

MINI-REVIEW

Vision-language models in transthoracic echocardiography: A narrative mini-review of promise and challenges

Francesca D'Auria^{1,2*}  and Danilo Flavio Santo³¹Cardiovascular Outpatient Clinic, Health Hub – Heart Hub, Teggiano, Salerno, Italy²Cardiovascular Department, Università degli Studi di Salerno, Salerno, Italy³Cardiovascular Outpatient Clinic, Santagostino Medical Center, Milano, Italy

Abstract

Echocardiography remains the cornerstone of non-invasive cardiac imaging. For patients with suspected or known cardiac disease—including heart failure, cardiomyopathies, valvular heart disease, or ischemic heart disease—transthoracic echocardiography (TTE) enables assessment of cardiac structure, function, valves, and hemodynamic. However, TTE interpretation is often labor-intensive, operator-dependent, and subject to inter- and intra-observer variability. Recently, advancements in deep learning (DL) and artificial intelligence (AI) have yielded automated or semi-automated tools for echocardiographic tasks, such as view classification, chamber segmentation, quantification of volumes and ejection fraction, and even detection of pathological patterns. The latest generation of AI—namely vision-language models (VLMs) and foundation models trained on vast datasets of echocardiogram videos paired with clinical reports—promise to push boundaries further: from isolated measurements to comprehensive, report-level interpretation of entire TTE studies. In this narrative mini-review, we examine whether, in patients undergoing TTE, the use of large VLMs vs standard echocardiographer interpretation without AI can improve diagnostic accuracy, report quality, efficiency, reduce variability, and maintain safety. We appraise current evidence, highlight strengths and limitations, and discuss future perspectives and challenges.

***Corresponding author:**
Francesca D'Auria
(f.dauria@hotmail.com)

Citation: D'Auria F, Santo DF.
Vision-language models in
transthoracic echocardiography:
A narrative mini-review of promise
and challenges. *Artif Intell Health*.
2026;3(4):025010120.
doi: 10.36922/AIH025010120

Received: December 31, 2025

Revised: March 16, 2026

Accepted: March 30, 2026

Published online: May 5, 2026

Copyright: © 2026 Author(s).
This is an Open-Access article
distributed under the terms of the
Creative Commons Attribution
License, permitting distribution,
and reproduction in any medium,
provided the original work is
properly cited.

Publisher's Note: AccScience
Publishing remains neutral with
regard to jurisdictional claims in
published maps and institutional
affiliations.

Keywords: Artificial intelligence; Transthoracic echocardiography; Deep learning; Vision-language models

1. Introduction

Transthoracic echocardiography (TTE) remains the most widely used non-invasive imaging modality for evaluating cardiac structure and function, guiding clinical decision-making across a spectrum of cardiovascular diseases. Traditional interpretation relies heavily on operator expertise, and even well-trained clinicians may demonstrate inter-observer variability in quantification and diagnostic reporting. Recent advances in artificial intelligence (AI) promise to address these limitations, with particular interest in vision-language models (VLMs) that integrate visual and textual clinical data to generate multimodal representations surpassing task-specific approaches. VLMs are designed to jointly interpret echocardiographic images or videos and associated clinical reports,

enabling broader diagnostic capabilities and potentially reducing the need for extensive manual labeling inherent to supervised learning. A notable example is EchoCLIP, a vision-language foundation model trained on over a million paired echocardiogram videos and expert textual interpretations, which can assess cardiac function and identify intracardiac devices without being explicitly trained on each individual task.¹ Beyond single-task performance, emerging models such as EchoPrime leverage multi-view and view-informed learning to synthesize comprehensive echocardiographic studies, demonstrating state-of-the-art performance across diverse interpretation benchmarks. Meanwhile, frameworks such as EchoVLM incorporate measurement-grounded contrastive learning to better align image features with clinically relevant quantitative assessments. These developments illustrate the potential of VLMs to standardize interpretation, support real-time reporting, and enhance reproducibility in clinical practice.² However, challenges persist, including data heterogeneity, the need for large annotated datasets, and the complexities of integrating AI models into routine echocardiographic workflows. This narrative mini-review outlines the current promise VLM in TTE and examines key obstacles to their safe and effective adoption. We examine whether, in patients undergoing TTE, the use of large VLMs vs standard echocardiographer interpretation without AI can improve diagnostic accuracy, report quality, efficiency, reduce variability, and maintain safety. We appraise current evidence, highlight strengths and limitations, and discuss future perspectives and challenges.

2. Material and methods

This paper is a narrative mini-review on VLM in TTE, rather than a systematic review, due to the highly heterogeneous and rapidly evolving nature of AI applications in echocardiography. The available literature encompasses a wide range of methodologies (*e.g.*, supervised and unsupervised learning, deep learning architectures), datasets, clinical endpoints, and performance metrics, which limits the feasibility of a structured quantitative synthesis. Furthermore, the field lacks standardized reporting frameworks and uniform validation strategies, making direct comparison across studies challenging. Our aim is therefore to provide a concise, clinically oriented overview of current developments, emerging applications, and future perspectives. The field of AI in TTE is still evolving, with limited standardization in study design, validation cohorts, and outcome reporting. In this context, a systematic review might risk premature conclusions, whereas a narrative mini-review allows a balanced and critical appraisal of ongoing developments. Our literature search and data extraction for eligible studies were

performed by consulting the Cochrane Central Register of Controlled Trials (CENTRAL Internet), MEDLINE, EMBASE, and official website of the main ultrasound manufacturers without date or language restrictions. Keywords and Medical Subject Headings terms pertinent to the exposure of interest were used in relevant combinations: “Artificial Intelligence,” “AI,” “Transthoracic Echocardiography,” “TTE,” “Deep Learning,” “DL,” “Vision-Language Models,” and “VLM.” A literature search was conducted from January 1, 2018 to December 31, 2025. In addition, we searched trial registries, and the reference lists were carefully analyzed for pertinent studies. Supplementary documents of the selected studies were also assessed, if available. Two primary investigators (F.D.A. and D.F.S.) independently screened the titles and abstracts and resolved any disagreements. After excluding non-relevant studies, the full texts of potentially relevant articles were screened for inclusion in the final narrative mini-review by the two primary investigators.

3. Results

3.1. Challenges of standard echocardiography interpretation

TTE interpretation relies heavily on operator skill: from image acquisition (proper acoustic windows, standardized views, *etc.*) to manual tracing, volume calculation, and clinical synthesis. Variability in experience and technique introduces substantial inter- and intra-observer variability, which can impact diagnostic consistency and comparability over time or across centers.^{3,4} Advanced measurements (*e.g.*, global longitudinal strain, volumetric analysis, 3D reconstructions) demand expertise, time, and consistent image quality, often limiting their availability in routine practice.^{4–6} The increasing demand for echocardiograms, which is driven by aging populations and rising prevalence of cardiac disease, strains human resources, potentially leading to delays, variability in report quality, and overburdened echocardiography labs. These challenges create a strong rationale for leveraging AI to support and possibly augment the echocardiographic workflow.

