Journal of Cardiothoracic and Vascular Anesthesia 38 (2024) 895–904

Contents lists available at ScienceDirect

Journal of Cardiothoracic and Vascular Anesthesia

journal homepage: www.jcvaonline.com

Original Article

Pragmatic Evaluation of a Deep-Learning Algorithm to Automate Ejection Fraction on Hand-Held, Point-of-Care Echocardiography in a Cardiac Surgical Operating Room

Emily J. MacKay, DO, MS[*,†,‡,§,1], Shyam Bharat, PhD[||], Rashid A. Mukaddim, PhD[||], Ramon Erkamp, PhD[||], Jonathan Sutton, PhD[||], Ather K. Muhammad, BS[¶], Joseph S. Savino, MD[*], Jiri Horak, MD[*]

[*]*Department of Anesthesiology and Critical Care, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA*
[†]*Penn Center for Perioperative Outcomes Research and Transformation, University of Pennsylvania, Philadelphia, PA*
[‡]*Penn's Cardiovascular Outcomes, Quality and Evaluative Research Center, University of Pennsylvania, Philadelphia, PA*
[§]*Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, PA*
[||]*Philips Research, North America, Cambridge, MA*
[¶]*Alabama College of Osteopathic Medicine, Dothan, AL*

Objective: To test the correlation of ejection fraction (EF) estimated by a deep-learning-based, automated algorithm (Auto EF) versus an EF estimated by Simpson's method.
Design: A prospective observational study.
Setting: A single-center study at the Hospital of the University of Pennsylvania.
Participants: Study participants were ≥18 years of age and scheduled to undergo valve, aortic, coronary artery bypass graft, heart, or lung transplant surgery.
Interventions: This noninterventional study involved acquiring apical 4-chamber transthoracic echocardiographic clips using the Philips hand-held ultrasound device, Lumify.
Measurements and Main Results: In the primary analysis of 54 clips, compared to Simpson's method for EF estimation, bias was similar for Auto EF (−10.17%) and the experienced reader-estimated EF (−9.82%), but the correlation was lower for Auto EF (r = 0.56) than the experienced reader-estimated EF (r = 0.80). In the secondary analyses, the correlation between EF estimated by Simpson's method and Auto EF increased when applied to 27 acquisitions classified as adequate (r = 0.86), but decreased when applied to 27 acquisitions classified as inadequate (r = 0.46).
Conclusions: Applied to acquisitions of adequate image quality, Auto EF produced a numerical EF estimate equivalent to Simpson's method. However, when applied to acquisitions of inadequate image quality, discrepancies arose between EF estimated by Auto EF and Simpson's method. Visual EF estimates by experienced readers correlated highly with Simpson's method in both variable and inadequate imaging conditions, emphasizing its enduring clinical utility.

This investigator-initiated research was funded by Philips Research, North America.
[1]*Address correspondence to* Emily J. MacKay, DO, MSHP, University of Pennsylvania, 310 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104.
E-mail addresses: emily.mackay@pennmedicine.upenn.edu, mackay.ej@gmail.com (E.J. MacKay).

https://doi.org/10.1053/j.jvca.2024.01.005
1053-0770/

IN RECENT YEARS, point-of-care (POC) echocardiography has become a ubiquitous tool across nearly all medical specialties.[1-4] Despite the limitations of a load-dependent measure of left ventricular (LV) systolic function, ejection fraction (EF) remains the primary, guideline-recommended parameter for quantification of LV systolic function.[5] Consequently, timely and accurate LV systolic function assessment is fundamental to bedside POC echocardiography.[3,5]

With advancements in both imaging technology and computational power, research on deep-learning applications to echocardiography has the potential to improve disease detection, increase reading efficiency, and reduce interobserver variability in the interpretation of echocardiographic images.[6-12] This previous work has centered on analyzing echocardiographic images acquired under optimal imaging conditions typical of an outpatient laboratory setting.[6-12] However, the performance of these algorithms under suboptimal imaging conditions (eg, supine positioning, positive-pressure ventilation, and obesity)[13,14] typically encountered in the perioperative setting is largely unknown.

Thus, the goal of this study was to pragmatically test the performance of EF estimates automatically computed by a novel artificial intelligence (AI) software called “Auto EF Quantification” (Philips North America, Cambridge, MA) in a preoperative clinical setting where suboptimal imaging conditions are encountered regularly. This study tested the hypothesis that the EF produced by Auto EF would correlate to EF estimated by conventional echocardiographic techniques—Simpson’s method and experienced visual estimate.

Methods

This blinded, prospective, observational study was preregistered on clinicaltrials.gov before the initiation of any study-related activities (NCT04943965). The trial protocol was reviewed and approved by the University of Pennsylvania Institutional Review Board (IRB). The trial was classified as a minimal-risk study by the IRB, given the nonrandomized, prospective observational design. The investigation was designed to compare 2 methods of EF estimation versus Simpson’s method of disc manual reference. One method was Auto EF Quantification software, and the second method was qualitative visual estimates of EF by experienced echocardiographers.

Population

Study participants were ≥18 years of age and scheduled to undergo valve, aortic, coronary artery bypass graft, heart, or lung transplant surgery. Patients were excluded from the study if they had documented persistent atrial fibrillation or flutter and/or were unable to consent for themselves. Because this study was classified as a minimal risk study by the University of Pennsylvania IRB, study enrollment was initiated in the preoperative area immediately before surgery, and written consent was obtained from all participants.

