Diagnostic performance of stress myocardial perfusion imaging for coronary artery disease: a systematic review and meta-analysis

Eur Radiol (2012) 22:1881–1895
DOI 10.1007/s00330-012-2434-1

CARDIAC

Diagnostic performance of stress myocardial perfusion imaging for coronary artery disease: a systematic review and meta-analysis

Marcus C. de Jong · Tessa S. S. Genders · Robert-Jan van Geuns · Adriaan Moelker · M. G. Myriam Hunink

Received: 11 October 2011 / Revised: 23 January 2012 / Accepted: 4 February 2012 / Published online: 19 April 2012

Abstract
Objectives To determine and compare the diagnostic performance of stress myocardial perfusion imaging (MPI) for the diagnosis of obstructive coronary artery disease (CAD), using conventional coronary angiography (CCA) as the reference standard.
Methods We searched Medline and Embase for literature that evaluated stress MPI for the diagnosis of obstructive CAD using magnetic resonance imaging (MRI), contrast-enhanced echocardiography (ECHO), single-photon emission computed tomography (SPECT) and positron emission tomography (PET).
Results All pooled analyses were based on random effects models. Articles on MRI yielded a total of 2,970 patients from 28 studies, articles on ECHO yielded a sample size of 795 from 10 studies, articles on SPECT yielded 1,323 from 13 studies. For CAD defined as either at least 50 %, at least 70 % or at least 75 % lumen diameter reduction on CCA, the natural logarithms of the diagnostic odds ratio (lnDOR) for MRI (3.63; 95 % CI 3.26–4.00) was significantly higher compared to that of SPECT (2.76; 95 % CI 2.28–3.25; P=0.006) and that of ECHO (2.83; 95 % CI 2.29–3.37; P=0.02). There was no significant difference between the lnDOR of SPECT and ECHO (P=0.52).
Conclusion Our results suggest that MRI is superior for the diagnosis of obstructive CAD compared with ECHO and SPECT. ECHO and SPECT demonstrated similar diagnostic performance.
Key Points
• *MRI can assess myocardial perfusion.*
• *MR perfusion diagnoses coronary artery disease better than echocardiography or SPECT.*
• *Echocardiography and SPECT have similar diagnostic performance.*
• *MRI can save coronary artery disease patients from more invasive tests.*
• *MRI and SPECT show evidence of publication bias, implying possible overestimation.*

Keywords Myocardial perfusion imaging · Diagnostic performance · Systematic review · Meta-analysis · Coronary artery disease

Electronic supplementary material The online version of this article (doi:10.1007/s00330-012-2434-1) contains supplementary material, which is available to authorized users.

M. C. de Jong · T. S. S. Genders · M. G. M. Hunink (✉)
Departments of Epidemiology and Radiology,
Erasmus MC – University Medical Center Rotterdam,
P.O. Box 2040, 3000 CA, Rotterdam, The Netherlands
e-mail: m.hunink@erasmusmc.nl

M. C. de Jong · T. S. S. Genders · R.-J. van Geuns · A. Moelker · M. G. M. Hunink
Department of Radiology,
Erasmus University Medical Center,
Rotterdam, The Netherlands

R.-J. van Geuns
Department of Cardiology,
Erasmus University Medical Center,
Rotterdam, The Netherlands

M. G. M. Hunink
Department of Health Policy and Management,
Harvard School of Public Health, Harvard University,
Boston, USA

Introduction

Coronary artery disease (CAD) is one of the major causes of mortality and morbidity throughout the world [1]. The initial assessment of a patient with chest pain usually consists of a stress ECG (electrocardiogram). However, its diagnostic accuracy is low [2] compared to conventional coronary angiography (CCA), which is the reference standard for diagnosing CAD. On the other hand, CCA is an invasive technique and carries a small risk of complications [3, 4]. Myocardial perfusion imaging (MPI) is a non-invasive technique that is used clinically as a gatekeeper test before CCA.

MPI can be conducted using stress magnetic resonance imaging (MRI), contrast-enhanced echocardiography (ECHO), single-photon emission computed tomography (SPECT), positron emission tomography (PET) and, under development, computed tomography (CT). The only available extensive study directly comparing two techniques is the MR-IMPACT study [5], a multicentre randomised trial which found that MRI is superior to SPECT. Systematic reviews and meta-analyses have been published for most of the techniques but none of these reviews compare MPI techniques [6–10]. The comparability between these different meta-analyses is questionable mainly because of differences in publication period, searching the literature, selection of the evidence, and analysis of the data. Furthermore, studies with verification bias are often included in these reports which may have overestimated the sensitivity and underestimated the specificity of the tests considered. To overcome these problems a systematic review of different MPI techniques is required using the same selection criteria and methods of analysis for all techniques and excluding studies with (potential) verification bias, to make a fair comparison between these imaging tests.

The aim of this study was to determine and compare the diagnostic performance of stress MPI tests for the diagnosis of obstructive CAD, with conventional CCA as the reference standard. We performed the review according to the PRISMA statement for such reviews [11, 12].

Materials and methods

Search strategy

We searched Medline and Embase for English-language literature published between January 2000 and May 2011 evaluating the presence of obstructive CAD by stress perfusion imaging tests, namely MRI, contrast-enhanced ECHO, SPECT and PET. In this meta-analysis we focus on functional imaging tests evaluating perfusion as a measure of haemodynamically significant myocardial ischaemia as opposed to anatomical imaging tests, such as coronary CT angiography, which evaluates structural abnormalities of the coronary arteries. We limited the search to publications from 2000 onwards to include only studies that evaluated state-of-the-art MPI techniques. This may have introduced a selection bias with respect to SPECT, because many SPECT studies were published before 2000. To deal with this problem we compare our results with a review of meta-analyses of SPECT studies by Heijenbrok-Kal et al. [13]. CT was excluded because it is still being developed technically. Review articles were checked for potential additional studies. The search included keywords corresponding to the four index tests (MRI, ECHO, SPECT and PET), the reference test (CCA), the target condition (CAD) and diagnostic performance. We used numerous synonyms including both 'text words' and MeSH (Medical Subject Headings) terms to maximise the sensitivity of our search. See Appendix A in the Electronic Supplementary Material for a detailed description of the search strategy.

Study selection

Two authors reviewed article titles and abstracts for eligibility. Discrepancies were resolved by consensus.

We included studies if they met all of the following criteria: (1) the study assessed diagnostic performance of stress perfusion MRI, stress perfusion contrast-enhanced ECHO, stress perfusion SPECT, or stress perfusion PET as a diagnostic test for CAD, (2) a prospective study design was used, (3) the study population consisted of known (previously diagnosed) or suspected adult CAD patients, (4) CCA was used as the reference standard test in all patients irrespective of the non-invasive test result, i.e. selective verification was not present, (5) obstructive CAD was defined as at least 1 vessel with at least 50 %, at least 70 % or at least 75 % lumen diameter reduction and, (6) absolute numbers of true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) were available at the patient level or could be derived adequately.

Studies were excluded if they met one of the following criteria: (1) the article was a review or meta-analysis, (2) patients had (suspected) acute coronary syndrome (ACS), (3) normal healthy volunteers or asymptomatic patients were included, (4) less than 30 patients were included (criterion to avoid TPs, FPs, TNs or FNs of zero), (5) (potentially) overlapping study populations were reported, (6) a very specific patient population (e.g. only patients with a heart transplant, left bundle branch block or aortic stenosis) was studied, (7) the study focused on in-stent or graft stenosis after percutaneous coronary intervention (PCI) or coronary artery bypass grafting (CABG).

Data extraction

Two authors independently extracted data on author, journal, year of publication, technique used, country, hospital

type, number of patients, mean age, percentage male, patient selection, brand of imaging device, magnetic field strength, radiotracer, contrast agent used, type of assessment (qualitative or quantitative), stressor used, CAD definition and the numbers of TP, FP, TN and FN. Discrepancies were resolved by consensus.

If a study reported pairs of sensitivities and specificities at different cut-off points, we extracted the pair with the highest sensitivity. When studies reported data for multiple CAD definitions (e.g. at least 50 % and at least 70 % stenosis), the highest sensitivity was used to calculate the overall estimates. This also applied when studies reported sensitivities and specificities for different observers.

Quality assessment

We used a modified QUADAS checklist (quality assessment of studies of diagnostic performance included in systematic reviews) [14] to assess the quality of included studies. Two authors independently assessed the study quality of the included articles. Discrepancies were resolved by consensus.

Statistical analysis and data synthesis

We analysed the data at the patient level using a bivariate random effects regression model [15]. The model assumes a binomial distribution of the within-study variability (variability between sensitivity and specificity within a study). The model furthermore assumes correlated normally distributed random effects between studies. The degree of correlation between the logit sensitivity and logit specificity corresponds to the inverse relation between sensitivity and specificity when the positivity criterion is varied. Additionally, meta-regression was performed to explore the effect of differences in patient selection and CAD disease definition, taking into account the possible interaction between differences in CAD disease definition and the techniques considered.

The data of each study were summarised in forest plots and summary estimates with a 95 % confidence interval of sensitivity and specificity for each imaging technique. Additionally, we summarised these numbers in receiver operator characteristic (ROC) spaces showing the summary estimates with a 95 % confidence region and a summary curve. To distinguish SPECT studies that used different protocols, we highlighted the studies that combined gated-SPECT with the use of ^{99m}technetium as a radiotracer (Fig. 4). Similarly, MRI studies that included the assessment of delayed contrast enhancement were highlighted. Figures were created using Cochrane's Review Manager (version 5, Copenhagen, Denmark).

To estimate the clinical utility of each technique we calculated the positive and negative likelihood ratios (LR + and LR−). The likelihood ratio is equivalent to the ratio of the likelihood of a certain test result in patients with the disease and the likelihood of the same test result in those without the disease. LR+ [=sensitivity/(1−specificity)] describes the likelihood when the test is positive and LR− [=(1−sensitivity)/specificity] describes the likelihood when the test is negative. To illustrate the clinical utility, we used the LRs to calculate post-test probabilities across the range of possible pre-test probabilities (Fig. 5).