3.2. Evolution of AI in echocardiography: From single-task AI to vision-language models

Early AI applications in echo focused on narrow tasks: segmentation, view classification, calculation of ejection fraction (EF), volumes, and strain, often on single views or single frames.^{4,5,7} Over time, more advanced systems emerged. There are models performing segmentation of cardiac chambers (*e.g.*, left ventricle, atria) on multiple views, enabling automated measurement of volumes, left ventricular (LV) mass, EF, even strain.^{4,6,8} AI workflows that process entire echo studies (multiple video loops, various

views, Doppler, *etc.*) and compute a comprehensive set of parameters.^{6,9,10} More recently, foundation-scale VLMs for echocardiography, trained on millions of video-report pairs, capable of holistic interpretation of a complete TTE study, integrating structural, functional, and pathological assessments—not just numeric parameters. An example is EchoPrime.¹¹ The availability of such VLMs raises the possibility that AI could transition from being a helper for single tasks to being an assistant (or even preliminary “reader”) for full echocardiographic interpretation. Here we review the empirical evidence (clinical and technical) of AI-based echocardiography, focusing on the main relevant outcomes like diagnostic accuracy, report quality, efficiency, variability, and safety.

3.3. Diagnostic accuracy and quantitative measurements

A landmark study describing PanEcho—an AI system for fully automated echocardiogram interpretation—evaluated 1.2 million echo videos from over 32,000 TTE studies (24,405 patients), across multiple centers. The AI performed 18 diagnostic classification tasks (*e.g.*, left or right ventricular dysfunction, valvular disease, structural abnormalities) and 21 parameter estimation tasks (volumes, EF, Doppler-derived measures, *etc.*). Median area under the curve (AUC) for classification tasks was 0.91 (IQR 0.88–0.93); median normalized mean absolute error (nMAE) for parameter estimation tasks was 0.13 (IQR 0.10–0.18). For example, for EF estimation the mean absolute error (MAE) was ~4.2% (internal) and 4.5% (external). Detection of moderate or worse LV systolic dysfunction had AUC up to 0.98–0.99; detection of severe aortic stenosis reached AUC of 0.98–1.00.⁹ These results demonstrate that AI systems can achieve high accuracy across a broad set of clinically relevant parameters and diagnoses, rivaling expert human interpretation—at least in retrospective and external-validation datasets. Based on these findings, AI can guarantee a comprehensive automated interpretation of the echocardiographic images. Several earlier systems based on deep learning (DL) employed convolutional neural networks (CNNs) to segment cardiac chambers across standard echo views and derive quantitative measurements. In one of the first large studies, segmentation models operating on five standard views produced LV ejection fraction, volumes, LV mass, and global longitudinal strain (GLS). In that cohort, the median absolute deviation (MAD) for left ventricular ejection fraction (LVEF) was ~9.7% across >6,400 studies; for volumes and mass MAD was 15–17%, and for strain MAD was 7.5–9.0%.^{5,11} A prospective monocentric study on 1,083 consecutive patients compared AI-based

EF measurement (automatic view selection and single workflow) with the modified biplane Simpson method (MBS) by humans. In ~81% of echocardiographic examinations, the AI successfully identified apical LV views and measured EF. Correlation between AI and MBS was strong ($R = 0.87$ – 0.93 for EF and volumes). Test-retest repeatability in a subset ($n = 40$) was better with AI than MBS (coefficient of variation 3.2% vs 5.9%).^{8,12} For GLS, a test-retest study ($n = 40$ and $n = 32$ at two centers) demonstrated that DL-based automated GLS reduced test-retest variability compared with manual analysis. Mean absolute difference was lower, minimal detectable change was smaller, and importantly, no systematic bias was found in AI measurements—as opposed to manual measurements, where bias was observed. Processing time was also very short ($\approx 7.9 \pm 2.8$ seconds per analysis).^{7,12} In a real-world cohort of patients treated with cardiotoxic therapy (*e.g.*, trastuzumab), the AI-derived LVEF and GLS showed moderate-to-strong correlation with conventional methods over serial assessments. However, feasibility was variable: successful GLS measurement only in ~14% of studies, LVEF in ~51%.¹³ A more recent study described a DL model for real-time echo quality scoring and LVEF estimation, across 175 participants from two centers. The model achieved high segmentation accuracy (IoU > 0.8), view classification accuracy ~0.96, and strong agreement with human experts (intraclass correlation coefficient, ICC = 0.821). For high-quality studies, correlation with human EF assessment was $R \approx 0.90$.^{14,15} Earlier automated pipelines demonstrated the ability to detect diseases such as hypertrophic cardiomyopathy, cardiac amyloidosis, and pulmonary arterial hypertension, with reasonable performance.¹⁶ In one large dataset of 8,666 routine echocardiograms, AI achieved automated measurements comparable to manual ones, and disease classification (based on CNNs) with good discrimination.⁵ More advanced models, such as foundation VLMs, attempt to interpret full echo studies—combining multiple views, videos, anatomical attention mechanisms, and report-level embeddings—potentially capturing complex pathologies, composite structural and functional abnormalities, or rare disease patterns. For example, EchoPrime was recently introduced as a multi-view/video-based VLM trained on over 12 million video-report pairs. The authors show that EchoPrime outperforms prior task-specific models across 23 benchmarks of cardiac form and function, including both common and rare diseases.¹¹ In summary, current evidence strongly supports that AI can match—and in many respects approach or even exceed—human expert performance in quantitative measurements (*e.g.*, EF, volumes, strain) and diagnostic classification in retrospective and external-validation datasets.¹⁶

3.4. Efficiency, workflow, and reproducibility

One of the major potential advantages of AI in echo lies in improving workflow efficiency, reproducibility, and consistency. In the PanEcho study, the capability to process comprehensive echo studies automatically suggests that such systems could drastically reduce the time required for interpretation, potentially allowing rapid screening (especially in resource-limited or point-of-care settings), and decrease dependence on expert human readers.⁹ In practice, some centers adopting AI—*e.g.*, using commercial systems—have reported increased number of exams per day, reduced operator fatigue, and a higher number of parameters reported per exam (from ~25 to ~85 parameters). In one recently reported experience, introduction of AI decreased the time per study (on average) and increased throughput, without compromising quality, as judged by final review. Reproducibility: as shown for GLS estimation, AI reduced test-retest variability and eliminated inter-reader bias, which may be particularly valuable for serial follow-up (*e.g.*, chemotherapy cardiotoxicity surveillance), where small changes over time are clinically relevant.⁷ Feasibility in routine settings: In the single-site study with 1,083 consecutive patients, AI achieved view classification and EF measurement in about 81% of echos—indicating that, while not yet perfect, AI can operate on a large fraction of real-world studies.⁸ Overall, these data support that AI-assisted (or AI-enabled) echocardiography can improve efficiency, throughput, and reproducibility, potentially relieving bottlenecks in echo labs and enabling wider access, especially where expert readers are scarce.¹⁶