The Device and Auto EF Quantification Software

Originally developed as “LVivo” (DiA Imaging Analysis Ltd, Be’er Sheva, Israel), the AI software program used in this study (“Auto EF Quantification” or Auto EF) was licensed by Philips and integrated into Philips Lumify handheld ultrasound imaging system (Philips North America). Auto EF operates on an apical 4-chamber (A4C) transthoracic echocardiographic view. When Auto EF is initiated, the underlying algorithm traces the endocardial border for all frames within the entire input loop and produces an EF estimate for each recorded cardiac beat within that acquisition. Then, the software automatically presents the estimated EF results for the second cardiac cycle unless there is only a single cardiac beat in each A4C acquisition (Fig 1). A video corresponding to how Auto EF appears on the Lumify tablet screen is illustrated in Supplementary Video S1. The current study evaluated the capability

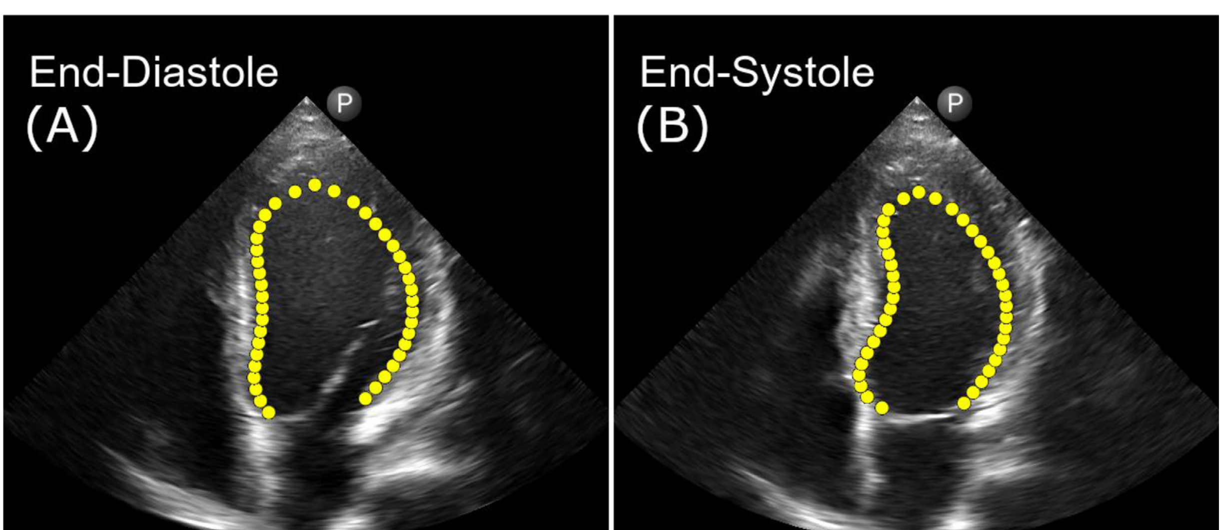

Fig 1. Illustration of the fully automated endocardial border tracing over a single cardiac cycle captured within an apical 4-chamber, cine loop acquired by Lumify. (A) End-diastole and (B) end-systole. The corresponding cine loop is provided in mp4 format and Supplementary Video S1. The illustration shown here is not representative of the border tracing appearance seen in the upcoming product version of automated ejection fraction quantification software.

of Auto EF as a fully automated tool with no manual endocardial border adjustments.

Study Design

A4C echocardiographic acquisitions were obtained prospectively by the principal investigator (PI) (E.M.) using the handheld Lumify device. The PI is certified by the National Board of Echocardiography and is an experienced POC echocardiographer. Images were acquired during the presurgical window (before and after induction of anesthesia) to obtain echocardiographic clips on a heterogeneous cardiac surgery patient population under varying hemodynamic and challenging imaging conditions. A summary of these imaging conditions is in Supplementary Table S1. Auto EF was not applied in real-time to keep the PI blinded to the EF estimates by the Auto EF software.

Simpson's Reference Method

Given its strong correlation with both visual and automated EF estimates,[8] the reference EF was defined as the average of 2 modified Simpson's monoplane method of discs applied to a single A4C acquisition.[5] Simpon's method of discs involves tracing the LV cavity at end-diastole and end-systole, and was completed by the PI using existing Philips TOMTEC Software (TOMTEC Imaging Systems, Unterschliessheim, Germany). Because the 2 Simpson's monoplane EF estimates used the same A4C echocardiographic acquisitions and the same cardiac cycle as Auto EF, they were conducted 1 month apart to reduce potential unconscious bias from the PI.

Visual EF Estimates by Cardiac Anesthesiologists

As a parallel analysis to the Auto EF versus Simpson's method, the study design involved comparing the bias and accuracy of the mean visual estimate of EF by 17 National Board of Echocardiography-certified (or testamured)[15] experienced echocardiographic readers versus the same Simpson's reference method. The protocol for visual estimation of EF by the 17 experienced echocardiographers involved each physician viewing the A4C POC echocardiographic clips individually and assigning a numerical value for EF based on that single A4C echocardiographic clip.

Statistical Analysis

EF estimated by Auto EF versus Simpson's method was compared using the Pearson correlation coefficient (r) set to 0.5, a type I (alpha) error of 0.05, and a type II (beta) error of 0.20. Using these parameters, the required sample size for this study was 29 subjects.[16] This was consistent with previous studies that compared 2 different echocardiographic imaging modalities, with sample sizes ranging from 17-to-31 subjects.[17-19] Thus, the goal enrollment was set at 35 subjects—a 20% increase over n = 29—as a contingency for potential lost data or inadequate imaging.

Baseline characteristics of the study cohort were analyzed using standard descriptive statistics, with categorical data presented as n (%) and continuous data presented as mean (±SD). For all analyses, the accuracy of the EF estimates produced by Auto EF was compared to Simpson's using linear regression

