Artificial Intelligence-Enabled Digital Auscultation and Phonocardiography for Clinically Significant Adult Valvular Heart Disease: A Systematic Review and Meta-Analysis

Yun-Chu Wang [1,2,3], Yu-Chuan Chen [1,2,3], Chih-Yu Yang[3,4,5,6,7]

[1] Department of Education, Taipei Veterans General Hospital, Taipei, Taiwan

[2] School of Medicine, National Yang Ming Chiao Tung University, Taipei, Taiwan

[3] Institute of Clinical Medicine, School of Medicine, National Yang Ming Chiao Tung University, Taipei, Taiwan

[4] Division of Nephrology, Department of Medicine, Taipei Veterans General Hospital, Taipei, Taiwan

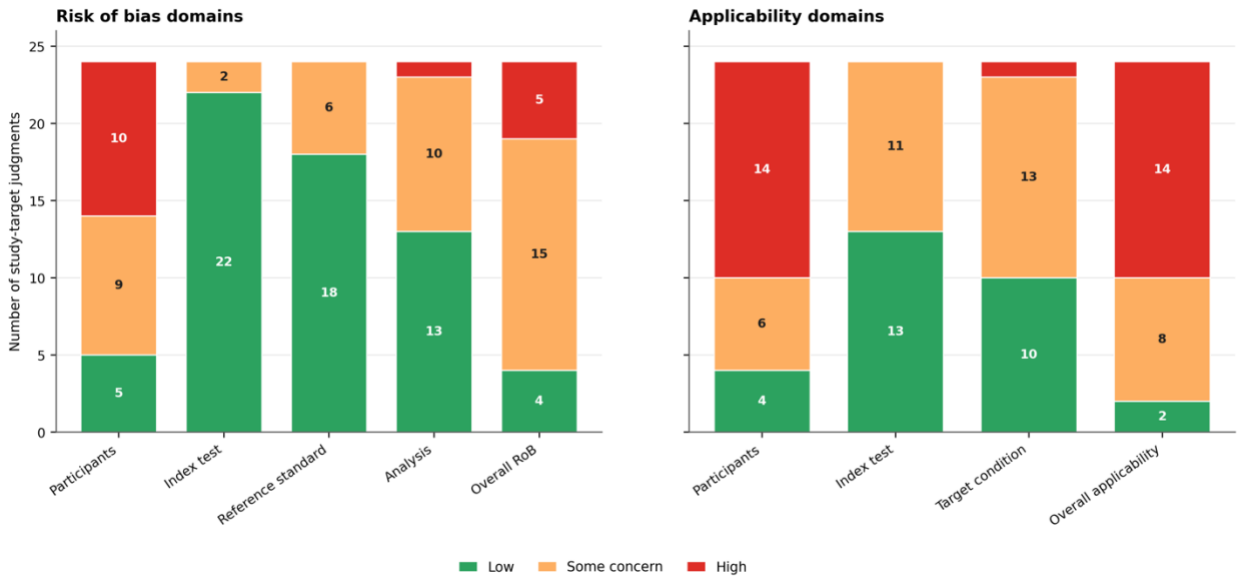
[5] Stem Cell Research Center, National Yang Ming Chiao Tung University, Taipei, Taiwan

[6] Center for Intelligent Drug Systems and Smart Bio-devices (IDS2B), Hsinchu, Taiwan

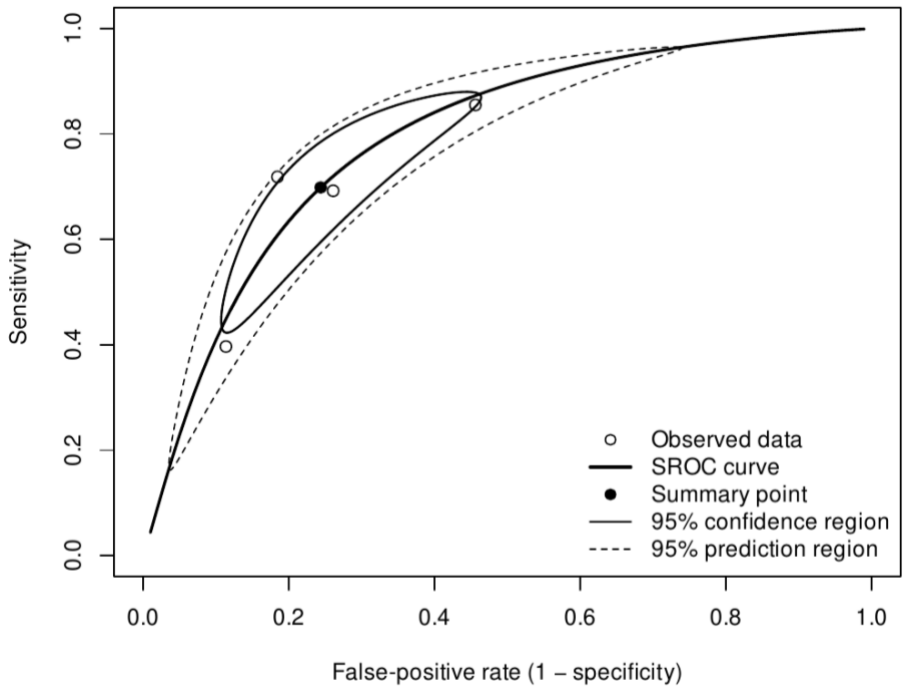
[7] International Medical Service Center, Taipei Veterans General Hospital, Taipei, Taiwan

