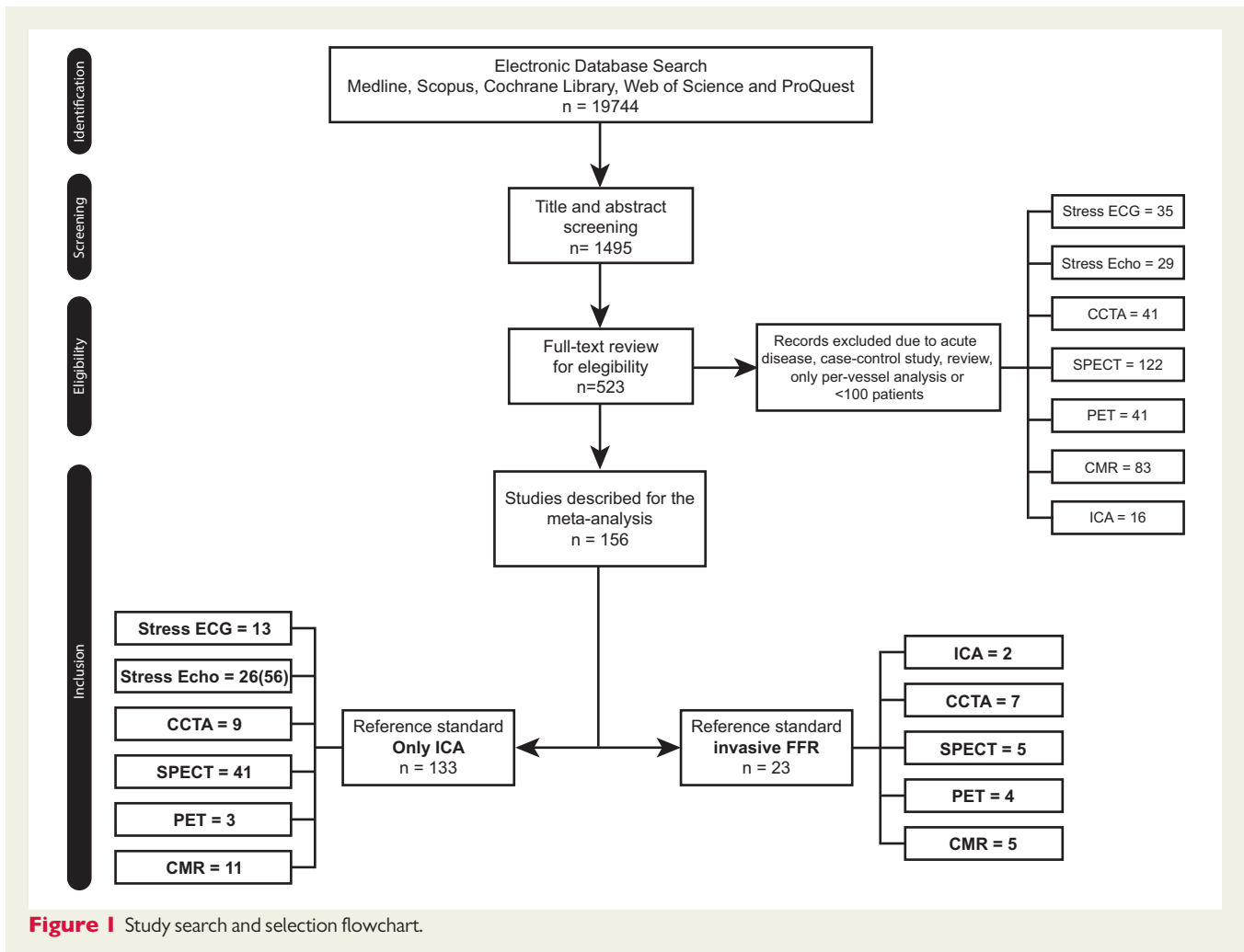
Figure 1 Study search and selection flowchart.

with stress ECG. The LR+ and LR- for the dobutamine stress echo-cardiography subgroup were 8.03 (4.98–12.95) and 0.27 (0.22–0.34), respectively (not shown in Table 1).

