INTERVENTIONAL CARDIOLOGY AND SURGERY

A meta-analysis comparing the prognostic accuracy of six diagnostic tests for predicting perioperative cardiac risk in patients undergoing major vascular surgery

M D Kertai, E Boersma, J J Bax, M H Heijenbrok-Kal, M G M Hunink, G J L'talien, J R T C Roelandt, H van Urk, D Poldermans

Heart 2003;**89**:1327–1334

See end of article for authors' affiliations

Correspondence to: Dr Don Poldermans, Department of Vascular Surgery, Room H921, Erasmus Medical Centre, Dr Molewaterplein 40, 3015 GD Rotterdam, Netherlands; d.poldermans@erasmusmc.nl

Accepted 25 March 2003

Objective: To evaluate the discriminatory value and compare the predictive performance of six non-invasive tests used for perioperative cardiac risk stratification in patients undergoing major vascular surgery.
Design: Meta-analysis of published reports.
Methods: Eight studies on ambulatory electrocardiography, seven on exercise electrocardiography, eight on radionuclide ventriculography, 23 on myocardial perfusion scintigraphy, eight on dobutamine stress echocardiography, and four on dipyridamole stress echocardiography were selected, using a systematic review of published reports on preoperative non-invasive tests from the Medline database (January 1975 and April 2001). Random effects models were used to calculate weighted sensitivity and specificity from the published results. Summary receiver operating characteristic (SROC) curve analysis was used to evaluate and compare the prognostic accuracy of each test. The relative diagnostic odds ratio was used to study the differences in diagnostic performance of the tests.
Results: In all, 8119 patients participated in the studies selected. Dobutamine stress echocardiography had the highest weighted sensitivity of 85% (95% confidence interval (CI) 74% to 97%) and a reasonable specificity of 70% (95% CI 62% to 79%) for predicting perioperative cardiac death and non-fatal myocardial infarction. On SROC analysis, there was a trend for dobutamine stress echocardiography to perform better than the other tests, but this only reached significance against myocardial perfusion scintigraphy (relative diagnostic odds ratio 5.5, 95% CI 2.0 to 14.9).
Conclusions: On meta-analysis of six non-invasive tests, dobutamine stress echocardiography showed a positive trend towards better diagnostic performance than the other tests, but this was only significant in the comparison with myocardial perfusion scintigraphy. However, dobutamine stress echocardiography may be the favoured test in situations where there is valvar or left ventricular dysfunction.

Patients undergoing major vascular surgery are at increased risk for cardiovascular complications such as cardiac death and non-fatal myocardial infarction because of underlying coronary artery disease.[1] These complications may occur during or directly after surgery. The aim of preoperative evaluation is to identify patients with significant coronary artery disease who are thus at increased risk of cardiac complications. Appropriate patient management would then include strategies to reduce this risk.

The identification of clinical risk factors and the role of non-invasive diagnostic testing to predict perioperative cardiac risk have been evaluated over recent decades. These have included multifactorial clinical scoring systems[2–4] based on non-invasive tests such as ambulatory electrocardiography,[5–12] exercise electrocardiography,[13–19] radionuclide ventriculography,[20–27] and myocardial perfusion scintigraphy.[10 28–49] It has been suggested that the most accurate information about the individual patient risk profile can be obtained by adding clinical characteristics to those obtained by myocardial perfusion scintigraphy.[33] Recently, the use of pharmacological stress echocardiography with either dipyridamole[50–53] or dobutamine[54–61] has also been proposed for risk stratification. Pharmacological stress echocardiography has proven to be a safe and sensitive technique for predicting perioperative cardiac events, with an excellent negative predictive power.

It is still uncertain, however, which of these tests shows the best prognostic accuracy. Limited data are available directly comparing the performance of these tests. In addition, variability in the pretest probability of coronary artery disease, the mixture of surgical procedures, and differences in criteria for positivity have made it difficult to compare the performance of the tests directly.

Our aim in this study was to evaluate the comparative performance of these six diagnostic tests under conditions which adjusted for variations in preoperative risk and diagnostic thresholds, using a meta-analytic design.

METHODS

Data extraction

An electronic search of published reports was undertaken to identify studies published between January 1975 and April 2001 in English language journals. A computer generated Medline search was applied, using the terms ''cardiac evaluation'', ''cardiac risk'', ''perioperative myocardial ischaemia'', ''perioperative cardiac morbidity'', ''myocardial infarction'', ''perioperative outcome predictors'', and ''major vascular surgery'', in conjunction with one of the following non-invasive tests used for detection of myocardial ischaemia: ''exercise electrocardiography'', ''continuous ambulatory electrocardiography monitoring'', ''radionuclide ventriculography'', ''dobutamine and dipyridamole stress echocardiography'', and ''myocardial perfusion scintigraphy''. Additional references were obtained from the bibliographies of review articles and original papers.

Studies were included if perioperative (30 day) data on cardiac death and non-fatal myocardial infarction or the composite were reported, and if the absolute numbers of true positive, false negative, true negative, and false positive observations were available (including positivity thresholds), or were derivable from the data presented. If several studies were done on overlapping patient populations then one report was selected which had the largest sample size. If several tests were studied simultaneously, data from each were extracted separately. Studies in which preoperative coronary revascularisation occurred as a result of a positive test result were only included if patients who underwent such procedures could be excluded or analysed separately.

Pertinent data from the selected studies were extracted independently by two of us (MDK and EB), using standardised spreadsheets. Discrepancies were resolved by consensus. Information extracted included reference data (first author, journal, institution), publication year, number of patients, mean age, proportion of male patients, type of vascular surgery, percentage of patients with a history of coronary artery disease (defined as either past or current angina pectoris, history of myocardial infarction, or heart failure) and diabetes mellitus, type of radionuclide used, and the type of exercise performed or the type of pharmacological stress agent used. Criteria for positivity were recorded. This information is shown in table 1.

Data analysis

One hundred and thirty one studies published between January 1975 and April 2001 were screened. Fifty eight met the inclusion criteria (table 2). Data on some explanatory variables were not specified in the studies included. In seven studies[14 15 18 22 37–39] the mean age was absent; in six[14 15 24 36 37 42] the sex distribution was absent; in three[10 13 14] the proportion of patients with a history of coronary artery disease was not specified; and in 20[7 10 13–16 20–23 25–31 35 37 40] the proportion of patients with diabetes was not given. Estimates for these variables were used, based on a best subset regression analysis, so that the maximum number of selected studies could be included. Weighted mean values for the missing data using sensitivity analysis or excluding studies from the analysis did not alter the results. Therefore, all selected studies were included for analyses.

Differences in baseline clinical characteristics between the study populations were evaluated using χ^2 statistics. To account for a possible source of heterogeneity in diagnostic threshold between studies, pooled results weighted by the sample size of each study were calculated using a random effect model, based on a single treatment effect and standard error for each of a set of studies.[62] Results are presented as percentage sensitivity and specificity with 95% confidence intervals (CI).