Finally, we calculated the natural logarithm of the diagnostic odds ratio (lnDOR). The lnDOR represents an overall summary estimate of diagnostic performance. The diagnostic odds ratio (DOR) is the odds of positive test results in patients with disease compared to the odds of positive test results in those without disease which equals the ratio of the positive and negative likelihood ratios.

We also created funnel plots to assess the presence of publication bias. The funnel plot shows the DOR horizontally and the standard error of the log transformed DOR vertically. Publication bias usually occurs when negative publications (in our case studies with a low DOR) with a small sample size are not published. An asymmetric funnel plot, for example one with fewer studies in the lower left part of the graph, suggests the presence of publication bias.

The statistical software package SAS (Proc NLMIXED, SAS v9.2, Raleigh, NC, USA) was used for the analyses.

Results

Medline (PubMed) and Embase searches yielded 1,649 unique studies (Fig. 1). On the basis of title and abstract we excluded 1,405 articles. On the basis of the full text, we excluded 202 for various reasons detailed in Fig. 1. Review

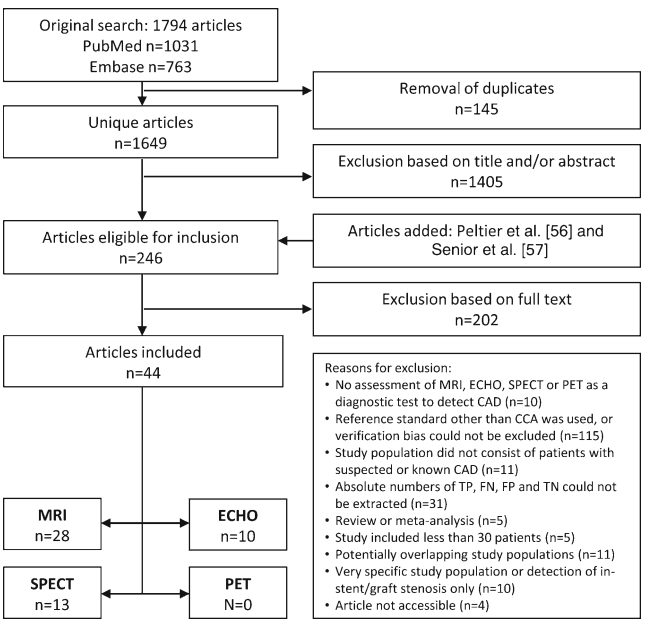


Fig. 1 Flow chart of systematic literature search

Table 1 Study characteristics

Author	Journal	Year	Technique	Country	Type	Patients (*n*)	Mean age	Age SD	% male	Patients[a]	Brand	Tesla	Perfusion tests	Assessment	Stressor	CAD definition
Arnold et al. [21]	JACC Cardiovasc Imaging	2010	MRI	UK	N	65	64	9	65	S&K	Siemens	3	Rest, stress, DE	Qualitative	Adenosine	≥50 %
Bernhardt et al. [22]	JACC Cardiovasc Imaging	2009	MRI	Germany and Canada	A	823	64	12	76	S&K	Philips	1.5	Stress, DE[b]	Qualitative	Adenosine	≥70 %
Cheng et al. [23]	J Am Coll Cardiol	2007	MRI	UK	A	61	64	8	75	S&K	Siemens	3	Rest, stress	Qualitative	Adenosine	≥50 %
Cury et al. [24]	Radiology	2006	MRI	Brasil	N	46	63	5	81	S&K	GE	1.5	Stress, DE	Qualitative	Dipyridamole	≥70 %
Donati et al. [25]	Am J Roentgenol	2010	MRI	Switzerland	A	65	64	9	81	S&K	Philips	1.5	Rest, stress, DE	Qualitative	Adenosine	>50 %
Doyle et al. [26]	J Cardiovasc Magn Reson	2003	MRI	USA	A	184	59	11	0	NS	Philips	1.5	Rest, stress	Semi-quantitative	Dipyridamole	≥70 %
Gebker et al. [27]	Radiology	2007	MRI	Germany	N	40	61	8	70	S&K	Philips	1.5	Rest, stress, DE	Qualitative	Adenosine	≥50 %
Gebker et al. [28]	Radiology	2008	MRI	Germany	N	101	62	8	70	S&K	Philips	3	Rest, stress, DE	Qualitative	Adenosine	≥50 %
Gebker et al. [29]	Int J Cardiol	2011	MRI	Germany	N	78	65	10	76	S&K	Philips	1.5	Rest, stress, DE	Qualitative	Dobutamine/ atropine	≥70 %
Giang et al. [30]	Eur Heart J	2004	MRI	Switzerland	A	44	58	NA	78	S&K	GE	1.5	Stress	Semi-quantitative	Adenosine	≥50 %
Kawase et al. [31]	Osaka City Med J	2004	MRI	Japan	N	50	67	12	58	NS	Philips	1.5	Rest, stress	Qualitative	Nicorandil	≥70 %
Kitagawa, et al. [32]	Eur Radiol	2008	MRI	Japan	A	50	65	9	72	S&K	GE	1.5	Rest, stress, DE	Qualitative	ATP	≥50 %
Klein et al. [33]	J Cardiovasc Magn Reson	2008	MRI	Germany	N	51	60	10	65	NS	Philips	1.5	Rest, stress, DE	Qualitative	Adenosine	>50 %
Klein et al. [34]	JACC Cardiovasc Imaging	2009	MRI	UK and Germany	A	78	66	8	90	K	Philips	1.5	Stress, DE	Qualitative	Adenosine	>50 %
Klem et al. [35]	J Am Coll Cardiol	2006	MRI	USA	A	92	58	12	49	S	Siemens	1.5	Rest, stress, DE[b]	Qualitative	Adenosine	≥50 % and ≥70 % (≥50 LM)
Klumpp et al. [36]	Eur Radiol	2010	MRI	Germany	A	57	62	11	82	S&K	Siemens	3	Rest, stress, DE	Qualitative	Adenosine	>70 %
Krittayaphong et al. [37]	Int J Cardiovasc Imaging	2009	MRI	Thailand	A	66	61	12	58	S	Philips	1.5	Rest, stress	Semi-quantitative	Adenosine	≥50 %
Merkle et al. [38]	Heart	2007	MRI	Germany	A	228	61	11	79	S&K	Philips	1.5	Rest, stress, DE	Qualitative	Adenosine	>50 % and >70 %
Meyer et al. [39]	Eur Radiol	2008	MRI	Germany	A	60	59	10	63	S&K	Philips	3	Rest, stress, DE	Qualitative	Adenosine	≥70 %
Nagel et al. [40]	Circulation	2003	MRI	Germany	A	84	63	8	81	S	Philips	1.5	Rest, stress	Semi-quantitative	Adenosine	≥75 %[‡]
Paetsch et al. [41]	Circulation	2004	MRI	Germany	N	79	61	9	66	S&K	Philips	1.5	Rest, stress	Qualitative	Adenosine	>50 %
Pilz et al. [42]	Clin Res Cardiol	2006	MRI	Germany	A	171	62	12	63	S&K	GE	1.5	Rest, stress, DE	Qualitative	Adenosine	>70 %
Pingitore et al. [43]	Am J Cardiol	2008	MRI	Italy	A	93	61	NA	70	S&K	GE	1.5	Rest, stress	Quantitative	Dipyridamole	>50 %
Plein et al. [44]	Radiology	2005	MRI	UK	N	82	58	NA	74	S&K	Philips	1.5	Rest, stress	Semi-quantitative	Adenosine	>70 %
Plein et al. [45]	Eur Heart J	2008	MRI	Switzerland	A	51	59	10	76	S&K	Philips	1.5	Stress, DE	Qualitative	Adenosine	>50 %
Plein et al. [46]	Radiology	2008	MRI	Switzerland and UK	A	33	58	11	73	S&K	Philips	3	Stress, DE	Qualitative	Adenosine	>50 %
Stolzmann et al. [47]	Int J Cardiovasc Imaging	2010	MRI	Switzerland	A	65	64	10	87	NS	Philips	1.5	Rest, stress, DE	Semi-quantitative	Adenosine	>50 %

Table 1 (continued)