When functionally significant CAD was considered as reference standard, LR– varied from 0.13 to 0.44. Coronary computed tomography angiography, PET, and stress CMR had the best and similar performance in ruling-out significant CAD [–LR = 0.13 (0.07–0.24)], while interestingly, ICA had the poorest. The LR+ of the available techniques varied from 1.97 to 7.10. The poorest performances in ruling-in an abnormal FFR were documented for CCTA [LR+ = 1.97 (1.28–3.03)] and ICA [LR+ = 2.49 (1.47–4.21)], while functional imaging tests conversely demonstrated the best performance (LR+ range: 3.87–7.1). We could not identify enough robust studies to pool estimates for stress ECG and stress echocardiography.

Effectiveness of non-invasive diagnostic techniques in ruling-in/out significant coronary artery disease

The Fagan nomogram is a useful tool to graphically apply LRs to a PTP to calculate the post-test probability. A parallel example of its use is depicted in *Figure 3*, which shows how one can calculate the post-test probabilities after a positive or negative test result starting from any PTP in an individual patient.

The same nomogram can be also utilized backwards so that we can assess the PTP values that will lead to a defined range of post-test probability for each diagnostic method. Therefore, using the data from the meta-analysis, we defined the ranges of PTP of CAD where the diagnostic techniques can confidently rule-in (by driving the post-test probability above 85%) and/or rule-out (by driving the post-test probability below 15%) significant CAD. This was done separately for both anatomically and functionally significant CAD. Such ranges are schematically shown along with their corresponding upper and lower limits in *Take home figure* and numerically reported in Supplementary material online, *Table S4*.

Finally, based on the obtained data described above, we transformed the PTP table from the 2013 ESC Guidelines on the management of stable CAD[4] into a supplemental guide that exemplifies how clinicians could implement the resulting estimates of performance in this report in order to select a diagnostic test that confidently rules-in or rules-out CAD (both anatomically and functionally significant CAD) at each patient PTP category (Supplementary material online, *Figure S2*A and B, respectively).
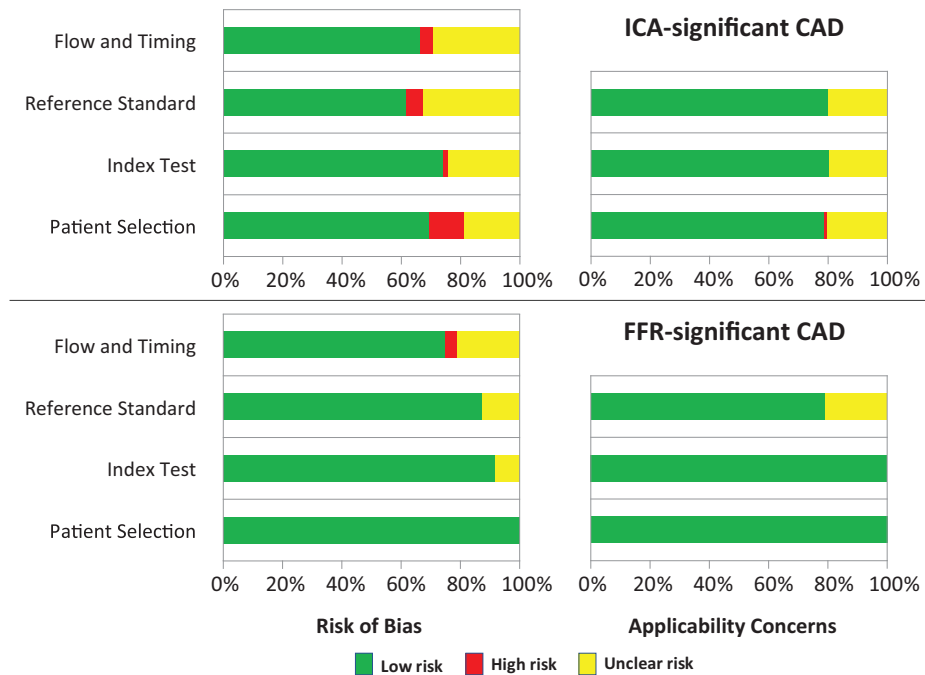
Figure 2 Quality Assessment of Diagnostic Accuracy Studies summary by type of reference standard for significant coronary artery disease.

Discussion

The present study analysed the evidence on the performance of different diagnostic techniques for the detection of either anatomically or functionally significant CAD. Beyond reporting traditional metrics, we also portrayed their performance as LRs and defined the optimal ranges of PTP for each test where they can reclassify patients from intermediate to either low or high post-test probability of CAD (i.e. rule-out or rule-in, respectively).