Table 1 Test positivity criteria used for each non-invasive preoperative test

Non-invasive test	Positivity criterion
Ambulatory ECG	ST segment depression of ⩾1 mm or ST elevation ⩾2 mm after J point (measured at 60 ms) lasting at least one minute
Exercise ECG	Development of exercise induced horizontal or downsloping ST depression of 1 mm or more
Radionuclide ventriculography	Ejection fraction ⩽35%
Myocardial perfusion scintigraphy	One or more fixed or reversible thallium-201 myocardial defects
Dipyridamole stress echocardiography	New or worsening ventricular wall motion abnormalities
Dobutamine stress echocardiography	New or worsening ventricular wall motion abnormalities

Summary receiver operating characteristic (SROC) curves were generated to describe diagnostic performance over a range of threshold values for each non-invasive test (see the appendix). Univariable and multivariable regression analyses were undertaken to study the influence of clinical and study characteristics on test performance, including the number of patients tested and operated on, the mean age of the patients, the proportion of men, the proportion of patients with a history of coronary artery disease, the proportion of patients with diabetes mellitus, and the year of publication.

Comparisons using SROC analysis were also undertaken to enable us to study diagnostic performance between separate tests. In each case we included all significant explanatory variables, along with the variable indicating the test comparison. We developed models with identical explanatory variables across all comparisons. The differences in diagnostic performance between separate tests are represented by the relative diagnostic odds ratios with 95% CI. The relative diagnostic odds ratio indicates the diagnostic performance of a test, with a value larger than 1 indicating better discriminatory power, a value equal to 1 indicating no difference, and values below 1 indicating reduced discriminatory ability. In order to adjust for the fact that multiple comparisons were made, a probability value of $p \leqslant 0.01$ was considered significant. All statistical analyses were done using ''meta'' and ''metareg'' commands for STATA 6.0 for Windows (STATA Corporation, Texas, USA).

RESULTS

Clinical and study characteristics

A summary of the clinical characteristics of the studies included in the meta-analysis is given in table 3. Mean age was similar between the studies. The majority of patients in the studies were male, with no significant difference between the studies ($p = 0.380$). Coronary artery disease was more common in patients who underwent radionuclide ventriculography, ambulatory ECG, and myocardial perfusion scintigraphy ($p = 0.03$) compared with the other tests. The prevalence of diabetes was less than that of coronary artery disease, with no significant difference between the studies ($p = 0.06$). Vascular surgery was not cancelled because of a positive test result after ambulatory ECG, radionuclide ventriculography, or dobutamine stress echocardiography. Preoperative revascularisation was undertaken following a positive test result in 16 patients who underwent exercise ECG, in 70 who underwent myocardial perfusion scintigraphy, and in 12 who underwent dipyridamole stress echocardiography. Patients undergoing preoperative coronary revascularisation were excluded or analysed separately in these studies. No operations were cancelled as a result of any exercise ECG abnormalities, but in 36 cases the operation was cancelled after a positive test result during myocardial perfusion scintigraphy.

Weighted pooled results

The diagnostic test performance for individual studies is outlined in table 3. In pooled data weighted by the number of patients with and without disease in each study, dobutamine stress echocardiography showed the highest sensitivity (true positive ratio) of 85% (95% CI 74% to 97%) with a specificity (1 − false positive ratio) of 70% (95% CI 62% to 79%) compared with the other tests (table 3).

Summary receiver operation characteristic analysis for each diagnostic test

In a univariable analysis for ambulatory ECG, none of the selected clinical risk factors was a significant predictor of the

Table 2 Clinical characteristics of the studies included in the meta-analysis

Reference	Mean age (years)	Male (%)	CAD (%)	DM (%)	Isotope used	Test used	Positivity criterion	Preop revascularisation (%)	Op cancelled (%)	TP	FN	TN	FP
Ambulatory electrocardiography													
Ouyang[5]	64	71	71	29	NA	Preop monitoring	ST depression >1 mm	0	0	2	0	9	13
Raby[6]	66	70	40	21	NA	Preop monitoring	ST depression >1 mm	0	0	3	1	143	29
Pasternack[7]	70	79	41	–	NA	Preop monitoring	ST depression >1 mm	0	0	9	0	73	118
Mangano[8]	69	100	51	25	NA	Preop monitoring	ST depression >1 mm/ST elevation >2 mm	0	0	1	5	109	25
Fleischer[9]	66	61	57	34	NA	Preop monitoring	ST depression >1 mm/ST elevation >2 mm	0	0	1	2	49	15
McPhail[10]	67	86	–	–	NA	Preop monitoring	ST depression >1 mm	0	0	5	4	62	29
Kirwin[11]	73	55	66	47	NA	Preop monitoring	ST depression >1 mm	0	0	1	14	73	8
Fleischer[12]	69	50	72	43	NA	Preop monitoring	ST depression >1 mm/ST elevation >2 mm	0	0	2	2	64	18
Exerciseelectrocardiography													
Cutler[13]	62	77	–	–	NA	Treadmill/arm ergometry	ST depression >1 mm	8	0	8	1	79	32
Gardine[14]	–	–	–	–	NA	Treadmill	ST depression >1 mm	0	0	1	0	10	6
von Knorring[15]	–	–	52	–	NA	Treadmill	ST depression >1 mm	0	0	2	1	78	24
Hanson[16]	62	73	20	–	NA	Arm ergometry	ST depression >1 mm	0	0	1	0	32	41
Leppo[17]	67	62	51	18	NA	Treadmill/arm ergometry	ST depression >1.5 mm	0	0	3	4	44	9
Kaaja[18]	–	67	29	22	NA	Bicycle	ST depression >1 mm	0	0	2	0	44	2
Urbinati[19]	62	74	15	46	NA	Bicycle	ST depression >1 mm	0	0	2	0	93	48
Radionuclide ventriculography													
Fiser[20]	61	100	25	–	Tc99m	NA	EF <30%	0	0	2	0	18	0
Pasternack[21]	70	96	34	–	Tc99m	NA	EF <35%	0	0	4	4	41	1
Mosley[22]	–	90	63	–	Tc99m	NA	EF <35%	0	0	3	1	36	1
Pasternack[23]	66	80	36	–	Tc99m	NA	EF <35%	1	0	6	8	84	1
Kazmers[24]	68	–	42	7	Tc99m	NA	EF <35%	0	0	1	4	46	9
Franco[25]	68	55	61	–	Tc99m	NA	EF <35%	0	0	3	12	58	12
McCann[26]	68	87	49	–	Tc99m	NA	EF <35%	0	0	1	2	83	18
Fletcher[27]	68	79	50	–	Tc99m	NA	EF <35%	0	0	3	0	62	7
Myocardial perfusion scintigraphy													
Boucher[28]	63	96	100	–	T201	Planar	Fixed or reversible perfusion defect	11	0	3	0	20	25
Cutler[29]	65	69	34	–	T201	Planar	Fixed or reversible perfusion defect	5	8	9	0	62	30
Fletcher[30]	64	72	27	–	T201	Planar	Fixed or reversible perfusion defect	7	0	2	0	57	3
Sachs[31]	60	83	30	–	T201	Planar	Reversible perfusion defect	0	0	2	0	32	12
Lane[32]	65	65	63	100	T201	Planar	Fixed or reversible perfusion defect	0	0	9	0	20	72
Eagle[33]	66	71	29	18	T201	Planar	Reversible perfusion defect	0	0	13	2	116	69
Younis[34]	65	72	42	31	T201	Planar	Fixed or reversible perfusion defect	4	15	8	0	51	48
McEnroe[35]	69	58	28	–	T201	Planar	Fixed or reversible perfusion defect	8	0	4	2	44	37
Strawn[36]	66	–	62	15	T201	Planar	Reversible perfusion defect	13	9	1	3	18	31
Waters[37]	–	–	100	–	T201	Planar	Fixed or reversible perfusion defect	4	4	3	0	11	12
Mangano[38]	–	98	45	25	T201	Planar/SPECT	Fixed or reversible perfusion defect	0	0	2	1	19	38
Hendel[39]	–	65	35	23	T201	Planar	Reversible perfusion defect	0	0	23	5	155	144
Madsen[40]	65	65	22	–	T201	Planar	Fixed or reversible perfusion defect	0	0	5	0	20	40
McPhail[10]	67	86	–	–	Tc99m	Planar	Fixed or reversible perfusion defect	0	0	6	3	67	24
Bry[41]	66	69	30	34	T201	Planar	Fixed or reversible perfusion defect	4	0	17	0	97	114
Baron[42]	63	–	36	9	T201	SPECT	Fixed or reversible perfusion defect	0	0	26	16	187	228
Ombrellaro[43]	68	62	38	23	T201/Tc99m	Planar	Fixed or reversible perfusion defect	7	0	2	2	125	33
Erickson[45]	67	88	26	15	T201	Planar	Fixed or reversible perfusion defect	3	0	6	1	55	80
Vanzetto[46]	65	94	36	13	T201	SPECT	Fixed or reversible perfusion defect	0	0	11	1	50	72
Klonaris[47]	67	87	36	32	T201	Planar	Fixed or reversible perfusion defect	0	0	7	1	43	116