Author	Journal	Year	Technique	Country	Type	Patients (*n*)	Mean age	Age SD	% male	Patients[a]	Brand	Tesla	Perfusion tests	Assessment	Stressor	CAD definition
Takase et al. [48]	Jpn Heart J	2004	MRI	Japan	N	102	66	9	83	S&K	GE	1.5	Rest, stress, DE	Qualitative	Dipyridamole	>50 %
Author	Journal	Year	Technique	Country	Type	Patients (*n*)	Mean age	Age SD	% male	Patients	Brand	Contrast agent	Perfusion tests	Assessment	Stressor	CAD definition
Aggeli et al. [49]	Am J Hypertens	2007	ECHO	Greece	A	50	67	5	68	S	Philips	SonoVue	Rest, stress	Qualitative	Adenosine	≥50 %
Arnold et al. [21]	JACC Cardiovasc Imaging	2010	ECHO	UK	N	65	64	9	65	S&K	Philips	Optison	Rest, stress	Qualitative	Adenosine	≥50 %
Chiou et al. [50]	Can J Cardiol	2004	ECHO	Taiwan	N	132	67	11	75	S&K	Philips	PESDA	Rest, stress	Qualitative	Dobutamine	≥50 % (≥40 % LM)
Jeetley et al. [51]	J Am Coll Cardiol	2006	ECHO	UK	A/N	123	62	12	71	S&K	Philips	Sonazoid	Rest, stress	Semi-quantitative	Dipyridamole	≥70 %
Kowatsch et al. [52]	J Am Soc Echocardiogr	2007	ECHO	Brasil	A	54	60	9	61	S&K	Philips	PESDA	Rest, stress	Quantitative	Adenosine	>50 %
Lipiec et al. [53]	J Am Soc Echocardiogr	2008	ECHO	Poland	A	103	58	9	63	S&K	Siemens	Optison	Rest, stress	Qualitative	Dipyridamole or atropine	≥70 %
Miszalski-Jamka et al. [54]	Int J Cardiol	2009	ECHO	Poland	A	103	58	NA	80	S&K	Philips	Sonovue	Rest, stress	Qualitative	Exercise	≥50 %
Moir et al. [55]	J Am Soc Echocardiogr	2005	ECHO	Australia	A	79	56	NA	80	S&K	Philips	Definity	Rest, stress	Quantitative	Dipyridamole, exercise	≥50 %
Peltier et al. [56]	J Am Coll Cardiol	2004	ECHO	Belgium	A	35	62	10	71	S&K	Agilent Technologies	PESDA	Rest, stress	Quantitative	Dipyridamole	>70 %
Senior et al. [57]	Am Heart J	2004	ECHO	UK, Germany, Belgium	N	54	61†	NA	82	NS	Philips	Sonazoid	Rest, stress	Qualitative	Dipyridamole	>50 %
Author	Journal	Year	Technique	Country	Type	Patients (*n*)	Mean age	Age SD	% male	Patients	Brand	Radiotracer	Perfusion tests	Assessment	Stressor	CAD definition
Aggeli et al. [49]	Am J Hypertens	2007	SPECT	Greece	A	48	67	5	68	S	GE	^{201}Tl	Rest, stress	Qualitative and qualitative	Adenosine	≥50 %
Astarita et al. [58]	J Hypertens	2001	SPECT	Italy	N	53	58	10	55	S	Picker	^{201}Tl	Rest, stress	Qualitative	Exercise	≥50 %
Budoff et al. [59]	Acad Radiol	2007	SPECT	USA	A	30	54	9	70	NS	NS	^{99m}Tc-MIBI	Rest, stress	Qualitative	Exercise	>70 % (>50 % LM)
Doyle et al. [26]	J Cardiovasc Magn Reson	2003	SPECT	USA	A	184	59	11	0	NS	ADAC	^{99m}Tc-MIBI/^{201}Tl	Rest, stress (gated)	Qualitative	Dipyridamole	≥70 %
Gonzalez et al. [60]	Rev Esp Med Nucl	2005	SPECT	Chile	A	145	60	12	68	S&K	GE	^{201}Tl	Rest, stress	Qualitative	Exercise (*n*=63), dipyridamole (*n*=82)	≥50 % and ≥75 %
Jeetley et al. [51]	J Am Coll Cardiol	2006	SPECT	UK	A/N	123	62	12	71	S&K	NS	^{99m}Tc-MIBI	Rest, stress	Semi-quantitative	Dipyridamole	≥70 %
Johansen et al. [61]	J Nucl Cardiol	2005	SPECT	Denmark	A	357	57	9	54	S	Picker	^{201}Tl(rest) and ^{99m}Tc-MIBI (stress)	Rest, stress (gated)	Semi-quantitative	Adenosine or dobutamine (*n*=180), exercise stress test (*n*=177)	≥50 %
Lipiec et al. [53]	J Am Soc Echocardiogr	2008	SPECT	Poland	A	103	58	9	63	S&K	GE	^{99m}Tc-MIBI	Rest, stress	Semi-quantitative	Dipyridamole	≥70 %
Peltier et al. [56]	J Am Coll Cardiol	2004	SPECT	Belgium	A	35	62	10	71	S&K	GE	^{99m}Tc-MIBI	Rest, stress	Qualitative	Dipyridamole	>70 %
Schepis et al. [62]	J Nucl Med	2007	SPECT	Switzerland	A	77	66	9	62	S	GE	^{99m}Tc-tetrofosmin	Rest, stress (gated)	Semi-quantitative	Adenosine	≥50 %
Senior et al. [57]	Am Heart J	2004	SPECT	UK, Germany, Belgium	N	53	61†	NA	82	NS	Amersham Health	^{99m}Tc-tetrofosmin	Rest, stress	Qualitative	Dipyridamole	>50 %

Table 1 (continued)

Author	Journal	Year	Technique	Country	Type	Patients (*n*)	Mean age	Age SD	% male	Patients[a]	Brand	Tesla	Perfusion tests	Assessment	Stressor	CAD definition
Yao et al. [63]	Nucl Med Commun	2000	SPECT	China	N	64	51	NA	95	S&K	Siemens	^{99m}Tc-MIBI	Rest, stress	Qualitative	Exercise	≥50 %
Yeih et al. [64]	J Formos Med Assoc	2007	SPECT	Taiwan	A	51	63	9	0	S&K	GE	^{201}Tl	Rest, stress	Qualitative	Dobutamine	≥50 %

A academic, *N* non-academic, *NS* not specified, *NA* not available, *PESDA* perfluorocarbon-exposed sonicated dextrose albumin, *S* suspected CAD only (and no history of myocardial infarction (MI), CABG or PCI), *S&K* patients with either suspected or known CAD, *K* known CAD, *^{99m}Tc* ^{99m}technetium, *-MIBI* sestamibi, *^{201}Tl* ^{201}thallium, *DE* delayed enhancement, *ATP* adenosine triphosphate, *LM* left main coronary artery

† Median; ‡ area reduction

[a] If there was any uncertainty about the study population it was not classified as suspected and/or known CAD but rather as "not specified"

[b] In Klem et al. [35] and Bernhardt et al. [22] the index test outcome was initially based on delayed enhancement (DE) images, considering perfusion images when DE was negative

of the study characteristics shows considerable differences between the included studies (Table 1).

Forty-four studies met the inclusion criteria. Articles on MRI yielded a total of 2,970 patients from 28 studies, articles on ECHO yielded a sample size of 795 from 10 studies, articles on SPECT yielded 1,323 from 13 studies. We could not include any PET studies, which is why PET was excluded from the analysis. The overview of the QUADAS checklist for all studies demonstrates some differences in terms of study quality (see Appendix B in the Electronic Supplementary Material). The funnel plots of MRI and SPECT suggest evidence for publication bias, whereas the funnel plot of ECHO shows no obvious evidence for publication bias (Fig. 2). The sensitivities and specificities of each study vary across studies with sample sizes ranging from 30 to 823 (Tables 2 and 3). The forest plots show the sensitivities and specificities of each study with their 95 % confidence intervals depicted as horizontal lines (Fig. 3), grouped by CAD definition and study population and then sorted by sensitivity.

Compared with coronary angiography the meta-analysis of the sensitivities and specificities of the different techniques (Table 3; Fig. 3) resulted for MRI in a sensitivity of 0.91 (95 % CI 0.88–0.93) and a specificity of 0.80 (95 % CI 0.76–0.83). Perfusion ECHO showed a sensitivity of 0.87 (95 % CI 0.81–0.91) and a specificity of 0.72 (95 % CI 0.56–0.83). SPECT demonstrated a sensitivity of 0.83 (95 % CI 0.73–0.89) and a specificity of 0.77 (95 % CI 0.64–0.86). The ROC spaces show the summary estimates for sensitivity and specificity of each technique two-dimensionally surrounded by its 95 % confidence area (Fig. 4). The sensitivity of MRI and SPECT differed significantly (P=0.03). In terms of specificity, no significant differences were found.

We found no effect of CAD definition on the sensitivities (Table 3; P=0.55). The disease definition greater than/at least 70 % stenosis compared to greater than/at least 50 % stenosis resulted in significantly lower specificities for SPECT (Table 3; P=0.045), but no significant differences for ECHO (P=0.39) and MRI (P=0.51). Furthermore, we found no effect of CAD definition on the lnDORs of MRI (P=0.24), ECHO (P=0.96) and SPECT (P=0.34) (Table 3).

Furthermore, MRI, ECHO and SPECT showed no significant differences in terms of sensitivity, specificity and lnDOR when comparing patients with suspected CAD without a prior history of CAD to patients with known or suspected CAD (all P values >0.05; Table 3).

We did not observe an association between the use of gated-SPECT in combination with ^{99m}technetium as radiotracer and the diagnostic performance of SPECT (Fig. 4). MRI studies that assessed delayed contrast enhancement were associated with high sensitivities albeit with a wide range of specificities (Fig. 4).

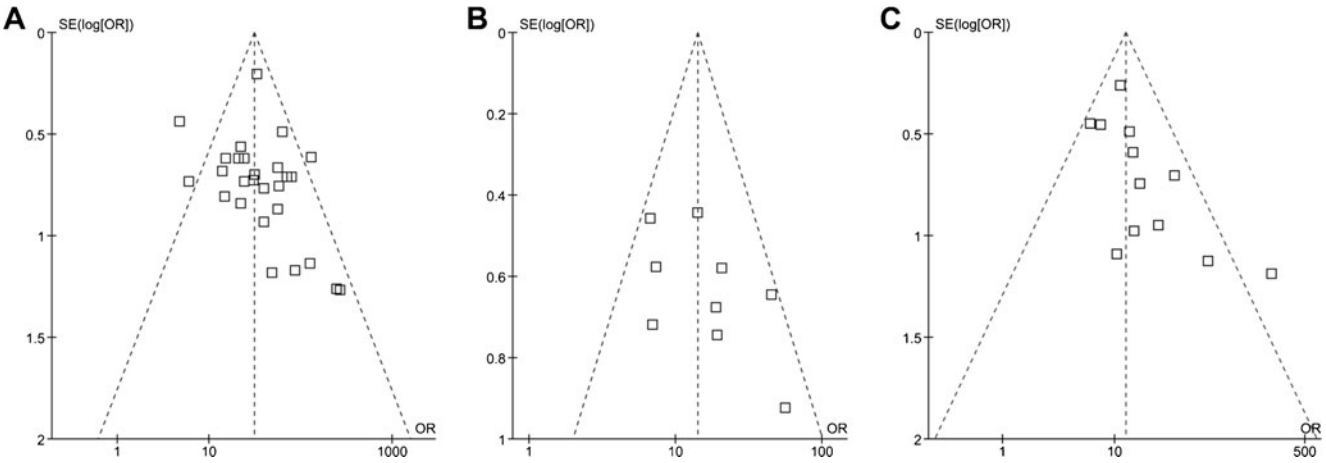


Fig. 2 Funnel plots. The diagnostic odds ratio (DOR) on the *x*-axis is plotted against the standard error (SE) of the log(DOR) on the *y*-axis. A symmetrical distribution of studies indicates the absence of publication bias. An asymmetrical distribution with, for example, relatively more smaller studies with a positive result (in the *lower right part* of the plot) would suggest the presence of publication bias. In the ECHO funnel plot Peltier et al. [56], in the SPECT funnel plot Astarita et al. [58] and in the MRI funnel plot Donati et al. [25] are not included, because their respective DORs could not be calculated (0 false negatives or false positives). **a** MRI, **b** ECHO, **c** SPECT

The positive likelihood ratios (LR+) of MRI, ECHO and SPECT were 4.43 (95 % CI 3.64–5.23), 3.08 (95 % CI 1.65–4.50) and 3.56 (95 % CI 2.07–5.04) respectively (Table 3). The negative likelihood ratios (LR-) for MRI, ECHO, and SPECT were 0.12 (95 % CI 0.08–0.15), 0.18 (95 % CI 0.13–0.24) and 0.22 (95 % CI 0.14–0.31), respectively. Figure 5 illustrates the revised probability of CAD after a positive and negative test. The lnDORs of MRI, ECHO and SPECT were 3.63 (95 % CI 3.26–4.00), 2.83 (95 % CI 2.29–3.37) and 2.76 (95 % CI 2.28–3.25), respectively (Table 3). We found significantly higher lnDORs for MRI in comparison with SPECT (P=0.006) and ECHO (P=0.02). There was no significant difference between the lnDOR of SPECT and ECHO (P=0.52).