From this analysis several main messages can be driven. Stress ECG appears to have very limited diagnostic power to rule-in or rule-out significant CAD. In fact, there was no single PTP value in which stress ECG can both define the diagnosis and exclude it. Moreover, even to confidently rule-out CAD, a very low PTP [≤19% (15–25)] is needed, while for ruling-in, a PTP ≥80% (76–83) is required.

As expected, the performance of imaging methods was clearly better than that of stress ECG. However, there appears to be also differences between them. A negative result in CCTA, which conveys a strong LR-, can exclude anatomically defined CAD in nearly all patients independently of their PTP. The performance was clearly poorer when FFR was considered the reference standard as CCTA could only exclude functionally significant CAD at a PTP ≤57% (40–72). Correspondingly, the rule-in power that was moderate to good when considering ICA as reference, also clearly deteriorated when FFR was used as reference standard.

The functional imaging techniques (PET, CMR, and SPECT), which had only moderate power in identifying anatomically significant CAD, performed much better when FFR was used as reference standard. This is in agreement with previous notions and a recently published

meta-analysis.[9,13] Positron emission tomography and stress CMR demonstrated the best diagnostic performance and offered reasonable range of pre-test probabilities where they could simultaneously rule-out or rule-in functionally significant CAD as shown in *Take home figure*. However, the comparison between functional imaging techniques must be done cautiously as not enough data was available for stress echocardiography and SPECT studies were older. Furthermore, in more recent studies, referral bias to reference technique is a common phenomenon with established techniques, which typically leads to underestimation of the test specificity. Also, the recent technical advances in were not accounted for as the data was heavily weighted by older studies. Therefore, the previously established tests may underperform in the present analysis.

We also assessed the performance of ICA itself in detecting functionally significant CAD even though it does not classify as a non-invasive test. ICA demonstrated the poorest ruling-out performance of all analysed techniques when the reference standard was FFR as a PTP ≤27% (24–31) was needed to rule-out functional CAD. Consistently, the PTP range to rule-in functionally significant CAD was rather modest [≥71% (59–81)] and only slightly superior to CCTA [≥75% (67–83)]. This behaviour fits well with the current recommendation that ICA should be used primarily in patients with high PTP.

Although a pooled evaluation of non-invasive imaging techniques for diagnosing functionally significant CAD has been performed recently,[14] the present study expands the evidence by also considering stress ECG performance, evaluating the competence of ICA alone in determining functionally significant CAD, conveying the practical ranges of application for the involved diagnostic techniques and parsing the determination of CAD both against anatomical and functional

Table I The performance of different tests for anatomically and functionally significant coronary artery disease

Anatomically significant CAD					Functionally significant CAD				
Test	Sensitivity (%), (95% CI)	Specificity (%), (95% CI)	+LR (95% CI)	−LR (95% CI)	Test	Sensitivity (%), (95% CI)	Specificity (%), (95% CI)	+LR (95% CI)	−LR (95% CI)
Stress ECG	58 (46–69)	62 (54–69)	1.53 (1.21–1.94)	0.68 (0.49–0.93)	ICA	68 (60–75)	73 (55–86)	2.49 (1.47–4.21)	0.44 (0.36–0.54)
Stress echo	85 (80–89)	82 (72–89)	4.67 (2.95–7.41)	0.18 (0.13–0.25)					
CCTA	97 (93–99)	78 (67–86)	4.44 (2.64–7.45)	0.04 (0.01–0.09)	CCTA	93 (89–96)	53 (37–68)	1.97 (1.28–3.03)	0.13 (0.06–0.25)
SPECT	87 (83–90)	70 (63–76)	2.88 (2.33–3.56)	0.19 (0.15–0.24)	SPECT	73 (62–82)	83 (71–90)	4.21 (2.62–6.76)	0.33 (0.24–0.46)
PET	90 (78–96)	85 (78–90)	5.87 (3.40–10.15)	0.12 (0.05–0.29)	PET	89 (82–93)	85 (81–88)	6.04 (4.29–8.51)	0.13 (0.08–0.22)
Stress CMR	90 (83–94)	80 (69–88)	4.54 (2.37–8.72)	0.13 (0.07–0.24)	Stress CMR	89 (85–92)	87 (83–91)	7.10 (5.07–9.95)	0.13 (0.09–0.18)

