ORIGINAL CONTRIBUTION

AEM
Education and Training

Use of artificial intelligence as a didactic tool to improve ejection fraction assessment in the emergency department: A randomized controlled pilot study

Ziv Dadon MD[1] | **Adi Butnaru MD**[1] | **David Rosenmann MD, MPH**[1] | **Liat Alper-Suissa MBA**[1] | **Michael Glikson MD**[2] | **Evan A. Alpert MD**[3]

[1]Jesselson Integrated Heart Center, Shaare Zedek Medical Center, Jerusalem, Israel

[2]Jesselson Integrated Heart Center, Shaare Zedek Medical Center, The Faculty of Medicine, Hebrew University of Jerusalem, Jerusalem, Israel

[3]Department of Emergency Medicine, Shaare Zedek Medical Center, The Faculty of Medicine, Hebrew University of Jerusalem, Jerusalem, Israel

Correspondence
Evan A. Alpert, MD, Shaare Zedek Medical Center, 12 Shmuel Bait St. PO Box 3235 Jerusalem 9103102, Israel.
Email: avraham.alpert@mail.huji.ac.il

Abstract

Objectives: Incorporating artificial intelligence (AI) into echocardiography operated by clinicians working in the emergency department to accurately assess left-ventricular ejection fraction (LVEF) may lead to better diagnostic decisions. This randomized controlled pilot study aimed to evaluate AI use as a didactic tool to improve noncardiologist clinicians' assessment of LVEF from the apical 4-chamber (A4ch) view.

Methods: This prospective randomized controlled pilot study tested the feasibility and acceptability of the incorporation of AI as a didactic tool by comparing the ability of 16 clinicians who work in the emergency department to assess LVEF before and after the introduction of an AI-based ultrasound application. Following a brief didactic course, participants were randomly equally divided into an intervention and a control group. In each of the first and second sessions, both groups were shown 10 echocardiography A4ch clips and asked to assess LVEF. Following each clip assessment, only the intervention group was shown the results of the AI-based tool. For the final session, both groups were presented with a new set of 40 clips and asked to evaluate the LVEF.

Results: In the "normal–abnormal" category evaluation, as related to own baseline accuracy assessment, the intervention group had an improvement in accuracy on 50 consecutive clip assessments compared with a decline in the control group (0.10 vs. −0.12, respectively, $p = 0.038$). In the "significantly reduced LVEF" category, the intervention group showed significantly less decline in clip assessment as compared to the control group (−0.03 vs. −0.12, respectively, $p = 0.050$).

Conclusions: A study involving AI incorporation as a didactic tool for clinicians working in the emergency department appears feasible and acceptable. The introduction of an AI-based tool to clinicians working in the emergency department improved the assessment accuracy of LVEF as compared to the control group.

INTRODUCTION

Point-of-care ultrasound (POCUS) has long been considered a necessary skill for the emergency physician, with many new applications introduced each year.[1] Diagnostic and procedural expertise in bedside ultrasound are core competencies for all emergency medicine (EM) residency graduates.[2] While there are strict requirements for residency graduates in terms of numbers and types of ultrasound scans, the expected results are often general or binary- abdominal fluid present or absent; pericardial fluid present or absent; ejection fraction normal or grossly abnormal.[3,4] As applications become more sophisticated, the learning curve becomes steeper, with more teaching and practice required to master the technique.[5-7] POCUS for bedside echocardiographic applications, in particular, is known to have a steep learning curve, especially for more advanced applications such as the identification of abnormalities of the right side of the heart, regional wall motion abnormality, and accurate, left ventricular ejection fraction (LVEF) measurement.[8]

To overcome this hurdle, artificial intelligence (AI) is used in general ultrasound and echocardiography imaging by noncardiologist clinicians to aid in image acquisition, diagnosis, and decision making.[9] In particular, it is also being used to aid nonexpert clinicians in cardiac image acquisition and LVEF estimation.[10]

This randomized controlled pilot study aims to evaluate AI use as a didactic tool to improve the accuracy assessment of LVEF from the apical 4-chamber (A4ch) view by clinicians who work in the emergency department (ED).