4. Comment

Despite promising data, substantial limitations remain—especially when considering a shift toward AI-assisted or AI-led echocardiographic interpretation. Many AI studies are retrospective or use curated datasets from high-quality echo labs. In real-world clinical practice, however, image quality varies—due to patient factors (body habitus, acoustic windows), operator skill, equipment type, and scan protocols. AI models trained on “ideal” images may perform less well on suboptimal data. Indeed, some studies report lower feasibility or success rate in “real-world” settings. For example, in the cardiotoxicity monitoring cohort, GLS by AI was only feasible in 14% of studies; LVEF in 51%.^{4,13} Some parameters remain challenging: right-sided heart measures, atrial volumes, certain Doppler-based metrics, and complex valvular lesions. In a study of pharmacological stress echocardiography (SE), AI-based volumetric evaluation showed strong performance for LV volumes and LVEF, but accuracy for some right-sided measurements and fractional area change (FAC) was

limited.¹⁵ While retrospective validation is valuable, the translation of AI performance into real-world, prospective, clinical workflows remains underexplored. For example, few (if any) studies employ randomized, blinded designs comparing AI-assisted vs standard human interpretation as initial read, then evaluating “final cardiologist decision + patient outcomes.” A recent review notes that despite technical advances, the evidence for AI-guided clinical decision-making and therapeutic interventions remains limited.^{3,15,17} There is a scarcity of data on how AI-assisted reporting affects clinical outcomes: diagnosis accuracy in routine practice, changes in patient management, time to treatment, morbidity/mortality, or cost-effectiveness. Deep-learning models (especially large multimodal ones) may provide outputs (measurements or diagnostic suggestions) without transparent reasoning. Especially for complex pathologies or rare conditions, lack of interpretability can impair clinician trust and hinder adoption. This epistemological concern—that AI may produce plausible but incorrect outputs—has been highlighted in recent reviews.¹⁷ For VLMs or foundation models trained on massive datasets, the risk of overfitting to biases in training data (*e.g.*, demographic biases, imaging equipment biases, selection bias) remains real. For example, a recent study using DL on echo images found that models could infer demographic traits (age, sex) from images—raising concerns that AI might exploit non-disease features correlated with demographic variables rather than purely anatomical or functional features. While quantification of EF, volumes, chamber size, and detection of some pathologies is now robust for many AI systems, many clinically important aspects remain challenging: complex valvular disease, diastolic function (which requires Doppler, flow, timing), 3D geometry, right heart hemodynamic, subtle regional wall motion abnormalities, congenital anomalies, prosthetic valves or devices, artifacts, arrhythmias, *etc.* For certain Doppler-based measures (*e.g.*, mitral inflow velocities, deceleration time), fully automated DL approaches exist but are relatively early in translation. For example, a proof-of-concept model for automated mitral inflow Doppler analysis showed excellent correlation with manual measurements in a small sample, but broader clinical validation remains sparse.² Even if AI provides accurate measurements and preliminary interpretations, the final clinical decision should remain under the responsibility of a qualified cardiologist. This “human-in-the-loop” approach is necessary to contextualize AI outputs with patient history, comorbidities, lab data, and clinical presentation. However, this reduces some of the theoretical gains in efficiency and may limit widespread adoption. Use of AI for diagnosis or decision support raises questions about liability, data

privacy, validation standards, regulatory approval, and standard of care. Currently, regulatory frameworks for fully automated interpretation of diagnostic imaging are still evolving. As mentioned, AI models may not generalize across populations, ultrasound machine types, protocols, or centers. There is a risk that models trained on specific datasets perform poorly in underrepresented populations (different ethnicities, body habitus), or in centers with lower image quality. Clinicians may distrust AI “black box” outputs, particularly in borderline, complex, or rare cases. Lack of explainability may limit adoption, especially in high-stakes clinical decisions. If AI becomes widely used, there is a risk that human operators may rely excessively on AI outputs, leading to loss of scanning or interpretation skills over time. This could impair training of new echocardiographers and reduce overall robustness of echo services if AI fails or is unavailable. Safety and clinical outcome data lacking: Ultimately, the goal is not just accurate measurements, but improved patient outcomes (more timely diagnoses, better management, fewer errors, improved prognoses). There is, to date, a dearth of data showing that AI-assisted echocardiography leads to better long-term clinical outcomes.² Table 1 summarizes the AI available on the market and/or ongoing research lines for the TTE main tasks purposes.

4.1. Actual commercial AI use in TTE

Contemporary commercial AI-enabled platforms for TTE differ substantially in their design philosophy, scope of automation, and integration into clinical workflows.

High-end ultrasound machines such as GE HealthCare’s Vivid series incorporate embedded DL-based tools (e.g., Easy AutoEF, Easy AFI) that automate contour detection, biplane ejection fraction, and speckle-tracking strain measurements directly at the point of acquisition, with the primary aim of reducing inter-operator variability and streamlining routine quantifications on the scanner console itself.¹⁷ In contrast, Philips’ EPIQ CVx and Affiniti platforms leverage AI for comprehensive 2D and 3D quantification including automated view classification, segmental wall motion analysis and dynamic heart modeling, with background view selection and advanced chamber/valve modeling integrated into broader workflow and post-processing environments.¹⁸ Siemens systems (e.g., Acuson Sequoia) offer AI-assisted measurement tools that expedite standard 2D/Doppler assessments but are typically positioned within a multi-specialty imaging platform, emphasizing reproducibility across varied applications.¹⁹ Commercial third-party software such as Us2.ai’s AI Echo Copilot extends fully automated analysis and guideline-based reporting to images acquired on any vendor’s hardware, generating zero-click reports with disease detection (including strain and chamber quantification) independently of the scanner’s native software.²⁰ These distinctions reflect divergent strategies: embedded vendor-specific DL tools optimized for on-device workflow enhancements versus vendor-agnostic AI solutions aimed at comprehensive, centralized interpretation and reporting, each with differing implications for clinical integration, reproducibility, and scalability.

Table 1. Summary of AI available on the market and/or ongoing research lines for the TTE main tasks

Task	What is available or near-clinical	What is emerging or in research stage
Quantitative measurements (volumes, EF, chamber size)	PanEcho; EchoConfidence; Us2.ai; segmentation CNNs; MSV-Mamba	Improved segmentation, more robust across anatomies/hardware
Comprehensive echo interpretation (structure + function + valvular + report)	PanEcho; Us2.ai; iCardio.ai; soon EchoPrime and EchoApex (in trial)	VLM-based full-reporting, multimodal integration
Doppler/hemodynamic/diastolic function automation	Research models (e.g., automated mitral inflow Doppler)	Further validation, expansion to other flows/valves
Disease detection (e.g., amyloidosis, aortic stenosis, structural HD)	Us2.ai (amyloidosis), iCardio.ai (aortic stenosis)	More pathology-specific classifiers, rare disease detection, prognostic AI
Workflow support/quality control/image-acquisition assistance	Many commercial AI-echo platforms; vendor-neutral AI included by hardware vendors	AI-guided acquisition, real-time feedback, remote/point-of-care echo

Abbreviations: AI: Artificial intelligence; CNNs: Convolutional neural networks; EF: Ejection fraction; HD: Heart disease; TTE: Transthoracic echocardiography; VLM: Vision-language model.

