Level in hierarchy of evidence based on Merlin *et al.*:[57]

1. Systematic review of level 2 studies.

2. Study of test accuracy and methodology, including an independent, blinded comparison with a valid reference standard, conducted among consecutive persons with a defined clinical presentation.

3a. Study of test accuracy, with an independent, blinded comparison with a valid reference standard, conducted among non-consecutive persons with a defined clinical presentation.

3b. Study comparing diagnosis with a reference standard that does not meet the criteria for level 2 or 3a.

3c. Diagnostic case–control study.

4. Study of diagnostic yield (no reference standard).

Study	**Author**	Acar *et al.*[62]
	Date	1991
	Pathology(ies) (for which accuracy measured)	Thrombosis, LA thrombi
	Population AF	44.9% AF
Study design	**Study design details**	Comparison of TTE against surgery for the diagnosis of LA thrombi in mitral stenosis (also some cases TOE and angiography) in patients who subsequently underwent mitral valve surgery
	Study design level in hierarchy[57]	3b comparison with reference standard
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	No
	Were the reference standard results interpreted without knowledge of the results of the index test?	No
	Were uninterpretable or indeterminate test results reported?	None reported, all cases used in analysis

Study	**Author**	Arques[63]
	Date	2005
	Pathology(ies) (for which accuracy measured)	CHF
	Population AF	No history of arrhythmia
Study design	**Study design details**	Case–control study, comparison of test accuracy of M-mode TTE and tissue Doppler TTE, with blinding of observers Cases = hypertensive patients with diastolic HF. Controls = gender- and age-matched hypertensive patients All assessments at time of admission
	Study design level in hierarchy[57]	3c
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	Yes

HF, heart failure.

Study	**Author**	Attenhofer Jost[64]
	Date	2000
	Pathology(ies) (for which accuracy measured)	Aortic stenosis, MVP, combined aortic and mitral valve disease, ventricular septal defect (also MR and AR, for which there is higher-level evidence available)
	Population AF	NR (all had heart murmur)
Study design	**Study design details**	Prospective comparison of accuracy, consecutive, blinded, clinical examination immediately before TTE, TTE as reference standard
	Study design level in hierarchy [57]	2
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	TTE as reference standard
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	None reported, all cases used

NR, not reported.

Study	**Author**	Barron *et al.*[65]
	Date	1988
	Pathology(ies) (for which accuracy measured)	MVP
	Population AF	NR
Study design	**Study design details**	Comparison of auscultation and echocardiography, consecutive patients, echocardiographer blinded to auscultatory findings, auscultation immediately prior to or after TTE
	Study design level in hierarchy [57]	2
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Unknown
	Were uninterpretable or indeterminate test results reported?	None reported, all cases used

NR, not reported.

Study	**Author**	Bova[66]
	Date	2003
	Pathology(ies) (for which accuracy measured)	PE
	Population AF	NR
Study design	**Study design details**	Prospective comparison of test accuracy of TTE with reference angiography, consecutive patients, blinded, TTE soon after reference standard
	Study design level in hierarchy [57]	2
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	Yes

NR, not reported.

Study	**Author**	Casella[67]
	Date	2009
	Pathology(ies) (for which accuracy measured)	Native valve infective endocarditis
	Population AF	No AF
Study design	**Study design details**	Blinded comparison in consecutive patients, TTE and TOE within 7 days
	Study design level in hierarchy[57]	2
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	Yes (all used in analysis, separate analysis excluding poor image quality)

Study	**Author**	Cassidy[68]
	Date	1992
	Pathology(ies) (for which accuracy measured)	Aortic stenosis (also MR and AR, for which there is higher-level evidence available)
	Population AF	NR (systolic murmur)
Study design	**Study design details**	Prospective comparison of accuracy, over two time periods unclear if consecutive within time period, blinded
	Study design level in hierarchy [57]	3a
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	TTE as reference standard
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	Yes

NR, not reported.