Discussion

In this systematic review and meta-analysis we compared the diagnostic performance of different stress MPI techniques. MRI showed the best diagnostic performance with the narrowest confidence intervals; the latter is explained by the large number of patients studied with MRI. We found a significantly higher sensitivity for MRI compared to SPECT and a significantly higher lnDOR for MRI compared to both ECHO and SPECT. In contrast to previous meta-analyses [9], we compared the different imaging techniques using the same search strategy and methods of analysing the data. Furthermore, we only included studies without verification bias.

In our review we paid special attention to the issue of verification bias. Sensitivity may be overestimated and specificity underestimated if patients with a positive test result are more likely to be verified with the reference standard test. Diagnostic odds ratios are generally not, or only minimally, affected by verification bias [16]. Underwood et al. [9] reviewed the diagnostic performance of SPECT and explained the overall low specificity (0.70–0.75 for high quality studies) of SPECT studies by verification bias. In their review of SPECT studies, Heijenbrok-Kal et al. [13] did not exclude studies with verification bias and demonstrated a sensitivity of 0.88 (95 % CI 0.87–0.90) and a specificity of 0.73 (95 % CI 0.69–0.74). By excluding studies with verification bias, we found a lower sensitivity of 0.83 (95 % CI 0.73–0.89), but a higher specificity of 0.77 (95 % CI 0.64–0.86). As pointed out above, the diagnostic odds ratios are less affected by verification bias and were the same for the previous and current review.

Nandalur et al. [7] and Hamon et al. [10] previously studied the diagnostic performance of myocardial perfusion MRI and found sensitivities of 91 % and 89 % respectively and specificities of 81 % and 80 % respectively, which is very similar to what we found. Unfortunately we could not include PET in the analysis, because no PET studies met our inclusion and exclusion criteria. Nandalur et al. [6] performed a meta-analysis of PET perfusion studies and they found a sensitivity of 0.92 and a specificity of 0.85. However, their analysis included studies with potential verification bias. Stress perfusion CT is an upcoming MPI technique, but we did not include this technique because of the low number of available studies and because perfusion CT is still in the technical development phase.

Other promising alternatives to CCA are non-invasive CT and MR coronary angiography. Schuetz et al. [17] compared CT and MR coronary angiography to CCA in a meta-analysis

Table 2 Source data for MRI, ECHO and SPECT

Author	Year	Technique	TP	FN	TN	FP	Sensitivity (%)	Specificity (%)	CAD definition
Arnold et al. [21]	2010	MRI	37	4	17	4	90.2	81.0	≥50 %
Bernhardt et al. [22]	2009	MRI	274	39	421	89	87.5	82.5	≥70 %
Cheng et al. [23]	2007	MRI	39	1	16	5	97.5	76.2	≥50 %
Cury et al. [24]	2006	MRI	29	1	12	4	96.7	75.0	≥70 %
Donati et al. [25]	2010	MRI	30	3	14	0	90.9	100	>50 %
Doyle et al. [26]	2003	MRI	15	11	123	35	57.7	77.8	≥70 %
Gebker et al. [27]	2007	MRI	19	3	14	4	86.4	77.8	≥50 %
Gebker et al. [28]	2008	MRI	63	7	22	9	90.0	71.0	≥50 %
Gebker et al. [29]	2011	MRI	48	4	19	4	92.3	82.6	≥70 %
Giang et al. [30]	2004	MRI	26	2	12	4	92.9	75.0	≥50 %
Kawase et al. [31]	2004	MRI	31	2	16	1	93.9	94.1	≥70 %
Kitagawa et al. [32]	2008	MRI	33	3	8	6	91.7	57.1	≥50 %
Klein et al. [33]	2008	MRI	22	3	23	3	88.0	88.5	>50 %
Klein et al. [34]	2009	MRI	36	18	21	3	66.7	87.5	>50 %
Klem et al. [35]	2006	MRI	34	10	42	6	77.3	87.5	≥50 %
			33	4	48	7	89.2	87.3	≥70 % (≥50 LM)
Klumpp et al. [36]	2010	MRI	40	1	14	2	97.6	87.5	>70 %
Krittayaphong et al. [37]	2009	MRI	34	4	22	6	89.5	78.6	≥50 %
Merkle et al. [38]	2007	MRI	160	12	48	8	93.0	85.7	>50 %
			147	6	54	21	96.1	72.0	>70 %
Meyer et al. [39]	2008	MRI	32	4	19	5	88.9	79.2	≥70 %
Nagel et al. [40]	2003	MRI	38	5	37	4	88.4	90.2	≥75 %
Paetsch et al. [41]	2004	MRI	48	5	16	10	90.6	61.5	>50 %
Pilz et al. [42]	2006	MRI	109	4	48	10	96.5	82.8	>70 %
Pingitore et al. [43]	2008	MRI	61	5	18	9	92.4	66.7	>50 %
Plein et al. [44]	2005	MRI	52	7	17	6	88.1	73.9	>70 %
Plein et al. [45]	2008	MRI	31	4	7	9	88.6	43.8	>50 %
Plein et al. [46]	2008	MRI	12	1	16	4	92.3	80.0	>50 %
Stolzmann et al. [47]	2010	MRI	28	8	21	3	77.8	87.5	>50 %
Takase et al. [48]	2004	MRI	71	5	22	4	93.4	84.6	>50 %
Aggeli et al. [49]	2007	ECHO	28	4	16	2	87.5	88.9	≥50 %
Arnold et al. [21]	2010	ECHO	35	6	16	5	85.4	76.2	≥50 %
Chiou et al. [50]	2004	ECHO	69	16	36	11	81.2	76.6	≥50 % (≥40 % LM)
Jeetley et al. [51]	2006	ECHO	74	11	19	19	87.1	50.0	≥70 %
Kowatsch et al. [52]	2007	ECHO	22	3	21	8	88.0	72.4	>50 %
Lipiec et al. [53]	2008	ECHO	69	10	18	6	87.3	75.0	≥70 %
Miszalski-Jamka et al. [54]	2009	ECHO	65	9	25	4	87.8	86.2	≥50 %
Moir et al. [55]	2005	ECHO	35	5	20	19	87.5	51.3	≥50 %
Peltier et al. [56]	2004	ECHO	22	0	10	3	100.0	76.9	>70 %
Senior et al. [57]	2004	ECHO	35	7	7	5	83.3	58.3	>50 %
Aggeli et al. [49]	2007	SPECT	24	6	17	1	80.0	94.4	≥50 %
Astarita et al. [58]	2001	SPECT	23	0	14	16	100.0	46.7	≥50 %
Budoff et al. [59]	2007	SPECT	17	4	7	2	81.0	77.8	>70 % (>50 % LM)
Doyle et al. [26]	2003	SPECT	16	10	130	28	61.5	82.3	≥70 %
Gonzalez et al. [60]	2005	SPECT	102	15	16	12	87.2	57.1	≥50 %
			91	7	24	23	92.9	51.1	≥75 %
Jeetley et al. [51]	2006	SPECT	73	12	19	19	85.9	50.0	≥70 %
Johansen et al. [61]	2005	SPECT	94	32	183	48	74.6	79.2	≥50 %
Lipiec et al. [53]	2008	SPECT	73	6	13	11	92.4	54.2	≥70 %

Table 2 (continued)

Author	Year	Technique	TP	FN	TN	FP	Sensitivity (%)	Specificity (%)	CAD definition
Peltier et al. [56]	2004	SPECT	18	4	11	2	81.8	84.6	>70 %
Schepis et al. [62]	2007	SPECT	32	10	32	3	76.2	91.4	≥50 %
Senior et al. [57]	2004	SPECT	20	21	11	1	48.8	91.7	>50 %
Yao et al. [63]	2000	SPECT	42	3	18	1	93.3	94.7	≥50 %
Yeih et al. [64]	2007	SPECT	20	8	20	3	71.4	87.0	≥50 %

resulting in a sensitivity and specificity of respectively 0.97 and 0.87 for CT, and 0.87 and 0.70 for MR, suggesting that CT angiography has a better diagnostic performance compared to the MPI techniques analysed in this article. However, drawbacks of CT angiography are the use of iodinated contrast material which poses a small risk of idiosyncratic reactions and nephrotoxicity and the lack of functional information [18].

Limitations

We focused on the diagnostic performance of MPI. However, an MPI examination can yield functional information as well (e.g. left ventricular function, presence of wall motion abnormalities, presence of scar tissue), rather than perfusion images alone. Our analysis does not take into account the possible impact of these parameters on the interpretation of the MPI test and the results of MRI are therefore likely to be even better than we estimated.

Also, it is important to note that in clinical practice a small proportion of patients will be unsuitable for MRI, either due to contraindications or claustrophobia. Likewise, an echocardiography procedure relies on an adequate acoustic window. Often, unsuitable patients were excluded from the original studies, which in turn could have resulted in an overestimation of the diagnostic performance in our analysis. Unfortunately, the included studies did not report sufficient information to explore these issues.