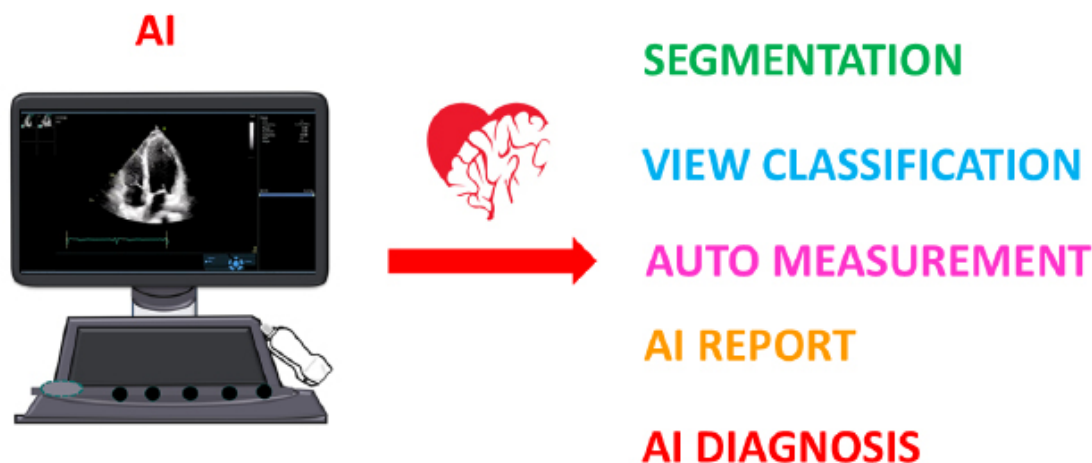


Figure 1. AI abilities in transthoracic echocardiography. Image created by the authors.

Figure 1 summarizes the actual abilities of AI in transthoracic echocardiography.

4.2. Role of vision-language models (VLMs) and foundation models: The new frontier

The recent development of VLMs for echocardiography—such as EchoPrime—represents a conceptual leap beyond single-task AI. These models aim to replicate, at least partially, the comprehensive reasoning performed by expert echocardiographers: integrating multiple views, anatomical structures, dynamic sequences, and generating a holistic, interpretive report.^{2,16} By considering the entire echo study (multiple views, 2D, Doppler, color, *etc.*), VLMs can contextualize structural and functional findings, valve pathology, chamber relationships, hemodynamic—approximating a human-level “clinical read.” Regarding scalability and efficiency, once trained and validated, VLMs could provide preliminary reports rapidly, anywhere, including resource-limited settings—possibly reducing disparities in access to expert echocardiography. Regarding standardization and reproducibility, reports generated by a VLM are likely to be more consistent, less operator-dependent, and less subject to human variability than those depending solely on human readers. Pertaining to support for rare diseases/underrecognized patterns, with training on large datasets that include rare pathologies, VLMs may identify patterns or combinations of features that humans seldom see—potentially uncovering “hidden” phenotypes or contributing to early detection. Regarding integration into clinical decision support, VLM outputs could be directly integrated into larger clinical workflows (electronic health records, risk stratification tools), enabling more seamless diagnosis, risk assessment, and follow-up. However, the promise comes with significant caveats—many of which echo the limitations already

outlined for narrower AI systems, and some that are more specific to large foundation models.^{2,21}

Based on current literature, the use of advanced AI—including but not limited to VLMs—in TTE fulfils many of the desired outcomes, albeit with important caveats. For many quantitative measurements (EF, volumes, GLS, chamber sizes), AI shows accuracy comparable to expert human readers, with good correlation and acceptable measurement error. AI reduces inter- and intra-observer variability, improving reproducibility—a critical issue especially for serial follow-up or multicenter studies. Comprehensive AI systems (*e.g.*, PanEcho) demonstrate the ability to perform multiple tasks simultaneously, across diverse echo parameters and diagnoses, suggesting scalability and broad applicability. AI can speed up interpretation, relieve workload on human experts, and potentially expand echo availability. Another important consideration is that holistic VLM-based interpretation (*e.g.*, EchoPrime) offers the possibility to generate complete reports that go beyond numeric outputs—approaching the interpretative reasoning of human readers. Most evidence remains retrospective or is derived from validation datasets; prospective data, especially regarding impact on clinical decision-making and patient outcomes, are limited to non-existent. Feasibility remains imperfect, in fact, some studies report that AI fails to process all studies (*e.g.*, low success rate for GLS in “real-world” oncology monitoring). Some clinically important parameters remain problematic for AI, for examples, right-sided heart measures, complex valvular disease, Doppler-derived flows, arrhythmias, and suboptimal images. Interpretability, transparency, and regulatory/ethical issues remain major hurdles for clinical adoption, there is a lack of large-scale, multicentric, long-term studies evaluating safety, reliability, effect on

management, and outcomes. Thus, the improvement due to AI/VLM-based echocardiography is context-dependent, and not uniformly guaranteed. In many cases, AI may serve best as a “first-read assistant” or decision-support tool, rather than as a stand-alone replacement for expert interpretation.^{16,21,22} The most recent studies on strength and limitation of the VLM in TTE are summarized in

Table 2.

The current body of literature demonstrates that AI in echocardiography has evolved from proof-of-concept image analysis toward increasingly integrated clinical decision-support systems.²³ Sakamoto *et al.*²⁴ consistently describe a transition from traditional machine learning

Table 2. Summary of key studies on strength and limitation of the VLM in TTE

Year	Model/Tool	Main function/use-case in echo/cardiac ultrasound	Strengths and/or key achievements	Limitations and considerations	Ref.
2025	PanEcho	Full multi-task DL system for TTE: diagnostic classification + parameter estimation (≈ 39 labels/tasks)	High performance: median AUC 0.91 for diagnostic tasks; median normalized MAE 0.13 for parameter estimation; EF MAE ~ 4.2 – 4.5% externally validated; robust across multiple sites and both complete and limited/abbreviated protocols.	Retrospective validation; still requires further prospective/clinical-workflow evaluation before widespread clinical adoption.	1,2,9
2025	EchoAgent	Framework combining specialized vision tools + LLM control: video-level reasoning, guideline-centric measurement and interpretation	Introduces measurement-feasibility prediction, automated frame selection, guideline-aligned interpretation, transparent outputs grounded in visual evidence—pushing toward trustworthy, interpretable AI echo automation.	Early-stage research; needs extensive clinical validation, demonstration of reliability across variable image quality/pathologies/settings.	10,13,14,16
2025	Real-time 3D CNN models for echo analysis (structural + functional parameters) — for real-time echo hardware compatibility	Enables echo interpretation (structure + function) in real-time on standard ultrasound machines; aims to bring AI assistance directly into echo acquisition/reading workflow	Potential to integrate AI directly into scan acquisition/interpretation, reducing need for offline processing; promising for real-time clinical echocardiography.	Real-time constraints impose trade-offs (speed vs accuracy); implementation depends on hardware compatibility; validation and regulatory clearance required; real-world performance under workflow stress needs evaluation.	19,20,22
2024	EchoPrime (a VLM/foundation model)	Holistic interpretation of a full echo study (multi-video, multi-view) + report-level outputs—closer to human style read.	Trained on > 12 million video-report pairs; achieves state-of-the-art performance on multiple “benchmarks of cardiac form and function” across form, function, disease detection.	As of now at preprint stage; needs peer-reviewed, real-world, external validation; real-world integration and regulatory acceptance pending.	11
2024	DL-based view classification (CNN for echo view identification)	Automatic identification of standard echo views (parasternal, apical, <i>etc.</i>)—prerequisite for downstream AI tasks	Very high accuracy: 2D-CNN accuracy 96.8%, AUC ~ 0.997 ; 3D-CNN accuracy $\sim 96.3\%$, AUC ~ 0.998 on comprehensive TTE; robustness also on POCUS datasets.	While view classification works well, correct view alone does not guarantee anatomical adequacy nor high image quality—further QA and human oversight needed if downstream measurements/interpretations are automated.	20

(Cont'd...)

