International Platform of Registered Systematic Review and Meta-analysis Protocols

INPLASY

INPLASY202650062

doi: 10.37766/inplasy2026.5.0062

Received: 10 May 2026

Published: 10 May 2026

Corresponding author:
Chih-Yu Yang

cyyang3@vghtpe.gov.tw

Author Affiliation:
Institute of Clinical Medicine, National Yang Ming Chiao Tung University.

Diagnostic accuracy and clinical yield of AI-enabled digital auscultation and phonocardiography for clinically significant valvular heart disease in adults: a systematic review and bivariate meta-analysis

Chen, YC; Wang, YC; Yang, CY.

ADMINISTRATIVE INFORMATION

Support - This review has received no specific financial support.

Review Stage at time of this submission - Preliminary searches.

Conflicts of interest - None declared.

INPLASY registration number: INPLASY202650062

Amendments - This protocol was registered with the International Platform of Registered Systematic Review and Meta-Analysis Protocols (INPLASY) on 10 May 2026 and was last updated on 10 May 2026.

INTRODUCTION

Review question / Objective Among adults aged 18 years or older, what is the diagnostic accuracy and clinical yield of artificial intelligence-enabled digital auscultation or phonocardiography for detecting echocardiography-confirmed clinically significant valvular heart disease?

The primary objective is to estimate the pooled sensitivity and specificity of AI-enabled digital auscultation or phonocardiography for detecting clinically significant valvular heart disease in adults, primarily defined as moderate-or-greater valvular disease on echocardiography.

Secondary objectives are to evaluate accuracy by lesion type, disease severity, clinical setting, index-test modality, validation design, failed-recording handling, and comparative performance versus clinician auscultation where paired data are available.

Condition being studied The condition being studied is adult valvular heart disease, including aortic stenosis, aortic regurgitation, mitral regurgitation, mitral stenosis, tricuspid regurgitation, tricuspid stenosis, and pulmonary valve disease. The primary target condition will be clinically significant valvular heart disease, operationally defined as moderate-or-greater valvular disease confirmed by transthoracic or transesophageal echocardiography, or author-defined clinically significant valvular disease requiring follow-up, referral, surveillance, or intervention. Severe valvular disease and lesion-specific targets will be evaluated as secondary outcomes.

METHODS

Participant or population Adults aged 18 years or older undergoing AI-enabled digital cardiac auscultation, phonocardiography, electronic/digital/smart stethoscope assessment, or smartphone-based heart sound recording, with echocardiographic assessment available as the

reference standard. Eligible settings will include primary care, community screening, cardiology outpatient clinics, echocardiography-referred populations, inpatient settings, and other adult clinical settings. Mixed adult and pediatric studies will be included only if adult data can be extracted separately.

Intervention The index test will be AI-enabled digital auscultation or phonocardiography, including electronic, digital, or smart stethoscope systems, phonocardiography-based machine learning or deep learning models, and smartphone-recorded heart sound AI systems used for cardiac auscultation. Multimodal AI stethoscope systems incorporating phonocardiography and single-lead ECG will be eligible and analyzed separately when the heart sound or phonocardiography component contributes to the index test.

Comparator The primary reference standard will be transthoracic or transesophageal echocardiography interpreted through clinical reports, expert cardiologist review, or core laboratory assessment. Where available, clinician auscultation, including general practitioner, primary care physician, or cardiologist auscultation, will be evaluated as a comparative index test. Human auscultation alone will not be accepted as the reference standard for valvular heart disease severity.

Study designs to be included Prospective or retrospective diagnostic accuracy studies, cross-sectional validation studies, cohort studies, nested diagnostic test accuracy studies within trials, and implementation studies reporting adult VHD diagnostic accuracy or clinical yield data. Case-control and internal validation studies will be included in secondary or sensitivity analyses where appropriate.

Eligibility criteria We will include studies of adults aged 18 years or older that evaluate AI-enabled digital auscultation, phonocardiography, electronic/digital/smart stethoscope systems, or smartphone-based heart sound AI for detecting echocardiography-confirmed valvular heart disease. Eligible studies must use transthoracic or transesophageal echocardiography as the reference standard and report sufficient information to extract or reconstruct participant-level TP, FP, FN, and TN values, or provide diagnostic accuracy or clinical yield data relevant to adult VHD.