Study	**Author**	Dittmann[69]
	Date	1987
	Pathology(ies) (for which accuracy measured)	AR in mitral valve disease
	Population AF	38% ($n = 21$)
Study design	**Study design details**	Comparison of pulsed Doppler echo, M-mode echo, clinical signs and cardiac catheterisation, consecutive patients, TTE 1 day before catheterisation
	Study design level in hierarchy [57]	3b comparison with reference standard
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Unknown
	Were the reference standard results interpreted without knowledge of the results of the index test?	Unknown
	Were uninterpretable or indeterminate test results reported?	Yes (states no exclusions for inadequate examinations)

Study	**Author**	Enia[70]
	Date	1989
	Pathology(ies) (for which accuracy measured)	Aortic dissection involving the ascending aorta
	Population AF	NR
Study design	**Study design details**	Case–control, prospective comparison of TTE and aortography in two groups of patients Cases = clinical suspicion of aortic dissection consecutive patients Controls = patients with TTE and aortography, consecutive
	Study design level in hierarchy [57]	3c
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	No
	Were the reference standard results interpreted without knowledge of the results of the index test?	No
	Were uninterpretable or indeterminate test results reported?	None reported, all tests used

NR, not reported.

Study	**Author**	Erbel[71]
	Date	1984
	Pathology(ies) (for which accuracy measured)	LV function
	Population AF	No AF
Study design	**Study design details**	Retrospective comparison of diagnostic accuracy of four echocardiography markers by catheterisation and echocardiography, TTE the day before catheterisation
	Study design level in hierarchy [57]	3b comparison with reference standard
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Unknown
	Were the reference standard results interpreted without knowledge of the results of the index test?	Unknown
	Were uninterpretable or indeterminate test results reported?	None reported, all used

Study	**Author**	Grossmann[72]
	Date	2002
	Pathology(ies) (for which accuracy measured)	MR
	Population AF	25% AF
Study design	**Study design details**	Comparison of TTE and TOE with the some patients having catheterisation for the detection and quantification of MR using the proximal flow convergence method. Consecutive patients, TTE and TOE performed during same examination
	Study design level in hierarchy [57]	3b comparison with reference standard
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes (if TOE reference standard, rather than catheterisation)
	Did patients receive the same reference standard regardless of the index test result?	Yes (if TOE reference standard, rather than catheterisation)
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	No
	Were the index test results interpreted without knowledge of the results of the reference standard?	No
	Were the reference standard results interpreted without knowledge of the results of the index test?	No
	Were uninterpretable or indeterminate test results reported?	None reported

Study	**Author**	Groves[73]
	Date	2004
	Pathology(ies) (for which accuracy measured)	Tricuspid regurgitation
	Population AF	NR
Study design	**Study design details**	Retrospective comparison of CT, TTE and RHC for the detection of tricuspid regurgitation; 61 selected patients (out of 86 consecutive); CT, TTE and RHC within 6 weeks of each other
	Study design level in hierarchy [57]	3a
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	TTE as reference standard
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	NA (selected for having usable examinations)

NA, not applicable; NR, not reported; RHC, right heart catheterisation.

Study	**Author**	Guyer[74]
	Date	1984
	Pathology(ies) (for which accuracy measured)	Rheumatic tricuspid stenosis
	Population AF	31/38 = 82%
Study design	**Study design details**	Retrospective comparison of echocardiography and cardiac catheterisation in selected patients with both examinations; catheterisation with 1 year of TTE
	Study design level in hierarchy [57]	3a
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	NA (selected for having both examinations)

NA, not applicable.

Study	**Author**	Helmcke[75]
	Date	1987
	Pathology(ies) (for which accuracy measured)	MR
	Population AF	31/82 with MR = 38%. None without MR (overall 21%)
Study design	**Study design details**	Comparison of colour Doppler echocardiography and cardiac catheterisation angiography in those with and without MR
	Study design level in hierarchy[57]	3c
Items from QUADAS[59]	**Were selection criteria clearly described?**	No
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	Yes

Study	**Author**	Jassal[76]
	Date	2007
	Pathology(ies) (for which accuracy measured)	Endocarditis
	Population AF	NR
Study design	**Study design details**	Prospective comparison of accuracy, selected population of likely endocarditis from consecutive patients, blinded, TTE within 24 hours of TOE
	Study design level in hierarchy[57]	2
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	Yes (indeterminate TTE included in analysis)

NR, not reported.