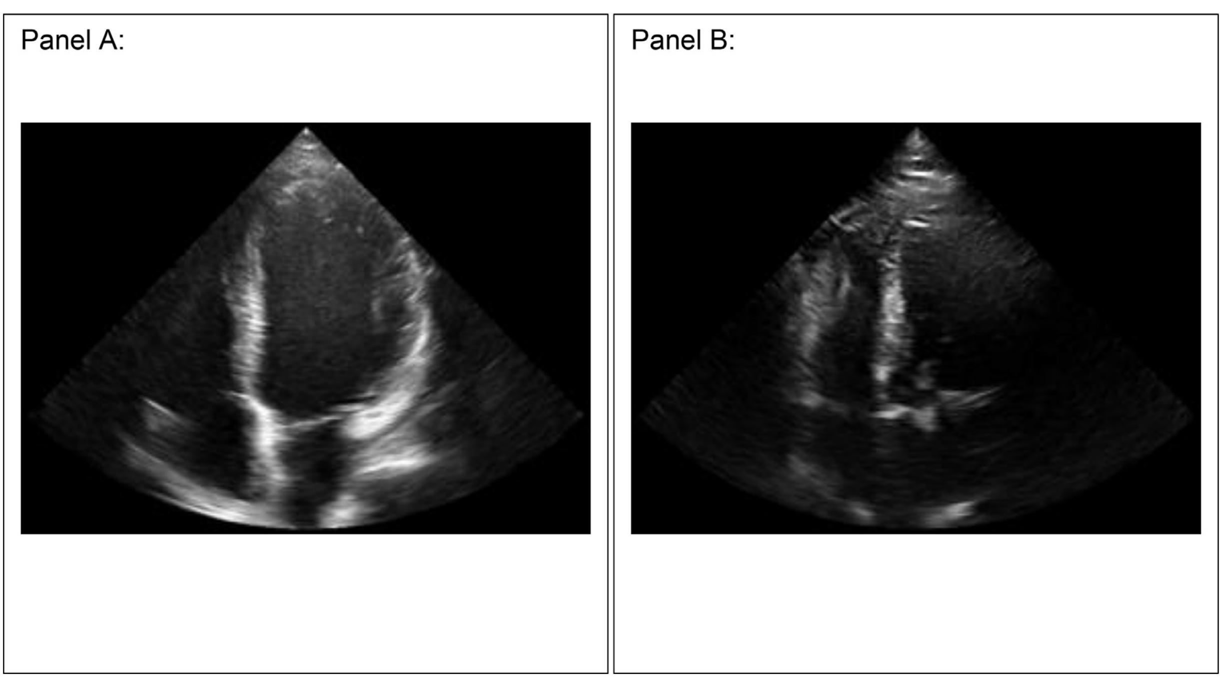

Fig 2. Examples of 2 apical 4-chamber transthoracic acquisitions. (A) An example of an acquisition that adhered to imaging quality adequacy criteria (ie, >80% of myocardium clearly captured, sharp endocardial border, and minimal reverberation artifact) and was classified as *adequate*. (B) An example of an acquisition that did not adhere to imaging quality adequacy criteria (ie, <80% of myocardium clearly captured, hazy and poor endocardial border demarcation, and excessive reverberation artifact obscuring the anterolateral myocardial wall) and was classified as *inadequate*. The corresponding cine clips are provided in mp4 format and Supplementary Video S2 (corresponding to panel A) and S3 (corresponding to panel B).

and the Pearson correlation coefficient. Agreement between EF estimates by Auto EF versus Simpson's and the mean of 17 experienced readers versus Simpson's was analyzed by Bland−Altman analyses.[20] The 95% limits of agreement (LOA) were calculated (±1.9 SD), and the mean difference (eg, bias) was calculated. Interobserver variability between Auto EF and the 17 experienced readers was calculated using the intraclass correlation coefficient (ICC). A secondary post hoc analysis was conducted on 2 distinct subsets of the 54 clips divided into 2 groups by study personnel blinded to the findings of the Auto EF analysis based on a set of imaging adequacy criteria (Fig 2, A; Supplementary Video S2, A) illustrates an example of an adequate acquisition (Fig 2, B; Supplementary Video S2, B) illustrates an example of an inadequate acquisition. Details on the criteria for adequacy classification are in Supplementary Table S2. Basic statistics were conducted using STATA 17.0 (StataCorp, LLC, College Station, TX). Linear regression, Bland−Altman, and ICC statistics were conducted using Python (Open Source, Python Software Foundation).

Results

The study population included 35 patients. Overall, the mean age was 58.7 years, 68.6% were male sex, 31.4% were female sex, and the mean body mass index was 28.6 kg/m^2. Table 1 summarizes the demographic and clinical data of the study population. Auto EF Quantification reported failures in the selected acquisitions in 2 out of 30 subjects (ie, producing a numerical EF estimate in 28 of the 30 subjects); a feasibility rate of 93%. Of those 28 patients, 54 selected A4C clips (2 per patient for 26 patients and 1 per patient for 2 patients) returned successful Auto EF Quantification results and were used for subsequent analysis.

Table 1
Patient Characteristics (n = 35)

Characteristics	Values
Demographics	
Age, mean ± SD, y	58.7 ± 13.3
Male sex, n (%)	24 (68.6)
Height, mean (±SD), cm	172.8 (±9.6)
Weight, mean (±SD), kg	85.5 (±18.6)
Medical history, n (%)	
Aortic aneurysm	10 (28.6)
Aortic dissection	1 (2.9)
End-stage cardiomyopathy	1 (2.9)
COPD	2 (5.7)
End-stage interstitial lung disease	1 (2.9)
Obesity	12 (34.3)
Pulmonary hypertension	4 (11.4)
Echocardiographic parameters (OR TEE)	
EF, mean (±SD), %*	56.2 ±13.5
Aortic regurgitation ≥moderate, n (%)	9 (25.7)
Aortic stenosis ≥moderate, n (%)	7 (20.0)
Mitral regurgitation ≥moderate, n (%)	8 (22.9)
Mitral stenosis ≥moderate, n (%)	2 (5.7)
Tricuspid regurgitation ≥moderate, n (%)	2 (5.7)
RV dilation, n (%)	8 (22.9)
Surgical procedure, n (%)	
Ascending aorta replacement	10 (28.6)
Ascending aorta and hemiarch replacement	5 (14.3)
Tricuspid (repair or replacement)	1 (2.9)
Ross procedure	2 (5.7)
CABG	7 (20.0)
Lung transplant	1 (2.9)
Heart transplant	1 (2.9)
Aortic valve replacement	7 (20.0)
Mitral valve replacement	2 (5.7)
Mitral valve repair	4 (11.4)

Abbreviations: CABG, coronary artery bypass graft; COPD, chronic obstructive pulmonary disease; EF, ejection fraction; OR, operating room; RV, right ventricle; TEE, transesophageal echocardiography.

* Please refer to Supplementary Fig S1 to view the EF distribution of the cohort.

Interobserver Variability Analysis

Interobserver variability was tested by plotting the range of estimated EF values among the 17 experienced readers. In 53 of 54 clips (98%), Auto EF estimates of EF fell within the range of the 17 visual estimates of EF (Fig 3). When considering Auto EF and the 17 visual estimates, the ICC was 0.60 (95% CI: 0.50-0.70), indicating moderate agreement.[21] The ICC did not significantly change when computed from only the 17 visual EF estimates (ICC = 0.61 [95% CI: 0.51-0.71]).