Table 2 Continued

Reference	Mean age (years)	Male (%)	CAD (%)	DM (%)	Isotope used	Test used	Positivity criterion	Preop revascularisation (%)	Op cancelled (%)	TP	FN	TN	FP
Huang[48]	68	89	12	27	T201	SPECT	Fixed or reversible perfusion defect	0	0	5	0	24	75
de Virgilio[49]	65	96	26	33	T201	Planar	Fixed or reversible perfusion defect	4	0	2	0	27	30
de Virgilio[50]	65	78	29	74	T201/Tc99m	SPECT	Fixed or reversible perfusion defect	0	0	3	1	35	41
Dipyridamole stress echocardiography													
Tischler[51]	68	61	34	29	NA	NA	New or worsening RWMA during test	0	0	4	0	100	5
Pasquet[52]	67	84	19	17	NA	NA	New or worsening RWMA during test	9	0	2	3	99	25
Rossi[53]	66	80	39	76	NA	NA	New or worsening RWMA during test	0	0	4	3	74	22
Sicari[54]	66	88	20	11	NA	NA	New or worsening RWMA during test	0	0	15	2	419	73
Dobutamine stress echocardiography													
Lalka[55]	64	78	30	18	NA	NA	New or worsening RWMA during test	0	0	8	1	21	30
Davila-Roman[56]	67	73	63	19	NA	NA	New or worsening RWMA during test	14	0	2	0	68	8
Langan[57]	69	84	31	10	NA	NA	New or worsening RWMA during test	5	6	3	0	31	40
Poldermans[58]	68	89	31	12	NA	NA	New or worsening RWMA during test	0	0	5	0	96	30
Eichelberger[59]	68	60	37	29	NA	NA	New or worsening RWMA during test	0	0	2	0	48	25
Poldermans[60]	67	85	31	11	NA	NA	New or worsening RWMA during test	0	0	17	0	228	55
Shafritz[61]	66	64	33	17	NA	NA	New or worsening RWMA during test	0	0	1	0	30	11
Boersma[62]	69	78	4 3	15	NA	NA	New or worsening RWMA during test	0	0	29	14	863	187

CAD, coronary artery disease; DM, diabetes mellitus; EF, ejection fraction; FN, false negative; FP, false positive; NA, not applicable; Op, operation; Preop, preoperative; RWMA, resting wall motion abnormality; TN, true negative; TP, true positive; -, missing values.

performance of the test. In fig 1, the SROC curve describes the test characteristics of ambulatory ECG monitoring in the studies included. When the selected clinical risk factors and study characteristics were tested in univariable analyses, none of the characteristics changed the SROC curve substantially. This was also observed when the performance of exercise ECG and dobutamine stress echocardiography was tested in separate univariable analyses—again analyses of exercise ECG and dobutamine stress echocardiography showed that none of the clinical and study characteristics changed the SROC curves (fig 1). Unlike the above described non-invasive tests, separate univariable SROC analyses for radionuclide ventriculography and myocardial perfusion scintigraphy showed that among the clinical risk factors and study characteristics only publication year changed the SROC curves significantly. The performance of the tests diminished with a later year of publication (fig 1). Remarkably, the estimates of sensitivity (true positive rates) and 1 − specificity (false positive rates) for dipyridamole stress echocardiography were inversely correlated when individual studies were plotted (fig 1). In this case only pooled sensitivity and specificity could be calculated. Thus the SROC analysis was not done and an SROC curve for dipyridamole stress echocardiography was not constructed.

Comparison of diagnostic tests

Table 4 shows the results of the comparison of SROC analyses with (table 4B) and without (table 4A) adjustment for publication year. In the present study only publication year was a significant predictor of test performance for some of the tests analysed (radionuclide ventriculography and myocardial perfusion scintigraphy). In order to compare differences in diagnostic performance between the studies, the variable ''publication year'' was also included in all comparisons.

Ambulatory ECG performed no better than exercise ECG or myocardial perfusion imaging. Although there was a trend for ambulatory ECG to have a lower predictive performance than radionuclide ventriculography ($p = 0.04$) or dobutamine stress echocardiography ($p = 0.03$), this did not reach significance in univariable analysis and after correcting for publication year. Indeed, after adjustment for publication year an inverse relation was observed between radionuclide ventriculography and ambulatory ECG, though this was not significant. Exercise ECG showed a trend for a better discriminative power than ambulatory ECG and myocardial perfusion scintigraphy, and a reduced discriminative ability compared with radionuclide ventriculography and dobutamine stress echocardiography. However, these differences were non-significant (table 4). Myocardial perfusion scintigraphy showed lower discriminatory ability than ambulatory ECG, exercise ECG, or radionuclide ventriculography, though the differences were also non-significant. However, myocardial perfusion scintigraphy had a substantially lower discriminatory power than dobutamine stress echocardiography ($p = 0.001$), and this difference remained significant after adjusting for publication year ($p = 0.002$). Finally, dobutamine stress echocardiography showed a positive trend towards a better diagnostic performance than all the other tests, but this only reached significance in comparison with myocardial perfusion scintigraphy.