Study	**Author**	Kaymaz[77]
	Date	2001
	Pathology(ies) (for which accuracy measured)	Thrombosis, LA thrombi
	Population AF	56.3% AF at time of study
Study design	**Study design details**	Comparison of TTE and TOE measurements of LA thrombi (before surgery) against intraoperative findings. Consecutive patients, TTE and TOE within 1–5 days prior to surgery
	Study design level in hierarchy[57]	3b comparison with reference standard
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	No
	Were the reference standard results interpreted without knowledge of the results of the index test?	No
	Were uninterpretable or indeterminate test results reported?	None reported (all included in analysis)

Study	**Author**	Kishon[78]
	Date	1993
	Pathology(ies) (for which accuracy measured)	VSD and papillary muscle rupture, post MI
	Population AF	NR (new systolic murmur in 68% VSD and 100% papillary rupture)
Study design	**Study design details**	Retrospective comparison of surgery and post-mortem examination against TTE and TOE data
	Study design level in hierarchy[57]	3b
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	No
	Were the reference standard results interpreted without knowledge of the results of the index test?	No
	Were uninterpretable or indeterminate test results reported?	Yes (included in analysis)

NR, not reported; VSD, ventricular septal defect.

Study	**Author**	Kitayama[79]
	Date	1997
	Pathology(ies) (for which accuracy measured)	RA thrombi and LA thrombi
	Population AF	100% CAF
Study design	**Study design details**	Comparison of TTE and CT, consecutive patients (unclear if blinded)
	Study design level in hierarchy[57]	3b
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	No (according to Kitayama *et al.*[79])
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Unknown
	Were the reference standard results interpreted without knowledge of the results of the index test?	Unknown
	Were uninterpretable or indeterminate test results reported?	Yes (included in analysis)

Study	**Author**	Lanzarini[80]
	Date	2005
	Pathology(ies) (for which accuracy measured)	Pulmonary hypertension
	Population AF	13% controlled AF
Study design	**Study design details**	Prospective comparison of test accuracy of TTE with reference cardiac catheterisation within 24 hours, consecutive patients, blinded
	Study design level in hierarchy[57]	2
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	None reported, all cases used

Study	**Author**	Maestre[81]
	Date	2009
	Pathology(ies) (for which accuracy measured)	LV dysfunction, heart failure
	Population AF	NR
Study design	**Study design details**	Comparison of clinical criteria and TTE, cross-sectional survey, 216 of 255 consecutive patients meeting criteria
	Study design level in hierarchy[57]	2
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	TTE as reference standard
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	None reported, all used

NR, not reported.

Study	**Author**	Mugge[82]
	Date	1995
	Pathology(ies) (for which accuracy measured)	ASA
	Population AF	14.4% in AF
Study design	**Study design details**	Database comparison of TOE and TTE, in patients with confirmed ASA (by TOE), TTE and TOE within 24 hours of each other
	Study design level in hierarchy[57]	3b comparison with reference standard
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	No
	Were the reference standard results interpreted without knowledge of the results of the index test?	No
	Were uninterpretable or indeterminate test results reported?	NA (selection for having both examinations)

ASA, atrial septal aneurysm; NA, not applicable.

Study	**Author**	Nienaber[83]
	Date	1993
	Pathology(ies) (for which accuracy measured)	Thoracic aortic dissection
	Population AF	NR
Study design	**Study design details**	Blinded comparison of TTE, TOE, CT, MRI validated against clinical findings to assess their reliability in diagnosis of dissection of the thoracic aorta. (All patients undergoing two imaging procedures, all patients validated by angiography, surgery or autopsy)
	Study design level in hierarchy[57]	2
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	None reported, all used

NR, not reported.

Study	**Author**	Nienaber[84]
	Date	1994
	Pathology(ies) (for which accuracy measured)	Aortic dissection
	Population AF	NR
Study design	**Study design details**	Comparison of the diagnostic accuracy of TTE and TOE with MRI for the exact morphological evaluation and anatomical mapping of the thoracic aorta, blinded
	Study design level in hierarchy[57]	3a
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	None reported, all used

NR, not reported.

Study	**Author**	Okura[85]
	Date	2006
	Pathology(ies) (for which accuracy measured)	Cardiomyopathy
	Population AF	NR
Study design	**Study design details**	Consecutive patients, non-blinded, TTE and angiography with 1 week of each other
	Study design level in hierarchy[57]	3b
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	No
	Were the reference standard results interpreted without knowledge of the results of the index test?	No
	Were uninterpretable or indeterminate test results reported?	Yes

NR, not reported.