Note: ICA itself was used as a reference standard for the anatomically significant CAD estimates but was included as a technique when FFR was used as the reference. Not every test had enough data using FFR as reference. CCTA, coronary computed tomography angiography; CI, confidence interval; CMR, stress cardiac magnetic resonance; ECG, electrocardiogram; ICA, invasive coronary angiography; LR, likelihood ratio; PET, positron emission tomography; SPECT, single-photon emission computed tomography (exercise stress SPECT with or without dipyridamole or adenosine); Stress echo, exercise stress echocardiography.

standards. This is timely and relevant considering that anatomical definition of CAD is still widely used in the daily clinical scenario in many healthcare centres around the world, while at the same time acknowledging that FFR indeed represents the currently most adequate reference standard.
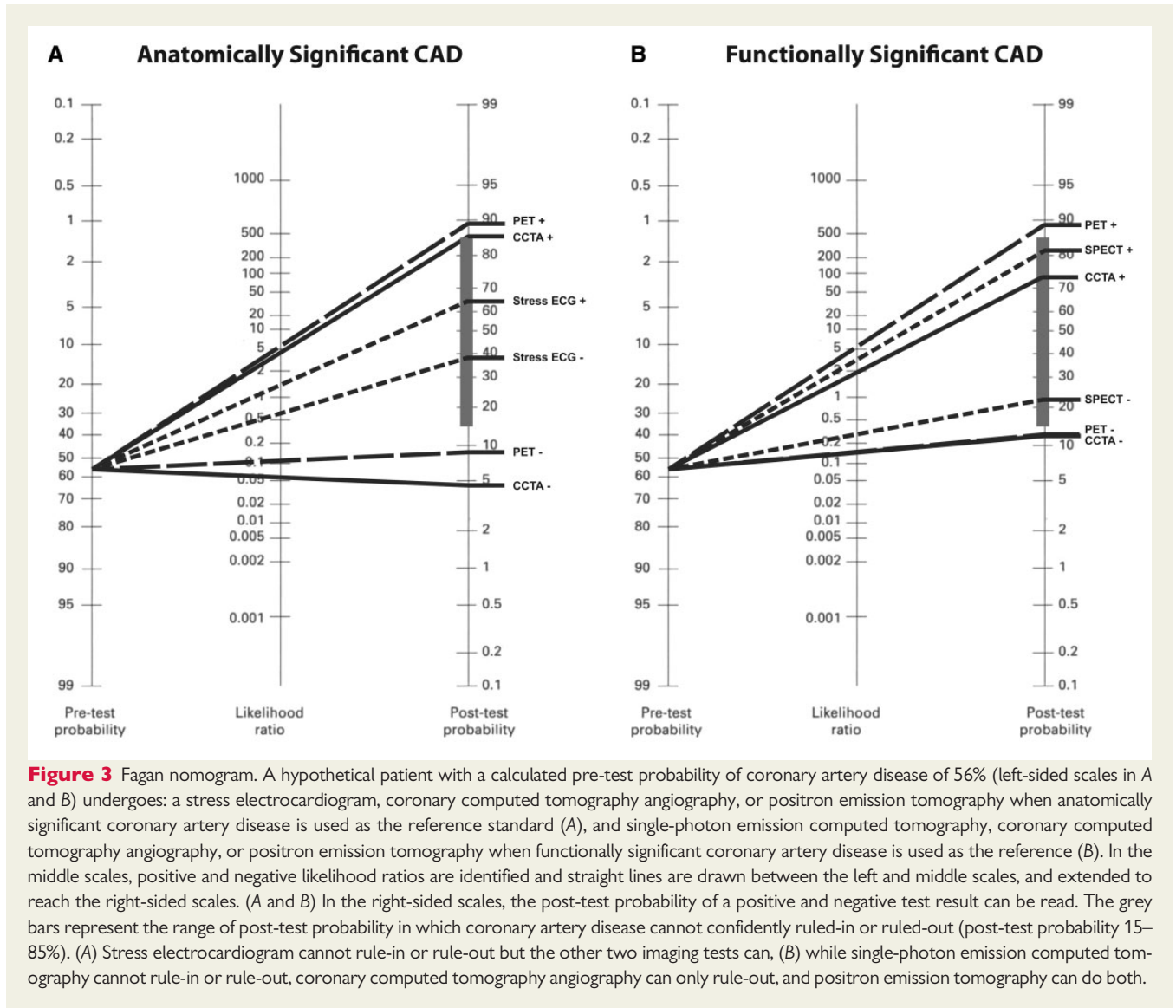
Clinical implications

Our clinical conclusions partly differ from those in the current clinical guidelines. For example, in ESC guidelines[4] stress ECG is recommended in patients with lower intermediate PTP (15–65%) of CAD. Our analysis argues against this statement as the practical utility of stress ECG in detecting CAD appears very limited (*Take home figure* and Supplementary material online, *Figure S2A*). However, exercise testing also provides complementary information beyond ECG changes, such as exercise capacity, arrhythmias, haemodynamic response, and symptoms during exercise, which are considered clinically useful. These, however, could not be taken into account in the present analysis.