Primary Analysis

The primary analysis of the 54 clips compared the correlation and bias between EF estimated by (1) Auto EF versus Simpson's, and (2) the mean of 17 experienced visual estimates versus Simpson's. Compared to Simpson's, Auto EF demonstrated moderate bias[22] (−10.17% [SD = 11.04%]; lower LOA: −31.82% and upper LOA: 11.46%) and modest correlation[23] (r = 0.56 [95% CI: 0.34-0.72]; $p < 0.0001$) (Fig 4, A and B). Compared to EF estimated by the Simpson's method, the mean of 17 visual estimates demonstrated moderate bias[22] (−9.82% [SD = 8.14%]; lower LOA: −25.77% and upper LOA: 6.13%) and strong correlation[23] (r = 0.80 [95% CI: 0.68-0.88]; $p < 0.0001$) (Fig 4, C and D).

Secondary Analyses

A secondary, post hoc analysis was undertaken to explore the modest correlation between the EF estimated by Simpson's versus Auto EF by dividing the 54 A4C clips into "adequate" and "inadequate" image quality cohorts based on criteria described in Supplementary Table S2. Bland−Altman and Linear regression analyses were repeated on the adequate and inadequate subgroups.

Analysis: Adequate Subgroup

Among the 27 adequate A4C clips, compared to EF estimated by Simpson's, Auto EF demonstrated low bias[22] (−3.57% [SD = 6.72%]; lower LOA: −16.74% and upper LOA: 9.60%) and strong correlation[23] (r = 0.86 [95% CI:

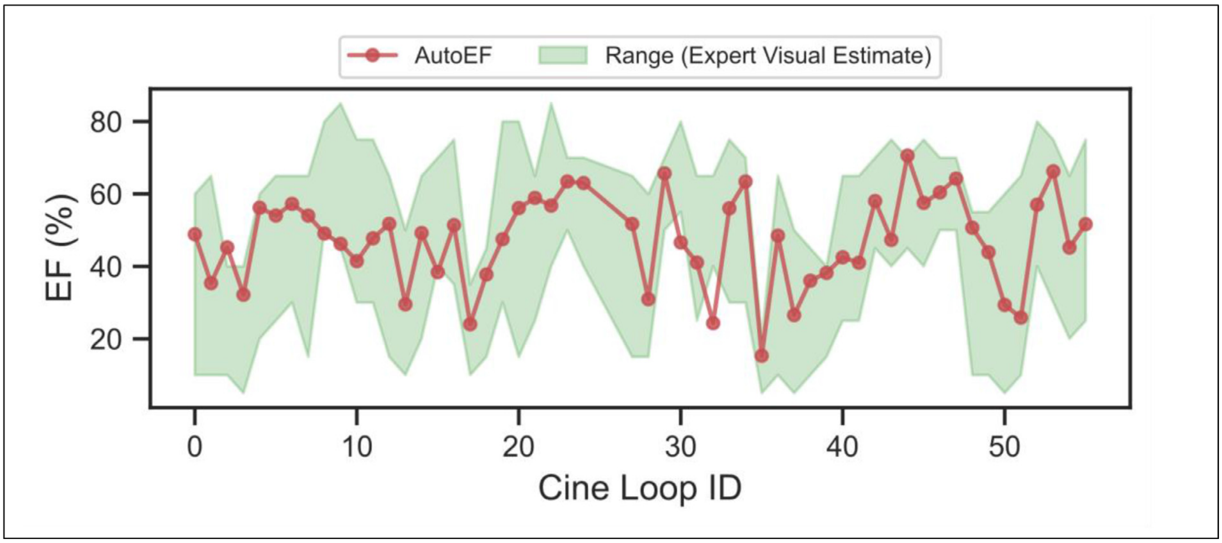

Fig 3. Data visualization of the ejection fraction (EF) estimates (y-axis) for all 54 cine clips (x-axis). The visual EF estimates (17 values for each of the 54 cine clips analyzed) are plotted as a range of values (*light green, shaded bar*). The automated EF estimates (1 value for each of the 54 cine clips analyzed) are plotted as a single value (*red line, single dots*). Automated EF produced numerical EF estimates that fell within the range of the EF estimated visually by 17 blinded experienced echocardiographers in 53 of 54 (98%) of the cine clips. Auto EF, automated ejection fraction.

0.71-0.93]; p < 0.0001) (Fig 5, A and B). Among those same 27 adequate A4C clips, compared to EF estimated by Simpson's, the mean of 17 visual estimates demonstrated moderate bias[22] (−6.39% [SD = 6.88%]; lower LOA: −19.89% and upper LOA: 7.10%) and a strong correlation[23] (r = 0.88 [95% CI: 0.75-0.94]; p < 0.0001) (Fig 5, C and D).

Analysis: Inadequate Subgroup

Among the 27 inadequate A4C clips, EF estimated by Simpson's, Auto EF demonstrated high bias[22] (−16.78% [SD = 10.61%]; lower LOA: −37.58% and upper LOA: 4.01%) and fair correlation[23] (r = 0.46 [95% CI: 0.10-0.72]; p = 0.02) (Fig 6, A and B]). Among those same 27 inadequate A4C clips, compared to EF estimated by Simpson's, the mean of 17 visual estimates demonstrated high bias[22] (−13.25% [SD = 7.95%]; lower LOA: −28.83% and upper LOA: 2.34%) and a strong correlation[23] (r = 0.80 [95% CI: 0.61-0.91]; p < 0.0001) (Fig 6, C and D).

Discussion

This study aimed to test and validate the performance of Auto EF as an AI-based, echocardiographic EF estimator in a clinical setting. This type of pragmatic validation testing is crucial to ensure these applications are used appropriately when deployed into clinical practice. On the primary analysis of 54 clips, compared to Simpson's method for EF estimation, bias was similar for Auto EF (−10.17%) and the experienced reader-estimated EF (−9.82%), but the correlation was lower for Auto EF (r = 0.56) than the experienced reader-estimated EF (r = 0.80). Secondary analyses on clips classified as adequate or inadequate indicated that the software is highly dependent on image quality. For instance, among the subset of 27 echocardiographic clips with adequate cine loop quality, Auto EF demonstrated a strong correlation (r = 0.86) to Simpson's method, which is comparable to existing work on the accuracy of deep learning algorithm-based estimates of EF.[6,24,25] Conversely, among the subset of 27 echocardiographic clips of inadequate quality, the correlation to Simpson's was lower for Auto EF (r = 0.46) than the mean visual EF estimate by experienced readers (r = 0.81).