Study	**Author**	Pochis[86]
	Date	1992
	Pathology(ies) (for which accuracy measured)	Atrial septal hypertrophy
	Population AF	53% AF or flutter, or paroxysmal atrial tachycardia
Study design	**Study design details**	Retrospective comparison of TTE and TOE in the detection of lipomatous hypertrophy of the atrial septum. Assessors blinded to other results
	Study design level in hierarchy[57]	3b – comparison with reference standard
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	Yes

Study	**Author**	Reichek[87]
	Date	1981
	Pathology(ies) (for which accuracy measured)	LV hypertrophy
	Population AF	NR
Study design	**Study design details**	Retrospective comparison of various diagnostic measures in patient groups
	Study design level in hierarchy[57]	3b comparison with reference standard
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Unknown
	Were the reference standard results interpreted without knowledge of the results of the index test?	Unknown
	Were uninterpretable or indeterminate test results reported?	None reported

NR, not reported.

Study	**Author**	Reichlin[88]
	Date	2004
	Pathology(ies) (for which accuracy measured)	Valvular heart disease
	Population AF	NR (all had heart murmur)
Study design	**Study design details**	Prospective comparison of initial clinical evaluation and TTE in the evaluation of systolic murmurs in the diagnosis of valvular heart disease; independent blinded assessors; 203 patients selected from 852 consecutive patients; TTE within 24 hours of clinical evaluation
	Study design level in hierarchy[57]	2
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	TTE as reference standard
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	NA (TTE as gold standard)

NA, not applicable; NR, not reported.

Study	**Author**	Roudaut[89]
	Date	1988
	Pathology(ies) (for which accuracy measured)	Aortic dissection
	Population AF	NR
Study design	**Study design details**	Retrospective comparison of TTE, angiography, CT or autopsy/surgery
	Study design level in hierarchy[57]	3b comparison with reference standard
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	No
	Were the reference standard results interpreted without knowledge of the results of the index test?	No
	Were uninterpretable or indeterminate test results reported?	Yes (excluded from analysis $n = 13$ of 673)

NR, not reported.

Study	**Author**	Saraste[90]
	Date	2005
	Pathology(ies) (for which accuracy measured)	Coronary artery stenosis
	Population AF	4% CAF
Study design	**Study design details**	Prospective comparison of diagnostic measures. Coronary angiography performed a day after TTE by a cardiologist blinded to results of TTE. TTE all performed by same physician
	Study design level in hierarchy[57]	3b – study of test accuracy, includes reference standard
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	None reported, all images used in calculation of sensitivity/specificity

Study	**Author**	Sharifi[91]
	Date	2007
	Pathology(ies) (for which accuracy measured)	Atrial thrombi
	Population AF	100% AF
Study design	**Study design details**	Blinded comparison of consecutive patients
	Study design level in hierarchy[57]	2
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	NA (selected for usable data)

NA, not applicable.

Study	**Author**	Sharma[92]
	Date	1992
	Pathology(ies) (for which accuracy measured)	Atrial septal defect (sinus venosus defect)
	Population AF	NR
Study design	**Study design details**	Retrospective comparison of TTE, TOE and cardiac catheterisation in the demonstration of sinus venosus defect
	Study design level in hierarchy[57]	3b
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	No
	Were the reference standard results interpreted without knowledge of the results of the index test?	No
	Were uninterpretable or indeterminate test results reported?	Yes (eight cases with inadequate TTE or angiography were excluded from analysis)

NR, not reported.

Study	**Author**	Sheiban[93]
	Date	1987
	Pathology(ies) (for which accuracy measured)	Intracardiac masses
	Population AF	NR
Study design	**Study design details**	Prospective comparison of 2D echocardiography and surgery
	Study design level in hierarchy[57]	3b comparison with reference standard
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Unknown
	Were the reference standard results interpreted without knowledge of the results of the index test?	Unknown
	Were uninterpretable or indeterminate test results reported?	None reported, all used

NR, not reported.

Study	**Author**	Shively[94]
	Date	1991
	Pathology(ies) (for which accuracy measured)	Endocarditis
	Population AF	NR
Study design	**Study design details**	Prospective comparison of TTE and TOE, using non-echocardiographic pathological data from the subsequent clinical course as the reference standard, blinded comparison in consecutive patients
	Study design level in hierarchy[57]	2 (blinded comparison in consecutive patients)
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	Yes (all included in analysis, poorer than average TTE image 18% tricuspid valve, 11% mitral valve, 32% aortic valve)

NR, not reported.