In the current review we included only studies that used the most advanced technology by searching for studies published from 2000 until 2011, which implies

Table 3 Measures of diagnostic performance for MRI, ECHO and SPECT, estimated using the bivariate random effects model

		Sensitivity	Specificity	LR+	LR-	DOR	lnDOR	CAD prevalence*
MRI	**Overall**	**0.91 (0.88 – 0.93)**	**0.80 (0.76 – 0.83)**	**4.43 (3.64 – 5.23)**	**0.12 (0.08 – 0.15)**	**37.69 (26.00 – 54.63)**	**3.63 (3.26 – 4.00)**	**54% (1603/2970)**
	Suspected	0.90 (0.78 – 0.96)	0.86 (0.74 – 0.93)	6.61 (2.23 – 10.99)	0.12 (0.03 – 0.22)	54.70 (20.07 – 149.07)	4.00 (3.00 – 5.00)	49% (118/242)
	CAD 50	0.89 (0.86 – 0.92)	0.79 (0.73 – 0.84)	4.25 (3.15 – 5.35)	0.13 (0.09 – 0.17)	31.84 (20.96 – 48.37)	3.46 (3.04 – 3.88)	66% (882/1338)
	CAD 70	0.91 (0.87 – 0.94)	0.82 (0.75 – 0.87)	4.97 (3.47 – 6.47)	0.11 (0.07 – 0.15)	46.40 (28.90 – 74.49)	3.84 (3.36 – 4.31)	48% (937/1952)
ECHO	**Overall**	**0.87 (0.81 – 0.91)**	**0.72 (0.56 – 0.83)**	**3.08 (1.65 – 4.50)**	**0.18 (0.13 – 0.24)**	**16.94 (9.84 – 29.15)**	**2.83 (2.29 – 3.37)**	**66% (525/795)**
	Suspected	0.88 (0.60 – 0.97)	0.89 (0.58 – 0.98)	8.35 (6.67 – 21.76)	0.13 (-0.05 – 0.32)	62.76 (7.37 – 534.54)	4.14 (2.00 – 6.28)	64% (32/50)
	CAD 50	0.86 (0.79 – 0.92)	0.74 (0.63 – 0.82)	3.28 (2.09 – 4.47)	0.19 (0.10 – 0.27)	17.59 (9.48 – 32.66)	2.87 (2.25 – 3.49)	63% (339/534)
	CAD 70	0.90 (0.80 – 0.96)	0.65 (0.46 – 0.80)	2.58 (1.32 – 3.84)	0.15 (0.04 – 0.26)	17.04 (6.60 – 44.04)	2.84 (1.89 – 3.79)	71% (186/261)
SPECT	**Overall**	**0.83 (0.73 – 0.89)**	**0.77 (0.64 – 0.86)**	**3.56 (2.07 – 5.04)**	**0.22 (0.14 – 0.31)**	**15.84 (9.74 – 25.77)**	**2.76 (2.28 – 3.25)**	**50% (666/1323)**
	Suspected	0.83 (0.70 – 0.91)	0.79 (0.66 – 0.87)	3.88 (2.03 – 5.73)	0.21 (0.09 – 0.34)	18.15 (8.34 – 39.52)	2.90 (2.12 – 3.68)	41% (221/535)
	CAD 50	0.81 (0.72 – 0.87)	0.81 (0.72 – 0.87)	4.15 (2.55 – 5.75)	0.24 (0.15 – 0.33)	17.24 (9.67 – 30.73)	2.85 (2.27 – 3.43)	53% (452/848)
	CAD 70	0.85 (0.76 – 0.91)	0.66 (0.54 – 0.77)	2.53 (1.69 – 3.37)	0.22 (0.12 – 0.33)	11.42 (6.04 – 21.59)	2.44 (1.80 – 3.07)	53% (331/620)

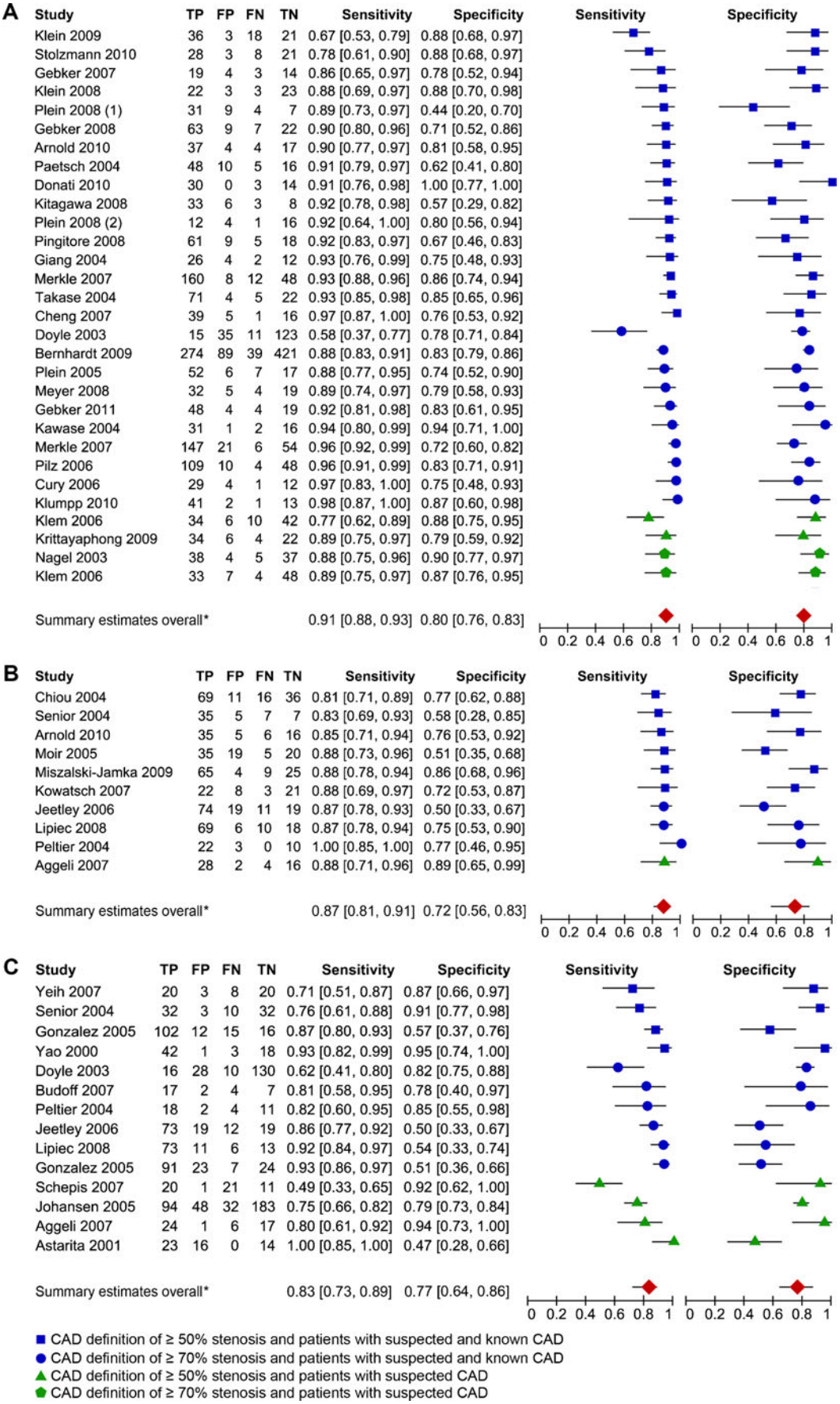
When data were available for both CAD definitions (≥50 % and ≥70 %) the overall estimates only include data from CAD ≥70 % stenosis

CAD 50 corresponds to the studies that defined obstructive CAD either as >50 % or ≥50 % stenosis

CAD 70 corresponds to the studies that defined obstructive CAD either as >70 %, ≥70 % or ≥75 % stenosis and studies that combined one of these with >50 % (or ≥50 %) stenosis in the left main coronary artery

"Suspected" refers to studies that only included patients with suspected CAD without a history of MI, PCI or CABG.

[a] The CAD prevalence defined by "CAD diagnosed by CCA" divided by the total sample size