Findings from this study highlighted an important inadequacy of using Auto EF as a fully automated tool (ie, no manual border adjustments by the user postanalysis). That is, although Auto EF may be reliable in clinical environments where there is a high likelihood of achieving adequate loop quality (eg, outpatient setting with echocardiograms performed by experienced transthoracic sonographers), Auto EF may not be reliable in clinical environments where it is harder to consistently acquire images with adequate loop quality (eg, perioperative or critical care setting with POC echocardiography performed by non-sonographers with less transthoracic imaging experience). If deployed in clinical environments in which challenging imaging conditions are commonplace, manual border adjustment or image adequacy screening may be required for Auto EF to produce reliable EF estimates.

This work adds to the existing published research testing both semiautomated (eg, manual endocardial border adjustment permitted)[26,27] and fully automated (eg, no manual endocardial border adjustment)[28,29] AI-based software tools for EF estimation by POC echocardiography.[26-29] The authors' study design and findings are most comparable to the 2 previous studies of fully automated EF estimates.[28,29] A 2021 investigation by Filipiak-Strzecka et al. tested the accuracy of EF estimated automatically by LVivo versus a 3-dimensional reference method applied to A4C images acquired using a hand-held ultrasound device in 96 patients.[28] In this study, among images classified as "acceptable" or better, a very strong (r = 0.92) correlation between EF estimated by Lvivo

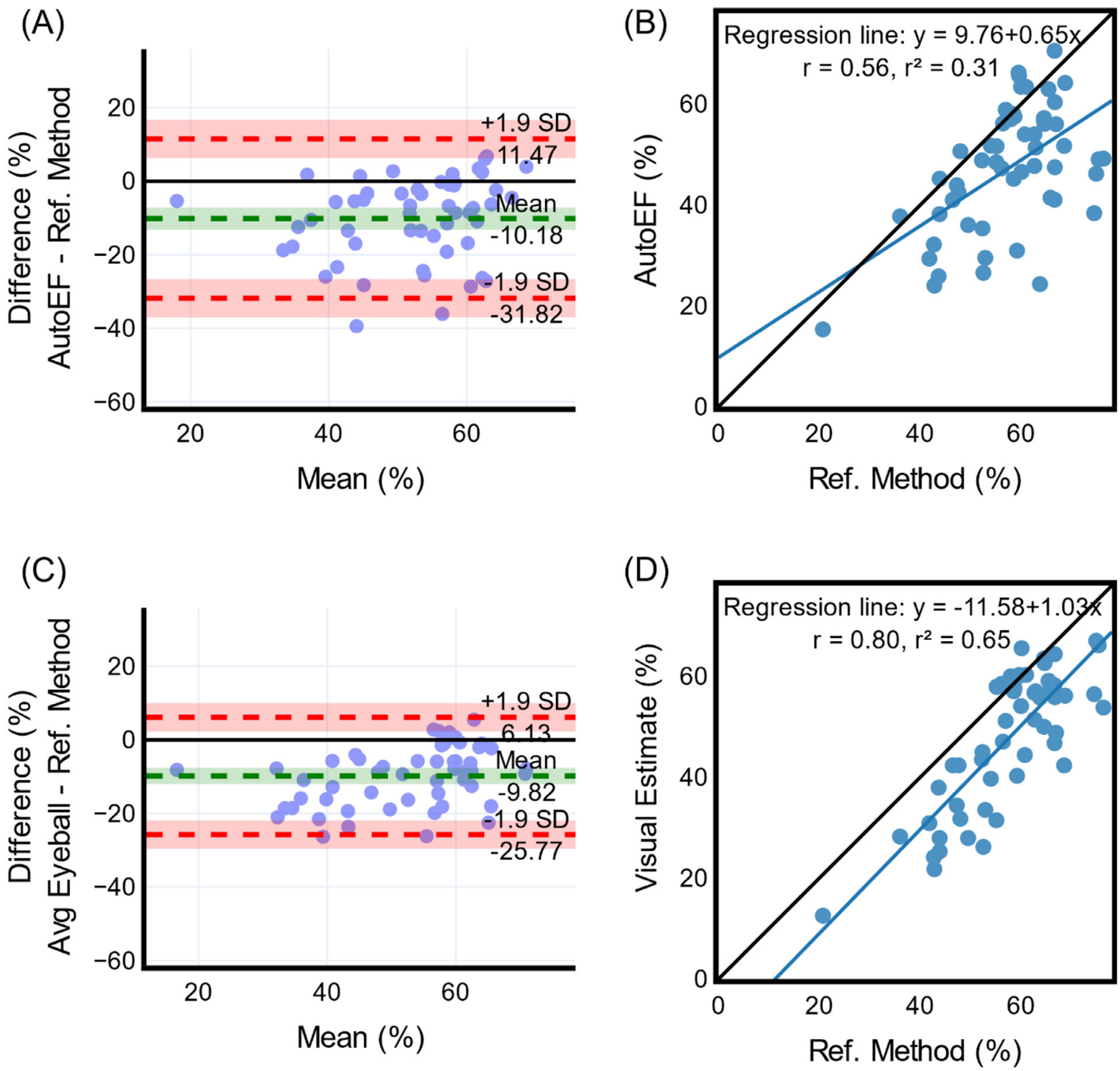

Fig 4. Bland−Altman (A and C) and linear regression (B and D) results for the primary analysis of 54 cine clips; Auto ejection fraction versus Simpson's reference method (*top*; A and B) and mean visual ejection fraction estimate from 17 experienced readers versus Simpson's (*bottom*; C and D). Auto EF, automated ejection fraction.

versus the 3-dimensional reference method was reported.[28] This result is only marginally higher than the finding of a very strong correlation between Auto EF and Simpson's method (r = 0.86) for EF quantification among images classified as "adequate." Moreover, both studies observed comparable bias between the automated EF estimate and the reference methods.[28] An older 2008 investigation by Rahmouni and colleagues tested the performance of an AI-based automated EF assessment software (Siemens Medical Solutions, Erlangen, Germany) versus 2 different reference methods (expert visual estimate and manual planimetry) applied to 2-dimensional (2D), A4C transthoracic echocardiographic images.[29] Fifteen years later, the authors' findings from the current study were remarkably similar to the findings from this previous investigation. Both studies found a modest correlation between the EF estimated by Auto EF versus Simpson's 2D reference method (Rahmouni et al.: r = 0.64[29] *v* this study: r = 0.56), both observed a very strong correlation between EF estimated by experienced readers versus Simpson's 2D reference method (Rahmouni et al.: r = 0.86[29] *v* this study: r = 0.80), and both observed narrower LOAs on Bland-Altman analysis with experienced readers versus Simpson's (Rahmouni et al.: LOA = −10% to 22%[29] *v* this study: LOA = −26% to 6%) compared to Auto EF versus Simpson's (Rahmouni et al.: LOA = −19% to 33%[29] *v* this study LOA: −32% to 11%).