DISCUSSION

Our report is a meta-analysis of contemporary papers on six non-invasive tests used for preoperative risk stratification in patients selected for vascular surgery. We used an innovative meta-analytic method to estimate the diagnostic accuracy of these tests from multiple studies. The accuracy of the tests is presented and compared using a summary ROC curve, and

Table 3 Summary of clinical characteristics and sensitivity and specificity of the studies included in the meta-analysis

Type of test	No. of studies	No. of patients	Mean age (years)	Proportion of men (%)	History of CAD (%)	Proportion of DM (%)	Sensitivity (%; 95% CI)	Specificity (%; 95% CI)
Radionuclide ventriculography	8	532	67.0	83	45	25	50 (32 to 69)	91 (87 to 96)
Ambulatory electrocardiography	8	893	68.0	72	55	32	52 (21 to 84)	70 (57 to 83)
Exercise electrocardiography	7	685	64.5	72	36	28	74 (60 to 88)	69 (60 to 78)
Dipyridamole stress echocardiography	4	850	66.8	78	28	33	74 (53 to 94)	86 (80 to 93)
Myocardial perfusion scintigraphy	23	3119	65.5	78	40	30	83 (77 to 89)	49 (41 to 57)
Dobutamine stress echocardiography	8	1877	67.3	76	37	16	85 (74 to 97)	70 (62 to 79)

Tests are sorted according to ascending sensitivities.
CAD, coronary artery disease; CI, confidence interval; DM, diabetes mellitus.

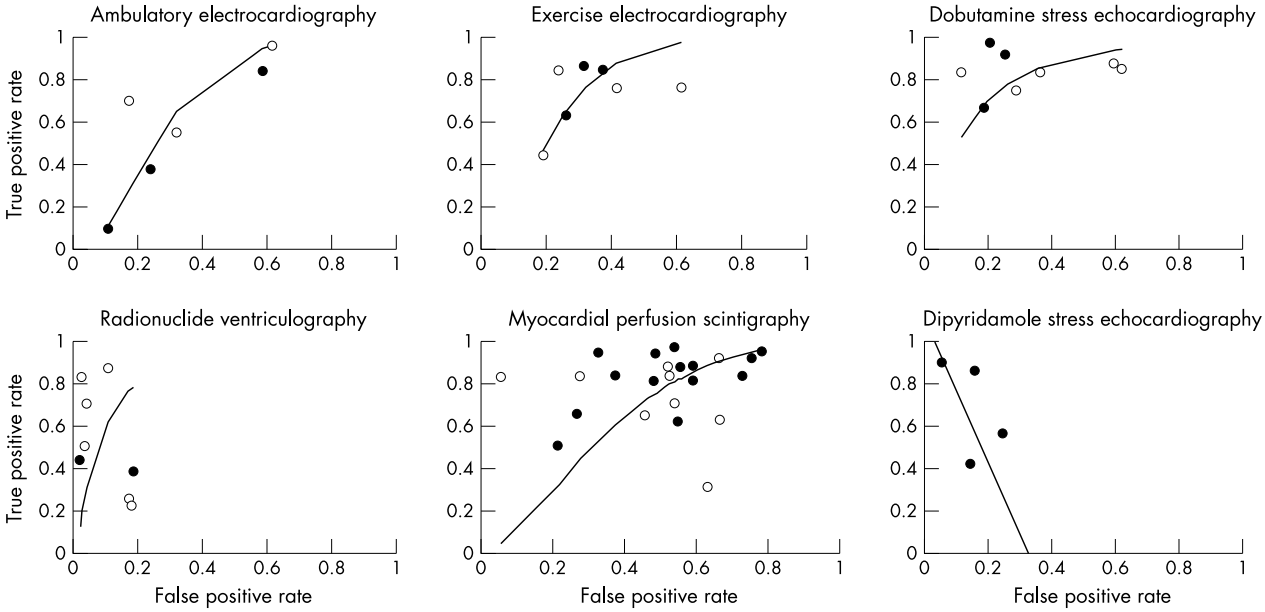

Figure 1 Graphs showing summary receiver operating characteristic (SROC) curves for ambulatory ECG, exercise ECG, dobutamine stress echocardiography, radionuclide ventriculography, and myocardial perfusion scintigraphy. The horizontal axis represents the false positive rate (1 − specificity) and the vertical axis the true positive rate (sensitivity). The graph for dipyridamole stress echocardiography represents a plot of the estimates of the true positive rate for dipyridamole stress echocardiography against the estimates of false positive rate, with a solid line representing the regression line. In all graphs, solid circles represent studies with more than 100 patients and open circles represent studies with less than 100 patients.

Table 4 Results of the comparison of summary receiver operating characteristic analyses between diagnostic tests

	Reference test				
	Ambulatory electrocardiography	Exercise electrocardiography	Radionuclide ventriculography	Myocardial perfusion scintigraphy	Dobutamine stress echocardiography
(A) Not adjusted for publication year					
Ambulatory electrocardiography		0.6 (0.2 to 1.8)	0.2 (0.0 to 1.0)	1.6 (0.5 to 5.0)	0.3 (0.1 to 1.0)
Exercise electrocardiography	1.6 (0.5 to 24.5)		0.5 (0.0 to 6.1)	2.7 (0.3 to 8.2)	0.6 (0.2 to 1.8)
Radionuclide ventriculography	5.5 (1.1 to 24.5)	2.2 (0.2 to 30.0)		5.5 (0.8 to 36.6)	0.9 (0.1 to 18.2)
Myocardial perfusion scintigraphy	0.6 (0.2 to 1.8)	0.4 (0.1 to 30.0)	0.2 (0.0 to 1.2)		0.3 (0.1 to 0.6)†
Dobutamine stress echocardiography	3.0 (1.2 to 7.4)	1.6 (0.5 to 4.5)	1.1 (0.1 to 20.1)	4.1 (1.6 to 10.0)†	
(B) Adjusted for publication year					
Ambulatory electrocardiography		0.7 (0.4 to 3.0)	1.3 (0.2 to 8.2)	1.5 (0.5 to 4.1)	0.4 (0.1 to 2.5)
Exercise electrocardiography	1.5 (0.3 to 6.7)		0.4 (0.0 to 4.1)	1.5 (0.4 to 5.5)	0.7 (0.1 to 6.7)
Radionuclide ventriculography	0.7 (0.1 to 4.5)	2.7 (0.2 to 30.0)		2.2 (0.3 to 13.5)	0.08 (0.0 to 5.5)
Myocardial perfusion scintigraphy	0.7 (0.3 to 2.0)	0.7 (0.2 to 2.5)	0.5 (0.1 to 3.0)		0.2 (0.1 to 0.5)†
Dobutamine stress echocardiography	2.5 (0.4 to 16.4)	1.5 (0.2 to 14.9)	12.2 (0.2 to 897.8)	5.5 (2.0 to 14.9)†	

The figures indicate relative diagnostic odds ratios for comparison between the reference test in the column *v* the test in the row; the relative diagnostic odds ratio indicates the diagnostic performance of a test, with a value larger than 1 indicating better discriminatory power, a value equal to 1 no difference, and values below 1 corresponding to reduced discriminatory ability. Figures in parenthesis are the 95% confidence intervals.
†$p<0.01$.

the performance of individual tests was corrected for selected patient and study characteristics. Our results show that pharmacological stress tests have a higher overall sensitivity and specificity than the other tests. In particular, dobutamine stress echocardiography showed a positive trend for better diagnostic performance than the other tests, but this was only significant in comparison with myocardial perfusion scintigraphy. Ambulatory ECG, exercise ECG, and radionuclide ventriculography yielded a lower sensitivity and reasonable specificity, but no significant difference in predictive performance.