*Corresponding author: email, cyyang3@vghtpe.gov.tw

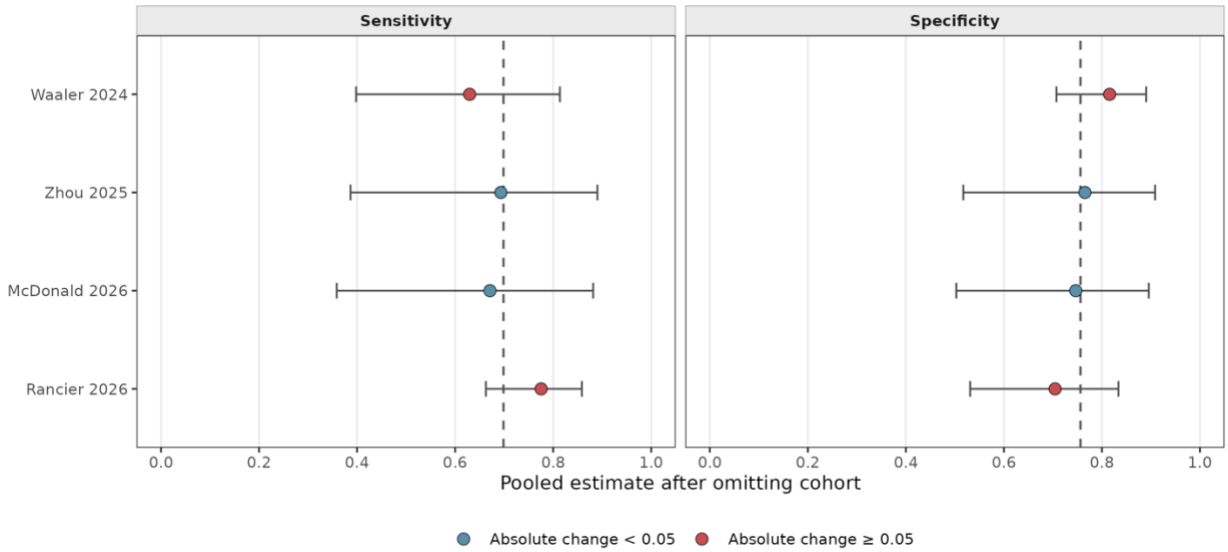
Supplementary Online Content



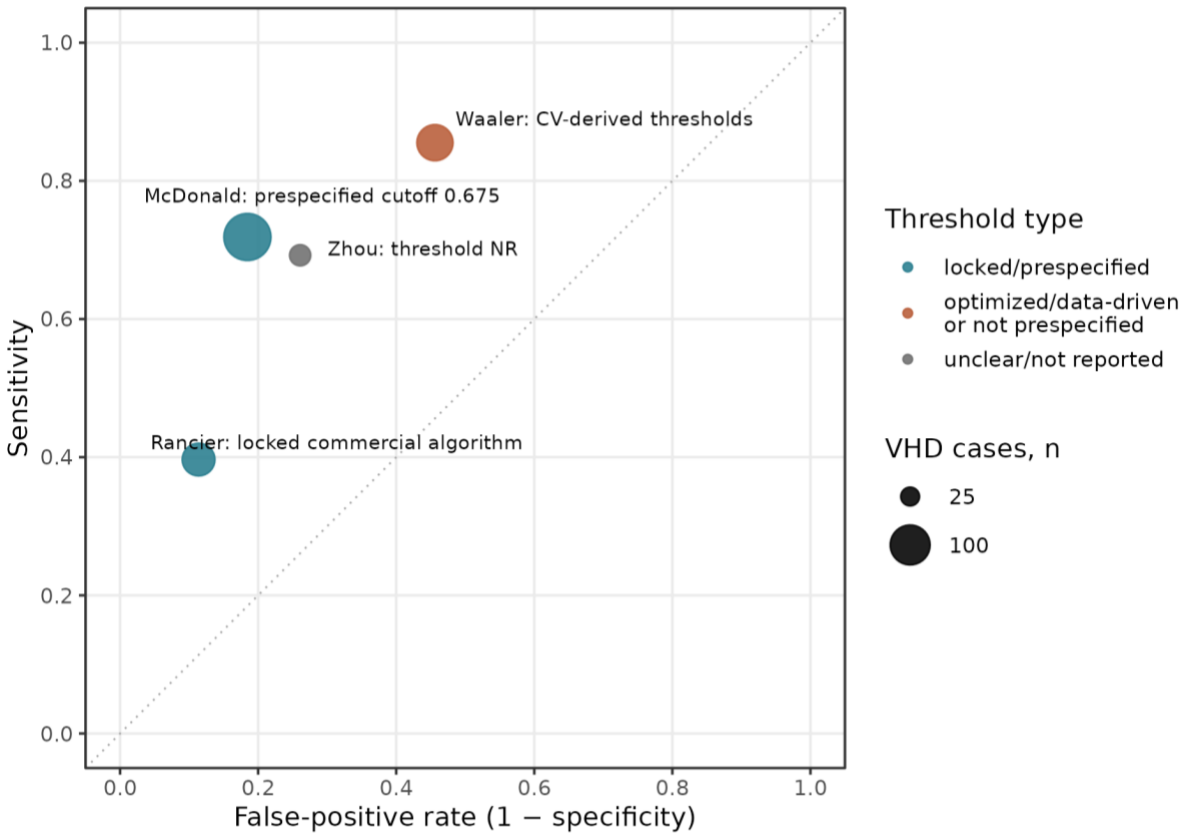
Supplementary Figure S1. Distribution of QUADAS-3 judgments across risk-of-bias and applicability domains. Stacked bars show the number of study-target judgments rated as low, some concern, or high for each QUADAS-3 domain. Counts are based on 24 study-target assessments across the clinically significant VHD, AS, and MR review questions.



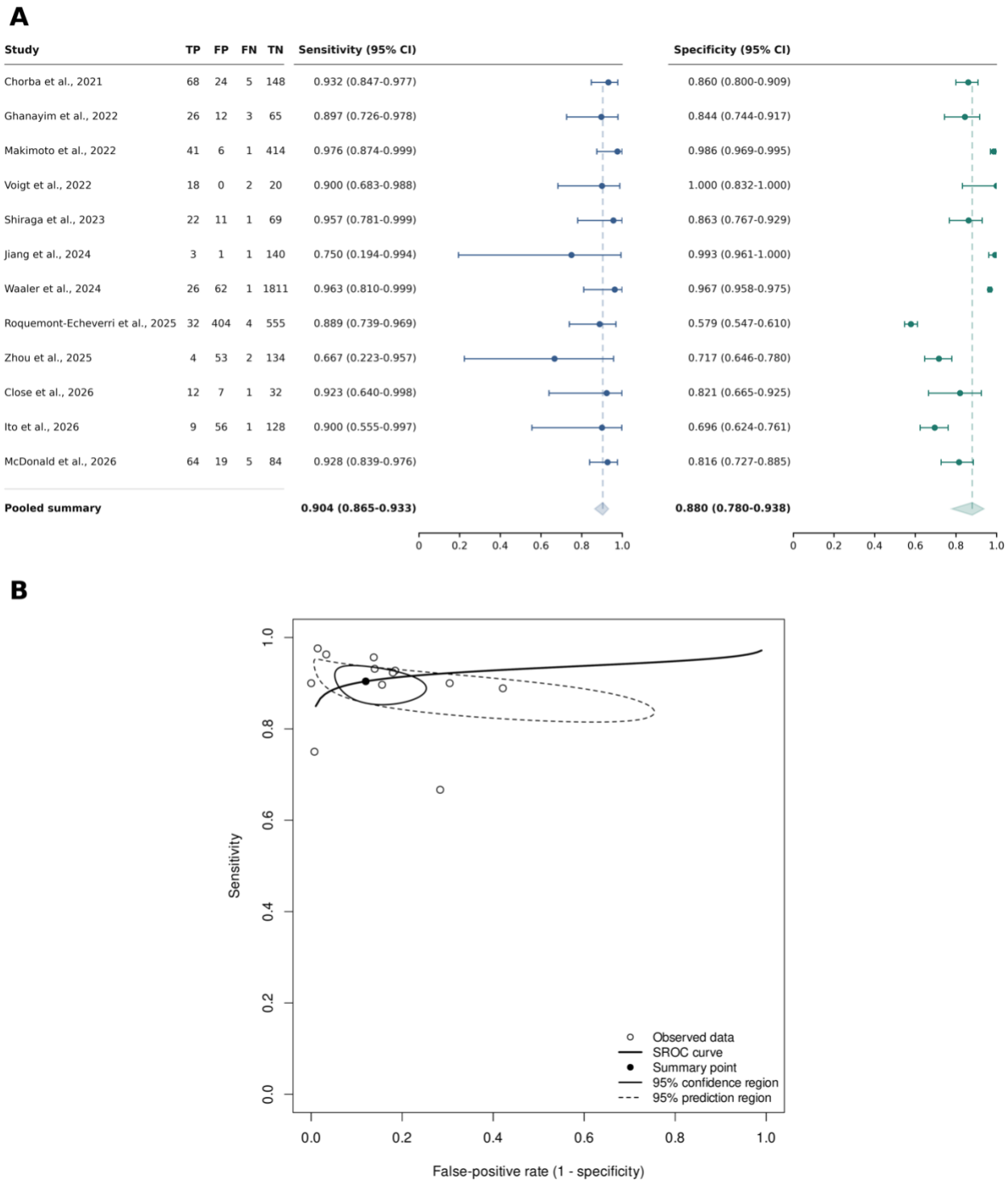
Supplementary Figure S2. Hierarchical summary receiver operating characteristic (HSROC) plot for clinically significant VHD. Open circles represent study estimates; the solid curve shows the SROC function, and the summary point is shown with 95% confidence and prediction regions. VHD, valvular heart disease.



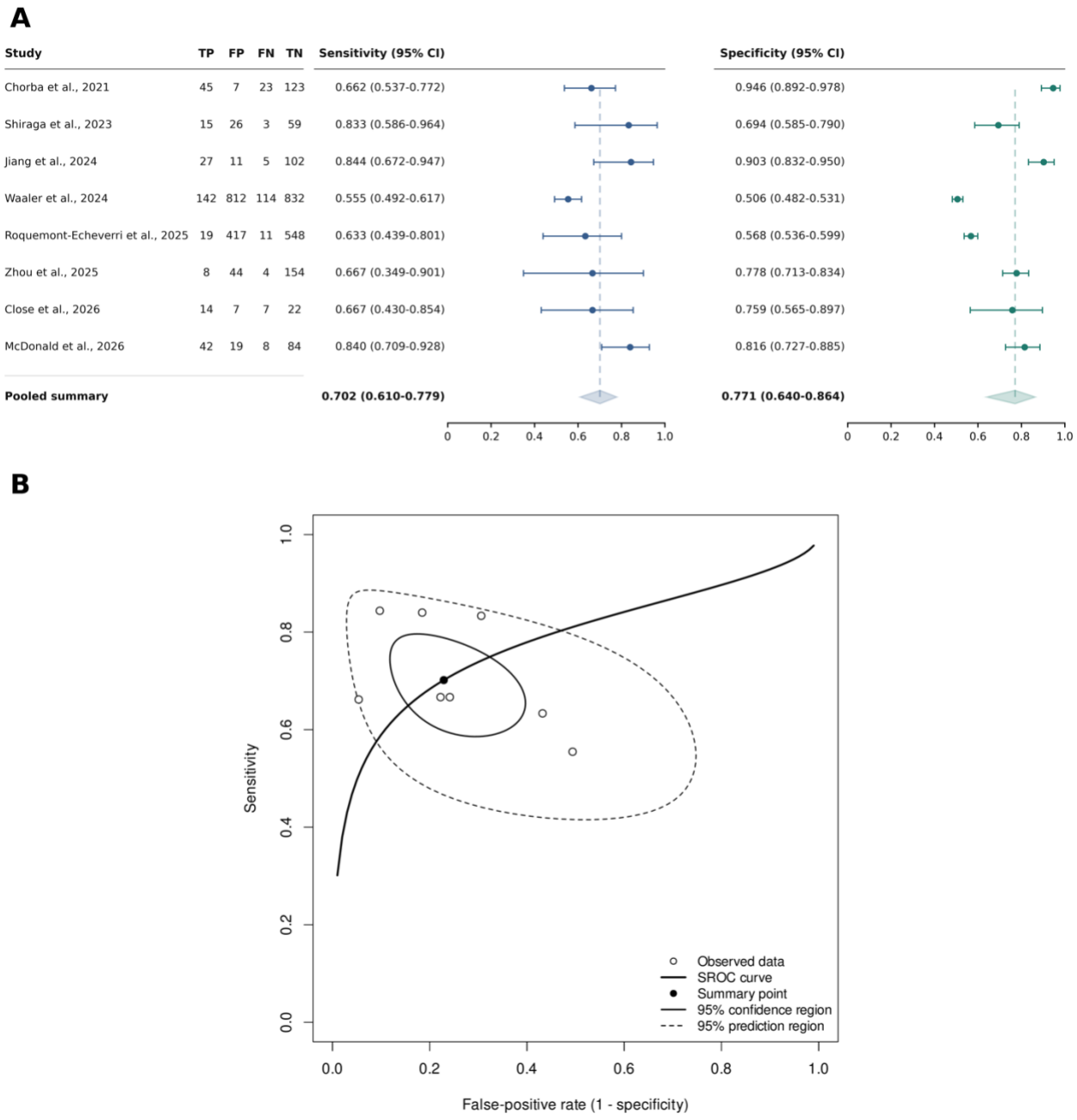
Supplementary Figure S3. Leave-one-out analysis of the primary diagnostic accuracy model. Points indicate the resulting pooled estimate for sensitivity or specificity, and horizontal bars indicate 95% CI. Dashed vertical lines indicate the corresponding primary pooled estimates from the full model including all eligible cohorts. Colors denote the absolute change from the full-model pooled estimate and are intended as a descriptive visual aid.