Coronary computed tomography angiography has rapidly gained popularity mainly based on its high negative predictive value. This was confirmed in the present analysis by the low LR-, which suggests that a negative result can reliably rule-out anatomic CAD virtually at any level of intermediate PTP (*Take home figure* and Supplementary material online, *Figure S2A*). However, with a high probability of CAD, exclusion of disease is clinically less beneficial because, statistically, most patients will have the disease, and in order to rule-out CAD in one patient, a considerably large number of patients must be investigated. Additionally, the rule-out power decreased when considering FFR as reference. A known limitation of CCTA is low specificity, especially in identifying functionally significant CAD (53%), and this links to our finding that a PTP ≥75% is required to rule it in (*Take home figure*).

Not surprisingly, non-invasive imaging methods that characterize the functional consequences of CAD (rather than the coronary atherosclerotic lesions themselves) perform better when FFR is used as a reference standard and outperform CCTA (*Take home figure*). Clearly, every technique has a particular diagnostic performance profile. The techniques focus on different levels of the ischaemic cascade including wall motion abnormalities (echocardiography and stress CMR), relative perfusion abnormalities (stress CMR and SPECT), and changes in physiological absolute regional myocardial perfusion (PET).

Out of the functional imaging tests, PET and stress CMR demonstrated good performance with optimal application ranges (for both ruling-in and ruling-out disease) for anatomic and functional CAD. Stress echocardiography and SPECT perfusion imaging performance numbers appeared moderate but direct comparison to other methods must be done cautiously, for the reasons explained above. In addition, as shown in Supplementary material online, *Figure S2*, the clinical impact of these differences in the utility of the various functional tests is modest although detectable. It is also important to remember that accessibility, simplicity, expertise, personnel, and costs are still important determinants for choosing a given test, and unfortunately, these variables could not be included in this analysis.

Finally, the 2016 update of the stable chest pain guideline, the National Institute for Health and Care Excellence (NICE)[15] has chosen not to include the assessment of PTP and rather recommended CCTA as the first-line diagnostic test and ischaemia testing