Because the authors were interested in studying the performance of Auto EF as a fully automated tool for estimating EF, their findings did not compare well to the 2 previous semiautomated studies that allowed for manual endocardial border adjustment post hoc when comparing EF estimated

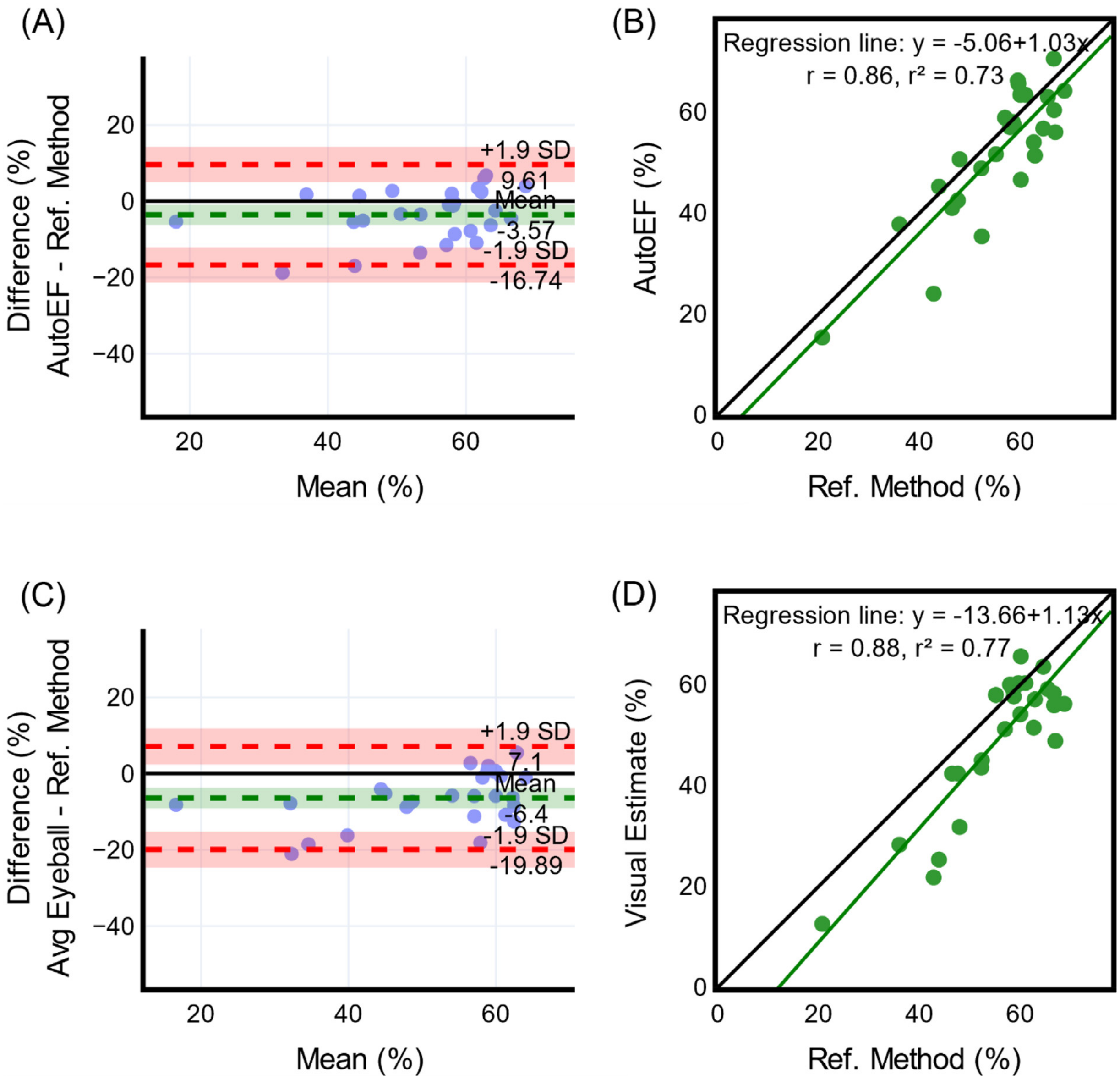

Fig 5. Bland−Altman (A and C) and linear regression (B and D) results for the secondary analysis of the 27 cine clips classified as adequate: Auto ejection fraction versus Simpson's reference method (*top*; A and B) and mean visual ejection fraction estimate from 17 experienced readers versus Simpson's (*bottom*; C and D). Auto EF, automated ejection fraction.

automatically versus various reference methods.[26,27] Nevertheless, the authors' secondary analysis of images classified as adequate did agree with previous work published in 2015 by Frederiksen and colleagues comparing semi-automated EF software (GE Healthcare, Horton, Norway) versus a reference method (Frederiksen et al.: r = 0.82[26] *v* this study: r = 0.86). In contrast, the authors' fully automated method was not able to reproduce the extremely strong correlation between EF estimated by AI-learned EF assessment software against the following 3 different reference methods: (1) versus a 2D manual planimetry reference (r = 0.98), (2) versus an experienced readers-estimated EF (r = 0.96), and (3) versus a cardiac magnetic resonance imaging reference (r = 0.95), as reported by Cannesson et al. in 2007.[27] This discordant result was likely due to the fact that 23% of the total analyzed cases in the Cannesson study required manual editing by experts and were acquired using a cart-based imaging system (Acuson Sequoia; Siemens, Mountain View, CA).[27]

Limitations

The findings presented in this study must be interpreted with awareness of its limitations. First, the small sample size used for analysis (eg, 54 clips from n = 30 patients) with correspondingly wide 95% CIs. Second, because all data acquisitions were performed by a single operator, the findings cannot be extrapolated to environments where acquisitions are performed by multiple operators. Third, although the authors had physically collected a substantial amount of data per patient, they elected to use only a small subset of the data to compare