Table 2. Continued

Year	Model/Tool	Main function/use-case in echo/cardiac ultrasound	Strengths and/or key achievements	Limitations and considerations	Ref.
2024	Commercial AI-echo platform US2.ai (or US2 AI-Echo)	Fully automated echo analysis: 2D and Doppler data, chamber quantification, function, valvular assessment, automated report generation	In comparative study: AI reduced measurement time by ~70% compared to manual, and report creation time drastically reduced (AI ~71s vs human ~429s). Also validated for TRJV measurement and pulmonary hypertension evaluation with high reliability (ICC 0.94) and similar diagnostic accuracy (AUC ~0.88) vs manual.	Performance dependent on image quality: in fair/poor quality studies, required more corrections and longer time; human confirmation still often needed for complex findings or suboptimal images.	17,18
2024	Real-world ML validation for POCUS (point-of-care ultrasound) EF estimation	Testing ML-based EF estimation on point-of-care echo (less ideal acquisition conditions)	Provides insights on features influencing ML performance in less controlled, real-world POCUS settings—useful for assessing limitations and robustness outside “academic echo lab” context.	Performance may be lower than in ideal lab settings; generalizability and reliability may suffer with variable acquisition, operator experience, and patient conditions.	20
2023–2025	General research and review-level work summarizing AI-Echo advances, challenges, future directions	Provide comprehensive mapping of AI capabilities, limitations, and readiness for clinical adoption	Useful for clinicians/researchers to understand landscape, recognize gaps (e.g., needs for prospective trials, standardization, regulatory issues), and plan research or deployment accordingly.	Reviews by nature are descriptive; actual performance and generalizability dependent on original studies; may lag behind the very latest models or unpublished innovations	3,7,8,12
2023	DL echo interpretation with prediction of clinical outcomes (structure/function ↔ future risk)	Use of DL-based echo quantification not only for diagnosis but prognostic prediction/risk stratification at scale	Demonstrates that model-derived measures predict clinically meaningful outcomes (heart failure, atrial fibrillation, myocardial infarction, death), supporting prognostic validity of AI-derived echo data.	Retrospective/observational; prognostic models need prospective validation; risk that models overfit to derivation cohorts; real-world use needs careful calibration and validation.	3,7,8,12
Various years (scoping review up to 2022)	Segmentation DL-models for 2D echo (generic)—many studies	Automatic segmentation of heart chambers (ventricles and atria) from 2D echo images	CNN-/DL-based segmentation enables automated quantification of volumes, EF, chamber sizes; reduces inter-operator variability; faster than manual tracing.	Generalizability remains a concern: many models developed on limited datasets, with homogeneous populations/equipment; echo image heterogeneity may impair performance.	15
2021	Automated echo analysis pipelines/general AI-echo review efforts	Overview and synthesis of many AI-echo tools/approaches, highlighting broad areas of application (segmentation, measurement, disease detection, workflow optimization)	Provides a comprehensive map of how AI has been applied in echo—from basic structural quantification to disease classification, helping clinicians/researchers understand strengths/weaknesses and research gaps.	Variability in methods, limited generalizability, need for standardization and larger datasets.	4

(Cont'd...)

Table 2. Continued

Year	Model/Tool	Main function/use-case in echo/cardiac ultrasound	Strengths and/or key achievements	Limitations and considerations	Ref.
2020	Automated Doppler analysis—e.g., Mitral Inflow Doppler with DL	Automatic detection of Mitral Inflow Doppler images, envelope detection, extraction of E/A velocities and deceleration time (diastolic function parameters)	Excellent correlation with operator measurements (E/A wave velocities: R ~0.99; deceleration time R ~0.82); shows feasibility of automating Doppler-based echo parameters.	Limited dataset; proof-of-concept—not yet widely validated; Doppler and hemodynamic measures remain more challenging than structural metrics, especially across variable acquisitions and patient/pathology variability.	4,15
Initial studies ~2018–2020; still referenced in recent literature	EchoNetDynamic (and related segmentation/EF models)	Semantic segmentation of LV + video-based EF estimation from echocardiography	Uses modern DL architectures (e.g., 3D/spatiotemporal CNN) to estimate EF and volumes; provides automated, frame-by-frame segmentation; demonstrates reasonable performance and utility for research and outcome prediction.	Originally designed for research datasets; performance may degrade with variable image quality/non-standard views; limited scope (mostly LV, volumes/EF)—less suited for full clinical report or complex valvular/valvular–hemodynamic evaluation.	4,5,15

Abbreviations: AI: Artificial intelligence; AUC: Area under the curve; CNN: Convolutional neural network; DL: Deep learning; EF: Ejection fraction; ICC: Intraclass correlation coefficient; LLM: Large language model; MAE: Mean absolute error; ML: Machine learning; POCUS: Point of care ultrasound; QA: Quality assurance; TRJV: Tricuspid regurgitant jet velocity; TTE: Transthoracic echocardiography; VLM: Vision-language model.

models to DL frameworks capable of automating view classification, segmentation, quantification, and even prognostic modeling. However, despite rapid technological maturation and exponential growth in scientific production, translational integration into routine clinical workflows remains heterogeneous and incompletely validated.²⁵ A central theme emerging from the literature concerns workflow optimization and operational scalability. Empirical evidence suggests that AI-assisted automation can reduce scan time and system interaction, potentially alleviating operator workload in high-volume laboratories.²⁶ Feasibility studies confirm good agreement between automated and human-derived measurements for standard parameters such as left ventricular ejection fraction, supporting the reproducibility and throughput advantages of AI-based quantification.²⁷ Nevertheless, performance variability in complex anatomical scenarios and suboptimal acoustic windows highlights the continued indispensability of expert oversight, reinforcing the conceptualization of AI as an augmentative rather than substitutive technology.²⁸ This distinction is crucial to avoid overreliance on algorithmic outputs in clinically ambiguous contexts.²⁹ Foundational algorithmic components, particularly view classification and segmentation, represent essential prerequisites for fully automated pipelines. Deep learning–based view classification models achieve high accuracy across canonical transthoracic projections, yet their robustness declines when confronted with atypical or low-quality images, indicating sensitivity to dataset