Ambulatory ECG monitoring showed low sensitivity and higher specificity in the present study, but no significant difference in predictive performance. The use of ambulatory ECG monitoring for perioperative cardiac risk assessment was first described by Raby and colleagues.[6] They reported a sensitivity of 75% and a specificity of 83% for predicting cardiac death and non-fatal myocardial infarction. In later studies the predictive value of the test was corroborated, but the sensitivity was less than reported here.[9 10] Variation in end point composition, surgical procedures, and the timing of events could explain the observed differences in sensitivity and specificity between studies. The advantages of the test are that it is cheap and widely available. However, the presence of resting ECG changes (bundle branch block, left ventricular hypertrophy, digitalis use) may preclude reliable ST segment analysis in 40% of patients.[63] Hence, the combination of low sensitivity and resting ECG changes limits the application of the technique.

In our study the pooled data showed reasonable sensitivity and specificity for exercise ECG compared with the other tests. However, SROC analyses showed that it did not perform better than the other non-invasive tests. This observation may be explained by type II error—that is, differences could be missed owing to lack of power. Conventional exercise ECG is considered the most physiological form of stress.[64] However, the test may not always be feasible because of exercise intolerance in such patients.

The value of radionuclide ventriculography for predicting perioperative cardiac complications has been assessed in many studies.[20–27] A preoperatively assessed low ejection fraction (<35%) showed a relatively poor sensitivity but a high specificity in our meta-analysis. The SROC analysis showed that resting left ventricular dysfunction, as determined by radionuclide ventriculography, did not have a significantly better predictive performance than the other tests. The observed limitation of the predictive performance of this test may be explained by the failure to detect severe underlying coronary artery disease, changes in predictive value over time, and improved anaesthetic and surgical perioperative care. Thus, radionuclide ventriculography may not be a suitable test for preoperative risk stratification. However, a low ejection fraction may be a useful predictor of long term survival.[64]

In patients unable to perform adequate exercise (and in most vascular surgery patients unable to exercise), a non-exercising test is mandatory. In this regard, myocardial perfusion scintigraphy—often combined with clinical risk assessment—is the most extensively studied non-invasive approach to cardiac risk stratification. In the present study a high sensitivity but a low specificity was observed for perioperative hard cardiac events, with a lower diagnostic accuracy than with dobutamine stress echocardiography. The earliest studies (between 1985 and 1987) showed that patients with fixed or reversible scintigraphic myocardial perfusion defects were at increased cardiac risk.[28 29] However, these results were later questioned by other investigators.[38 42] The poor prognostic value observed in these studies may reflect small sample sizes.

Comparison of myocardial perfusion scintigraphy with stress echocardiography in our study showed that dobutamine stress echocardiography performed significantly better. There are several possible explanations for these findings. First, myocardial perfusion scintigraphy is more widely used in consecutive patients presenting for vascular surgery than in selected patients with clinical risk factors; second, unblinded test results are available to clinicians, thus influencing perioperative care; third, repeat imaging 3–4 hours after thallium infusion may not allow sufficient time for thallium redistribution; and finally, thallium uptake may be uniformly restricted in patients with severe and diffuse coronary artery disease. Nevertheless myocardial perfusion scintigraphy is a valuable test for cardiac risk assessment, especially in patients with contraindications to stress echocardiography, with a reported complication rate of 3.9% in the studies included in the present meta-analysis. However, myocardial perfusion scintigraphy should be avoided in patients with significant bronchospasm, critical carotid disease, or on regular theophylline treatment.[64]

Dipyridamole stress echocardiography has been proposed for cardiac risk stratification in patients undergoing major vascular surgery.[50–53] The low false positive rates and extremely high negative predictive values can make this test a useful predictor in low risk scenarios. Indeed, in the present meta-analysis dipyridamole stress echocardiography had a higher specificity than myocardial perfusion scintigraphy or dobutamine stress echocardiography. However, there have been limited numbers of studies published to date, and in these studies the true positive and false positive rates for dipyridamole stress echocardiography showed a negative correlation when they were plotted graphically. Hence only weighted pooled sensitivity and specificity could be calculated and the performance of the test could not be studied in further detail. The high specificity of dipyridamole stress echocardiography in clinical practice may indicate that the test can identify patients at less severe ischaemic responses. These patients may not need to undergo further testing and can proceed to major vascular surgery directly. However, in clinical practice it is more valuable to have a sensitive test that can identify patients at increased risk of perioperative cardiac events. Further conclusions about dipyridamole stress echocardiography are limited by the reported differences in sensitivity and specificity between the few studies reported to date.

Dobutamine stress echocardiography showed a similar high sensitivity, but a significantly higher specificity, compared with myocardial perfusion scintigraphy. Comparison using SROC analysis showed a trend towards better performance than the other tests, but this was only significant versus myocardial perfusion scintigraphy. As with dipyridamole, dobutamine stress echocardiography has been proposed for cardiac risk stratification owing to its high negative predictive value.[54–61] Moreover, in the present study this test showed significantly higher sensitivity than dipyridamole stress echocardiography. The role of dobutamine stress echocardiography in cardiac risk stratification has been studied extensively in recent reports.[59 61 65] The results of these studies suggest that the investigation can be done safely and with reasonable patient tolerance and may provide additional information about valvar dysfunction, in contrast to myocardial perfusion scintigraphy. The test has certain limitations—for example, it should not be used in patients with serious arrhythmias, severe hypertension, or hypotension. However, in a recent report of the use of dobutamine stress echocardiography in 6595 patients it was found that the incidence of cardiac arrhythmias and hypotension was 8% and 3%, respectively, and these complications were well tolerated and rarely required treatment.[66]

Study limitations

The study has several limitations. Meta-analyses are subject to publication bias—that is, studies with a significant result are more likely to be submitted. Heterogeneity of study design is another aspect that may influence the interpretation of the results. The predictive value of a given test can be influenced by at least two important factors: patient selection and the blinding of test results. A test done in consecutive patients may have a lower diagnostic accuracy than the same test done in a selected group of patients. This is commonly referred to as selection bias and it occurs when consecutive cohorts of patients with a high prior probability of coronary artery disease, and therefore of adverse cardiac outcome after surgery, are more likely to get the test. The predictive value of a given test may also be influenced by the availability of the test result to the treating physicians. In this case patients with positive test results may undergo less invasive operations or receive better perioperative care, such as cardioprotective drugs, invasive haemodynamic monitoring, and admission to an intensive care unit. Studies in which patients are enrolled consecutively and treating physicians are blinded to test results are more likely to provide a relatively unbiased estimate for a given test.

Conclusions

Our meta-analysis shows that ambulatory ECG monitoring has a relatively low sensitivity and a low specificity, with no incremental value over the other tests. Furthermore, resting ECG changes frequently preclude reliable assessment of the ambulatory ECG and this test is therefore not recommended for perioperative risk assessment. Radionuclide ventriculography had the highest specificity but a relatively low sensitivity, with a limited predictive performance for perioperative events. This test should not therefore be considered as a tool for preoperative cardiac risk assessment. The test of choice in most ambulatory patients is an exercise ECG done according to the American College of Cardiology/American Heart Association guideline.[64] However, most vascular surgical candidates have important abnormalities on their resting ECG and are unable to perform adequate exercise. In such patients stress echocardiography or myocardial perfusion scintigraphy should be considered. In the current study dobutamine stress echocardiography showed a similar sensitivity to myocardial perfusion scintigraphy but a higher specificity and a better overall predictive performance. Moreover, dobutamine stress echocardiography is the favoured test if there is an additional question of valvar or left ventricular dysfunction. Dipyridamole stress echocardiography had a lower sensitivity and a higher specificity than myocardial perfusion scintigraphy. However, further conclusions about dipyridamole stress echocardiography are limited by the reported differences between studies and the limited number of studies reported to date.