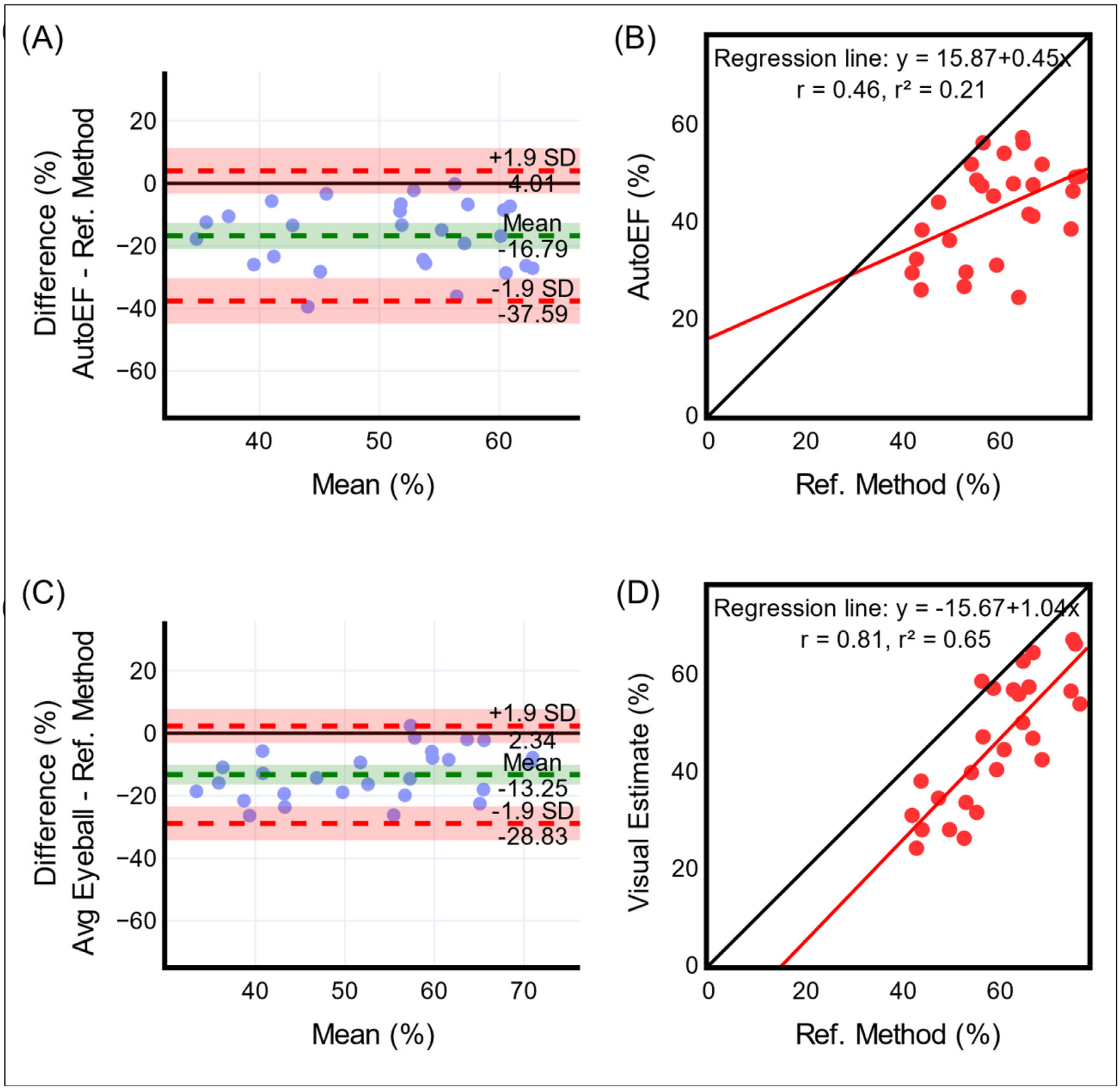

Fig 6. Bland−Altman (A and C) and linear regression (B and D) results for the secondary analysis of the 27 cine clips classified as inadequate: automated ejection fraction versus Simpson's reference method (*top*; A and B) and mean visual EF estimate from 17 experienced readers versus Simpson's (*bottom*; C and D). Auto EF, automated ejection fraction.

EF estimates from the same cardiac cycle to avoid the inevitable variability of LVEF (as a load-dependent measure of systolic function) over time. Fourth, among the 30 selected cardiac surgical patients, only 2 had an EF of <35%.

Consequently, the authors' findings may not be generalizable to patients with severely decreased cardiac function. Fifth, the most significant limitation of this study (a limitation noted in all 4 previously published, similarly designed studies)[26-29] was that no true gold standard for EF (ie, a ground truth reference) was compared. For this study, the reference EF was calculated by a single echocardiographer applying Simpson's method of disks at 2 time points with a 1-month timespan in between measurements to limit bias. However, this reference method for EF estimation is not considered to be a "ground truth" for comparison. This lack of a "ground truth" reference EF limitation potentially could be addressed in future research by the addition of multibeat averaging over a clinically meaningful timespan.

Conclusions

Applied to acquisitions of adequate image quality, Auto EF produced a numerical EF estimate equivalent to Simpson's method. However, when applied to acquisitions of inadequate image quality, discrepancies arose between EF estimated by Auto EF and Simpson's method. Under variable or inadequate imaging conditions, visual EF estimates by experienced readers correlated highly with Simpson's method. Therefore, the use of Auto EF in a fully automated manner (ie, without the manual border editing) is not recommended on echocardiographic clips of inadequate image quality. Continued work to improve deep learning applications to echocardiography, such as loop quality feedback during image acquisition, is critical for making objective, accurate, reliable, and reproducible echocardiography-based assessments.

Declaration of competing interest

None.

CRediT authorship contribution statement

Emily J. MacKay: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing − original draft, Writing − review & editing. **Shyam Bharat:** Data curation, Formal analysis, Methodology, Project administration, Visualization, Writing − review & editing. **Rashid A. Mukaddim:** Data curation, Formal analysis, Methodology, Visualization, Writing − review & editing. **Ramon Erkamp:** Data curation, Formal analysis, Software, Writing − review & editing. **Jonathan Sutton:** Conceptualization, Data curation, Methodology, Project administration, Writing − review & editing. **Ather K. Muhammad:** Data curation, Software, Writing − review & editing. **Jiri Horak:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Resources, Supervision, Writing − review & editing.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1053/j.jvca.2024.01.005.