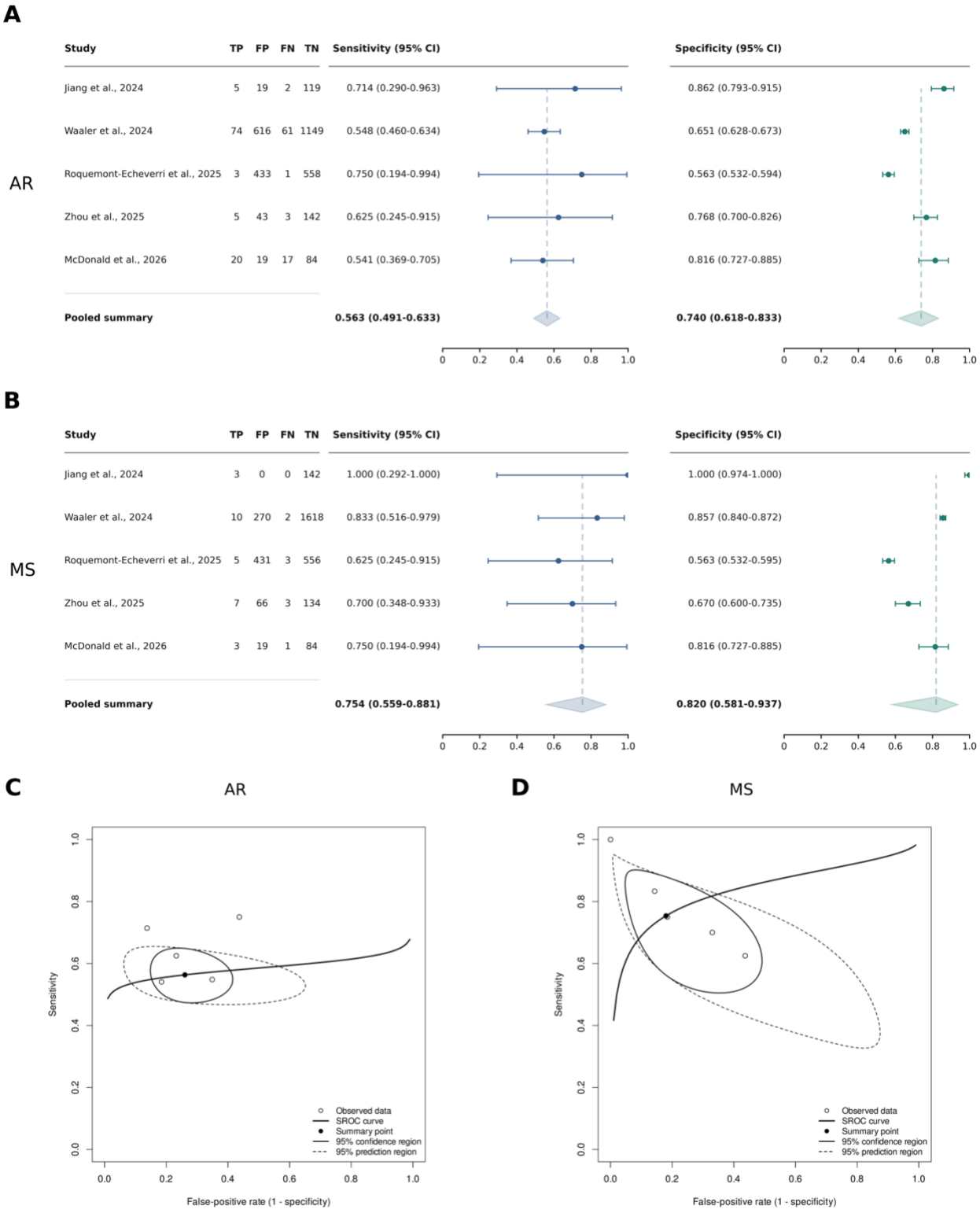
Supplementary Figure S4. Exploratory threshold-effect assessment in the primary diagnostic accuracy analysis. Each point represents one eligible validation cohort plotted at its reported operating point, with sensitivity on the y-axis and false-positive rate on the x-axis. Colours indicate threshold type, and point size reflects the number of VHD cases. Text labels show the study and reported thresholding rule. The dotted diagonal line indicates the line of no discrimination. This analysis was exploratory and was not used to select an optimal threshold. CV, cross-validation; NR, not reported; VHD, valvular heart disease.



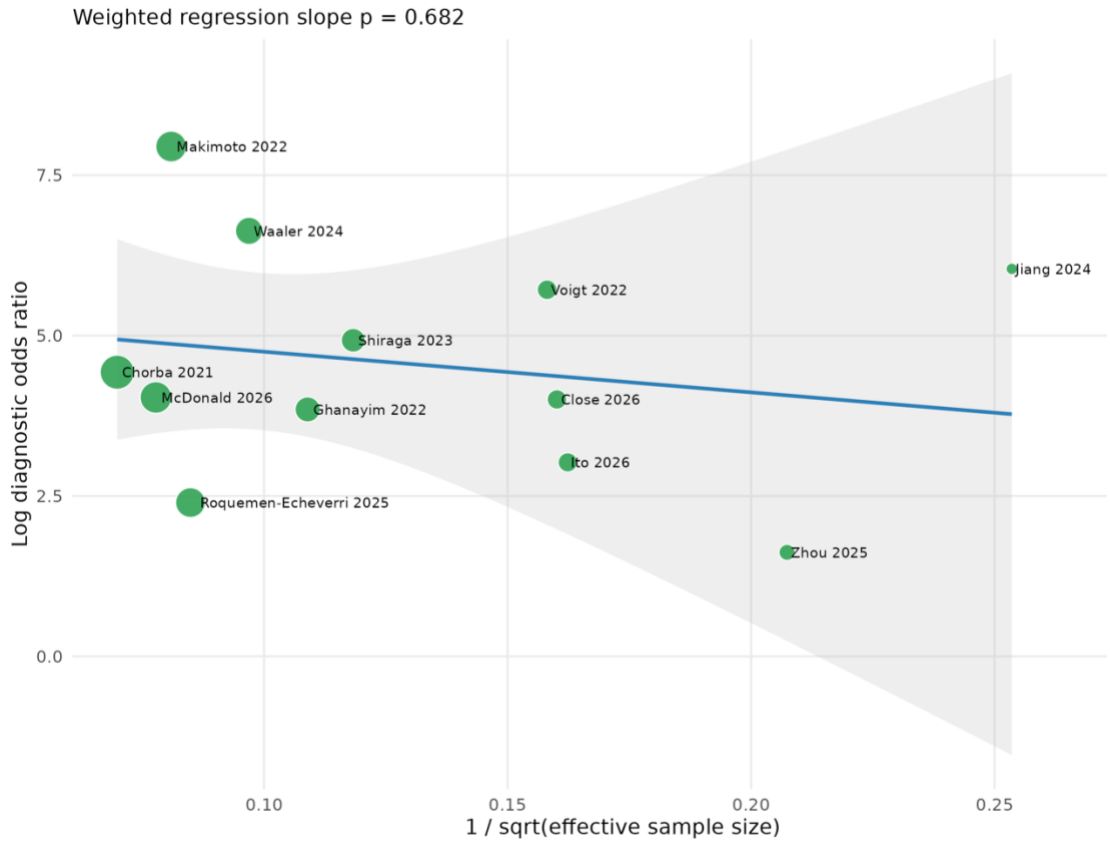
Supplementary Figure S5. Forest and SROC plots for AS. (A) Coupled forest plots for sensitivity and specificity; (B) SROC plot with summary point and 95% confidence and prediction regions. AS, aortic stenosis.



