INPLASY

Diagnostic Accuracy of Artificial Intelligence Applications in Echocardiography for Detecting Left Ventricular Systolic Dysfunction: A Systematic Review and Meta-analysis

INPLASY202620071

doi: 10.37766/inplasy2026.2.0071

Received: 24 February 2026

Published: 24 February 2026

Dang, HNN; Luong, VT; Marwick, TH; Huynh, Q.

Corresponding author:
Hai Nguyen Ngoc Dang

hai.dang@utas.edu.au

Author Affiliation:
Menzies Institute for Medical Research - University of Tasmania.

ADMINISTRATIVE INFORMATION

Support - None.

Review Stage at time of this submission - Formal screening of search results against eligibility criteria.

Conflicts of interest - None declared.

INPLASY registration number: INPLASY202620071

Amendments - This protocol was registered with the International Platform of Registered Systematic Review and Meta-Analysis Protocols (INPLASY) on 24 February 2026 and was last updated on 24 February 2026.

INTRODUCTION

Review question / Objective In patients undergoing transthoracic echocardiography, what is the diagnostic accuracy of artificial intelligence–guided echocardiography algorithms for detecting left ventricular systolic dysfunction?

Rationale Heart failure is a rapidly growing global health burden, with prevalence increasing as populations age and survival improves. Left ventricular systolic dysfunction (LVSD), commonly operationalized by reduced left ventricular ejection fraction (LVEF), is central to heart failure phenotyping and guides major management decisions; echocardiography is the cornerstone modality for evaluating LVEF in clinical practice. However, echocardiographic acquisition and LVEF quantification are operator-dependent and subject to measurement variability, which can limit scalability and timely detection in settings with constrained expert capacity. Recent advances in artificial intelligence (AI) have enabled "AI-guided echocardiography" solutions that support image acquisition and automate LVEF estimation on handheld/point-of-care platforms, with early clinical studies suggesting that such tools may help non-expert users identify reduced LVEF/LVSD for screening or triage. Nonetheless, the available evidence remains fragmented across heterogeneous AI approaches, devices, clinical settings, LVSD thresholds, and reference standards, leaving uncertainty about overall diagnostic accuracy and generalisability. Therefore, a systematic review and diagnostic test accuracy meta-analysis is needed to provide pooled estimates of sensitivity and specificity, and inform safe clinical implementation and future research.

Condition being studied Left ventricular systolic dysfunction (LVSD) refers to impaired left ventricular contractile function and is most commonly operationalised as reduced LVEF on

echocardiography. LVSD is clinically important because reduced LVEF is used to classify heart failure phenotypes and guide evidence-based therapy. LVSD may be defined using thresholds such as LVEF ≤40% (HFrEF).

METHODS

Search strategy We will search PubMed/MEDLINE, CENTRAL, Embase, LILACS and Scopus from inception to the search date. We will also search ClinicalTrials.gov and the WHO International Clinical Trials Registry Platform (ICTRP) for completed/ongoing studies. Search concepts will combine: (1) AI terms (artificial intelligence; machine learning; deep learning; neural network*; automated), (2) echocardiography terms (echocardiograph*; echocardiogram*; transthoracic; handheld; point-of-care ultrasound/POCUS), (3) target condition terms (left ventricular systolic dysfunction; reduced ejection fraction; LVEF), and (4) diagnostic accuracy terms (sensitivity; specificity; ROC; AUC; validation). We will adapt syntax for each database, use both controlled vocabulary (e.g., MeSH) and free-text, and apply no language limits at search stage. We will also screen reference lists of included studies and relevant reviews.

Participant or population Human participants (any age) undergoing transthoracic echocardiography (standard, handheld, or point-of-care) in any clinical setting (e.g., outpatient, inpatient, emergency/ICU, community screening). Studies must evaluate detection/classification of LVSD (reduced LVEF) in these participants.

Intervention Index test: AI-guided echocardiography applications that assist acquisition and/or automatically estimate or classify LVEF/LV systolic function from transthoracic echocardiographic images/videos (including handheld or point-of-care devices).

Comparator Reference standard: expert echocardiography interpretation with quantitative LVEF measurement (preferably biplane Simpson's method), and/or cardiac MRI where available, or other prespecified accepted reference standards reported by the study.

Study designs to be included Prospective or retrospective diagnostic test accuracy studies (cross-sectional or cohort) reporting sufficient data to derive TP/FP/FN/TN for AI-guided echocardiography versus a reference standard.