Figure 3 Fagan nomogram. A hypothetical patient with a calculated pre-test probability of coronary artery disease of 56% (left-sided scales in *A* and *B*) undergoes: a stress electrocardiogram, coronary computed tomography angiography, or positron emission tomography when anatomically significant coronary artery disease is used as the reference standard (*A*), and single-photon emission computed tomography, coronary computed tomography angiography, or positron emission tomography when functionally significant coronary artery disease is used as the reference (*B*). In the middle scales, positive and negative likelihood ratios are identified and straight lines are drawn between the left and middle scales, and extended to reach the right-sided scales. (*A* and *B*) In the right-sided scales, the post-test probability of a positive and negative test result can be read. The grey bars represent the range of post-test probability in which coronary artery disease cannot confidently ruled-in or ruled-out (post-test probability 15–85%). (*A*) Stress electrocardiogram cannot rule-in or rule-out but the other two imaging tests can, (*B*) while single-photon emission computed tomography cannot rule-in or rule-out, coronary computed tomography angiography can only rule-out, and positron emission tomography can do both.
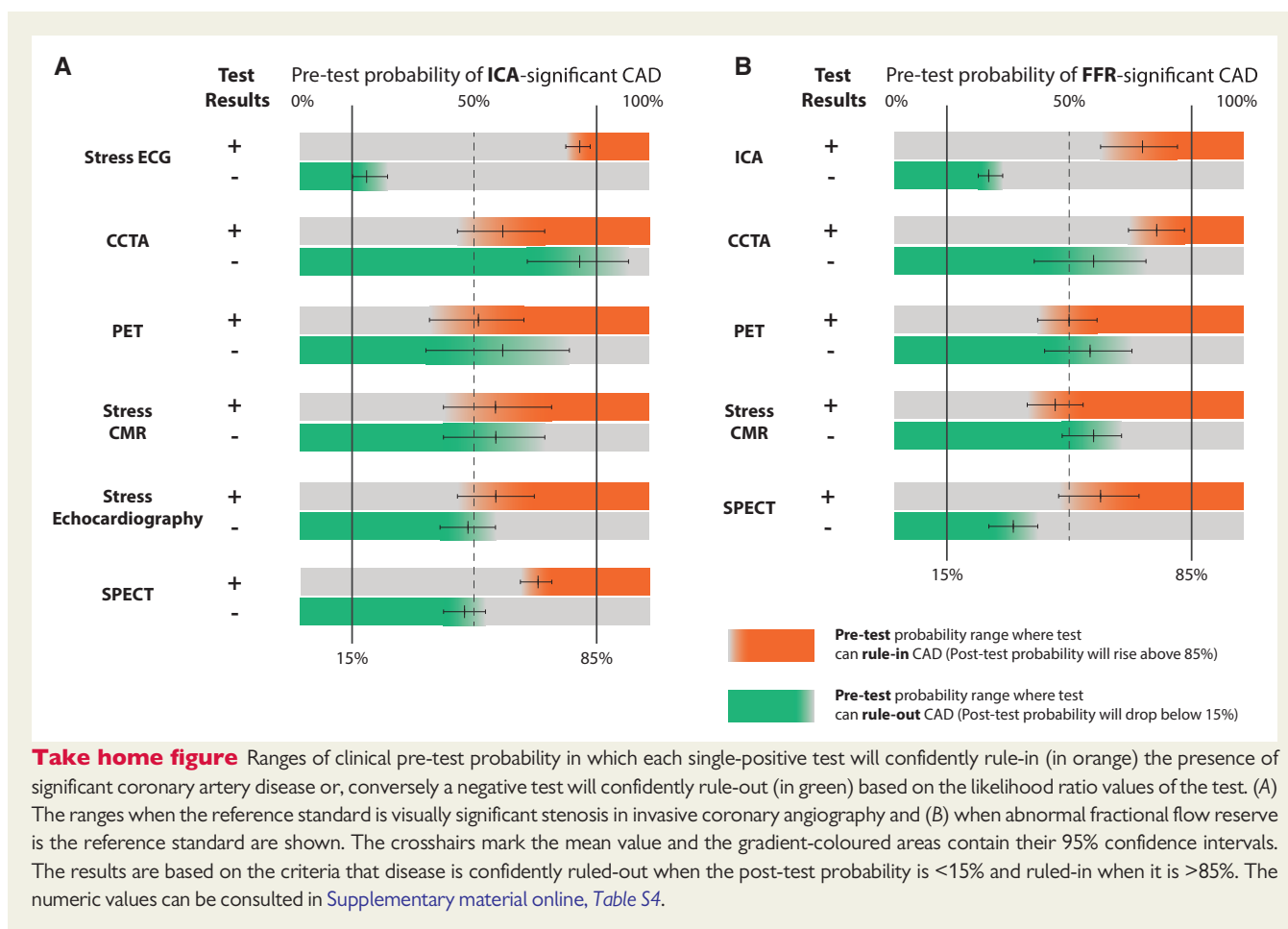
as second step in those with suspected anatomically-relevant CAD. Our analysis does not argue against this approach, but we would like to underline that such rationale will depend on the actual prevalence of CAD in the population. The PTP tables currently included in the guidelines are based on reasonably old data while the prevalence of CAD is continuously decreasing. With low prevalence of CAD the primary first task of imaging may be the accurate exclusion of anatomic CAD, for which CCTA has demonstrated a strong role. The proposed sequential utilization of functional imaging tests may indeed be relevant but it must be kept in mind that the evidence is still limited although prognostic utility and overall safety appears to be excellent.[16]

Limitations

The performance of a given test in different publications varies due to numerous reasons such as population selection and referral bias. Age, gender, or participants with history of MI may effect on the estimates of diagnostic accuracy but analyses of these characteristics on

a group level may lead to spurious results due to the risk of ecological fallacy bias. We did not have access to individual patient level data or subgroup data that are needed to validly analyse these characteristics. Another potentially important source of variation or bias is study selection based on prior test results or known CAD. Although we excluded case-control studies, we do not know whether study selection was restricted to participants with specific prior test results. The inconsistency between studies lowers the confidence in the summary estimates and future studies should aim to dissect sources of bias and variation.