heterogeneity.³⁰ This limitation underscores the broader issue of generalizability, a recurrent concern throughout the literature.^{25,31} In the domain of cardiac segmentation, comparative methodological analysis reveals a progressive shift from CNN-based architectures to transformer-based models. CNNs have historically dominated echocardiographic image analysis due to their capacity for hierarchical feature extraction and spatial invariance, enabling reliable boundary detection and volumetric quantification in structured datasets.³² Their relative computational efficiency and established optimization pipelines have facilitated widespread adoption in early clinical AI prototypes.^{28,32} However, CNNs inherently rely on localized receptive fields, potentially limiting their ability to capture long-range spatial dependencies within dynamic cardiac structures, particularly in frames characterized by noise or incomplete visualization.³² Transformer-based architectures have been introduced to address these limitations through self-attention mechanisms capable of modeling global contextual relationships across the entire image.³³ Empirical evidence demonstrates that transformer-based left ventricular segmentation improves boundary delineation and contextual feature integration compared with conventional CNN approaches, especially in challenging frames.^{33,34} Similarly, shifted windows vision transformer architectures attempt to balance global context modeling with computational feasibility, achieving competitive segmentation accuracy while reducing the quadratic complexity typically associated with full

attention mechanisms.³⁴ Despite these advantages, transformer-based models introduce new constraints. They often require larger annotated datasets to achieve stable convergence and are associated with increased computational demands, which may limit deployment in resource-constrained clinical environments.^{32–34} Furthermore, while transformer architectures show superior theoretical capacity for global representation learning, external validation across heterogeneous, real-world echocardiographic datasets remains limited, raising concerns about overfitting and reproducibility.³³ In contrast, CNN-based models, although potentially less expressive in global context modeling, benefit from extensive prior validation and optimized implementation pipelines that may currently render them more pragmatic for immediate clinical integration.³² Beyond performance metrics, explainability emerges as a critical comparative dimension. The increasing architectural complexity of both CNN- and transformer-based models exacerbates the “black-box” problem, potentially undermining clinician trust and regulatory acceptance.³¹ Post-hoc interpretability tools such as saliency maps and attention visualization provide partial insight into model reasoning; however, these explanations often lack direct alignment with cardiological pathophysiology, limiting their clinical interpretability.^{29,31} Transformer models, by design, offer attention weights that can be visualized as indicators of spatial focus, yet whether these attention-maps correspond to clinically meaningful reasoning remains debated.³³ Consequently, explainability should be evaluated not solely on technical transparency but on its capacity to enhance diagnostic accountability and decision traceability in real-world settings.³¹

As it is summarized in **Table 3**, Transformer is a DL architecture originally developed for natural language processing that is based on a mechanism called self-attention, which allows the model to weigh relationships between all elements of the input simultaneously. In echocardiography, a transformer processes ultrasound images (or sequences of frames) by modeling global

spatial relationships across the entire image rather than focusing only on local neighborhoods. This is particularly useful for LV segmentation, wall motion assessment, temporal modeling of cardiac cycles, and detection of subtle structural abnormalities. Unlike traditional CNNs, which extract features using localized filters, transformers can directly capture long-range dependencies—such as the geometric relationship between distant myocardial segments—through attention mechanisms. This global context modeling can improve performance in noisy or partially obscured ultrasound images, where local texture alone may be insufficient for accurate interpretation. In practical echocardiography applications, transformers are used for cardiac chamber segmentation, EF estimation, view classification, and automated structural measurements. However, they typically require large annotated datasets and substantial computational resources, which may limit their immediate deployment in routine clinical environments.^{33–36}

4.3. Ethical issues and biases

Another critical dimension concerns ethical governance and bias mitigation. Reviews consistently warn that retrospective training datasets may embed demographic or institutional biases, which could propagate inequities when algorithms are deployed at scale.²⁴ Bibliometric analyses reveal geographic concentration of research production, suggesting potential underrepresentation of diverse patient populations in training cohorts.²⁵ Both CNN- and transformer-based systems are vulnerable to such biases, reinforcing the necessity of multicenter, prospective validation frameworks.²⁹ Ultimately, the comparative synthesis suggests that transformer-based models represent a methodological advancement in capturing global cardiac dynamics and contextual relationships, particularly for complex segmentation tasks.³³ Nevertheless, CNN-based architectures remain clinically relevant due to their computational efficiency, established validation history, and operational feasibility.²⁸ The optimal future direction may not involve exclusive reliance on one paradigm but

Table 3. Transformer and vision transformer in echocardiography

Year	AI analysis	Architecture/Model	Dataset/Task	Ref.
2025	Myocardial strain segmentation	Video Swin Transformer (3D)	Segmentation over time + strain estimation	35
2024	Feature learning/embedding extraction	R3D transformer	Vector embeddings for EF classification	36
2024	LV segmentation	Swin Transformer + UNet (shifted windows)	Semantic segmentation 2D echocardiographic frames	34
2023	LV segmentation	Pure Transformer variants (Swin Transformer + KNet; SegFormer)	EchoNetDynamic LV segment classification	33

Abbreviations: EF: Ejection fraction; LV: Left ventricular.

rather hybrid architectures integrating convolutional inductive biases with transformer-based global attention mechanisms.³⁴ In summary, the literature converges on the perspective that AI in echocardiography has achieved substantial technical progress but faces persistent translational, ethical, and interpretability challenges.²⁹ Sustainable implementation will require rigorous prospective validation, interdisciplinary collaboration, workflow-sensitive integration, and clinically meaningful explainability frameworks.³⁷ AI should therefore be positioned not as an autonomous diagnostic authority but as a collaborative instrument designed to enhance precision, reproducibility, and efficiency while preserving the central role of expert clinical judgment.²⁴

4.4. Limitation and future direction

Several limitations characterize both the current literature and the scope of this review. First, the majority of available evidence derives from retrospective, single-center, or technically curated datasets, which may overestimate algorithmic performance compared with real-world clinical environments. Image quality variability, heterogeneous acquisition protocols, differences in ultrasound vendors, and operator-dependent factors remain substantial sources of performance degradation that are insufficiently addressed in controlled validation settings. Second, the hierarchy of evidence remains imbalanced. Most AI systems in TTE—including CNN-based tools, transformer architectures, and emerging VLMs—have been validated primarily using technical accuracy metrics (*e.g.*, Dice coefficient, correlation with manual measurements, AUC). Prospective, randomized, and outcome-based studies evaluating the impact of AI-assisted interpretation on patient management, morbidity, mortality, or cost-effectiveness are scarce. Consequently, claims regarding clinical transformation remain largely inferential rather than empirically demonstrated. Third, feasibility remains imperfect. Even when measurement accuracy is high in ideal conditions, real-world studies report incomplete processing rates, particularly for strain analysis, right-sided measurements, Doppler-derived indices, and complex valvular assessments. These gaps highlight that AI performance is task-dependent and that comprehensive echocardiographic interpretation extends beyond automated quantification of left ventricular volumes and ejection fraction. Fourth, interpretability and transparency remain critical barriers. Deep learning systems—especially large multimodal or foundation models—often generate outputs without fully transparent reasoning pathways. This “black-box” characteristic may undermine clinician trust, particularly in borderline, rare, or high-stakes clinical scenarios. Furthermore, algorithmic