A
Study TP FP FN TN Sensitivity Specificity
Klein 2009 36 3 18 21 0.67 [0.53, 0.79] 0.88 [0.68, 0.97]
Stolzmann 2010 28 3 8 21 0.78 [0.61, 0.90] 0.88 [0.68, 0.97]
Gebker 2007 19 4 3 14 0.86 [0.65, 0.97] 0.78 [0.52, 0.94]
Klein 2008 22 3 3 23 0.88 [0.69, 0.97] 0.88 [0.70, 0.98]
Plein 2008 (1) 31 9 4 7 0.89 [0.73, 0.97] 0.44 [0.20, 0.70]
Gebker 2008 63 9 7 22 0.90 [0.80, 0.96] 0.71 [0.52, 0.86]
Arnold 2010 37 4 4 17 0.90 [0.77, 0.97] 0.81 [0.58, 0.95]
Paetsch 2004 48 10 5 16 0.91 [0.79, 0.97] 0.62 [0.41, 0.80]
Donati 2010 30 0 3 14 0.91 [0.76, 0.98] 1.00 [0.77, 1.00]
Kitagawa 2008 33 6 3 8 0.92 [0.78, 0.98] 0.57 [0.29, 0.82]
Plein 2008 (2) 12 4 1 16 0.92 [0.64, 1.00] 0.80 [0.56, 0.94]
Pingitore 2008 61 9 5 18 0.92 [0.83, 0.97] 0.67 [0.46, 0.83]
Giang 2004 26 4 2 12 0.93 [0.76, 0.99] 0.75 [0.48, 0.93]
Merkle 2007 160 8 12 48 0.93 [0.88, 0.96] 0.86 [0.74, 0.94]
Takase 2004 71 4 5 22 0.93 [0.85, 0.98] 0.85 [0.65, 0.96]
Cheng 2007 39 5 1 16 0.97 [0.87, 1.00] 0.76 [0.53, 0.92]
Doyle 2003 15 35 11 123 0.58 [0.37, 0.77] 0.78 [0.71, 0.84]
Bernhardt 2009 274 89 39 421 0.88 [0.83, 0.91] 0.83 [0.79, 0.86]
Plein 2005 52 6 7 17 0.88 [0.77, 0.95] 0.74 [0.52, 0.90]
Meyer 2008 32 5 4 19 0.89 [0.74, 0.97] 0.79 [0.58, 0.93]
Gebker 2011 48 4 4 19 0.92 [0.81, 0.98] 0.83 [0.61, 0.95]
Kawase 2004 31 1 2 16 0.94 [0.80, 0.99] 0.94 [0.71, 1.00]
Merkle 2007 147 21 6 54 0.96 [0.92, 0.99] 0.72 [0.60, 0.82]
Pilz 2006 109 10 4 48 0.96 [0.91, 0.99] 0.83 [0.71, 0.91]
Cury 2006 29 4 1 12 0.97 [0.83, 1.00] 0.75 [0.48, 0.93]
Klumpp 2010 41 2 1 13 0.98 [0.87, 1.00] 0.87 [0.60, 0.98]
Klem 2006 34 6 10 42 0.77 [0.62, 0.89] 0.88 [0.75, 0.95]
Krittayaphong 2009 34 6 4 22 0.89 [0.75, 0.97] 0.79 [0.59, 0.92]
Nagel 2003 38 4 5 37 0.88 [0.75, 0.96] 0.90 [0.77, 0.97]
Klem 2006 33 7 4 48 0.89 [0.75, 0.97] 0.87 [0.76, 0.95]
Summary estimates overall* 0.91 [0.88, 0.93] 0.80 [0.76, 0.83]
Sensitivity
Specificity
0 0.2 0.4 0.6 0.8 1
0 0.2 0.4 0.6 0.8 1
B
Study TP FP FN TN Sensitivity Specificity
Chiou 2004 69 11 16 36 0.81 [0.71, 0.89] 0.77 [0.62, 0.88]
Senior 2004 35 5 7 7 0.83 [0.69, 0.93] 0.58 [0.28, 0.85]
Arnold 2010 35 5 6 16 0.85 [0.71, 0.94] 0.76 [0.53, 0.92]
Moir 2005 35 19 5 20 0.88 [0.73, 0.96] 0.51 [0.35, 0.68]
Miszalski-Jamka 2009 65 4 9 25 0.88 [0.78, 0.94] 0.86 [0.68, 0.96]
Kowatsch 2007 22 8 3 21 0.88 [0.69, 0.97] 0.72 [0.53, 0.87]
Jeetley 2006 74 19 11 19 0.87 [0.78, 0.93] 0.50 [0.33, 0.67]
Lipiec 2008 69 6 10 18 0.87 [0.78, 0.94] 0.75 [0.53, 0.90]
Peltier 2004 22 3 0 10 1.00 [0.85, 1.00] 0.77 [0.46, 0.95]
Aggeli 2007 28 2 4 16 0.88 [0.71, 0.96] 0.89 [0.65, 0.99]
Summary estimates overall* 0.87 [0.81, 0.91] 0.72 [0.56, 0.83]
Sensitivity
Specificity
0 0.2 0.4 0.6 0.8 1
0 0.2 0.4 0.6 0.8 1
C
Study TP FP FN TN Sensitivity Specificity
Yeih 2007 20 3 8 20 0.71 [0.51, 0.87] 0.87 [0.66, 0.97]
Senior 2004 32 3 10 32 0.76 [0.61, 0.88] 0.91 [0.77, 0.98]
Gonzalez 2005 102 12 15 16 0.87 [0.80, 0.93] 0.57 [0.37, 0.76]
Yao 2000 42 1 3 18 0.93 [0.82, 0.99] 0.95 [0.74, 1.00]
Doyle 2003 16 28 10 130 0.62 [0.41, 0.80] 0.82 [0.75, 0.88]
Budoff 2007 17 2 4 7 0.81 [0.58, 0.95] 0.78 [0.40, 0.97]
Peltier 2004 18 2 4 11 0.82 [0.60, 0.95] 0.85 [0.55, 0.98]
Jeetley 2006 73 19 12 19 0.86 [0.77, 0.92] 0.50 [0.33, 0.67]
Lipiec 2008 73 11 6 13 0.92 [0.84, 0.97] 0.54 [0.33, 0.74]
Gonzalez 2005 91 23 7 24 0.93 [0.86, 0.97] 0.51 [0.36, 0.66]
Schepis 2007 20 1 21 11 0.49 [0.33, 0.65] 0.92 [0.62, 1.00]
Johansen 2005 94 48 32 183 0.75 [0.66, 0.82] 0.79 [0.73, 0.84]
Aggeli 2007 24 1 6 17 0.80 [0.61, 0.92] 0.94 [0.73, 1.00]
Astarita 2001 23 16 0 14 1.00 [0.85, 1.00] 0.47 [0.28, 0.66]
Summary estimates overall* 0.83 [0.73, 0.89] 0.77 [0.64, 0.86]
Sensitivity
Specificity
0 0.2 0.4 0.6 0.8 1
0 0.2 0.4 0.6 0.8 1
CAD definition of ≥ 50% stenosis and patients with suspected and known CAD
CAD definition of ≥ 70% stenosis and patients with suspected and known CAD
CAD definition of ≥ 50% stenosis and patients with suspected CAD
CAD definition of ≥ 70% stenosis and patients with suspected CAD

◀ **Fig. 3** Forest plots. The data are sorted by suspected and known CAD versus suspected CAD and CAD definition of ≥50 % versus ≥70 % stenosis from lowest to highest sensitivity and data are reported at the patient level. **a** MRI, **b** ECHO, **c** SPECT. *When data were available for both CAD definitions (≥50 % and ≥70 %) the summary estimates only include data from CAD ≥70 % stenosis

that some large landmark SPECT studies performed in the 1980s and 1990s were excluded from our analysis. A previously published comprehensive systematic review sheds light on the effect of this exclusion criterion [13]. In the previous review 103 SPECT studies with a total of 11,977 patients published between 1984 and 2002 were analysed. There is no overlap with the SPECT studies that we included. The diagnostic odds ratios for SPECT found in the previous review and in the current review are the same: they found an lnDOR

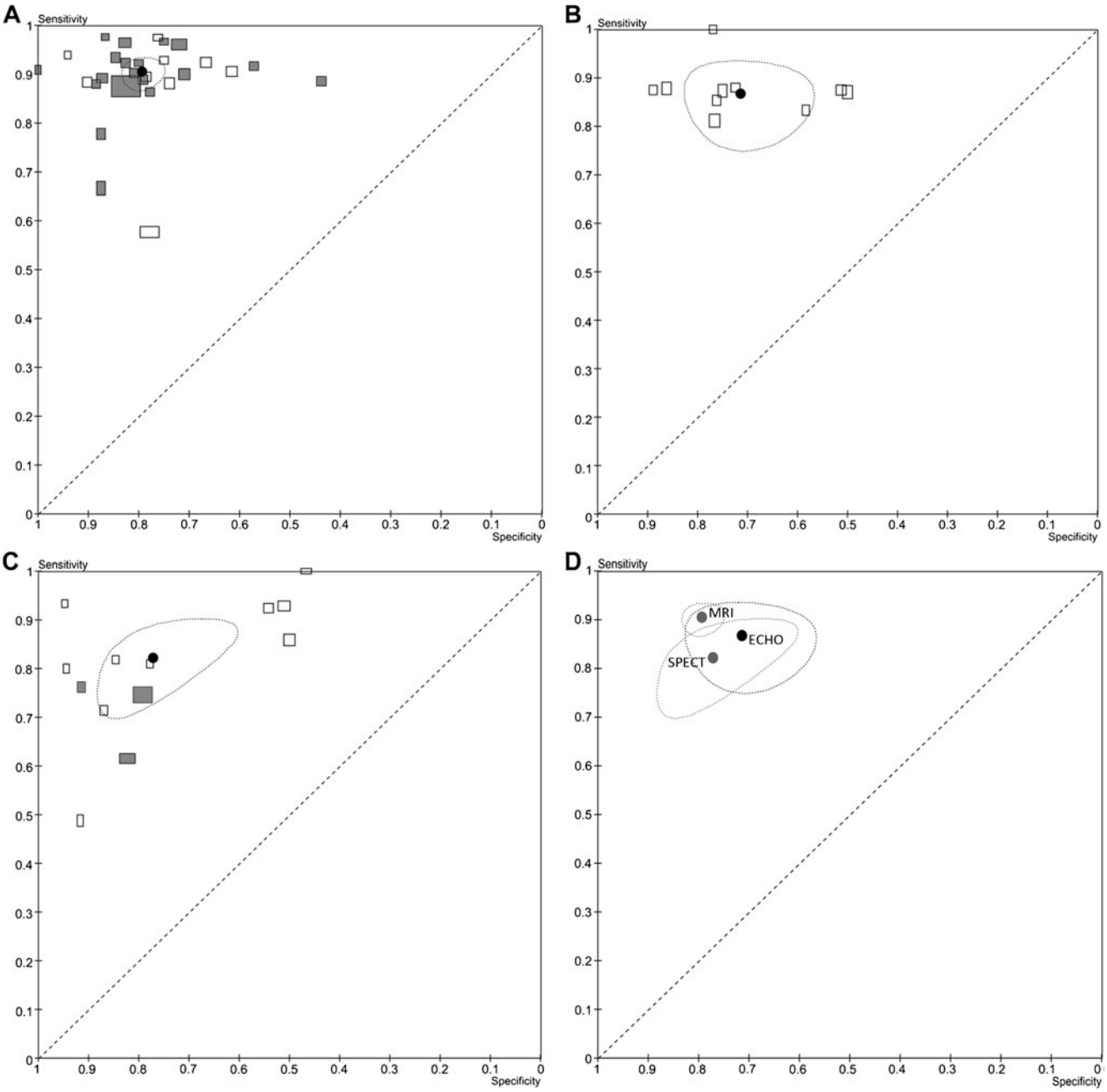


Fig. 4 ROC space with summary estimates for each technique with 95 % confidence areas. This figure shows the diagnostic performance of studies relative to each other with specificity (plotted in reverse) on the *x*-axis and sensitivity on the *y*-axis. Perfect diagnostic accuracy is in the *upper left corner*, where sensitivity and specificity are both 1. **a** MRI, **b** ECHO, **c** SPECT, **d** All three techniques. The *grey rectangles* in **a** refer to the studies using delayed contrast enhancement and in **c** they refer to the studies using gated SPECT with radiotracer ^{99m}Tc. The size of the *rectangles* corresponds with the inverse standard error of sensitivity and specificity, which correlates with the size of the study