Study	**Author**	Shrestha[95]
	Date	1983
	Pathology(ies) (for which accuracy measured)	LA thrombus (in rheumatic heart disease)
	Population AF	NR for whole population, for those with thrombus 45/51 = 88%
Study design	**Study design details**	Retrospective comparison of 2D echocardiography and surgical findings of LA thrombi, surgery within 1 week of TTE
	Study design level in hierarchy[57]	3b
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes (video recordings reviewed by blinded observer)
	Were the reference standard results interpreted without knowledge of the results of the index test?	Unknown
	Were uninterpretable or indeterminate test results reported?	None reported

NR, not reported.

Study	**Author**	Shub[96]
	Date	1983
	Pathology(ies) (for which accuracy measured)	Atrial septal defect
	Population AF	NR
Study design	**Study design details**	Retrospective comparison of 2D echocardiography against surgery/ catheterisation from 171 patients, 154 entered study (nine excluded for poor TTE, eight patients had incomplete examination)
	Study design level in hierarchy[57]	3b comparison with reference standard
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	No
	Were the reference standard results interpreted without knowledge of the results of the index test?	No
	Were uninterpretable or indeterminate test results reported?	Yes (9 of 171 patients excluded for poor image quality)

NR, not reported.

Study	**Author**	Shyu[97]
	Date	1992
	Pathology(ies) (for which accuracy measured)	Ruptured chordae tendineae
	Population AF	Some AF
Study design	**Study design details**	Diagnostic case–control study, blinded Cases = ruptured chordae tendineae Control subjects = MR due to other causes, most catheterisations within 1 week of echocardiography studies 37/40 cases and 18/20 control subjects had catheterisations
	Study design level in hierarchy[57]	3c
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	None reported (all used in analysis)

Study	**Author**	Smith[98]
	Date	1985
	Pathology(ies) (for which accuracy measured)	Ventricular septal rupture (in patients with AMI)
	Population AF	NR
Study design	**Study design details**	Comparison with reference standard, 13 patients excluded for not having reference standard
	Study design level in hierarchy[57]	3b comparison with reference standard
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	No
	Were the reference standard results interpreted without knowledge of the results of the index test?	No
	Were uninterpretable or indeterminate test results reported?	None reported, all used

NR, not reported.

Study	**Author**	Sparrow[99]
	Date	2003
	Pathology(ies) (for which accuracy measured)	LV systolic dysfunction
	Population AF	NR
Study design	**Study design details**	Prospective comparison of accuracy, cross-section not consecutive, blinded
	Study design level in hierarchy[57]	3a
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	TTE as reference standard
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	Yes (13% excluded from study owing to inadequate TTE images)

NR, not reported.

Study	**Author**	Stratton[100]
	Date	1982
	Pathology(ies) (for which accuracy measured)	LV thrombus
	Population AF	Percentage NR but some patients had AF
Study design	**Study design details**	Retrospective comparison of 2D echocardiography and indium-111 platelet imaging and surgical findings. Assessors blinded
	Study design level in hierarchy[57]	3b comparison with reference standard
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	Yes (excluded from analysis)

NR, not reported.

Study	**Author**	Veyrat[101]
	Date	1983
	Pathology(ies) (for which accuracy measured)	AR
	Population AF	38/95 = 40% overall
Study design	**Study design details**	Retrospective comparison of echocardiography against aortic root angiography (some surgical findings)
	Study design level in hierarchy[57]	3b comparison with reference standard
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Unknown
	Were the reference standard results interpreted without knowledge of the results of the index test?	Unknown
	Were uninterpretable or indeterminate test results reported?	None reported, all used

Study	**Author**	Vigna[102]
	Date	1993
	Pathology(ies) (for which accuracy measured)	LA thrombus
	Population AF	59% in AF at time of study
Study design	**Study design details**	Comparison of TTE and TOE, consecutive patients, blinded ('two observers who were unaware of TTE findings') TTE and TOE within 24 hours of each other
	Study design level in hierarchy[57]	2
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	None reported, all used

Study	**Author**	Wong[103]
	Date	1983
	Pathology(ies) (for which accuracy measured)	Mitral and aortic valve stenosis, valvular calcification
	Population AF	NR
Study design	**Study design details**	Prospective comparison of 2D echocardiography and cinefluorography for detection of valvular calcification, blinding, non-consecutive
	Study design level in hierarchy[57]	3a comparison with reference standard
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
	Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
	Were uninterpretable or indeterminate test results reported?	None reported, all used

NR, not reported.