bias related to demographic, institutional, or equipment-specific characteristics may limit generalizability and propagate inequities if not systematically addressed. Fifth, regulatory, medico-legal, and ethical frameworks are still evolving. Questions regarding liability, standard-of-care integration, data governance, and cross-platform validation remain unresolved. The absence of standardized validation benchmarks across vendors and institutions further complicates broad clinical deployment. Finally, the rapid expansion of commercial AI solutions introduces heterogeneity in design philosophy, transparency, and validation standards. Marketing claims may outpace published peer-reviewed evidence, reinforcing the need for cautious interpretation and clear stratification of evidentiary strength. Future research should prioritize prospective, multicenter validation studies that assess AI performance across diverse patient populations, imaging platforms, and acquisition protocols. Such studies should move beyond technical accuracy and incorporate clinically meaningful endpoints, including changes in diagnostic confidence, therapeutic decision-making, time to intervention, and long-term patient outcomes. Randomized controlled designs comparing AI-assisted interpretation versus standard human interpretation would provide critical evidence regarding true translational benefit. A standardized hierarchy of evidence for cardiovascular imaging AI should be adopted, explicitly distinguishing preclinical technical validation, retrospective external validation, prospective workflow studies, and outcome-based clinical trials. This stratification would facilitate transparent assessment of clinical readiness and prevent overgeneralization from early-stage validation studies. From a methodological standpoint, hybrid architectures integrating convolutional inductive biases with transformer-based global attention mechanisms may represent a balanced path forward, combining computational efficiency with enhanced contextual modeling. In parallel, research into interpretable AI frameworks—such as clinically grounded attention visualization, uncertainty quantification, and explainability aligned with pathophysiological reasoning—should be prioritized to improve trust and accountability. Large-scale, diverse training datasets are essential to mitigate demographic and institutional bias. Collaborative international registries and federated learning approaches may help overcome data fragmentation while preserving privacy. Regulatory agencies and professional societies should also contribute to establishing consensus standards for validation, reporting, and post-market surveillance of AI-enabled echocardiographic tools. Importantly, future investigations must address the human–AI interaction paradigm. Rather than positioning AI as a replacement for

expert interpretation, research should explore optimized “human-in-the-loop” frameworks that preserve clinician oversight while enhancing efficiency and reproducibility. Educational strategies should ensure that increasing automation does not erode core echocardiographic skills among trainees.

5. Conclusion

Translating AI echocardiography tools—particularly VLM-based ones—into widespread clinical practice faces significant hurdles. Beyond technical validation, deployment requires standardization across echo lab to face and harmonize the diverse ultrasound machines, protocols, acquisition settings, and image quality. AI must be robust to these variations. AI-based diagnostic/reporting systems must meet regulatory standards, demonstrate safety, reliability, transparency, and data privacy compliance.

Healthcare workers must be able to trust AI outputs, supported by seamless integration into existing PACS (Picture Archiving and Communication System), reporting systems, and electronic health records, as well as clear delineation of responsibility and liability. Regarding training, maintenance, quality control, AI models may drift over time (e.g., as equipment changes), requiring local retraining, calibration, and continuous quality assurance. In patients undergoing TTE for suspected or known cardiac disease, current evidence demonstrates that AI, including advanced DL systems, substantially improves quantitative accuracy, reproducibility, and efficiency of measurements (e.g., EF, volumes, strain), and shows strong potential for disease classification and diagnostic tasks. Emerging vision-language and foundation models (VLMs) offer a compelling vision for future echocardiography by a holistic, automated interpretation of entire studies, bridging imaging data and clinical report generation. However, despite technical promise, the evidence is not yet sufficient to declare that AI (or VLMs) should replace human echocardiographers in routine clinical practice. Rather, the most appropriate current paradigm is a hybrid, human-in-the-loop workflow, where AI serves as an assistive tool, enhancing efficiency, standardization, and reproducibility—while final clinical interpretation, decision-making, and responsibility remain with qualified physicians. For AI-enabled echocardiography to realize its full potential, future efforts must focus on: large-scale prospective validation, rigorous external testing across diverse settings, demonstration of impact on patient outcomes, transparent and interpretable AI models, regulatory frameworks, and safe integration into clinical workflows. The confluence of high-quality echocardiographic imaging, vast clinical workloads, and the maturation of DL and multimodal foundation models

suggests we stand at the threshold of a paradigm shift in echocardiography. However, as with all powerful tools, caution is warranted. The promise of AI/VLM-assisted TTE is real—but must be balanced with careful validation, human oversight, ethical deployment, and commitment to patient safety and clinical value. In our view, the most plausible near-term scenario is a collaborative human–AI model, where AI enhances human capacity, reduces variability, and increases efficiency, but does not replace human judgement and responsibility.

Acknowledgments

None.

Funding

None.

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: All authors

Writing—original draft: Francesca D’Auria

Writing—review & editing: All authors

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

Not applicable.

References

1. Christensen M, Vukadinovic M, Yuan N, *et al.* Vision-language foundation model for echocardiogram interpretation. *Nat Med.* 2024;30(5):1481–1488.
doi: 10.1038/s41591-024-02959-y
2. Vukadinovic M, Chiu IM, Tang X, *et al.* Comprehensive echocardiogram evaluation with view-primed vision language AI. *Nature.* 2026;650:970–977.
doi: 10.1038/s41586-025-09850-x
3. Aziz D, Maganti K, Yanamala N, Sengupta P. The role of artificial intelligence in echocardiography: a clinical update. *Curr Cardiol Rep.* 2023;25(12):1897–1907.
doi: 10.1007/s11886-023-02005-2
4. Akkus Z, Aly YH, Attia IZ, *et al.* Artificial intelligence (AI)-empowered echocardiography interpretation: a state-of-the-