Meta-analysis of six non-invasive tests showed a positive trend for dobutamine stress echocardiography to have a better diagnostic performance than the other tests, but this only reached significant difference in comparison with myocardial perfusion scintigraphy.

To see the appendix, visit the Heart website—www.heartjnl.com/supplemental

Authors' affiliations

M D Kertai, E Boersma, J J Bax, J R T C Roelandt, Department of Cardiology, Erasmus Medical Centre Rotterdam, Rotterdam, Netherlands
M H Heijenbrok-Kal, M G M Hunink, Departments of Epidemiology and Statistics, Erasmus Medical Centre Rotterdam
H van Urk, D Poldermans, Department of Vascular Surgery, Erasmus Medical Centre Rotterdam
G J L' Italien, Bristol-Myers Squibb, Pharmaceutical Research Institute, Wallingford, Connecticut, USA

REFERENCES

1 **Mangano DT**. Perioperative cardiac morbidity. *Anesthesiology* 1990;**72**:153–84.
2 **Goldman L**, Caldera DL, Nussbaum SR, *et al*. Multifactorial index of cardiac risk in noncardiac surgical procedures. *N Engl J Med* 1977;**297**:845–50.
3 **Detsky AS**, Abrams HB, McLaughlin JR, *et al*. Predicting cardiac complications in patients undergoing non-cardiac surgery. *J Gen Intern Med* 1986;**1**:211–19.
4 **L'Italien GJ**, Paul SD, Hendel RC, *et al*. Development and validation of a Bayesian model for perioperative cardiac risk assessment in a cohort of 1081 vascular surgical candidates. *J Am Coll Cardiol* 1996;**27**:779–86.
5 **Ouyang P**, Gerstenblith G, Furman WR, *et al*. Frequency and significance of early postoperative silent myocardial ischemia in patients having peripheral vascular surgery. *Am J Cardiol* 1989;**64**:1113–16.
6 **Raby KE**, Goldman L, Creager MA, *et al*. Correlation between preoperative ischemia and major cardiac events after peripheral vascular surgery [see comments]. *N Engl J Med* 1989;**321**:1296–300.
7 **Pasternack PF**, Grossi EA, Baumann FG, *et al*. The value of silent myocardial ischemia monitoring in the prediction of perioperative myocardial infarction in patients undergoing peripheral vascular surgery. *J Vasc Surg* 1989;**10**:617–25.
8 **Mangano DT**, Browner WS, Hollenberg M, *et al*. Association of perioperative myocardial ischemia with cardiac morbidity and mortality in men undergoing noncardiac surgery. The study of perioperative ischemia research group. *N Engl J Med* 1990;**323**:1781–8.
9 **Fleisher LA**, Rosenbaum SH, Nelson AH, *et al*. The predictive value of preoperative silent ischemia for postoperative ischemic cardiac events in vascular and nonvascular surgery patients. *Am Heart J* 1991;**122**:980–6.
10 **McPhail NV**, Ruddy TD, Barber GG, *et al*. Cardiac risk stratification using dipyridamole myocardial perfusion imaging and ambulatory ECG monitoring prior to vascular surgery. *Eur J Vasc Surg* 1993;**7**:151–5.
11 **Kirwin JD**, Ascer E, Gennaro M, *et al*. Silent myocardial ischemia is not predictive of myocardial infarction in peripheral vascular surgery patients. *Ann Vasc Surg* 1993;**7**:27–32.
12 **Fleisher LA**, Rosenbaum SH, Nelson AH, *et al*. Preoperative dipyridamole thallium imaging and ambulatory electrocardiographic monitoring as a predictor of perioperative cardiac events and long-term outcome. *Anesthesiology* 1995;**83**:906–17.
13 **Cutler BS**, Wheeler HB, Paraskos JA, *et al*. Applicability and interpretation of electrocardiographic stress testing in patients with peripheral vascular disease. *Am J Surg* 1981;**141**:501–6.
14 **Gardine RL**, McBride K, Greenberg H, *et al*. The value of cardiac monitoring during peripheral arterial stress testing in the surgical management of peripheral vascular disease. *J Cardiovasc Surg (Torino)* 1985;**26**:258–61.
15 **von Knorring J**, Lepantalo M. Prediction of perioperative cardiac complications by electrocardiographic monitoring during treadmill exercise testing before peripheral vascular surgery. *Surgery* 1986;**99**:610–13.
16 **Hanson P**, Pease M, Berkoff H, *et al*. Arm exercise testing for coronary artery disease in patients with peripheral vascular disease. *Clin Cardiol* 1988;**11**:70–4.
17 **Leppo J**, Plaja J, Gionet M, *et al*. Noninvasive evaluation of cardiac risk before elective vascular surgery. *J Am Coll Cardiol* 1987;**9**:269–76.
18 **Kaaja R**, Sell H, Erkola O, *et al*. Predictive value of manual ECG-monitored exercise test before abdominal aortic or peripheral vascular surgery. *Angiology* 1993;**44**:11–15.
19 **Urbinati S**, Di Pasquale G, Andreoli A, *et al*. Preoperative noninvasive coronary risk stratification in candidates for carotid endarterectomy. *Stroke* 1994;**25**:2022–7.
20 **Fiser WP**, Thompson BW, Thompson AR, *et al*. Nuclear cardiac ejection fraction and cardiac index in abdominal aortic surgery. *Surgery* 1983;**94**:736–9.
21 **Pasternack PF**, Imparato AM, Bear G, *et al*. The value of radionuclide angiography as a predictor of perioperative myocardial infarction in patients undergoing abdominal aortic aneurysm resection. *J Vasc Surg* 1984;**1**:320–5.
22 **Mosley JG**, Clarke JM, Ell PJ, *et al*. Assessment of myocardial function before aortic surgery by radionuclide angiocardiography. *Br J Surg* 1985;**72**:886–7.
23 **Pasternack PF**, Imparato AM, Riles TS, *et al*. The value of the radionuclide angiogram in the prediction of perioperative myocardial infarction in patients undergoing lower extremity revascularization procedures. *Circulation* 1985;**72**:II13–17.
24 **Kazmers A**, Cerqueira MD, Zierler RE. The role of preoperative radionuclide left ventricular ejection fraction for risk assessment in carotid surgery. *Arch Surg* 1988;**123**:416–9.
25 **Franco CD**, Goldsmith J, Veith FJ, *et al*. Resting gated pool ejection fraction: a poor predictor of perioperative myocardial infarction in patients undergoing vascular surgery for infrainguinal bypass grafting. *J Vasc Surg* 1989;**10**:656–61.
26 **McCann RL**, Wolfe WG. Resection of abdominal aortic aneurysm in patients with low ejection fractions. *J Vasc Surg* 1989;**10**:240–4.