Furthermore, the present study considers visual analysis alone for the determination of significant CAD through ICA. Advances in ICA evaluation, such as quantitative coronary angiography and the implementation of intravascular ultrasound and optical coherence tomography,[17] could improve identification of haemodynamically-significant lesions. However, clinical practice in many centres currently relies on direct visual ICA evaluation and, therefore, our results on technique performances are likely to be widely applicable. The cut-off of 50% in

Take home figure Ranges of clinical pre-test probability in which each single-positive test will confidently rule-in (in orange) the presence of significant coronary artery disease or, conversely a negative test will confidently rule-out (in green) based on the likelihood ratio values of the test. (*A*) The ranges when the reference standard is visually significant stenosis in invasive coronary angiography and (*B*) when abnormal fractional flow reserve is the reference standard are shown. The crosshairs mark the mean value and the gradient-coloured areas contain their 95% confidence intervals. The results are based on the criteria that disease is confidently ruled-out when the post-test probability is <15% and ruled-in when it is >85%. The numeric values can be consulted in Supplementary material online, *Table S4*.

ICA was used as this was available in all studies. In addition to known pitfalls of ICA, FFR is not without limitation as it is highly dependent on achieving hyperaemia through maximal decrease in microvascular resistances.

As the data was available only at the study-level in several reports, we cannot evaluate how the different techniques can assess the extent and severity of the disease, which are important factors in guiding therapies. As there are limited data on direct comparisons between modalities, differences could not be comprehensively tested.

With regard to analyses using FFR as the reference standard, the low number of identified studies did not allow analysing all modalities. In addition, our summary estimates were vastly derived from single test accuracy studies, providing indirect evidence to compare test modalities. Due to the very low number of comparative studies identified, no consistency check could be performed between direct and indirect summary estimates. Therefore, small differences between techniques and summary estimates should be interpreted cautiously and considered as directional only. Coronary computed tomography angiography derived FFR has been investigated recently but this method is not yet well standardized and we decided not to include this method in the current analysis. It is also possible that the best diagnostic performance could be achieved when the tests are applied sequentially.[16] The relevance of complementary features in different techniques warrants further investigation. The supplemental technique selection guide (Supplementary material online, *Figure S2*) was based on the PTP values published in 2013 ESC guidelines and is naturally susceptible to change when updated PTP values are available.

Conclusions

The various diagnostic modalities have different optimal performance ranges for the detection of anatomically and functionally significant CAD. Stress ECG appears to have limited diagnostic value at any level of PTP. Imaging methods perform generally better but also have different strengths and weaknesses. Coronary computed tomography angiography performs best against anatomical reference standard and functional tests perform better than CCTA or ICA for functionally significant CAD.

The selection of a diagnostic technique for any given patient to rule-in or rule-out CAD should be based on the optimal PTP range for each test. Using LRs, we were able to create individual pre-test ranges for each test to rule-in and/or rule-out anatomic or functional CAD, and these can be used in aiding in the selection of a diagnostic technique for a given patient.

Supplementary material

Supplementary material is available at *European Heart Journal* online.

Acknowledgements

Funding

References

1. Tonino PAL, Fearon WF, De Bruyne B, Oldroyd KG, Leesar MA, Ver Lee PN, MacCarthy PA, van't Veer M, Pijls NHJ. Angiographic versus functional severity of coronary artery stenoses in the FAME study fractional flow reserve versus angiography in multivessel evaluation. *J Am Coll Cardiol* 2010;**55**:2816–2821.

2. Fihn SD, Blankenship JC, Alexander KP, Bittl JA, Byrne JG, Fletcher BJ, Fonarow GC, Lange RA, Levine GN, Maddox TM, Naidu SS, Ohman EM, Smith PK. 2014 ACC/AHA/AATS/PCNA/SCAI/STS focused update of the guideline for the diagnosis and management of patients with stable ischemic heart disease: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines, and the American Association for Thoracic Surgery, Preventive Cardiovascular Nurses Association, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. *J Am Coll Cardiol* 2014;**64**:1929–1949.

3. Pothineni NV, Shah NN, Rochlani Y, Nairooz R, Raina S, Leesar MA, Uretsky BF, Hakeem A. U.S. Trends in inpatient utilization of fractional flow reserve and percutaneous coronary intervention. *J Am Coll Cardiol* 2016;**67**:732–733.