METHODS

Study design

This prospective randomized controlled pilot study involved 16 physicians (including EM attendings, EM residents, and internal medicine residents) and physician assistants who work in the emergency department of a tertiary care medical center. The chosen participants routinely work at the front lines of evaluating patients with potential acute cardiac pathology in this institution's ED. Ultrasound machines that they use daily include the Sparq (Philips Healthcare; Amsterdam, Netherlands) and the Venue (GE Healthcare, Chicago, Illinois, United States).

The study used archived ultrasound clips acquired by a certified echocardiographic technician (equivalent to a Registered Diagnostic Cardiac Sonographer). The clips, taken from adult patients, were all deidentified to remove any patient's details. All clips were evaluated by two fellowship-trained echocardiographic cardiologists for LVEF whose assessment was set as the ground truth for further study requirements. In case of a ≤10% difference in their assessments, the mean LVEF was used, and if the difference was higher, we used Simpson's methods for exact LVEF calculation.

The clips were also evaluated with LVivoEF (DiA Imaging Analysis Ltd, Israel), a patented AI algorithm that evaluates the LVEF from the A4ch view. The algorithm results were compared and verified with the ground truth.

Study protocol

The study design is presented in Figure 1. The study took place after work in the late afternoon/early evening. After explaining the aims and design, study participants underwent a frontal didactic lecture delivered by a cardiology fellow (ZD) with 2 years of echocardiography experience. The lecture focused on LVEF and regional wall motion abnormality, including theory and clinical importance. All 17 segments were thoroughly explained and identified on both parasternal short-axis and A4ch views via illustrations and actual echocardiographic clips of deidentified patients. After this, eight A4ch view video clips and nine parasternal short-axis video clips were projected and the following details were explained for each: left ventricular function, normal or reduced; exact LVEF value; type of dysfunction (global or regional); and identifying segment, if any, that demonstrates reduced contraction. This study only evaluated the ability of the AI tool to aid clinicians working in the emergency department to assess LVEF from the A4ch view.

During the presentation, all participants were given time before presenting the results to assess the above parameters and if not
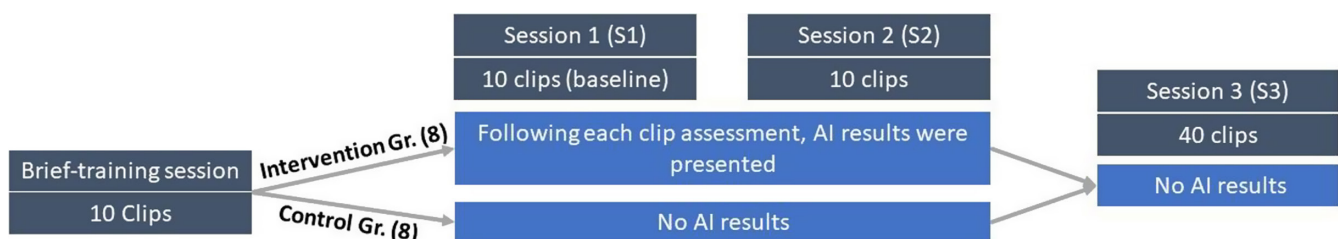
FIGURE 1 Flowchart depicting the study design. Gr.; group

corrected later, the examination was readdressed with a real-time focused explanation.

Following the training session, the participants were randomly divided into two groups, according to clinician subgroups: an intervention group and a control group. The groups were later separated into different rooms for the first trial session. The session was held by two separate nonmedical study members and was supervised by a physician (ZD and EAA) for verification of protocol adherence.

At the beginning of the first session, the intervention group was given a 10-min introductory session on the AI-based tool and how to interpret its report. The understanding was confirmed using a sampled report. Each participant was given an answer sheet with a unique identification number. There was a short series of questions that focused on demographics and prior POCUS training and experience. Then, for the first two sessions, both groups were presented with a total of 20 identical serial A4ch clips (10 clips for each session) and were requested to individually fill out a written form estimating the LVEF. Participants in both groups were given 50 s for each clip evaluation and documentation of their assessment. Following each clip assessment, only the intervention group was shown the AI-based tool assessment of the clip (35 s). The control group was blinded to the A4ch clip results. The clips were played continuously and no questions or talking were allowed.