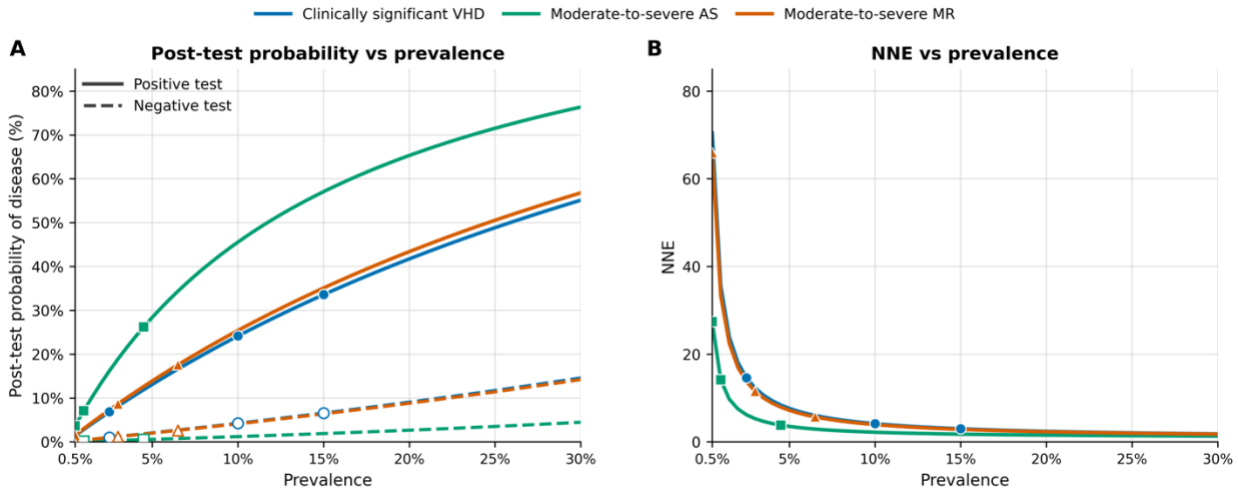
Supplementary Figure S6. Forest and SROC plots for MR. (A) Coupled forest plots for sensitivity and specificity; (B) SROC plot with summary point and 95% confidence and prediction regions. MR, mitral regurgitation.



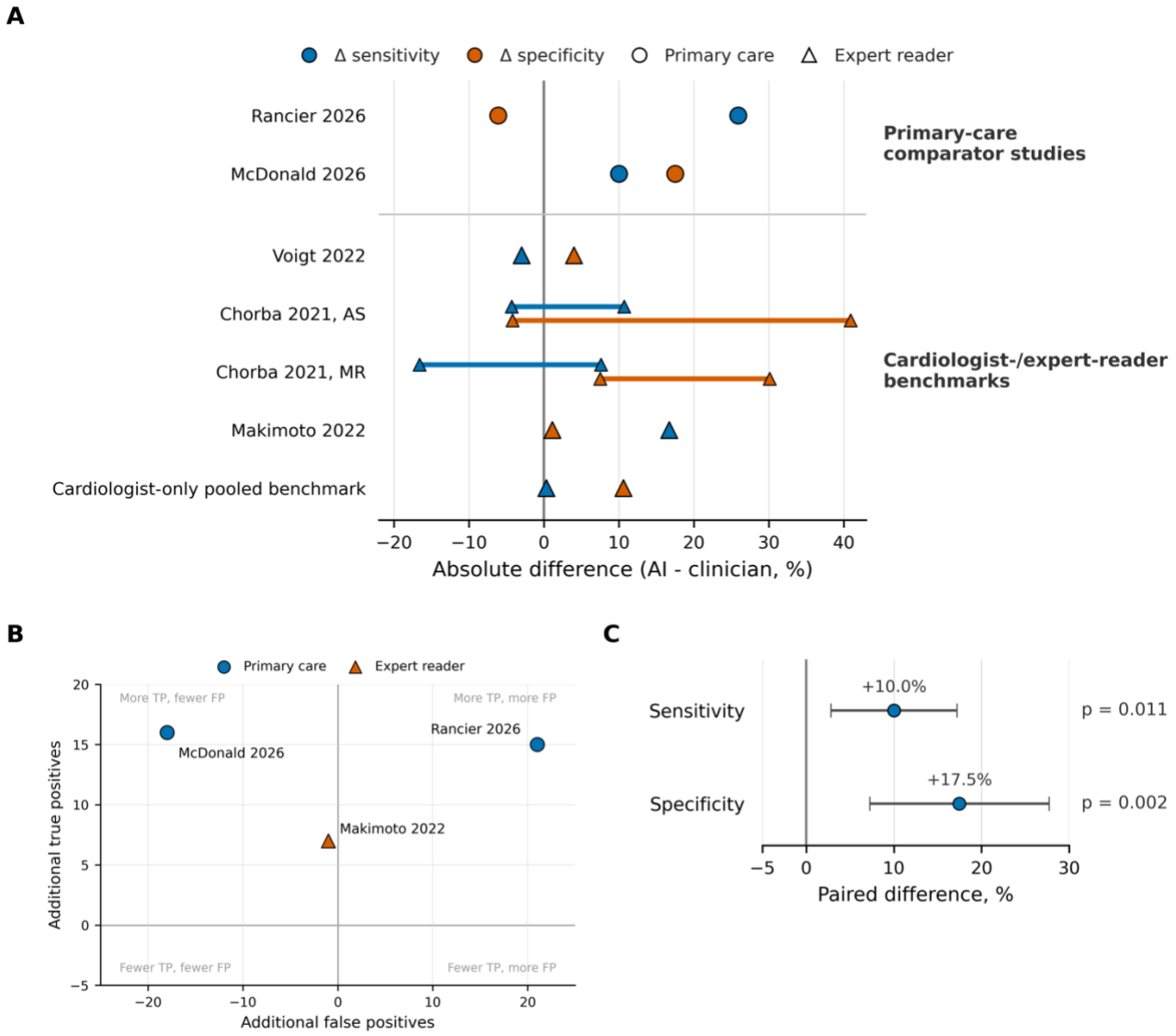
Supplementary Figure S7. Forest and SROC plots for AR and MS. (A,B) Forest plots for sensitivity and specificity of AR and MS; (C,D) SROC plots with summary points and 95% confidence and prediction regions. AR, aortic regurgitation; MS, mitral stenosis.



Supplementary Figure S8. Deeks' funnel-plot asymmetry test for AS. Each point represents a study; point size reflects effective sample size, and the regression line assesses small-study effects. AS, aortic stenosis.



Supplementary Figure S9. Post-test probability and number needed to echo (NNE) across disease prevalence. (A) Post-test probability of disease after AI-positive and AI-negative results across prevalence values. Solid lines indicate AI-positive results, and dashed lines indicate AI-negative results. (B) NNE, defined as the number of AI-positive echocardiography referrals required to confirm one true-positive case. Markers indicate the representative prevalence settings used in the clinical-yield analysis. VHD, valvular heart disease; AS, aortic stenosis; MR, mitral regurgitation; NNE, number needed to echo.



Supplementary Figure S10. AI-enabled auscultation versus clinician comparators.
(A) Absolute differences in sensitivity and specificity for AI-enabled auscultation versus clinician comparators (AI – clinician). Horizontal bars for Chorba et al. show ranges across cardiologist readers.
(B) Additional true positives versus additional false positives among studies with comparable 2×2 data.
(C) Paired comparison from McDonald et al., showing absolute paired differences with approximate 95% CI; p values were calculated using McNemar tests.

Supplementary Table S1. Full search strategy.

Database	Keyword	Date	Results
PubMed	("Phonocardiography"[Mesh] OR "Heart Auscultation"[Mesh] OR "Auscultation"[Mesh] OR "Stethoscopes"[Mesh] OR "Heart Sounds"[Mesh] OR phonocardiograph*[tiab] OR phonocardiogram*[tiab] OR PCG[tiab] OR "heart sound"[tiab] OR "heart sounds"[tiab] OR "heart-sound"[tiab] OR "heart-sounds"[tiab] OR "cardiac sound"[tiab] OR "cardiac sounds"[tiab] OR "cardiac acoustic"[tiab] OR "cardiac acoustics"[tiab] OR "cardioacoustic"[tiab] OR "cardio-acoustic"[tiab] OR "acoustic cardiography"[tiab] OR "heart sound signal"[tiab] OR "heart sound signals"[tiab] OR "cardiac acoustic signal"[tiab] OR "cardiac acoustic signals"[tiab] OR auscultat*[tiab] OR stethoscop*[tiab] OR "digital stethoscope"[tiab] OR "digital stethoscopes"[tiab] OR "electronic stethoscope"[tiab] OR "electronic stethoscopes"[tiab] OR "smart stethoscope"[tiab] OR "smart stethoscopes"[tiab] OR "intelligent stethoscope"[tiab] OR "intelligent stethoscopes"[tiab] OR "digital auscultation"[tiab] OR "electronic auscultation"[tiab] OR "automated auscultation"[tiab] OR "automatic auscultation"[tiab] OR "computer-assisted auscultation"[tiab] OR "computer assisted auscultation"[tiab] OR "computer-aided auscultation"[tiab] OR "computer aided auscultation"[tiab] OR infrasound[tiab]) AND ("Artificial Intelligence"[Mesh] OR "Machine Learning"[Mesh] OR "Deep Learning"[Mesh] OR "Neural Networks, Computer"[Mesh] OR "Algorithms"[Mesh] OR "Diagnosis, Computer-Assisted"[Mesh] OR "Pattern Recognition, Automated"[Mesh] OR "Signal Processing, Computer-Assisted"[Mesh] OR AI[tiab] OR "artificial intelligence"[tiab] OR "machine learning"[tiab] OR "deep learning"[tiab] OR "neural network"[tiab] OR "neural networks"[tiab] OR "artificial neural network"[tiab] OR "artificial neural networks"[tiab] OR "convolutional neural network"[tiab] OR "convolutional neural networks"[tiab] OR CNN[tiab] OR "recurrent neural network"[tiab] OR "recurrent neural networks"[tiab] OR RNN[tiab] OR transformer[tiab] OR transformers[tiab] OR "self-attention"[tiab] OR "support vector machine"[tiab] OR "support vector machines"[tiab] OR SVM[tiab] OR "random forest"[tiab] OR "random forests"[tiab] OR "gradient boosting"[tiab] OR XGBoost[tiab] OR classifier*[tiab] OR classification[tiab] OR algorithm*[tiab] OR automated[tiab] OR automatic[tiab] OR "computer-aided"[tiab] OR "computer aided"[tiab] OR "computer-assisted"[tiab] OR "computer assisted"[tiab] OR intelligent[tiab] OR "pattern recognition"[tiab] OR "signal processing"[tiab] OR "transfer learning"[tiab] OR "feature extraction"[tiab]) AND	06/30/2026	969