- art review. *J Clin Med.* 2021;10(7):1391.
doi: 10.3390/jcm10071391
5. Zhang J, Gajjala S, Agrawal P, *et al.* Fully automated echocardiogram interpretation in clinical practice. *Circulation.* 2018;138(16):1623-1635.
doi: 10.1161/circulationaha.118.034338
6. Shiokawa N, Izumo M, Shimamura T, *et al.* Accuracy and efficacy of artificial intelligence-derived automatic measurements of transthoracic echocardiography in routine clinical practice. *J Clin Med.* 2024;13(7):1861.
doi: 10.3390/jcm13071861
7. Salte IM, Østvik A, Olaisen SH, *et al.* Deep learning for improved precision and reproducibility of left ventricular strain in echocardiography: a test-retest study. *J Am Soc Echocardiogr.* 2023;36(7):788-799.
doi: 10.1016/j.echo.2023.02.017
8. Sveric KM, Botan R, Dindane Z, *et al.* Single-site experience with an automated artificial intelligence application for left ventricular ejection fraction measurement in echocardiography. *Diagnostics.* 2023;13(7):1298.
doi: 10.3390/diagnostics13071298
9. Holste G, Oikonomou EK, Tokodi M, *et al.* Complete AI-enabled echocardiography interpretation with multitask deep learning. *JAMA.* 2025;334(4):306-318.
doi: 10.1001/jama.2025.8731
10. Baggiano A, Mushtaq S, Fusini L, *et al.* Artificial intelligence in cardiovascular imaging: current applications and new horizons. *J Cardiovasc Echogr.* 2025;35(2):97-107.
doi: 10.4103/jcecho.jcecho_62_25
11. Vukadinovic M, Tang X, Yuan N, Cheng P, Li D, Cheng S, *et al.* EchoPrime: A Multi-Video View-Informed Vision-Language Model for Comprehensive Echocardiography Interpretation. *arXiv.* Preprint posted online 2024.
doi: 10.48550/arXiv.2410.09704
12. Jiang J, Liu B, Li YW, Hothi SS. Clinical service evaluation of the feasibility and reproducibility of novel artificial intelligence-based echocardiographic quantification of global longitudinal strain and left ventricular ejection fraction in trastuzumab-treated patients. *Front Cardiovasc Med.* 2023;10:1250311.
doi: 10.3389/fcvm.2023.1250311
13. Li X, Liao L, Wu K, *et al.* An automatic and real-time echocardiography quality scoring system based on deep learning to improve reproducible assessment of left ventricular ejection fraction. *Quant Imaging Med Surg.* 2025;15(1):770-785.
doi: 10.21037/qims-24-512
14. Karuzas A, Ciampi Q, Kazukauskienė I, *et al.* Stress Echo 2030 Study Group of the Italian Society of Echocardiography and Cardiovascular Imaging. Artificial intelligence implementation in automated heart chambers quantification during pharmacological stress echocardiography. *Eur Heart J Digit Health.* 2025;6:ztaf121.
doi: 10.1093/ehjdh/ztaf121
15. Karakuş G, Değirmencioglu A, Nanda NC. Artificial intelligence in echocardiography: review and limitations including epistemological concerns. *Echocardiography.* 2022;39(8):1044-1053.
doi: 10.1111/echo.15417
16. Sahashi Y, Ieki H, Yuan V, *et al.* Artificial intelligence automation of echocardiographic measurements. *medRxiv.* Preprint posted online March 19, 2025.
doi: 10.1016/j.jacc.2025.07.053
17. GEHealthCare. The role of artificial intelligence in streamlining echocardiography quantification. Published 2022. Available from: https://www.gehealthcare.com/-/jssmedia/gehc/us/files/products/ultrasound/vivid/redesign/resources/whitepaper_the-role-of-artificial-intelligence__jb20789xx.pdf?srsltid=AfmBOoqKbWqidwnxAiBIu9R1yZPdzn1KD55Ez_RPj94u5E03ig48-yTa [Last accessed on February 28, 2026].
18. Philips Healthcare. AI ultrasound technology: technical description. Published 2025. Available from: <https://www.philips.com/eg/healthcare/technology/ai-ultrasound> [Last accessed on February 28, 2026].
19. Venneri L, Aimo A, Porcari A, *et al.* Artificial intelligence-based echocardiographic assessment for monitoring disease progression in transthyretin cardiac amyloidosis. *Eur J Heart Fail.* 2025; 27(12):3392-3400.
doi: 10.1002/ehf.70073
20. Chiu IM, Vukadinovic M, Sahashi Y, *et al.* Automated evaluation for pericardial effusion and cardiac tamponade with echocardiographic artificial intelligence. *medRxiv.* Preprint posted online 2024.
doi: 10.1101/2024.11.27.24318110
21. Ieki H, Sahashi Y, Vukadinovic M, *et al.* Comprehensive aortic stenosis characterization using multi-view deep learning. *medRxiv.* Preprint posted online 2025.
doi: 10.1101/2025.09.26.25336778
22. van der Boon RMA, Karper JC. Transforming echocardiography with artificial intelligence. *Neth Heart J.* 2025;33(10):319-321.
doi: 10.1007/s12471-025-01987-8
23. Raissi-Dehkordi N, Raissi-Dehkordi N, Xu B. Contemporary applications of artificial intelligence and machine learning in echocardiography. *npj Cardiovasc Health.* 2025;2:30.
doi: 10.1038/s44325-025-00064-8

24. Sakamoto A, Kaneko T, Sato E, *et al.* Artificial intelligence in echocardiography: current applications and future perspectives. *J Echocardiogr.* 2025;23(4):231-240.
doi: 10.1007/s12574-025-00703-0
25. Liu Y, Zhao L, Tu B, Wang J, He Y, Jiang R, *et al.* Application of artificial intelligence in echocardiography from 2009 to 2024: a bibliometric analysis. *Front Med.* 2025;12:1587364.
doi: 10.3389/fmed.2025.1587364
26. Hollitt KJ, Milanese S, Joseph M, Perry R. Can automation and artificial intelligence reduce echocardiography scan time and ultrasound system interaction? *Echo Res Pract.* 2025;12(1):11.
doi: 10.1186/s44156-025-00077-0
27. Lafitte S, Lafitte L, Jonveaux M, Pascual Z, Ternacle J, Dijos M, Bonnet G, Reant P, Bernard A. Integrating artificial intelligence into an echocardiography department: Feasibility and comparative study of automated versus human measurements in a high-volume clinical setting. *Arch Cardiovasc Dis.* 2025;118(8-9):477-488.
doi: 10.1016/j.acvd.2025.04.051
28. Hirata Y, Kusunose K. AI in echocardiography: state-of-the-art automated measurement techniques and clinical applications. *JMA J.* 2025;8(1):141-150.
doi: 10.31662/jmaj.2024-0180
29. Maturi B, Dulal S, Sayana SB, Ibrahim A, Ramakrishna M, Chinta V, Sharma A, Ravipati H. Revolutionizing Cardiology: The Role of Artificial Intelligence in Echocardiography. *J Clin Med.* 2025;14(2):625.
doi: 10.3390/jcm14020625
30. Naser JA, Lee E, Pislaru SV, *et al.* Artificial intelligence-based classification of echocardiographic views. *Eur Heart J Digit Health.* 2024;5(3):260-269.
doi: 10.1093/ehjdh/ztae015
31. Hu X, Zhu Y, Zhang Z, *et al.* Explainable artificial intelligence in echocardiography. *Adv Ultrasound Diagn Ther.* 2025;9(4):409-425.
doi: 10.26599/AUDT.2025.250089
32. Qayyum S. A comprehensive review of applications of artificial intelligence in echocardiography. *Curr Probl Cardiol.* 2024;49(2):102250.
doi: 10.1016/j.cpcardiol.2023.102250
33. Liao M, Lian Y, Yao Y, *et al.* Left ventricle segmentation in echocardiography with transformer. *Diagnostics.* 2023;13(14):2365.
doi: 10.3390/diagnostics13142365
34. Nemri S, Duong L. Automatic segmentation of echocardiographic images using a shifted windows vision transformer architecture. *Biomed Phys Eng Express.* 2024;10(6):065017.
doi: 10.1088/2057-1976/ad7594
35. Huang KC, Lin CE, Lin DSH, *et al.* Video transformer for segmentation of echocardiography images in myocardial strain measurement. *J Imaging Informatics Med.* 2025.
doi: 10.1007/s10278-025-01682-5
36. Chung DJ, Lee SM, Kaker V, *et al.* Echocardiogram vector embeddings via R3D transformer for the advancement of automated echocardiography. *JACC Adv.* 2024;3(9):101196.
doi: 10.1016/j.jacadv.2024.101196
37. Seetharam K, Thyagaturu H, Lora Ferreira G, *et al.* Broadening perspectives of artificial intelligence in echocardiography. *Cardiol Ther.* 2024;13(2):267-279.
doi: 10.1007/s40119-024-00368-3