For the third session, all participants rejoined in a single room. All participants were presented with a different set of 40 A4ch clips over 50 s per clip and were required to enter the same information as in the first session.

AI-based tool

The LVivo EF™ tool provides an objective automated AI-based EF analysis from the A4ch view. Within a few seconds, the tool outlines the internal borders of the LV with marked tracings that visually appear on the screen throughout the entire cardiac cycle with the results of an exact LVEF value next to it (Figure 2).

Measures

Feasibility

The feasibility of this randomized controlled pilot trial was assessed by the following: (1) breadth of recruitment strategies needed to achieve the required study sample size, (2) response rates to invitations to participate (calculated as actual study participants divided by entire potential target group individuals), and (3) retention rate throughout the trial.

Acceptability

Acceptability to this trial was assessed by measuring the rate of fully assessed clips out of the total presented clips. The participants were asked not to blindly guess and if they felt that they had no idea of the LVEF, to leave that assessment blank. To estimate whether the allocation to each of the trial arms affected the active assessments, we calculated this rate for the intervention and control groups, and compared the two for the difference.

LVEF assessment

LVEF assessments were compared to the ground truth (evaluation by two expert echocardiographers) according to categories: normal-abnormal LVEF ("normal-abnormal," <50%) and significant LVEF
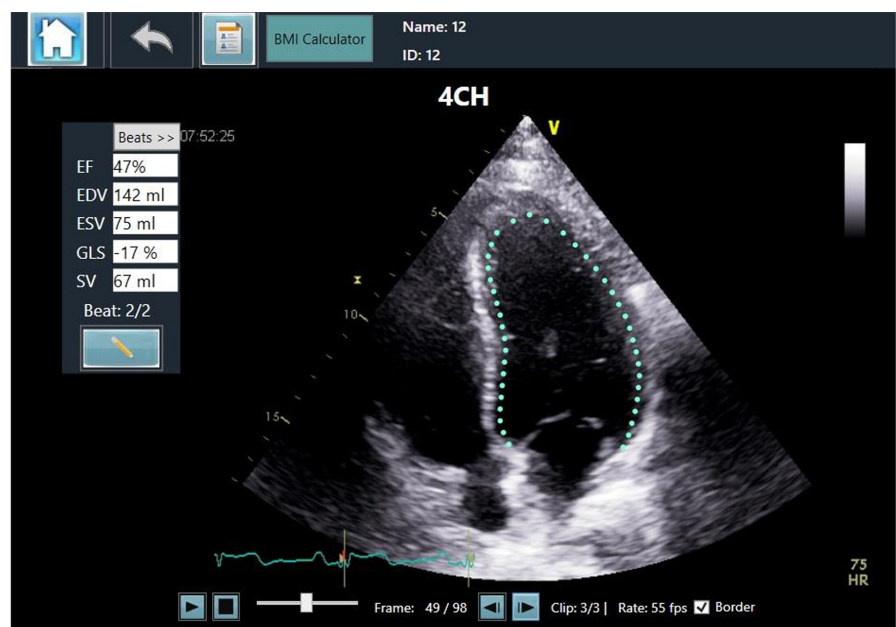
FIGURE 2 LVivoEF: Artificial intelligence-based tool for automated left ventricle function evaluation from an apical 4-chamber view echocardiographic clip

reduction ("significantly reduced LVEF," ≤40%). Each participant was scored for the number of correct assessments ("assessment accuracy") in each of the three sessions. Subsequently, the accuracy rate for each group was set for each session. For progress calculations, each group's accuracy rate in Session 1 was set as the baseline achievement. The percentage change of assessment accuracy for each of the two groups was calculated comparing the accuracy in Sessions 2 and 3 and combined Sessions 2 and 3 versus the accuracy in Session 1.

Finally, the progress was compared between the two groups according to the two predefined categories: "normal-abnormal" and "significantly reduced LVEF."