	("Heart Valve Diseases"[Mesh] OR "Aortic Valve Stenosis"[Mesh] OR "Aortic Valve Insufficiency"[Mesh] OR "Mitral Valve Insufficiency"[Mesh] OR "Mitral Valve Stenosis"[Mesh] OR "Pulmonary Valve Stenosis"[Mesh] OR "Pulmonary Valve Insufficiency"[Mesh] OR "Tricuspid Valve Insufficiency"[Mesh] OR "Tricuspid Valve Stenosis"[Mesh] OR "Rheumatic Heart Disease"[Mesh] OR "Heart Murmurs"[Mesh] OR "valvular heart disease"[tiab] OR "valvular heart diseases"[tiab] OR VHD[tiab] OR HVD[tiab] OR "heart valve disease"[tiab] OR "heart valve diseases"[tiab] OR "valve disease"[tiab] OR "valve diseases"[tiab] OR "valvular disease"[tiab] OR "valvular diseases"[tiab] OR "valvular lesion"[tiab] OR "valvular lesions"[tiab] OR "valve lesion"[tiab] OR "valve lesions"[tiab] OR valvulopath*[tiab] OR "valvular dysfunction"[tiab] OR "valve dysfunction"[tiab] OR "cardiac valve disease"[tiab] OR "cardiac valve diseases"[tiab] OR "aortic stenosis"[tiab] OR "aortic valve stenosis"[tiab] OR "aortic sclerosis"[tiab] OR "aortic valve sclerosis"[tiab] OR "aortic regurgitation"[tiab] OR "aortic valve regurgitation"[tiab] OR "aortic insufficiency"[tiab] OR "aortic valve insufficiency"[tiab] OR "mitral regurgitation"[tiab] OR "mitral valve regurgitation"[tiab] OR "mitral insufficiency"[tiab] OR "mitral valve insufficiency"[tiab] OR "mitral stenosis"[tiab] OR "mitral valve stenosis"[tiab] OR "tricuspid regurgitation"[tiab] OR "tricuspid valve regurgitation"[tiab] OR "tricuspid insufficiency"[tiab] OR "tricuspid stenosis"[tiab] OR "tricuspid valve stenosis"[tiab] OR "pulmonary regurgitation"[tiab] OR "pulmonic regurgitation"[tiab] OR "pulmonary valve regurgitation"[tiab] OR "pulmonary insufficiency"[tiab] OR "pulmonic insufficiency"[tiab] OR "pulmonary stenosis"[tiab] OR "pulmonic stenosis"[tiab] OR "pulmonary valve stenosis"[tiab] OR "rheumatic heart disease"[tiab] OR "rheumatic valvular"[tiab] OR murmur*[tiab] OR "structural heart disease"[tiab] OR "structural murmur"[tiab] OR "structural murmurs"[tiab] OR "pathologic murmur"[tiab] OR "pathologic murmurs"[tiab] OR "pathological murmur"[tiab] OR "pathological murmurs"[tiab]) NOT (animals[mh] NOT humans[mh])		
Web of Science	TS=((phonocardiograph* OR phonocardiogram* OR PCG OR ((heart OR cardiac) NEAR/1 sound*) OR (cardiac NEAR/2 acoustic*) OR cardioacoustic* OR "cardio-acoustic" OR "acoustic cardiography" OR ((heart OR cardiac) NEAR/2 auscultat*) OR auscultat* OR stethoscop* OR ((digital OR electronic OR smart OR intelligent) NEAR/3 stethoscop*) OR ((digital OR electronic OR automated OR automatic OR "computer-assisted" OR "computer assisted" OR "computer-aided" OR "computer aided") NEAR/3 auscultat*) OR infrasound)	06/30/2026	1742

	AND ("artificial intelligence" OR AI OR "machine learning" OR "deep learning" OR "neural network*" OR "artificial neural network*" OR "convolutional neural network*" OR CNN OR "recurrent neural network*" OR RNN OR transformer* OR "self-attention" OR "support vector machine*" OR SVM OR "random forest*" OR "gradient boosting" OR XGBoost OR classifier* OR classification OR algorithm* OR automated OR automatic OR "computer-aided" OR "computer aided" OR "computer-assisted" OR "computer assisted" OR intelligent OR "pattern recognition" OR "signal processing" OR "transfer learning" OR "feature extraction") AND ("valvular heart disease*" OR VHD OR HVD OR "heart valve disease*" OR "valve disease*" OR "valvular disease*" OR "valve lesion*" OR "valvular lesion*" OR valvulopath* OR "valvular dysfunction" OR "valve dysfunction" OR "cardiac valve disease*" OR ((aortic OR mitral OR tricuspid OR pulmonary OR pulmonic) NEAR/3 (stenos* OR regurgitat* OR insufficien* OR scleros*)) OR "rheumatic heart disease" OR (rheumatic NEAR/3 valv*) OR murmur* OR "structural heart disease" OR "structural murmur*" OR "pathologic murmur*" OR "pathological murmur*"))		
Cochrane CENTRAL	#1 MeSH descriptor: [Phonocardiography] explode all trees #2 MeSH descriptor: [Heart Auscultation] explode all trees #3 MeSH descriptor: [Auscultation] explode all trees #4 MeSH descriptor: [Stethoscopes] explode all trees #5 MeSH descriptor: [Heart Sounds] explode all trees #6 (phonocardiograph* OR phonocardiogram* OR PCG OR stethoscop* OR auscultat* OR cardioacoustic* OR "cardio-acoustic" OR "acoustic cardiography" OR infrasound):ti,ab,kw #7 ("heart sound" OR "heart sounds" OR "heart-sound" OR "heart-sounds" OR "cardiac sound" OR "cardiac sounds" OR "cardiac acoustic" OR "cardiac acoustics" OR "heart sound signal" OR "heart sound signals" OR "cardiac acoustic signal" OR "cardiac acoustic signals"):ti,ab,kw #8 ("digital stethoscope" OR "digital stethoscopes" OR "electronic stethoscope" OR "electronic stethoscopes" OR "smart stethoscope" OR "smart stethoscopes" OR "intelligent stethoscope" OR "intelligent stethoscopes"):ti,ab,kw	06/30/2026	40

	#9 ("digital auscultation" OR "electronic auscultation" OR "automated auscultation" OR "automatic auscultation" OR "computer-assisted auscultation" OR "computer assisted auscultation" OR "computer-aided auscultation" OR "computer aided auscultation"):ti,ab,kw #10 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 #11 MeSH descriptor: [Artificial Intelligence] explode all trees #12 MeSH descriptor: [Machine Learning] explode all trees #13 MeSH descriptor: [Deep Learning] explode all trees #14 MeSH descriptor: [Neural Networks, Computer] explode all trees #15 MeSH descriptor: [Algorithms] explode all trees #16 MeSH descriptor: [Diagnosis, Computer-Assisted] explode all trees #17 MeSH descriptor: [Pattern Recognition, Automated] explode all trees #18 MeSH descriptor: [Signal Processing, Computer-Assisted] explode all trees #19 ("artificial intelligence" OR AI OR "machine learning" OR "deep learning" OR "neural network" OR "neural networks" OR "artificial neural network" OR "artificial neural networks" OR "convolutional neural network" OR "convolutional neural networks" OR CNN OR "recurrent neural network" OR "recurrent neural networks" OR RNN OR transformer OR transformers OR "self-attention" OR "support vector machine" OR "support vector machines" OR SVM OR "random forest" OR "random forests" OR "gradient boosting" OR XGBoost OR classifier* OR classification OR algorithm* OR automated OR automatic OR "computer-aided" OR "computer aided" OR "computer-assisted" OR "computer assisted" OR intelligent OR "pattern recognition" OR "signal processing" OR "transfer learning" OR "feature extraction"):ti,ab,kw #20 #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 #21 MeSH descriptor: [Heart Valve Diseases] explode all trees #22 MeSH descriptor: [Aortic Valve Stenosis] explode all trees #23 MeSH descriptor: [Aortic Valve Insufficiency] explode all trees #24 MeSH descriptor: [Mitral Valve Insufficiency] explode all trees #25 MeSH descriptor: [Mitral Valve Stenosis] explode all trees #26 MeSH descriptor: [Pulmonary Valve Stenosis] explode all trees #27 MeSH descriptor: [Pulmonary Valve Insufficiency] explode all trees #28 MeSH descriptor: [Tricuspid Valve Insufficiency] explode all trees #29 MeSH descriptor: [Tricuspid Valve Stenosis] explode all trees #30 MeSH descriptor: [Rheumatic Heart Disease] explode all trees #31 MeSH descriptor: [Heart Murmurs] explode all trees		