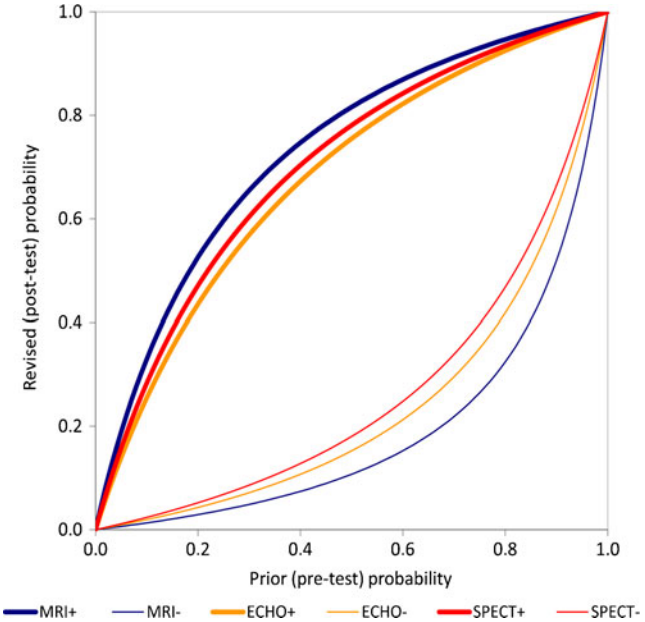


Fig. 5 Revised probability of CAD. This figure shows the revised (post-test) probability of CAD (*y*-axis) as a function of prior (pre-test) probability (*x*-axis) of CAD for positive and negative MPI results, based on the likelihood ratios presented in Table 3 (overall analysis). MRI+, ECHO+ and SPECT+ represent the lines for a positive test result and MRI−, ECHO− and SPECT− represent the lines for a negative test result

of 2.8 (95 % CI 2.6–3.0) compared to our lnDOR of 2.8 (95 % CI 2.3–3.3).

The funnel plot for MRI and SPECT suggests that there is evidence of publication bias, which implies that our summary measures may be overestimated. Nevertheless, the overestimation applies to both MRI and SPECT. The funnel plot for ECHO does not suggest evidence of publication bias.

Heterogeneity across studies is a limitation of meta-analyses of diagnostic performance. Across studies differences exist with respect to imaging techniques, assessment methods, stressors, radiotracers, contrast media, CAD definition (lumen diameter reduction of at least 50 %, at least 70 % or at least 75 %), CAD prevalence, percentage male patients, patient inclusion criteria, setting and country. Although we were able to analyse the effect of using different CAD definitions and patient inclusion criteria, sample size limitations did not allow us to do subset analyses for the other cross-study variations. Due to chance there will always be variability between studies, but there may also be different types of biases influencing the results. We used a random effects model which adjusts the estimates and confidence intervals to account for between-study variations. Nevertheless, heterogeneity across studies remains an important limitation.

For calculation and precision purposes, we excluded studies with less than 30 patients. In this way, we minimised the number of studies with for example zero FPs or FNs. This exclusion criterion may have introduced a selection bias.

Another limitation of meta-analyses is the dependence on the level of detail reported in the original papers. For example, data on the individual territories were generally not available. Furthermore, most studies included a mix of known and suspected CAD patients or did not report the test characteristics for the subgroup of patients with suspected CAD without a prior history of MI, PCI or CABG. Therefore, our subgroup analysis of suspected CAD was limited due to a small sample size. Nevertheless, our analysis did suggest that the diagnostic performance of MPI tests is not substantially affected by including patients with known CAD.

Although our results show that all tests are reasonably accurate, the likelihood ratios suggest that neither one of them is suitable to rule out or rule in the presence of disease [19]. This can also be seen in Fig. 5, where the post-test probability after a positive test rarely exceeds 90 %, and the post-test probability of disease after a negative test may still be substantial. Since MPI is intended as a gatekeeper test, ruling out disease is more important than ruling in disease. MRI performs quite well in this respect with an LR− of 0.12 (0.08–0.15). SPECT and ECHO demonstrate less favourable LRs (Table 3).

The reference standard test for diagnosing CAD is CCA. Innovative technological developments in diagnosing CAD are most often compared with CCA. The limitation of CCA is that it evaluates the lumen diameter reduction of the coronary arteries, but for instance a 50 % vessel diameter reduction does not always result in the same reduction in blood flow and does not necessarily lead to myocardial ischaemia. There are alternative techniques such as fractional flow reserve (FFR) that measure the pressure difference across a coronary stenosis. It is even possible that the imaging techniques we evaluated are better diagnostic tools than CCA to begin with, since they measure myocardial perfusion which is the physiological basis of myocardial function. Thus, the less than perfect sensitivity and specificity could in part be attributed to imperfections of CCA instead of the limitations of perfusion imaging.

Clinical implications

The results of our systematic review and meta-analysis suggest that MRI is superior to ECHO and SPECT in diagnosing CAD. This statement is strengthened firstly by the findings of the MR-IMPACT study [5]—a multicentre randomised trial—which suggested that MRI is superior to SPECT and secondly by the findings of the EuroCMR registry

[20], which demonstrated that in patients who underwent stress MRI for the diagnostic workup of suspected CAD, invasive angiography could be avoided in nearly one-half of the patients. All in all, the results suggest that stress perfusion MRI is potentially useful as a gatekeeper test before CCA in patients with low to intermediate prior probability of CAD but this needs to be confirmed with a comparative cost-effectiveness analysis. Furthermore, more research of the diagnostic performance of stress perfusion ECHO, PET and CT is required to evaluate their clinical usefulness.

In conclusion, our results suggest that stress perfusion MRI is superior for the diagnosis of obstructive CAD compared to stress perfusion contrast-enhanced echocardiography and SPECT, and that echocardiography and SPECT are similar in terms of diagnostic performance.

References

1. Murray CJ, Lopez AD (1997) Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. Lancet 349:1436–1442
2. Kwok Y, Kim C, Grady D, Segal M, Redberg R (1999) Meta-analysis of exercise testing to detect coronary artery disease in women. Am J Cardiol 83:660–666
3. Noto TJ Jr, Johnson LW, Krone R et al (1991) Cardiac catheterization 1990: a report of the Registry of the Society for Cardiac Angiography and Interventions (SCA&I). Cathet Cardiovasc Diagn 24:75–83
4. Scanlon PJ, Faxon DP, Audet AM et al (1999) ACC/AHA guidelines for coronary angiography. A report of the American College of Cardiology/American Heart Association Task Force on practice guidelines (Committee on Coronary Angiography). Developed in collaboration with the Society for Cardiac Angiography and Interventions. J Am Coll Cardiol 33:1756–1824
5. Schwitter J, Wacker CM, van Rossum AC et al (2008) MR-IMPACT: comparison of perfusion-cardiac magnetic resonance with single-photon emission computed tomography for the detection of coronary artery disease in a multicentre, multivendor, randomized trial. Eur Heart J 29:480–489
6. Nandalur KR, Dwamena BA, Choudhri AF, Nandalur SR, Reddy P, Carlos RC (2008) Diagnostic performance of positron emission tomography in the detection of coronary artery disease: a meta-analysis. Acad Radiol 15:444–451
7. Nandalur KR, Dwamena BA, Choudhri AF, Nandalur MR, Carlos RC (2007) Diagnostic performance of stress cardiac magnetic resonance imaging in the detection of coronary artery disease: a meta-analysis. J Am Coll Cardiol 50:1343–1353
8. Abdelmoneim SS, Dhoble A, Bernier M et al (2009) Quantitative myocardial contrast echocardiography during pharmacological stress for diagnosis of coronary artery disease: a systematic review and meta-analysis of diagnostic accuracy studies. Eur J Echocardiogr 10:813–825
9. Underwood SR, Anagnostopoulos C, Cerqueira M et al (2004) Myocardial perfusion scintigraphy: the evidence. Eur J Nucl Med Mol Imaging 31:261–291
10. Hamon M, Fau G, Nee G, Ehtisham J, Morello R (2010) Meta-analysis of the diagnostic performance of stress perfusion cardiovascular magnetic resonance for detection of coronary artery disease. J Cardiovasc Magn Reson 12:29
11. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 339:b2535
12. Liberati A, Altman DG, Tetzlaff J et al (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 339:b2700
13. Heijenbrok-Kal MH, Fleischmann KE, Hunink MG (2007) Stress echocardiography, stress single-photon-emission computed tomography and electron beam computed tomography for the assessment of coronary artery disease: a meta-analysis of diagnostic performance. Am Heart J 154:415–423
14. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J (2003) The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol 3:25
15. Arends LR, Hamza TH, van Houwelingen JC, Heijenbrok-Kal MH, Hunink MG, Stijnen T (2008) Bivariate random effects meta-analysis of ROC curves. Med Decis Making 28:621–638
16. Knottnerus JA (1987) The effects of disease verification and referral on the relationship between symptoms and diseases. Med Decis Making 7:139–148
17. Schuetz GM, Zacharopoulou NM, Schlattmann P, Dewey M (2010) Meta-analysis: noninvasive coronary angiography using computed tomography versus magnetic resonance imaging. Ann Intern Med 152:167–177
18. Einstein AJ, Henzlova MJ, Rajagopalan S (2007) Estimating risk of cancer associated with radiation exposure from 64-slice computed tomography coronary angiography. JAMA 298:317–323
19. Jaeschke R, Guyatt GH, Sackett DL (1994) Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. JAMA 271:703–707
20. Bruder O, Schneider S, Nothnagel D et al (2009) EuroCMR (European Cardiovascular Magnetic Resonance) registry: results of the German pilot phase. J Am Coll Cardiol 54:1457–1466
21. Arnold JR, Karamitsos TD, Pegg TJ et al (2010) Adenosine stress myocardial contrast echocardiography for the detection of coronary artery disease: a comparison with coronary angiography and cardiac magnetic resonance. JACC Cardiovasc Imaging 3:934–943
22. Bernhardt P, Spiess J, Levenson B et al (2009) Combined assessment of myocardial perfusion and late gadolinium enhancement in patients after percutaneous coronary intervention or bypass grafts: a multicenter study of an integrated cardiovascular magnetic resonance protocol. JACC Cardiovasc Imaging 2:1292–1300
23. Cheng AS, Pegg TJ, Karamitsos TD et al (2007) Cardiovascular magnetic resonance perfusion imaging at 3-tesla for the detection of coronary artery disease: a comparison with 1.5-tesla. J Am Coll Cardiol 49:2440–2449
24. Cury RC, Cattani CA, Gabure LA et al (2006) Diagnostic performance of stress perfusion and delayed-enhancement MR imaging in patients with coronary artery disease. Radiology 240:39–45
25. Donati OF, Scheffel H, Stolzmann P et al (2010) Combined cardiac CT and MRI for the comprehensive workup of hemodynamically relevant coronary stenoses. AJR Am J Roentgenol 194:920–926