We will exclude pediatric-only studies, congenital heart disease-only pediatric cohorts, animal studies, simulation-only studies, public dataset-only classification studies without adult clinical VHD reference standard, human auscultation without AI, AI-ECG studies without phonocardiography or heart sound input, AI echocardiography studies, lung sound AI studies, reviews, editorials, comments, and technical papers without clinical validation. Studies reporting only normal versus abnormal heart sound classification without VHD-specific diagnosis will be excluded from the primary analysis.

Information sources Electronic databases will include MEDLINE via PubMed, Web of Science Core Collection, and Cochrane CENTRAL. Additional sources will include reference lists of included studies and relevant reviews, forward citation chasing in Web of Science, ClinicalTrials.gov, WHO International Clinical Trials Registry Platform where appropriate, and author contact for missing or unclear diagnostic 2×2 data. Preprints and conference abstracts will be considered for narrative synthesis or sensitivity analysis if sufficient diagnostic accuracy data and reference standard information are available.

Main outcome(s) The main outcomes will be pooled sensitivity and specificity of AI-enabled digital auscultation or phonocardiography for detecting echocardiography-confirmed clinically significant valvular heart disease in adults. Clinically significant VHD will primarily be defined as moderate-or-greater valvular disease on echocardiography. Participant-level TP, FP, FN, and TN data will be extracted whenever possible. Echocardiography performed on the same day or within 30 days of the index test will be prioritized; a window up to 90 days will be accepted if no valve intervention or major clinical event occurred.

Quality assessment / Risk of bias analysis Risk of bias and applicability will be assessed using the QUADAS-2/QUADAS-3 framework for diagnostic accuracy studies. Domains will include patient selection, index test, reference standard, and flow/timing, with attention to applicability for adult clinically significant VHD and echocardiography triage. Comparative studies evaluating AI-enabled auscultation against clinician auscultation will additionally be assessed using QUADAS-C principles where appropriate. AI-specific reporting and applicability items adapted from STARD-AI will be extracted, including patient-level data splitting, external validation, locked thresholds, algorithm version, non-analyzable recordings, demographic fairness, and manufacturer involvement.

Strategy of data synthesis For each eligible study and target condition, participant-level TP, FP, FN, and TN values will be extracted or reconstructed. The primary meta-analysis will use a bivariate random-effects model to jointly pool sensitivity and specificity for clinically significant VHD. HSROC models will be used where thresholds vary across studies or where multiple operating points are reported. Summary estimates will include pooled sensitivity, specificity, 95% confidence intervals, 95% prediction regions where possible, LR+, LR–, diagnostic odds ratio, and HSROC plots.

Quantitative pooling will be attempted when at least four independent cohorts report sufficient data for the same target condition. If fewer studies are available or model convergence fails, results will be summarized narratively with paired forest plots. Clinical yield outcomes will be summarized by setting and, where appropriate, pooled using random-effects models for proportions. PPV and NPV will be interpreted according to disease prevalence and setting.

Subgroup analysis Pre-specified subgroup analyses will examine target condition, including any clinically significant VHD, left-sided VHD, aortic stenosis, severe aortic stenosis, mitral regurgitation, aortic regurgitation, mitral stenosis, tricuspid valve disease, and audible structural murmur; disease severity, including moderate-or-greater versus severe disease; clinical setting, including primary care/community screening, cardiology clinic, echo-referred populations, and inpatient settings; index-test modality, including PCG-only, digital stethoscope, PCG plus single-lead ECG, PCG plus clinical variables, and smartphone-based PCG; validation design, including prospective external validation, retrospective external validation, temporal validation, internal validation, cross-validation, and case-control design; and commercial versus academic AI systems.

Sensitivity analysis Sensitivity analyses will exclude studies at high risk of bias, case-control designs, training-set or resubstitution estimates, cross-validation-only studies, studies without independent validation, studies without blinded reference standard interpretation, studies with unclear index-reference timing, studies lacking participant-level diagnostic data, studies using murmur or expert auscultation as part of the ground truth, studies without echo-confirmed moderate-or-greater VHD, conference abstracts or preprints, and studies with industry employees as authors. Additional sensitivity analyses will include only prospective studies, only external validation studies, only PCG-only models, and analyses incorporating failed or non-analyzable recordings.

Country(ies) involved Taiwan.

Keywords valvular heart disease; digital auscultation; phonocardiography; artificial intelligence; diagnostic accuracy; meta-analysis.

Contributions of each author

Author 1 - Yu Chuan Chen.
Email: ycchen.md08@nycu.edu.tw
Author 2 - Yun Chu Wang.
Email: yunzuwang@gmail.com
Author 3 - Chih Yu Yang.
Email: cyyang3@vghtpe.gov.tw