	#32 ("valvular heart disease" OR "valvular heart diseases" OR VHD OR HVD OR "heart valve disease" OR "heart valve diseases" OR "valve disease" OR "valve diseases" OR "valvular disease" OR "valvular diseases" OR "valvular lesion" OR "valvular lesions" OR "valve lesion" OR "valve lesions" OR valvulopath* OR "valvular dysfunction" OR "valve dysfunction" OR "cardiac valve disease" OR "cardiac valve diseases" OR "aortic stenosis" OR "aortic valve stenosis" OR "aortic sclerosis" OR "aortic valve sclerosis" OR "aortic regurgitation" OR "aortic valve regurgitation" OR "aortic insufficiency" OR "aortic valve insufficiency" OR "mitral regurgitation" OR "mitral valve regurgitation" OR "mitral insufficiency" OR "mitral valve insufficiency" OR "mitral stenosis" OR "mitral valve stenosis" OR "tricuspid regurgitation" OR "tricuspid valve regurgitation" OR "tricuspid insufficiency" OR "tricuspid stenosis" OR "tricuspid valve stenosis" OR "pulmonary regurgitation" OR "pulmonic regurgitation" OR "pulmonary valve regurgitation" OR "pulmonary insufficiency" OR "pulmonic insufficiency" OR "pulmonary stenosis" OR "pulmonic stenosis" OR "pulmonary valve stenosis" OR "rheumatic heart disease" OR "rheumatic valvular" OR murmur* OR "structural heart disease" OR "structural murmur" OR "structural murmurs" OR "pathologic murmur" OR "pathologic murmurs" OR "pathological murmur" OR "pathological murmurs"):ti,ab,kw #33 #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 #34 #10 AND #20 AND #33		
Embase	('phonocardiography'/exp OR 'auscultation'/exp OR 'stethoscope'/exp OR (phonocardiograph* OR phonocardiogram* OR PCG OR ((heart OR cardiac) NEAR/1 sound*) OR (cardiac NEAR/2 acoustic*) OR cardioacoustic* OR 'cardio-acoustic' OR 'acoustic cardiography' OR auscultat* OR stethoscop* OR ((digital OR electronic OR smart OR intelligent) NEAR/3 stethoscop*) OR (	06/30/2026	2577

	(digital OR electronic OR automated OR automatic OR 'computer-assisted' OR 'computer assisted' OR 'computer-aided' OR 'computer aided') NEAR/3 auscultat*) OR infrasound):ti,ab,kw) AND ('artificial intelligence'/exp OR 'machine learning'/exp OR 'deep learning'/exp OR 'artificial neural network'/exp OR 'algorithm'/exp OR 'computer assisted diagnosis'/exp OR (AI OR 'artificial intelligence' OR 'machine learning' OR 'deep learning' OR 'neural network' OR 'neural networks' OR 'artificial neural network' OR 'artificial neural networks' OR 'convolutional neural network' OR 'convolutional neural networks' OR CNN OR 'recurrent neural network' OR 'recurrent neural networks' OR RNN OR transformer* OR 'self-attention' OR		

	'support vector machine' OR 'support vector machines' OR SVM OR 'random forest' OR 'random forests' OR 'gradient boosting' OR XGBoost OR classifier* OR classification OR algorithm* OR automat* OR 'computer-aided' OR 'computer aided' OR 'computer-assisted' OR 'computer assisted' OR intelligent OR 'pattern recognition' OR 'signal processing' OR 'transfer learning' OR 'feature extraction'):ti,ab,kw) AND ('valvular heart disease'/exp OR 'heart murmur'/exp OR 'rheumatic heart disease'/exp OR ('valvular heart disease' OR 'valvular heart diseases' OR VHD OR HVD OR 'heart valve disease' OR		

	'heart valve diseases' OR 'valve disease' OR 'valve diseases' OR 'valvular disease' OR 'valvular diseases' OR 'valvular lesion' OR 'valvular lesions' OR 'valve lesion' OR 'valve lesions' OR valvulopath* OR 'valvular dysfunction' OR 'valve dysfunction' OR 'cardiac valve disease' OR 'cardiac valve diseases' OR ((aortic OR mitral OR tricuspid OR pulmonary OR pulmonic) NEAR/3 (stenos* OR regurgitat* OR insufficien* OR scleros*)) OR 'rheumatic heart disease' OR (rheumatic NEAR/3 valv*) OR murmur* OR 'structural heart disease' OR ((structural OR pathologic* OR pathological*) NEAR/2 murmur*)):ti,ab,kw) NOT ('animal'/exp NOT 'human'/exp)		

Supplementary Table S2. Descriptive summary of right-sided valve lesions not included in bivariate pooling.

Lesion	Study/cohort	Participants	Cases	TP	FP	FN	TN	Sensitivity, n/N %	Specificity, n/N %	Notes
TR	Jiang 2024	145	6	6	23	0	116	6/6 (100.0%)	116/139 (83.5%)	zero-cell row present; very small number of cases
TR	Roquemen-Echeverri 2025	995	49	30	406	19	540	30/49 (61.2%)	540/946 (57.1%)	fewer than four independent cohorts; no bivariate pooled estimate calculated
TR	McDonald 2026	156	53	33	19	20	84	33/53 (62.3%)	84/103 (81.6%)	
PR	Roquemen-Echeverri 2025	995	4	1	435	3	556	1/4 (25.0%)	556/991 (56.1%)	very small number of cases; fewer than four independent cohorts; no bivariate pooled estimate calculated
PR	McDonald 2026	109	6	3	19	3	84	3/6 (50.0%)	84/103 (81.6%)	
PS	Roquemen-Echeverri 2025	995	2	2	434	0	559	2/2 (100.0%)	559/993 (56.3%)	zero-cell row present; very small number of cases; fewer than four independent cohorts; no bivariate pooled estimate calculated

TP, true positive; FP, false positive; FN, false negative; TN, true negative; TR, tricuspid regurgitation; PR, pulmonary regurgitation; PS, pulmonary stenosis.

Supplementary Table S3. Detailed Meta-regression coefficients pairwise contrasts and subgroup tests.

Valve lesion	Domain	Outcome	Test type	Contrast	df	χ^2	p-value	β	95% CI
AS	Severity	sensitivity	Global subgroup Wald test		2	3.336	0.189	NA	NA
			Wald contrast	Moderate-or-greater AS vs Any AS / mild-or-greater	1	2.079	0.149	0.445	-0.160 to 1.049

				Severe AS vs Any AS / mild-or-greater	1	3.319	0.068	0.734	-0.056 to 1.524
				Severe AS vs Moderate-or-greater AS	1	1.645	0.2	0.289	-0.153 to 0.732
			Ordinal trend Wald test	ordinal severity rank per one-level increase	1	3.539	0.06	0.361	-0.015 to 0.737
		specificity	Global subgroup Wald test		2	1.037	0.595	NA	NA
			Wald contrast	Moderate-or-greater AS vs Any AS / mild-or-greater	1	0.844	0.358	0.444	-0.504 to 1.393
				Severe AS vs Any AS / mild-or-greater	1	0.006	0.94	0.05	-1.254 to 1.355
				Severe AS vs Moderate-or-greater AS	1	0.435	0.509	-0.394	-1.565 to 0.777
			Ordinal trend Wald test	ordinal severity rank per one-level increase	1	0.001	0.974	-0.011	-0.652 to 0.631
		joint	Joint LRT		4	3.116	0.539	NA	NA
	Location	sensitivity	Global subgroup Wald test		4	38.071	**<0.001**	NA	NA
			Wald contrast	Mitral / apex vs Aortic	1	1.933	0.164	-0.505	-1.218 to 0.207
				Multi-site vs Aortic	1	16.063	**<0.001**	1.356	0.693 to 2.020
				Multi-site vs Mitral / apex	1	26.488	**<0.001**	1.862	1.153 to 2.571
				Multi-site vs Pulmonic	1	20.599	**<0.001**	1.601	0.910 to 2.293