Sample size calculation

Sample size calculations were designed to meet the study endpoint and were performed using PS: Power and Sample Size Software (version 3.1.2. NCSS, LLC, Utah). We planned a study of the independent intervention group and controls with a 1:1 randomization ratio. Lacking previous data regarding the LVivoEF usage as a didactic tool, we assumed that the introduction of an AI-based tool to clinicians working in the emergency department will improve the assessment accuracy of LVEF by 12% as compared with participants without such exposure. Based on the prior data[11] and these assumptions, we calculated that data accrued from eight intervention group participants and eight controls would suffice to reject the null hypothesis with a probability (power) of 0.8. The type I error was calculated as 0.05.

Statistical analyses

Descriptive statistics were first used to describe baseline characteristics and the averaged scoring of the two groups in the different sessions to study variable frequencies and distributions to determine the need for nonparametric methods. The unadjusted association of baseline characteristics between the two groups was initially studied using the Fisher exact test. Echocardiography LVEF assessments and the progress between the two groups for Sessions 2 and 3 and combined Sessions 2 and 3 were assessed

using the t-test or Mann–Whitney U test for normally distributed and nonnormally distributed continuous variables, respectively. The combined progress of Sessions 2 and 3 was set as the main indicator for progress. AI-based tool assessment was compared to the ground truth for linear correlation using the Pearson correlation coefficient. The r values <0.3, 0.3 to 0.5, 0.5 to 0.7, and ≥0.7 were considered to represent poor, poor to fair, fair to good, and excellent agreement, respectively. The sensitivity and specificity of the AI-based tool and the ground truth were examined using a cutoff of 40% as the LVEF. Finally, the two expert echocardiographers' assessments were assessed for interobserver correlation using the Pearson correlation coefficient. Covariates with p values of less than 0.05 were considered statistically significant. Statistical analyses were performed using SPSS Statistics for Windows version 21 (SPSS Inc., Chicago, IL).

All participants provided written consent. The study received approval from the hospital's Institutional Review Board.

RESULTS

Participants' characteristics

Sixteen participants, 11 physicians and 5 physician assistants, were recruited to the trial (5 females, all of whom were physicians, 31.3%) and randomized to the two study groups. Baseline characteristics between the two groups did not differ (Table 1).

Feasibility

The study target sample size ($n = 16$) was achieved by directly approaching potential participants and by sending an invitation via email and WhatsApp groups. The rate of participation was 40% (potential participants included 40 physicians and physician assistants). Two participants were assigned to the trial but were unable to attend due to last-minute schedule limitations. Retention rates throughout the trial were excellent with no participant dropouts (Retention rate = 100%).

TABLE 1 Clinician characteristics and prior POCUS exposure of intervention group vs control group

Variable	All n = 16	Intervention group n = 8	Control group n = 8
Female sex, n (%)	5 (31.3)	3 (37.5)	2 (25.0)
Physician medical profession, n (%)	11 (68.8)	6 (75.0)	5 (62.5)
Residents, n (%)	9 (56.3)	5 (62.5)	4 (50.0)
Attendings, n (%)	2 (12.5)	1 (12.5)	1 (12.5)
Physician assistants, n (%)	5 (31.3)	2 (25.0)	3 (37.5)
Prior POCUS course, n (%)	9 (56.3)	4 (50.0)	5 (62.5)
Prior POCUS practice, n (%)	1 (6.3)	0 (0.0)	1 (12.5)

Note: No statistically significant difference was demonstrated between the two groups.

Abbreviations: n, number; POCUS, point-of-care ultrasound.

Acceptability

Among the total cohort, the rates of assessed clips for each of the sessions were 94.4% for session 1, 98.1% for session 2, and 98.9% for session 3. The rates among the intervention group were 97.5% for session 1, 98.8% for session 2, and 100% for session 3, and among the control group were 91.2% for session 1, 97.5% for session 2, and 97.8% for session 3. The rates of assessed clips did not differ between the two trial groups (p values of 0.438, 0.554, and 0.277, respectively).

LVEF assessments

The ground truth and the mean assessment of both groups and the AI are presented in Table 2. In each of the sessions, both groups underestimated the LVEF as related to the ground truth ($p \leq 0.017$), whereas AI did not differ from the ground truth in

Sessions 1 or 2 ($p = 0.102$ and $p = 0.093$, respectively, data not shown).