27 **Fletcher JP**, Antico VF, Gruenewald S, *et al*. Risk of aortic aneurysm surgery as assessed by preoperative gated heart pool scan. *Br J Surg* 1989;**76**:26–8.
28 **Boucher CA**, Brewster DC, Darling RC, *et al*. Determination of cardiac risk by dipyridamole-thallium imaging before peripheral vascular surgery. *N Engl J Med* 1985;**312**:389–94.
29 **Cutler BS**, Leppo JA. Dipyridamole thallium 201 scintigraphy to detect coronary artery disease before abdominal aortic surgery. *J Vasc Surg* 1987;**5**:91–100.
30 **Fletcher JP**, Antico VF, Gruenewald S, *et al*. Dipyridamole-thallium scan for screening of coronary artery disease prior to vascular surgery. *J Cardiovasc Surg (Torino)* 1988;**29**:666–9.
31 **Sachs RN**, Tellier P, Larmignat P, *et al*. Assessment by dipyridamole-thallium-201 myocardial scintigraphy of coronary risk before peripheral vascular surgery. *Surgery* 1988;**103**:584–7.
32 **Lane SE**, Lewis SM, Pippin JJ, *et al*. Predictive value of quantitative dipyridamole-thallium scintigraphy in assessing cardiovascular risk after vascular surgery in diabetes mellitus. *Am J Cardiol* 1989;**64**:1275–9.
33 **Eagle KA**, Coley CM, Newell JB, *et al*. Combining clinical and thallium data optimizes preoperative assessment of cardiac risk before major vascular surgery. *Ann Intern Med* 1989;**110**:859–66.
34 **Younis LT**, Aguirre F, Byers S, *et al*. Perioperative and long-term prognostic value of intravenous dipyridamole thallium scintigraphy in patients with peripheral vascular disease. *Am Heart J* 1990;**119**:1287–92.
35 **McEnroe CS**, O'Donnell TFJ, Yeager A, *et al*. Comparison of ejection fraction and Goldman risk factor analysis to dipyridamole-thallium 201 studies in the evaluation of cardiac morbidity after aortic aneurysm surgery. *J Vasc Surg* 1990;**11**:497–504.
36 **Strawn DJ**, Guernsey JM. Dipyridamole thallium scanning in the evaluation of coronary artery disease in elective abdominal aortic surgery. *Arch Surg* 1991;**126**:880–4.
37 **Watters TA**, Botvinick EH, Dae MW, *et al*. Comparison of the findings on preoperative dipyridamole perfusion scintigraphy and intraoperative transesophageal echocardiography: implications regarding the identification of myocardium at ischemic risk. *J Am Coll Cardiol* 1991;**18**:93–100.
38 **Mangano DT**, London MJ, Tubau JF, *et al*. Dipyridamole thallium-201 scintigraphy as a preoperative screening test. A reexamination of its predictive potential.Study of Perioperative Ischemia Research Group. *Circulation* 1991;**84**:493–502.
39 **Hendel RC**, Whitfield SS, Villegas BJ, *et al*. Prediction of late cardiac events by dipyridamole thallium imaging in patients undergoing elective vascular surgery. *Am J Cardiol* 1992;**70**:1243–9.
40 **Madsen PV**, Vissing M, Munck O, *et al*. A comparison of dipyridamole thallium 201 scintigraphy and clinical examination in the determination of cardiac risk before arterial reconstruction. *Angiology* 1992;**43**:306–11.
41 **Bry JD**, Belkin M, O'Donnell TFJ, *et al*. An assessment of the positive predictive value and cost–effectiveness of dipyridamole myocardial scintigraphy in patients undergoing vascular surgery. *J Vasc Surg* 1994;**19**:112–21.
42 **Baron JF**, Mundler O, Bertrand M, *et al*. Dipyridamole-thallium scintigraphy and gated radionuclide angiography to assess cardiac risk before abdominal aortic surgery. *N Engl J Med* 1994;**330**:663–9.
43 **Ombrellaro MP**, Dieter RA, Freeman M, *et al*. Role of dipyridamole myocardial scintigraphy in carotid artery surgery. *J Am Coll Surg* 1995;**181**:451–8.
44 **Erickson CA**, Carballo RE, Freischlag JA, *et al*. Using dipyridamole-thallium imaging to reduce cardiac risk in aortic reconstruction. *J Surg Res* 1996;**60**:422–8.
45 **Vanzetto G**, Machecourt J, Blendea D, *et al*. Additive value of thallium single-photon emission computed tomography myocardial imaging for prediction of perioperative events in clinically selected high cardiac risk patients having abdominal aortic surgery. *Am J Cardiol* 1996;**77**:143–8.
46 **Klonaris CN**, Bastounis EA, Xiromeritis NC, *et al*. The predictive value of dipyridamole-thallium scintigraphy for cardiac risk assessment before major vascular surgery. *Int Angiol* 1998;**17**:171–8.
47 **Huang Z**, Komori S, Sawanobori T, *et al*. Dipyridamole thallium-201 single-photon emission computed tomography for prediction of perioperative cardiac events in patients with arteriosclerosis obliterans undergoing vascular surgery. *Jpn Circ J* 1998;**62**:274–8.
48 **de Virgilio C**, Kirby L, Lewis RJ, *et al*. Limited utility of dipyridamole-thallium for predicting adverse cardiac events after vascular surgery. *Vasc Surg* 1998;**32**:275–9.
49 **de Virgilio C**, Toosie K, Elbassir M, *et al*. Dipyridamole-thallium/sestamibi before vascular surgery: a prospective blinded study in moderate-risk patients. *J Vasc Surg* 2000;**32**:77–89.
50 **Tischler MD**, Lee TH, Hirsch AT, *et al*. Prediction of major cardiac events after peripheral vascular surgery using dipyridamole echocardiography. *Am J Cardiol* 1991;**68**:593–7.
51 **Pasquet A**, D'Hondt AM, Verhelst R, *et al*. Comparison of dipyridamole stress echocardiography and perfusion scintigraphy for cardiac risk stratification in vascular surgery patients. *Am J Cardiol* 1998;**82**:1468–74.
52 **Rossi E**, Citterio F, Vescio MF, *et al*. Risk stratification of patients undergoing peripheral vascular revascularization by combined resting and dipyridamole echocardiography. *Am J Cardiol* 1998;**82**:306–10.
53 **Sicari R**, Ripoli A, Picano E, *et al*. Perioperative prognostic value of dipyridamole echocardiography in vascular surgery: a large-scale multicenter study in 509 patients. EPIC (Echo Persantine International Cooperative) Study Group. *Circulation* 1999;**100**:II269–74.
54 **Lalka SG**, Sawada SG, Dalsing MC, *et al*. Dobutamine stress echocardiography as a predictor of cardiac events associated with aortic surgery. *J Vasc Surg* 1992;**15**:831–40.
55 **Davila-Roman VG**, Waggoner AD, Sicard GA, *et al*. Dobutamine stress echocardiography predicts surgical outcome in patients with an aortic aneurysm and peripheral vascular disease. *J Am Coll Cardiol* 1993;**21**:957–63.
56 **Langan EM**, Youkey JR, Franklin DP, *et al*. Dobutamine stress echocardiography for cardiac risk assessment before aortic surgery. *J Vasc Surg* 1993;**18**:905–11.
57 **Poldermans D**, Fioretti PM, Forster T, *et al*. Dobutamine stress echocardiography for assessment of perioperative cardiac risk in patients undergoing major vascular surgery. *Circulation* 1993;**87**:1506–12.
58 **Eichelberger JP**, Schwarz KQ, Black ER, *et al*. Predictive value of dobutamine echocardiography just before noncardiac vascular surgery. *Am J Cardiol* 1993;**72**:602–7.
59 **Poldermans D**, Arnese M, Fioretti PM, *et al*. Improved cardiac risk stratification in major vascular surgery with dobutamine-atropine stress echocardiography. *J Am Coll Cardiol* 1995;**26**:648–53.
60 **Shafritz R**, Ciocca RG, Gosin JS, *et al*. The utility of dobutamine echocardiography in preoperative evaluation for elective aortic surgery. *Am J Surg* 1997;**174**:121–5.
61 **Boersma E**, Poldermans D, Bax JJ, *et al*. The combination of clinical characteristics, dobutamine echocardiography and beta-blocker use optimizes perioperative management of patients scheduled for major vascular surgery. *JAMA* 2001;**285**:1865–73.
62 **DerSimonian R**, Laird N. Meta-analysis in clinical trials. *Controlled Clin Trials* 1986;**7**:177–88.
63 **Rose EL**, Liu XJ, Henley M, *et al*. Prognostic value of noninvasive cardiac tests in the assessment of patients with peripheral vascular disease. *Am J Cardiol* 1993;**71**:40–4.
64 **Eagle KA**, Berger PB, Calkins H, *et al*. ACC/AHA guideline update for perioperative cardiovascular evaluation for noncardiac surgery: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (committee to update the 1996 guidelines on perioperative cardiovascular evaluation for noncardiac surgery). *J Am Coll Cardiol* 2002;**39**:542–53.
65 **Poldermans D**, Bax JJ, Thomson IR, *et al*. Role of dobutamine stress echocardiography for preoperative cardiac risk assessment before major vascular surgery: a diagnostic tool comes of age. *Echocardiography* 2000;**17**:79–91.
66 **Poldermans D**, Rambaldi R, Bax JJ, *et al*. Safety and utility of atropine addition during dobutamine stress echocardiography for the assessment of viable myocardium in patients with severe left ventricular dysfunction. *Eur Heart J* 1998;**19**:1712–18.