				Multi-site vs Tricuspid / left sternal border	1	23.497	**<0.001**	1.757	1.046 to 2.467
				Pulmonic vs Aortic	1	0.476	0.49	-0.245	-0.939 to 0.450
				Tricuspid / left sternal border vs Aortic	1	1.209	0.272	-0.4	-1.114 to 0.313
		specificity	Global subgroup Wald test		4	2.24	0.692	NA	NA
			Wald contrast	Mitral / apex vs Aortic	1	0.472	0.492	-0.528	-2.033 to 0.978
				Multi-site vs Aortic	1	0.09	0.764	0.181	-0.998 to 1.360
				Multi-site vs Mitral / apex	1	1.188	0.276	0.708	-0.565 to 1.981
				Multi-site vs Pulmonic	1	0.024	0.877	0.103	-1.193 to 1.399
				Multi-site vs Tricuspid / left sternal border	1	1.455	0.228	0.765	-0.478 to 2.007
				Pulmonic vs Aortic	1	0.01	0.92	0.078	-1.447 to 1.603
				Tricuspid / left sternal border vs Aortic	1	0.598	0.439	-0.584	-2.064 to 0.896
		joint	Joint LRT		8	29.33	**<0.001**	NA	NA
	Clinical setting	sensitivity	Global subgroup Wald test		1	0.817	NA	NA	NA
			Wald contrast	Screening / primary care vs Enriched clinical / echo-referred	1	0.817	NA	0.94	-1.099 to 2.978

		specificity	Global subgroup Wald test		1	1.274	NA	NA	NA
			Wald contrast	Screening / primary care vs Enriched clinical / echo-referred	1	1.274	NA	1.47	-1.083 to 4.023
		joint	Joint LRT		2	2.347	NA	NA	NA
	Commercial status	sensitivity	Global subgroup Wald test		1	0.315	0.575	NA	NA
			Wald contrast	Commercial / industry-associated vs Academic / prototype	1	0.315	0.575	0.231	-0.575 to 1.036
		specificity	Global subgroup Wald test		1	0.621	0.431	NA	NA
			Wald contrast	Commercial / industry-associated vs Academic / prototype	1	0.621	0.431	-0.644	-2.246 to 0.958
		joint	Joint LRT		2	0.969	0.616	NA	NA
	Modality	sensitivity	Global subgroup Wald test		1	0.538	NA	NA	NA
			Wald contrast	PCG-only vs Multimodal	1	0.538	NA	-0.766	-2.810 to 1.279
		specificity	Global subgroup Wald test		1	0.029	NA	NA	NA
			Wald contrast	PCG-only vs Multimodal	1	0.029	NA	0.247	-2.609 to 3.103
		joint	Joint LRT		2	0.567	NA	NA	NA
	Validation credibility	sensitivity	Global subgroup Wald test		1	0.084	0.771	NA	NA
			Wald contrast	Lower credibility vs Higher credibility	1	0.084	0.771	0.151	-0.866 to 1.168

		specificity	Global subgroup Wald test		1	0.032	0.857	NA	NA
			Wald contrast	Lower credibility vs Higher credibility	1	0.032	0.857	0.16	-1.578 to 1.897
		joint	Joint LRT		2	0.126	0.939	NA	NA
MR	Severity	sensitivity	Global subgroup Wald test		2	2.573	NA	NA	NA
				Moderate-or-greater MR vs Mild-or-greater MR	1	0.041	NA	0.139	-1.200 to 1.479
				Severe MR vs Mild-or-greater MR	1	1.512	NA	0.968	-0.575 to 2.511
				Severe MR vs Moderate-or-greater MR	1	2.595	0.107	0.829	-0.180 to 1.837
			Ordinal trend Wald test	Ordinal severity rank per one-level increase	1	7.308	**0.007**	0.565	0.155 to 0.975
		specificity	Global subgroup Wald test		2	0.438	NA	NA	NA
				Moderate-or-greater MR vs Mild-or-greater MR	1	0.019	NA	0.156	-2.032 to 2.343
				Severe MR vs Mild-or-greater MR	1	0.062	NA	-0.296	-2.614 to 2.023
				Severe MR vs Moderate-or-greater MR	1	0.432	0.511	-0.451	-1.797 to 0.894
			Ordinal trend Wald test	Ordinal severity rank per one-level increase	1	1.077	0.299	-0.264	-0.761 to 0.234
		joint	Joint LRT		4	3.505	NA	NA	NA
	Location	sensitivity	Global subgroup Wald test		4	12.749	**0.013**	NA	NA

			Wald contrast	Mitral / apex vs Aortic	1	0.228	0.633	0.291	-0.902 to 1.483
				Multi-site / combined locations vs Aortic	1	7.660	**0.006**	1.373	0.401 to 2.345
				Multi-site / combined locations vs Mitral / apex	1	4.737	**0.030**	1.082	0.108 to 2.057
				Multi-site / combined locations vs Pulmonic	1	5.637	**0.018**	1.184	0.207 to 2.162
				Multi-site / combined locations vs Tricuspid / left sternal border	1	3.747	0.053	0.962	-0.012 to 1.936
				Pulmonic vs Aortic	1	0.096	0.757	0.189	-1.006 to 1.384
				Tricuspid / left sternal border vs Aortic	1	0.457	0.499	0.411	-0.781 to 1.603
		specificity	Global subgroup Wald test		4	0.769	0.943	NA	NA
			Wald contrast	Mitral / apex vs Aortic	1	0.000	1.000	0.000	-1.404 to 1.404
				Multi-site / combined locations vs Aortic	1	0.314	0.575	-0.319	-1.436 to 0.798
				Multi-site / combined locations vs Mitral / apex	1	0.314	0.575	-0.319	-1.436 to 0.799
				Multi-site / combined locations vs Pulmonic	1	0.314	0.575	-0.319	-1.436 to 0.800
				Multi-site / combined locations vs Tricuspid / left sternal border	1	0.314	0.575	-0.319	-1.436 to 0.801
				Pulmonic vs Aortic	1	0.000	1.000	0.000	-1.404 to 1.404

				Tricuspid / left sternal border vs Aortic	1	0.000	1.000	0.000	-1.404 to 1.404
		joint	Joint LRT		8	14.085	0.080	NA	NA
	Clinical setting	sensitivity	Global subgroup Wald test		1	3.993	NA	NA	NA
			Wald contrast	Screening / primary care vs Enriched clinical / echo-referred	1	3.993	NA	-0.802	-1.588 to -0.015
		specificity	Global subgroup Wald test		1	2.439	NA	NA	NA
			Wald contrast	Screening / primary care vs Enriched clinical / echo-referred	1	2.439	NA	-1.375	-3.101 to 0.350
		joint	Joint LRT		2	7.832	NA	NA	NA
	Commercial status	sensitivity	Global subgroup Wald test		1	7.084	**0.008**	NA	NA
			Wald contrast	Commercial / industry-associated vs Academic / prototype	1	7.084	**0.008**	-0.93	-1.616 to -0.245
		specificity	Global subgroup Wald test		1	0.979	0.323	NA	NA
			Wald contrast	Commercial / industry-associated vs Academic / prototype	1	0.979	0.323	-0.681	-2.030 to 0.668
		joint	Joint LRT		2	6.788	0.034	NA	NA
	Modality	sensitivity	Global subgroup Wald test		1	1.045	NA	NA	NA
			Wald contrast	PCG-only vs Multimodal	1	1.045	NA	-0.799	-2.330 to 0.733
		specificity	Global subgroup Wald test		1	0.19	NA	NA	NA

			Wald contrast	PCG-only vs Multimodal	1	0.19	NA	0.464	-1.624 to 2.552
		joint	Joint LRT		2	1.324	NA	NA	NA
	Validation credibility	sensitivity	Global subgroup Wald test		1	3.762	0.052	NA	NA
			Wald contrast	Lower credibility vs Higher credibility	1	3.762	0.052	-0.717	-1.441 to 0.008
		specificity	Global subgroup Wald test		1	1.508	0.219	NA	NA
			Wald contrast	Lower credibility vs Higher credibility	1	1.508	0.219	-0.918	-2.383 to 0.547
		joint	Joint LRT		2	6.283	**0.043**	NA	NA

CI, confidence interval; LRT, likelihood-ratio test; NA, not applicable.

Supplementary Table S4. Prevalence assumptions for modelled clinical yield.

Target	Setting	Prevalence used	Age / clinical anchor	Rationale
Clinically significant VHD	General adult/community	2.50%	General adult population	Nkomo population-based study reported age-sex adjusted prevalence of moderate/severe valve disease of 2.5%.[1]
	Primary care/general outpatient	10.00%	Older primary-care population	OxVALVE reported 6.4% undiagnosed clinically significant VHD plus 4.9% pre-existing VHD, total 11.3%; rounded to 10.0%.[4]
	Cardiology/echo-referred	15.00%	Symptomatic echo-referred adults	Large UK referred-echo cohort reported 14.1% moderate/severe valve disease; rounded to 15.0%.[29]
Moderate-to-severe AS	General adult/community	0.50%	General adult/community	Published estimates report calcific AS prevalence of approximately 0.4% in the general population and 1.7% among adults >65 years; population-based data also show a steep age-related increase in AS prevalence. Therefore, 0.5% was used as a rounded low-prevalence adult-community scenario rather than an older screening scenario.[30-31]
	Primary care/general outpatient	1.00%	Older primary-care/outpatient	OxVALVE reported AS present in 1.3% of adults ≥65 in primary care; rounded down to 1.0% for moderate-to-severe AS modelling.[4]
	Cardiology/echo-referred	4.50%	Older or AS-enriched referred population	Enriched scenario; severe AS prevalence alone is reported around 3.4% in adults ≥75, so 4.5% is reasonable for moderate-to-severe AS in referred/enriched settings.[32]
Moderate-to-severe MR	General adult/community	0.50%	General adult/community	Global moderate-to-severe MR prevalence reported as 0.67%; rounded to 0.5%.[33]
	Primary care/general outpatient	3.00%	Older outpatient, approximately age 70	Same MR meta-analysis estimated prevalence around 2.85% at age 70; rounded to 3.0%.[33]
	Cardiology/echo-referred	6.50%	Very elderly or high-prevalence referred setting	Same MR meta-analysis estimated 6.45% by age 90; rounded to 6.5%.[33]

VHD, valvular heart disease; AS, aortic stenosis; MR, mitral regurgitation.