Study participants and the ground truth comparison

Comparing the percentage change of assessment accuracy in the "normal-abnormal" category as compared with the control group, the intervention group improved in Session 2 vs. Session 1 (intervention: +0.17, control: −0.26, $p = 0.010$), had a nonsignificant improvement in Session 3 vs. Session 1 (intervention: +0.09, control: −0.08, $p = 0.083$), and improved in combined Sessions 2+3 vs. Session 1 (intervention: +0.10, control: −0.12, $p = 0.038$) (Figure 3).

Comparing the percentage change of accuracy assessment in the "significantly reduced LVEF" category as compared with the control group, the intervention group had a statistically similar improvement in Session 2 vs. Session 1 (intervention: +0.06, control: +0.11, $p = 0.645$), less decline in Session 3 vs. Session 1 (intervention: −0.05,
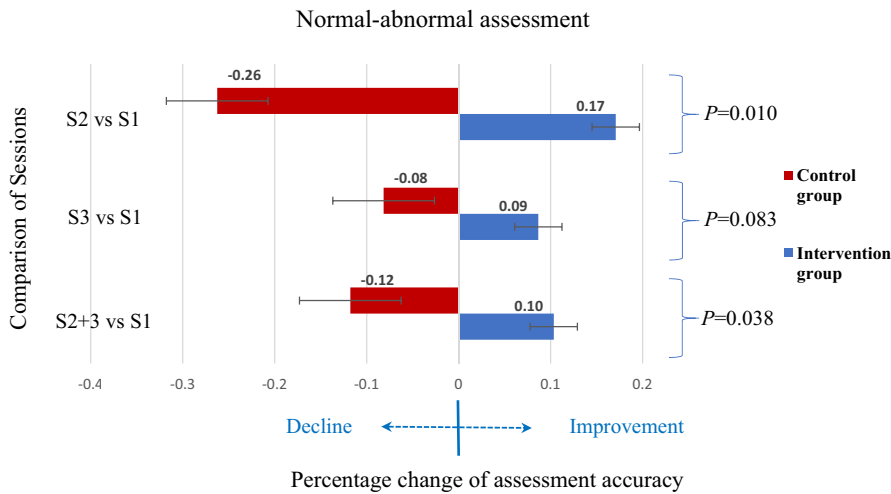
TABLE 2 Echocardiography EF assessment in each session according to expert echocardiographers (ground truth), intervention group, control group, and AI-based tool

| Variable	Session 1 $n = 10$	Session 2 $n = 10$	Session 3 $n = 40$	p value[a]
Gold Standard				
Ejection fraction, % (mean ± SD)	50.6 ± 13.6	53.3 ± 14.1	47.1 ± 12.3	0.648
Systolic function classification				
Normal/preserved (EF ≥ 50%), n (%)	5 (50.0%)	7 (70.0%)	18 (45.0%)	
Mildly reduced (40% <EF < 50%), n (%)	3 (30.0%)	1 (10.0%)	9 (22.5%)	
Significantly reduced (EF ≤ 40%), n (%)	2 (20.0%)	2 (20.0%)	13 (32.5%)	
Intervention group assessment[b]				
Ejection fraction, % (mean ± SD)	43.2 ± 12.1	47.3 ± 16.5	42.9 ± 13.9	0.891
Systolic function classification				
Normal/preserved (EF ≥ 50%), n (%)	3 (30.0%)	5 (50.0%)	16 (40.0%)	
Mildly reduced (40% < EF < 50%), n (%)	4 (40.0%)	2 (20.0%)	9 (22.5%)	
Significantly reduced (EF ≤ 40%), n (%)	3 (30.0%)	5 (50.0%)	15 (37.5%)	
Control group assessment[b]				
Ejection fraction, % (mean ± SD)	42.9 ± 9.0	44.1 ± 11.7	40.9 ± 7.4	0.557
Systolic function classification				
Normal/preserved (EF ≥ 50%), n (%)	4 (40.0)	3 (30.0%)	5 (12.5%)	
Mildly reduced (40% < EF < 50%), n (%)	2 (20.0%)	3 (30%)	14 (35%)	
Significantly reduced (EF ≤ 40%), n (%)	4 (40.0%)	4 (40.0%)	21 (52.5%)	
AI-based tool assessment				
Ejection fraction, % (mean ± SD)	45.0 ± 20.5	46.9 ± 21.1		0.661
Systolic function classification				
Normal/preserved (EF ≥ 50%), n (%)	6 (60.0%)	4 (40.0%)		
Mildly reduced (40% < EF < 50%), n (%)	2 (20.0%)	3 (30.0%)		
Significantly reduced (EF ≤ 40%), n (%)	2 (20.0%)	3 (30.0%)		