In appendix

67 **Hanley JA**, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;**143**:29–36.
68 **Littenberg B**, Moses LE. Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method. *Med Decis Making* 1993;**13**:313–21.
69 **de Vries SO**, Hunink MG, Polak JF. Summary receiver operating characteristic curves as a technique for meta-analysis of the diagnostic performance of duplex ultrasonography in peripheral arterial disease. *Acad Radiol* 1996;**3**:361–9.
70 **Lijmer JG**, Mol BW, Heisterkamp S, *et al*. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999;**282**:1061–6.
71 **Midgette AS**, Stukel TA, Littenberg B. A meta-analytic method for summarizing diagnostic test performances: receiver-operating-characteristic-summary point estimates. *Med Decis Making* 1993;**13**:253–7.

A meta-analysis comparing the prognostic accuracy of six diagnostic tests for prediction perioperative cardiac risk in patients undergoing major vascular surgery

Appendix

Summary receiver characteristics curves

Conventional receiver operating characteristic curves (ROC) graphically represent the true positive and false positive rates for a diagnostic test when the threshold for a positive test result is varied.[67] The area under the curve summarises the overall diagnostic performance of the test, with larger areas corresponding to a more discriminating test. Most reports of the performance of a diagnostic test provide only single estimates of the true positive (TP) and false positive (FP) rates for one fixed threshold value. A possible source of heterogeneity in a meta-analysis of a diagnostic test is the variability of threshold values across studies. Variability in TP and FP rates derived from published reports may represent different operating points on one common underlying ROC curve. Hence, pooling TP and FP rates from different sources would underestimate the diagnostic performance.[68]

Summary ROC (SROC) analysis assumes that part of the reported variability in test performance reflects the difference in the cut off points or positivity criteria used, and aims to adjust for these differences. SROC analysis also allows adjustment for important clinical covariates and for comparison between different types of tests.[69] In this case, a variable corresponding to the type of non-invasive testing used is entered into the regression equation. The differences in diagnostic performance of the tests are represented by the regression coefficients and can be interpreted after antilogarithm transformation as relative diagnostic odds ratios. They indicate the diagnostic performance of a test, with a value larger than 1 indicating better discriminatory power, whereas a value equal to 1 indicates no difference, and values below 1 indicate reduced discriminatory ability.[70]

First, the estimates of sensitivity (true positive rate) against the estimates of 1specificity (false positive rate) were plotted. If there was a positive relation between true positive rates and false positive rates (that is, both true positive and false positive rates increase), then this association was identified using the non-parametric Spearmen correlation test.[71] If the true positive rates and false positive rates were positively correlated, then summary ROC analysis was undertaken. In case of a negative correlation only summary point estimates that is, weighted

sensitivity and specificity of the test were given.

Construction of a summary receiver characteristics curve

The construction of a SROC curve requires that the number of true positive (TP), false positive (FP), true negative (TN), and false negative (FN) observations in each study be available.[69]

Computational steps involved in the construction of summary ROC curve

Computation of the difference (D) and the sum (S) of the logit transforms of TP and FP rates

$$D = \ln((TP+1/2)/(FN+1/2)) \ln((FP+1/2)/(TN+1/2))$$

$$S = \ln((TP+1/2)/(FN+1/2))+\ln((FP+1/2)/(TN+1/2))$$

where TP, FN, TN, and FP are corrected by one half to ensure that D and S would not be undefined if TP, FN, TN, or FP equals zero.

Computation of the asymptotic variance (VAR) of D

$$VAR (D) = ((1/TP+1.2)+(1/FN+1/2)+(1/TN+1/2)+(1/FP+1/2)$$

Weighted least square regression analysis, using weights proportional to the inverse of VAR (D)

In our study, the following equation was used to evaluate the prognostic accuracy of individual tests:

$$D = \alpha + \beta_s*S + \varepsilon$$

where D is the dependent variable from equation 3, α, β_s, are regression coefficients, S is the independent variable for equation 4, and ε is the error term.

The following equation was used for between-test comparisons to compare the prognostic accuracy of two tests:

$$D = \alpha + \beta_s*S + \beta_t*T + \varepsilon$$

where D is the dependent variable from equation 3, α, β_s, β_t are regression coefficients, S is the independent variable for equation 4, and T is a dummy variable indicating the type of test for between-test comparison (with T=0 for one test and T=1 for the other).

The regression lines were converted to the ROC space using the following formula:

$$TPRp = [1+e^{(intercept)/(1-slope)}*(1-FPR/FPR)^{(1+slope)/(1-slope)}]^{-1}$$

where TRPp is the predicted value of the TP rate for a given FP rate (FPR), the intercepts represents intercept of the regression line in the (S, D) space, and slope is the slope of this line.