References

1 Moore CL, Copel JA. Point-of-care ultrasonography. N Engl J Med 2011;364:749–57.
2 Díaz-Gómez JL, Mayo PH, Koenig SJ. Point-of-care ultrasonography. N Engl J Med 2021;385:1593–602.
3 Kirkpatrick JN, Grimm R, Johri AM, et al. Recommendations for echocardiography laboratories participating in cardiac point of care cardiac ultrasound (POCUS) and critical care echocardiography training: Report from the American Society of Echocardiography. J Am Soc Echocardiogr 2020;33;409-22.e404.
4 Levitov A, Frankel HL, Blaivas M, et al. Guidelines for the appropriate use of bedside general and cardiac ultrasonography in the evaluation of critically ill patients—part II: Cardiac ultrasonography. Crit Care Med 2016;44:1206–27.
5 Lang RM, Badano LP, Mor-Avi V, et al. Recommendations for cardiac chamber quantification by echocardiography in adults: An update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. Eur Heart J Cardiovasc Imaging 2015;16:233–71.
6 Ouyang D, He B, Ghorbani A, et al. Video-based AI for beat-to-beat assessment of cardiac function. Nature 2020;580:252–6.
7 Tromp J, Bauer D, Claggett BL, et al. A formal validation of a deep learning-based automated workflow for the interpretation of the echocardiogram. Nat Commun 2022;13:6776.
8 Knackstedt C, Bekkers SC, Schummers G, et al. Fully automated versus standard tracking of left ventricular ejection fraction and longitudinal strain: The FAST-EFs multicenter study. J Am Coll Cardiol 2015;66:1456–66.
9 Kusunose K, Abe T, Haga A, et al. A deep learning approach for assessment of regional wall motion abnormality from echocardiographic images. JACC Cardiovasc Imaging 2020;13:374–81.
10 Medvedofsky D, Mor-Avi V, Amzulescu M, et al. Three-dimensional echocardiographic quantification of the left-heart chambers using an automated adaptive analytics algorithm: Multicentre validation study. Eur Heart J Cardiovasc Imaging 2018;19:47–58.
11 Narang A, Mor-Avi V, Prado A, et al. Machine learning based automated dynamic quantification of left heart chamber volumes. Eur Heart J Cardiovasc Imaging 2019;20:541–9.
12 Narula S, Shameer K, Salem Omar AM, et al. Machine-learning algorithms to automate morphological and functional assessments in 2D echocardiography. J Am Coll Cardiol 2016;68:2287–95.
13 Subramaniam K, Boisen ML, Yehushua L, et al. Perioperative transthoracic echocardiography practice by cardiac anesthesiologists—report of a "start-up" experience. J Cardiothorac Vasc Anesth 2021;35:222–32.
14 Duffy G, Cheng PP, Yuan N, et al. High-throughput precision phenotyping of left ventricular hypertrophy with cardiovascular deep learning. JAMA Cardiol 2022;7:386–95.
15 Quiñones MA, Douglas PS, Foster E, et al. ACC/AHA clinical competence statement on echocardiography: A Report of the American College of Cardiology/American Heart Association/American College of Physicians−American Society of Internal Medicine Task Force on Clinical Competence Developed in Collaboration with the American Society of Echocardiography, the society of cardiovascular anesthesiologists, and the society of pediatric echocardiography. J Am Coll Cardiol 2003;41:687–708.
16 Newman TB, Browner WS, Cummings SR, Hulley SB. Designing studies of medical tests. In: Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB, editors. Designing Clinical Research, Third Edition, Philadelphia, PA: Lippincott Williams & Wilkins; 2007. p. 183–205.
17 Kengne C, Ouattara B, Angoran I, et al. Interobserver variability in estimation of the left ventricular systolic function of dilated cardiopathies: Ejection fraction and global longitudinal strain. Arch Cardiovasc Dis Suppl 2018;10:36.
18 Sugeng L, Mor-Avi V, Weinert L, et al. Quantitative assessment of left ventricular size and function: Side-by-side comparison of real-time three-dimensional echocardiography and computed tomography with magnetic resonance reference. Circulation 2006;114:654–61.
19 Sugeng L, Mor-Avi V, Weinert L, et al. Multimodality comparison of quantitative volumetric analysis of the right ventricle. JACC Cardiovasc Imaging 2010;3:10–1.
20 Bland JM, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;327:307–10.
21 Bobak CA, Barr PJ, O'Malley AJ. Estimation of an inter-rater intra-class correlation coefficient that overcomes common assumption violations in the assessment of health measurement scales. BMC Med Res Methodol 2018;18:1–11.
22 Bunting KV, Steeds RP, Slater LT, et al. A practical guide to assess the reproducibility of echocardiographic measurements. J Am Soc Echocardiogr 2019;32:1505–15.
23 Akoglu H. User's guide to correlation coefficients. Turk J Emerg Med 2018;18:91–3.
24 Zhang J, Gajjala S, Agrawal P, et al. Fully automated echocardiogram interpretation in clinical practice: Feasibility and diagnostic accuracy. Circulation 2018;138:1623–35.
25 Ghorbani A, Ouyang D, Abid A, et al. Deep learning interpretation of echocardiograms. NPJ Digit Med 2020;3:10.

26 Frederiksen CA, Juhl-Olsen P, Hermansen JF, et al. Clinical utility of semi-automated estimation of ejection fraction at the point-of-care. Heart Lung Vessel 2015;7:208.

27 Cannesson M, Tanabe M, Suffoletto MS, et al. A novel two-dimensional echocardiographic image analysis system using artificial intelligence-learned pattern recognition for rapid automated ejection fraction. J Am Coll Cardiol 2007;49:217–26.

28 Filipiak-Strzecka D, Kasprzak JD, Wejner-Mik P, et al. Artificial Intelligence-Powered measurement of left ventricular ejection fraction using a handheld ultrasound device. Ultrasound Med Biol 2021;47:1120–5.

29 Rahmouni HW, Ky B, Plappert T, et al. Clinical utility of automated assessment of left ventricular ejection fraction using artificial intelligence−assisted border detection. Am Heart J 2008;155:562–70.