Abbreviations: AI, artificial intelligence; EF, ejection fraction; n, number; SD, standard deviation.
[a]p value refers to the comparison between Sessions 2+3 and Session 1.
[b]Averaged assessment is presented in each session.

Normal-abnormal assessment

FIGURE 3 Comparison of the percentage change of assessment accuracy for normal-abnormal echocardiography category between intervention and control groups: Further accuracy rates in Sessions 2 (S2) and 3 (S3) are compared to baseline accuracy rate in Session 1 (S1)

Significantly reduced ejection fraction assessment
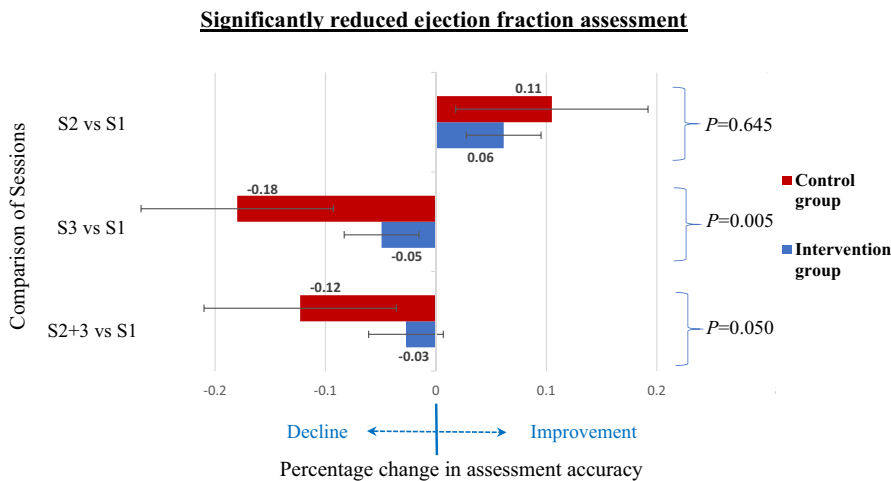
FIGURE 4 Comparison of percentage change of assessment accuracy for significantly reduced left ventricular ejection fraction echocardiography category between intervention and control groups: Further accuracy rates in Sessions 2 (S2) and 3 (S3) are compared to baseline accuracy rate in Session 1 (S1)

control: −0.18, $p = 0.005$), and less decline in combined Sessions 2+3 vs. Session 1 (intervention: −0.03, control: −0.12, $p = 0.050$) (Figure 4).

AI assessment and the ground truth comparison

Comparing AI assessment to the ground truth for the clips presented at Sessions 1+2 showed an almost perfect correlation ($r = 0.889$, $p < 0.001$) (Figure 5). The sensitivity and specificity of the AI-based tool and the ground truth examining LVEF cutoff of 40% were 0.86 and 0.85, respectively (data not shown).

Expert echocardiographers interobserver comparison

Comparing the two expert echocardiographers' assessments showed an almost perfect correlation ($r = 0.826$, $p < 0.001$).

DISCUSSION

This randomized controlled pilot study demonstrated that research involving AI incorporation as a didactic tool appears to be feasible and

acceptable. In the "normal-abnormal" category evaluation, as related to own baseline assessment accuracy, the intervention group (i.e. the group that was shown the AI results) had an improvement in accuracy on 50 consecutive clip assessments compared with a decline in the control group. Moreover, in the "significantly reduced LVEF" category evaluation, the intervention group had a significantly less decline on the 50 consecutive clips assessment compared to the control group.