26. Doyle M, Fuisz A, Kortright E et al (2003) The impact of myocardial flow reserve on the detection of coronary artery disease by perfusion imaging methods: an NHLBI WISE study. J Cardiovasc Magn Reson 5:475–485
27. Gebker R, Jahnke C, Paetsch I et al (2007) MR myocardial perfusion imaging with k-space and time broad-use linear acquisition speed-up technique: feasibility study. Radiology 245:863–871
28. Gebker R, Jahnke C, Paetsch I et al (2008) Diagnostic performance of myocardial perfusion MR at 3 T in patients with coronary artery disease. Radiology 247:57–63
29. Gebker R, Jahnke C, Manka R et al (2011) High spatial resolution myocardial perfusion imaging during high dose dobutamine/atropine stress magnetic resonance using k-t SENSE. Int J Cardiol. doi:10.1016/j.ijcard.2011.01.060
30. Giang TH, Nanz D, Coulden R et al (2004) Detection of coronary artery disease by magnetic resonance myocardial perfusion imaging with various contrast medium doses: first European multicentre experience. Eur Heart J 25:1657–1665
31. Kawase Y, Nishimoto M, Hato K, Okajima K, Yoshikawa J (2004) Assessment of coronary artery disease with nicorandil stress magnetic resonance imaging. Osaka City Med J 50:87–94
32. Kitagawa K, Sakuma H, Nagata M et al (2008) Diagnostic accuracy of stress myocardial perfusion MRI and late gadolinium-enhanced MRI for detecting flow-limiting coronary artery disease: a multicenter study. Eur Radiol 18:2808–2816
33. Klein C, Gebker R, Kokocinski T et al (2008) Combined magnetic resonance coronary artery imaging, myocardial perfusion and late gadolinium enhancement in patients with suspected coronary artery disease. J Cardiovasc Magn Reson 10:45
34. Klein C, Nagel E, Gebker R et al (2009) Magnetic resonance adenosine perfusion imaging in patients after coronary artery bypass graft surgery. JACC Cardiovasc Imaging 2:437–445
35. Klem I, Heitner JF, Shah DJ et al (2006) Improved detection of coronary artery disease by stress perfusion cardiovascular magnetic resonance with the use of delayed enhancement infarction imaging. J Am Coll Cardiol 47:1630–1638
36. Klumpp BD, Seeger A, Doesch C et al (2010) High resolution myocardial magnetic resonance stress perfusion imaging at 3 T using a 1 M contrast agent. Eur Radiol 20:533–541
37. Krittayaphong R, Boonyasirinant T, Saiviroonporn P et al (2009) Myocardial perfusion cardiac magnetic resonance for the diagnosis of coronary artery disease: do we need rest images? Int J Cardiovasc Imaging 25:139–148
38. Merkle N, Wohrle J, Grebe O et al (2007) Assessment of myocardial perfusion for detection of coronary artery stenoses by steady-state, free-precession magnetic resonance first-pass imaging. Heart 93:1381–1385
39. Meyer C, Strach K, Thomas D et al (2008) High-resolution myocardial stress perfusion at 3 T in patients with suspected coronary artery disease. Eur Radiol 18:226–233
40. Nagel E, Klein C, Paetsch I et al (2003) Magnetic resonance perfusion measurements for the noninvasive detection of coronary artery disease. Circulation 108:432–437
41. Paetsch I, Jahnke C, Wahl A et al (2004) Comparison of dobutamine stress magnetic resonance, adenosine stress magnetic resonance, and adenosine stress magnetic resonance perfusion. Circulation 110:835–842
42. Pilz G, Bernhardt P, Klos M, Ali E, Wild M, Hofling B (2006) Clinical implication of adenosine-stress cardiac magnetic resonance imaging as potential gatekeeper prior to invasive examination in patients with AHA/ACC class II indication for coronary angiography. Clin Res Cardiol 95:531–538
43. Pingitore A, Lombardi M, Scattini B et al (2008) Head to head comparison between perfusion and function during accelerated high-dose dipyridamole magnetic resonance stress for the detection of coronary artery disease. Am J Cardiol 101:8–14
44. Plein S, Radjenovic A, Ridgway JP et al (2005) Coronary artery disease: myocardial perfusion MR imaging with sensitivity encoding versus conventional angiography. Radiology 235:423–430
45. Plein S, Kozerke S, Suerder D et al (2008) High spatial resolution myocardial perfusion cardiac magnetic resonance for the detection of coronary artery disease. Eur Heart J 29:2148–2155
46. Plein S, Schwitter J, Suerder D, Greenwood JP, Boesiger P, Kozerke S (2008) k-Space and time sensitivity encoding-accelerated myocardial perfusion. Radiology 249:493–500
47. Stolzmann P, Alkadhi H, Scheffel H et al (2010) Combining cardiac magnetic resonance and computed tomography coronary calcium scoring: added value for the assessment of morphological coronary disease? Int J Cardiovasc Imaging 27:969–977
48. Takase B, Nagata M, Kihara T et al (2004) Whole-heart dipyridamole stress first-pass myocardial perfusion MRI for the detection of coronary artery disease. Jpn Heart J 45:475–486
49. Aggeli C, Christoforatou E, Giannopoulos G et al (2007) The diagnostic value of adenosine stress-contrast echocardiography for diagnosis of coronary artery disease in hypertensive patients: comparison to Tl-201 single-photon emission computed tomography. Am J Hypertens 20:533–538
50. Chiou KR, Huang WC, Lin SL et al (2004) Real-time dobutamine stress myocardial contrast echocardiography for detecting coronary artery disease: correlating abnormal wall motion and disturbed perfusion. Can J Cardiol 20:1237–1243
51. Jeetley P, Hickman M, Kamp O et al (2006) Myocardial contrast echocardiography for the detection of coronary artery stenosis: a prospective multicenter study in comparison with single-photon emission computed tomography. J Am Coll Cardiol 47:141–145
52. Kowatsch I, Tsutsui JM, Osorio AF et al (2007) Head-to-head comparison of dobutamine and adenosine stress real-time myocardial perfusion echocardiography for the detection of coronary artery disease. J Am Soc Echocardiogr 20:1109–1117
53. Lipiec P, Wejner-Mik P, Krzeminska-Pakula M et al (2008) Accelerated stress real-time myocardial contrast echocardiography for the detection of coronary artery disease: comparison with 99mTc single photon emission computed tomography. J Am Soc Echocardiogr 21:941–947
54. Miszalski-Jamka T, Kuntz-Hehner S, Schmidt H et al (2009) Impact of previous myocardial infarction on the incremental value of myocardial contrast to two-dimensional supine bicycle stress echocardiography in evaluation of coronary artery disease. Int J Cardiol 136:47–55
55. Moir S, Haluska BA, Jenkins C, McNab D, Marwick TH (2005) Myocardial blood volume and perfusion reserve responses to combined dipyridamole and exercise stress: a quantitative approach to contrast stress echocardiography. J Am Soc Echocardiogr 18:1187–1193
56. Peltier M, Vancraeynest D, Pasquet A et al (2004) Assessment of the physiologic significance of coronary disease with dipyridamole real-time myocardial contrast echocardiography. Comparison with technetium-99 m sestamibi single-photon emission computed tomography and quantitative coronary angiography. J Am Coll Cardiol 43:257–264
57. Senior R, Lepper W, Pasquet A et al (2004) Myocardial perfusion assessment in patients with medium probability of coronary artery disease and no prior myocardial infarction: comparison of myocardial contrast echocardiography with 99mTc single-photon emission computed tomography. Am Heart J 147:1100–1105
58. Astarita C, Palinkas A, Nicolai E, Maresca FS, Varga A, Picano E (2001) Dipyridamole-atropine stress echocardiography versus exercise SPECT scintigraphy for detection of coronary artery disease

in hypertensives with positive exercise test. J Hypertens 19:495–502
59. Budoff MJ, Rasouli ML, Shavelle DM et al (2007) Cardiac CT angiography (CTA) and nuclear myocardial perfusion imaging (MPI)-a comparison in detecting significant coronary artery disease. Acad Radiol 14:252–257
60. Gonzalez P, Massardo T, Jofre MJ et al (2005) 201Tl myocardial SPECT detects significant coronary artery disease between 50% and 75% angiogram stenosis. Rev Esp Med Nucl 24:305–311
61. Johansen A, Hoilund-Carlsen PF, Christensen HW et al (2005) Diagnostic accuracy of myocardial perfusion imaging in a study population without post-test referral bias. J Nucl Cardiol 12:530–537
62. Schepis T, Gaemperli O, Koepfli P et al (2007) Added value of coronary artery calcium score as an adjunct to gated SPECT for the evaluation of coronary artery disease in an intermediate-risk population. J Nucl Med 48:1424–1430
63. Yao Z, Liu XJ, Shi RF et al (2000) A comparison of 99Tcm-MIBI myocardial SPET and electron beam computed tomography in the assessment of coronary artery disease in two different age groups. Nucl Med Commun 21:43–48
64. Yeih DF, Huang PJ, Ho YL (2007) Enhanced diagnosis of coronary artery disease in women by dobutamine thallium-201 ST-segment/heart rate slope and thallium-201 myocardial SPECT. J Formos Med Assoc 106:832–839