Eligibility criteria Inclusion: full-text human studies evaluating AI-guided transthoracic echocardiography for detecting LVSD and providing (or allowing derivation of) TP/FP/FN/TN at a defined threshold. Exclusion: conference abstracts/abstract-only reports; trial registrations without results; studies not assessing diagnostic accuracy for LVSD (wrong scope/target condition); wrong population/setting (non-TTE or non-human); wrong index test (AI not eligible or not echocardiography-based); wrong reference standard; and studies with insufficient data for 2×2 tables.

Information sources Electronic databases: PubMed/MEDLINE, CENTRAL, Embase, LILACS and Scopus. Trial registries: ClinicalTrials.gov and WHO ICTRP. Additional sources: reference lists and forward citation tracking of included studies; contacting corresponding authors when key 2×2 data are missing; and searching preprint servers/grey literature sources if needed to identify unpublished full-text reports (with eligibility applied at full text).

Main outcome(s) Primary outcomes: pooled sensitivity and specificity of AI-guided echocardiography for detecting LVSD at the study-defined threshold(s), with 95% confidence intervals and prediction regions. Effect measures will be derived from 2×2 tables (TP/FP/FN/TN).

Additional outcome(s) Secondary outcomes: positive and negative likelihood ratios (LR+/LR–), diagnostic odds ratio (DOR), area under the ROC curve (AUC)/SROC parameters, and agreement of continuous LVEF estimates with the reference standard when reported (e.g., mean difference or correlation) as supportive evidence.

Data management Records will be imported into Covidence for automatic de-duplication, and then managed within Covidence for title/abstract and full-text screening. Data extraction will be performed using a piloted extraction form in Covidence, and the extracted dataset will be exported to Excel for analysis, capturing study design, setting, participant characteristics, AI model details, LVSD threshold, reference standard, and 2×2 outcomes (TP/FP/FN/TN).

Quality assessment / Risk of bias analysis Risk of bias and applicability will be assessed independently by two reviewers using QUADAS-2 across four domains (patient selection, index test, reference standard, flow and timing). Disagreements will be resolved by discussion or a

third reviewer. Results will be presented in tabular and graphical summaries.

Strategy of data synthesis If ≥4 clinically comparable studies are available, we will pool sensitivity and specificity using hierarchical random-effects models (bivariate and/or HSROC), generating summary points and SROC curves with 95% confidence and prediction regions. When multiple thresholds are reported, we will prioritise the prespecified LVSD threshold (or the most commonly used threshold) and explore threshold effects using HSROC or meta-regression. If meta-analysis is inappropriate due to extreme heterogeneity or sparse data, we will provide a structured narrative synthesis with paired estimates of sensitivity and specificity. Analyses will be conducted in R using appropriate DTA meta-analysis packages.

Subgroup analysis Planned subgroup/meta-regression analyses (as data permit): clinical setting (ED/ICU vs outpatient/community); device type (handheld/POCUS vs standard systems); operator expertise (novice vs expert); AI task/type (acquisition guidance vs automated LVEF estimation/classification; ML vs deep learning); LVSD definition/threshold (e.g., LVEF ≤40% vs <50%); and reference standard (echo vs CMR).

Sensitivity analysis Sensitivity analyses (as data permit): excluding studies at high risk of bias in key QUADAS-2 domains; excluding studies with non-standard reference standards; using alternative thresholds where multiple cut-offs are reported; including only studies with prospective design or external validation; and leave-one-out analyses to assess influence of individual studies.

Language restriction No language restrictions will be applied at the search stage.

Country(ies) involved Australia, Vietnam.

Other relevant information Reporting will follow PRISMA-DTA, and search reporting will align with PRISMA-S. Where possible, we will obtain unpublished 2×2 data by contacting authors. Certainty of evidence may be assessed using GRADE for diagnostic tests/strategies.

Keywords Artificial intelligence; echocardiography; left ventricular systolic dysfunction; ejection fraction; diagnostic accuracy; meta-analysis.

Contributions of each author

Author 1 - Hai Nguyen Ngoc Dang.
Email: hai.dang@utas.edu.au
Author 2 - Thang Viet Luong.
Email: vietthang.luong@utas.edu.au
Author 3 - Thomas H Marwick.
Author 4 - Quan Huynh.