The use of AI is currently in its infancy in EM and is being used and studied mostly for operational improvement and clinical prediction modeling.[12,13] While AI is currently used in echocardiography interpretation, it is becoming more widespread in medical education across specialties. It is currently being used mostly for learning support.[14] One recent study supports the use of AI to assist medical students in identifying hip fractures.[15] As radiology is going to be positively impacted by the AI revolution, residents are already being prepared for its integration into their practice.[16] Students on an ophthalmology rotation have used AI to improve their understanding of congenital cataracts.[17] AI is even being integrated into surgical education in simulated virtual operation training.[18] In the field of cardiology, AI is being studied to improve image acquisition and aid in diagnostic evaluation such as estimating LVEF.[10] This study shows that AI can be a tool to help physicians and physician assistants working in the emergency

FIGURE 5 Correlation of AI-based tool assessment and the ground truth on echocardiographic clips in sessions 1 and 2
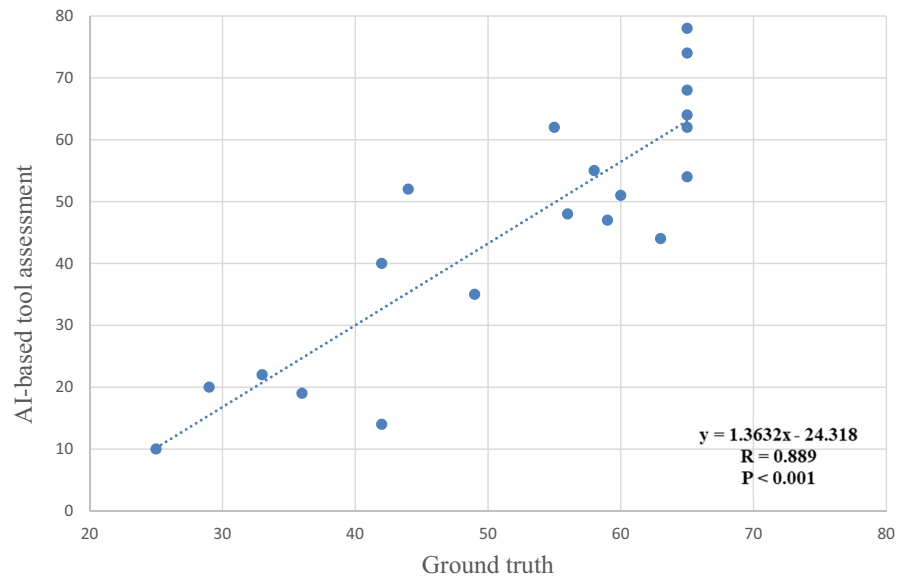
department acquire the skills to make clinically important estimations of LVEF. An important advantage of this AI-based tool is the concomitant visual tracing of LV internal borders, appearing throughout the cardiac cycle and thus presenting to the observer the LV dynamic contraction and relaxation. Also, other parameters are being evaluated by this tool (including LV systolic and diastolic volumes, stroke volume, and global longitudinal strain). A meaningful challenge currently not addressed by this tool is the requirement for correct clip acquisition that may have an impact on the LVEF assessment necessitating proper manual teaching before POCUS use. The AI-based tool (LVivoEF) is available for real-time use in POCUS settings, and though not studied specifically in this current trial, the AI-based tool may potentially improve the learning curve of LVEF assessment accuracy during routine clinical use.

One of the most common applications for the knowledge of the LVEF is diagnosing and managing exacerbations of congestive heart failure.[19] However, other concerns such as undifferentiated shock, chest pain, or shortness of breath may include in their differential diagnosis entities such as myocarditis or sepsis where it is also important to evaluate the LVEF. General knowledge of the LVEF is also critical in guiding antiarrhythmic drug choice for the patient with atrial fibrillation. For example, certain medications such as propafenone and flecainide would be contraindicated in the patient with a low ejection fraction.[20]

Part of the success of the group exposed to the AI program was that it resulted in less cognitive decline compared to the control group. The challenges of long-term knowledge retention, skill mastery, and cognitive fatigue have been well studied in the setting of medical education.[21,22] Part of the challenge of this study may have been that it was conducted late in the afternoon and evening hours after the groups participated in a full day of work and as such may represent underdiagnosis of the two groups' improvements in further sessions and may explain the decline in the consecutive third session among the control group. In this respect, this study may reflect the real world of clinicians working long hours in the busy and chaotic emergency department.