Study	**Author**	Zanolla[104]
	Date	1982
	Pathology(ies) (for which accuracy measured)	Mitral stenosis, mitral valve calcification
	Population AF	NR
Study design	**Study design details**	Retrospective comparison of 2D echocardiography and surgical findings, non-consecutive
	Study design level in hierarchy[57]	3b comparison with reference standard
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	Unknown
	Were the reference standard results interpreted without knowledge of the results of the index test?	Unknown
	Were uninterpretable or indeterminate test results reported?	None reported, all used

NR, not reported.

Study	**Author**	Zotz[105]
	Date	1993
	Pathology(ies) (for which accuracy measured)	Ventricular septal rupture (in patients with AMI)
	Population AF	NR
Study design	**Study design details**	comparison with reference standard, not blinded, investigated consecutively
	Study design level in hierarchy[57]	3b comparison with reference standard
Items from QUADAS[59]	**Were selection criteria clearly described?**	Yes
	Is the reference standard likely to correctly classify the target condition?	Yes
	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
	Did the whole sample (rather than a random selection of the sample) receive verification using a reference standard of diagnosis?	Yes
	Did patients receive the same reference standard regardless of the index test result?	Yes
	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes
	Were the index test results interpreted without knowledge of the results of the reference standard?	No
	Were the reference standard results interpreted without knowledge of the results of the index test?	No
	Were uninterpretable or indeterminate test results reported?	None reported (all images used in analysis)

NR, not reported.

Prognostic studies: quality assessment

Level in hierarchy of evidence based on Merlin *et al.*:[57]

1. Systematic review of level 2 studies.

2. Prospective cohort study.

3a. All or none study.

3b. Analysis of prognostic factors amongst persons in a single arm of a randomised controlled trial (RCT).

3c. Retrospective cohort study.

4. Case series or cohort study of persons at different stages of disease.

Study	**Author**	Atrial Fibrillation Investigators[106]
	Date	1998
	Pathology(ies) (for which accuracy measured)	LV dysfunction, LAD, MVP, MR
	Population AF	All participants non-valvular AF
Study design	**Study design details**	Review of 3 (prospective) RCTs, using data from single arm of each (placebo/control), with outcome of subsequent stroke, also looked at clinical criteria for risk of stroke
	Study design level in hierarchy[57]	3b (review of level 3b)

Study	**Author**	Klem[107]
	Date	2003
	Pathology(ies) (for which accuracy measured)	Reduced LV function, LAD valvular abnormality
	Population AF	A total of 336 patients with non-rheumatic AF and 73 patients with non-rheumatic AF and also diabetes (for both groups, selected from 409 eligible of 474 consecutive patients)
Study design	**Study design details**	Prospective cohort study
	Study design level in hierarchy[57]	2

Study	**Author**	Miyaska[108]
	Date	2000
	Pathology(ies) (for which accuracy measured)	MR
	Population AF	All participants non-rheumatic AF
Study design	**Study design details**	Retrospective database study
	Study design level in hierarchy[57]	3c retrospective cohort study

Study	**Author**	Nakagami[109]
	Date	1998
	Pathology(ies) (for which accuracy measured)	Degree of MR and LAD
	Population AF	A total of 290 patients with non-rheumatic AF
Study design	**Study design details**	Retrospective cohort
	Study design level in hierarchy[57]	3c

Study	**Author**	The Stroke Prevention in Atrial Fibrillation (SPAF) Investigators[110]
	Date	1992
	Pathology(ies) (for which accuracy measured)	Mitral annular calcification, severe MR, LV dysfunction and LAD
	Population AF	A total of 568 non-rheumatic AF, inpatient or outpatient, placebo arm of RCT (SPAF study)
Study design	**Study design details**	Cohort study of placebo arm of RCT
	Study design level in hierarchy[57]	3b analysis of prognostic factors amongst persons in a single arm of a RCT