4. Task Force Members, Montalescot G, Sechtem U, Achenbach S, Andreotti F, Arden C, Budaj A, Bugiardini R, Crea F, Cuisset T, Di Mario C, Ferreira JR, Gersh BJ, Gitt AK, Hulot J-S, Marx N, Opie LH, Pfisterer M, Prescott E, Ruschitzka F, Sabaté M, Senior R, Taggart DP, van der Wall EE, Vrints CJM; ESC Committee for Practice Guidelines, Zamorano JL, Achenbach S, Baumgartner H, Bax JJ, Bueno H, Dean V, Deaton C, Erol C, Fagard R, Ferrari R, Hasdai D, Hoes AW, Kirchhof P, Knuuti J, Kolh P, Lancellotti P, Linhart A, Nihoyannopoulos P, Piepoli MF, Ponikowski P, Sirnes PA, Tamargo JL, Tendera M, Torbicki A, Wijns W, Windecker S; Document Reviewers, Knuuti J, Valgimigli M, Bueno H, Claeys MJ, Donner-Banzhoff N, Erol C, Frank H, Funck-Brentano C, Gaemperli O, Gonzalez-Juanatey JR, Hamilos M, Hasdai D, Husted S, James SK, Kervinen K, Kolh P, Kristensen SD, Lancellotti P, Maggioni A, Pietro Piepoli MF, Pries AR, Romeo F, Rydén L, Simoons ML, Sirnes PA, Steg PG, Timmis A, Wijns W, Windecker S, Yildirir A, Zamorano JL. 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.

5. Diamond GA, Kaul S. Gone fishing!: on the "real-world" accuracy of computed tomographic coronary angiography: comment on the "Ontario multidetector computed tomographic coronary angiography study". *Arch Intern Med* 2011;**171**:1029–1031.

6. Miller TD, Roger VL, Hodge DO, Gibbons RJ. A simple clinical score accurately predicts outcome in a community-based population undergoing stress testing. *Am J Med* 2005;**118**:866–872.

7. Nakazato R, Berman DS, Alexanderson E, Slomka P. Myocardial perfusion imaging with PET. *Imaging Med* 2013;**5**:35–46.

8. Steurer J, Fischer JE, Bachmann LM, Koller M, ter RG. Communicating accuracy of tests to general practitioners: a controlled study. *BMJ* 2002;**324**:824–826.

9. Danad I, Szymonifka J, Twisk JWR, Norgaard BL, Zarins CK, Knaapen P, Min JK. Diagnostic performance of cardiac imaging methods to diagnose ischaemia-causing coronary artery disease when directly compared with fractional flow reserve as a reference standard: a meta-analysis. *Eur Heart J* 2017;**38**:991–998.

10. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;**6**:e1000097.

11. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000;**283**:2008–2012.

12. Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005;**58**:982–990.

13. Carlsson M. The impacts on healthcare when coronary angiography as the reference method for diagnostic accuracy of coronary artery disease is replaced by fractional flow reserve! *Eur Heart J* 2017;**38**:999–1001.

14. Takx RAP, Blomberg BA, Aidi HE, Habets J, de Jong PA, Nagel E, Hoffmann U, Leiner T. Diagnostic accuracy of stress myocardial perfusion imaging compared to invasive coronary angiography with fractional flow reserve meta-analysis. *Circ Cardiovasc Imaging* 2015;**8**:e002666.

15. National Institute for Health and Clinical Excellence. *Chest Pain of Recent Onset: Assessment and Diagnosis of Recent Onset Chest Pain or Discomfort of Suspected Cardiac Origin.* London: Royal College of Physicians (UK); 2010.

16. Maaniitty T, Stenström I, Bax JJ, Uusitalo V, Ukkonen H, Kajander S, Mäki M, Saraste A, Knuuti J. Prognostic value of coronary CT angiography with selective PET perfusion imaging in coronary artery disease. *JACC Cardiovasc Imaging* 2017;**10**:1361–1370.

17. Pyxaras SA, Tu S, Barbato E, Barbati G, Di Serafino L, De Vroey F, Toth G, Mangiacapra F, Sinagra G, De Bruyne B, Reiber JHC, Wijns W. Quantitative angiography and optical coherence tomography for the functional assessment of nonobstructive coronary stenoses: comparison with fractional flow reserve. *Am Heart J* 2013;**166**:1010–1018.e1.