Supplementary Table S5. Modelled diagnostic consequences per 1,000 adults tested.

Target	Prevalence scenario	Prevalence	Sensitivity	Specificity	TP / 1,000	FN / 1,000	FP / 1,000	AI-positive echo referrals / 1,000	PPV	NNE
Clinically significant VHD	General adult/community	2.50%	69.90%	75.60%	17	8	238	255	6.80%	14.6
	Primary care/general outpatient	10.00%			70	30	220	290	24.10%	4.1
	Cardiology/echo-referred	15.00%			105	45	207	312	33.60%	3.0
Moderate-to-severe AS	General adult/community	0.50%	90.40%	88.00%	5	0–1	119	124	3.60%	27.4
	Primary care/general outpatient	1.00%			9	1	119	128	7.10%	14.1
	Cardiology/echo-referred	4.50%			41	4	115	155	26.20%	3.8
Moderate-to-severe MR	General adult/community	0.50%	70.20%	77.10%	4	1–2	228	231	1.50%	65.9
	Primary care/general outpatient	3.00%			21	9	222	243	8.70%	11.5
	Cardiology/echo-referred	6.50%			46	19	214	260	17.60%	5.7

PPV, positive predictive value; NNE, number needed to echo.

Supplementary Table S6. GRADE-DTA evidence profiles and certainty-of-evidence assessments.

GRADE, Grading of Recommendations Assessment, Development and Evaluation; DTA: diagnostic test accuracy; CI: confidence interval; QoE: quality of evidence; VHD: valvular heart disease; AS: aortic stenosis; MR: mitral regurgitation.

Review question 1: AI/digital stethoscope for clinically significant VHD detection.

Pooled sensitivity	69.9% (47.9–85.4%)
Pooled specificity	75.6% (58.4-87.3%)
Assumed prevalence / pre-test probability	General adult/community 2.5%; Primary care/general outpatient 10%; Cardiology/echo-referred 15%

Outcome	Number of studies and patients	Study design	Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Effect per 1,000 patients tested	Test accuracy QoE
True positives	4 studies; 2,730 patients	Diagnostic accuracy cohort/cross-sectional studies	Serious[a]	Serious[b]	Serious[c]	Very serious[d]	No serious concern[f]	Community (2.5%): 17 (12-21) Primary care (10%): 70 (48-85) Cardiology (15%): 105 (72-128)	⨁○○○ VERY LOW
False negatives	4 studies; 2,730 patients	Diagnostic accuracy cohort/cross-sectional studies	Serious[a]	Serious[b]	Serious[c]	Very serious[d]	No serious concern[f]	Community (2.5%): 8 (4-13) Primary care (10%): 30 (15-52) Cardiology (15%): 45 (22-78)	⨁○○○ VERY LOW

True negatives	4 studies; 2,730 patients	Diagnostic accuracy cohort/cross-sectional studies	No serious concern	Serious[b]	Serious[c]	Serious[e]	No serious concern[f]	Community (2.5%): 737 (570-851) Primary care (10%): 681 (526-785) Cardiology (15%): 643 (497-742)	⨁○○○ VERY LOW
False positives	4 studies; 2,730 patients	Diagnostic accuracy cohort/cross-sectional studies	No serious concern	Serious[b]	Serious[c]	Serious[e]	No serious concern[f]	Community (2.5%): 238 (124-405) Primary care (10%): 219 (115-374) Cardiology (15%): 207 (108-353)	⨁○○○ VERY LOW

[a] Risk of bias downgraded one level for sensitivity-related outcomes because influential studies had high or some-concern judgments.

[b] Indirectness downgraded one levels because high applicability concerns had substantial weighted influence.

[c] Inconsistency downgraded one level because single-study estimates varied materially (sensitivity range 0.46; specificity range 0.34).

[d] Imprecision downgraded two level because confidence intervals were wide (sensitivity CI width 0.34; specificity CI width 0.23).

[e] Imprecision downgraded one level for specificity-related outcomes because applicability concerns were present but did not justify a two-level downgrade.

[f] Publication bias was not downgraded; direct evidence was not identified, but small/developmental/proprietary algorithms warrant caution.

Review question 2: AI/digital stethoscope for AS detection.

Pooled sensitivity	90.4% (86.5-93.3%)
Pooled specificity	88.0% (78.0-93.8%)
Assumed prevalence / pre-test probability	General adult/community 0.5%; Primary care/general outpatient 1%; Cardiology/echo-referred 4.5%

Outcome	Number of studies and patients	Study design	Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Effect per 1,000 patients tested	Test accuracy QoE
True positives	12 studies; 4,607 patients	Diagnostic accuracy cohort/cross-sectional studies	Serious[a]	Serious[b]	No serious concern	No serious concern	No serious concern[d]	Community (0.5%): 5 (4-5) Primary care (1%): 9 (9-9) Cardiology (4.5%): 41 (39 to 42)	⨁⨁◯◯ LOW
False negatives	12 studies; 4,607 patients	Diagnostic accuracy cohort/cross-sectional studies	Serious[a]	Serious[b]	No serious concern	No serious concern	No serious concern[d]	Community (0.5%): 0 (0 to 1) Primary care (1%): 1 (1 to 1) Cardiology (4.5%): 4 (3 to 6)	⨁⨁◯◯ LOW
True negatives	12 studies; 4,607 patients	Diagnostic accuracy cohort/cross-sectional studies	No serious concern	Serious[b]	Serious[c]	No serious concern	No serious concern[d]	Community (0.5%): 876 (776 to 934) Primary care (1%): 871 (772 to 929) Cardiology	⨁⨁◯◯ LOW

								(4.5%): 841 (745 to 896)	
False positives	12 studies; 4,607 patients	Diagnostic accuracy cohort/cross-sectional studies	No serious concern	Serious[b]	Serious[c]	No serious concern	No serious concern[d]	Community (0.5%): 119 (61 to 219) Primary care (1%): 119 (61 to 218) Cardiology (4.5%): 114 (59 to 210)	⨁⨁○○ LOW

[a] Risk of bias downgraded one level for sensitivity-related outcomes because influential AS studies had high or some-concern judgments.

[b] Indirectness downgraded one level because the evidence body included applicability concerns in population, setting, index test, or endpoint definitions.

[c] Inconsistency downgraded one level for specificity-related outcomes because single-study estimates varied materially (specificity range 0.42).

[d] Publication bias was not downgraded; direct evidence was not identified, but small/developmental/proprietary algorithms warrant caution.

Review question 3: AI/digital stethoscope for MR detection.

Pooled sensitivity	70.2% (61.0-77.9%)
Pooled specificity	77.1% (64.0-86.4%)
Assumed prevalence / pre-test probability	General adult/community 0.5%; Primary care/general outpatient 3%; Cardiology/echo-referred 6.5%

Outcome	Number of studies and patients	Study design	Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Effect per 1,000 patients tested	Test accuracy QoE
True positives	8 studies; 3,754 patients	Diagnostic accuracy cohort/cross-sectional studies	Serious[a]	Serious[b]	Serious[c]	No serious concern	No serious concern[f]	Community (0.5%): 4 (3 to 4) Primary care (3%): 21 (18 to 23) Cardiology (6.5%): 46 (40 to 51)	⨁○○○ VERY LOW
False negatives	8 studies; 3,754 patients	Diagnostic accuracy cohort/cross-sectional studies	Serious[a]	Serious[b]	Serious[c]	No serious concern	No serious concern[f]	Community (0.5%): 1 (1 to 2) Primary care (3%): 9 (7 to 12) Cardiology (6.5%): 19 (14 to 25)	⨁○○○ VERY LOW
True negatives	8 studies; 3,754 patients	Diagnostic accuracy cohort/cross-sectional studies	Serious[d]	Serious[b]	Serious[c]	Serious[e]	No serious concern[f]	Community (0.5%): 767 (637 to 860) Primary care (3%): 748 (621 to 838)	⨁○○○ VERY LOW

								Cardiology (6.5%): 721 (599 to 808)	
False positives	8 studies; 3,754 patients	Diagnostic accuracy cohort/cross-sectional studies	Serious[d]	Serious[b]	Serious[c]	Serious[e]	No serious concern[f]	Community (0.5%): 228 (135 to 358) Primary care (3%): 222 (132 to 349) Cardiology (6.5%): 214 (127 to 336)	⨁○○○ VERY LOW

[a] Risk of bias downgraded one level for sensitivity-related outcomes because influential studies had high or some-concern judgments.

[b] Indirectness downgraded one level because applicability concerns were present but did not justify a two-level downgrade.

[c] Inconsistency downgraded one level because single-study estimates varied materially (sensitivity range 0.29; specificity range 0.44).

[d] Risk of bias downgraded one level for specificity-related outcomes because the mixed evidence body had material high/some-concern contribution.

[e] Imprecision downgraded one level for specificity-related outcomes because the specificity CI was wide (CI width 0.22).

[f] Publication bias was not downgraded; direct evidence was not identified, but small/developmental/proprietary algorithms warrant caution.