Other methods of training nonexpert clinicians have been described. One such device for self-training is a simulator.[23] However, this may incur a significant expense. As AI becomes more ubiquitous in clinical medicine, the algorithms can be programmed into existing POCUS platforms.

Limitations

This study only evaluated the ability of the didactic tool to aid the participant to identify LVEF based on previously collected video clips. It does not reflect the real-world setting whereby the clinician has to collect the images and then interpret them. The study incorporated all staff clinicians in this ED including EM attendings, EM residents, internal medicine residents, and physician assistants. This reflects the real-world ED at this institution as well as most ED in the country where this study took place. The results could perhaps be different if only certain subgroups were evaluated. Also, the study involved a relatively small number of participants from a single medical center, which should be taken into account when drawing conclusions. The participants were generally aware that they were part of a study to evaluate AI as a didactic tool, thereby raising the issue of bias. To minimize this bias, we did not compare absolute accuracy rates between the groups, but each participant was compared to their own achievements and only the progress rate was then compared between the groups. Lastly, the assessment accuracy was evaluated only on A4ch clips and the generalizability of its conclusions on other views needs to be demonstrated.

CONCLUSION

This randomized controlled pilot study demonstrated that research involving AI incorporation as a didactic tool appears to be feasible and acceptable. The introduction of an AI-based tool to clinicians

working in the emergency department improved the assessment accuracy of LVEF as compared to the control group. Studies should be conducted with a larger sample size of participants in a real-world setting with the hands-on acquisition of images taken from multiple views. Also, studies should be conducted on other clinically important cardiac POCUS assessment skills such as the identification of regional wall motion abnormality and right ventricular function.

CONFLICT OF INTEREST

There are no conflicts of interest on the part of any of the authors.

AUTHOR CONTRIBUTIONS

ZD contributed to the study concept and design, acquisition of the data, analysis and interpretation of the data, drafting of the manuscript, critical revision of the manuscript for important intellectual content, and statistical expertise. AB contributed to the acquisition of the data and critical revision of the manuscript for important intellectual content. DR contributed to the acquisition of the data and critical revision of the manuscript for important intellectual content. LAS contributed to the acquisition of the data and critical revision of the manuscript for important intellectual content. MG contributed to the study concept and design, analysis and interpretation of the data, and critical revision of the manuscript for important intellectual content. EAA contributed to the study concept and design, acquisition of the data, analysis and interpretation of the data, drafting of the manuscript, and critical revision of the manuscript for important intellectual content.

REFERENCES

1. Reardon R, Heegaard B, Plummer D, Clinton J, Cook T, Tayal V. Ultrasound is a necessary skill for emergency physicians. *Acad Emerg Med.* 2006;13(3):334-336. doi:10.1197/j.aem.2006.01.003
2. Beeson MS, Ankel F, Bhat R, et al. 2019 EM Model Review Task Force, Keehbauch JN; American Board of Emergency Medicine. The 2019 Model of the Clinical Practice of Emergency Medicine. *J Emerg Med.* 2020;59(1):96-120. doi:10.1016/j.jemermed.2020.03.018
3. Henneberry RJ, Hanson A, Healey A, et al. Use of point of care sonography by emergency physicians. *Can J Emerg Med.* 2012;14:106-112. doi:10.2310/8000.CAEPPS
4. Nilsson PM, Todsen T, Subhi Y, Graumann O, Nolsøe CP, Tolsgaard MG. Cost-effectiveness of mobile app-guided training in Extended Focused Assessment with Sonography for Trauma (eFAST): A randomized trial. *Ultraschall Med.* 2017;38:642-647. doi:10.1055/s-0043-119354
5. Gómez Betancourt M, Moreno-Montoya J, Barragán González AM, Ovalle JC, Bustos Martínez YF. Learning process and improvement of point-of-care ultrasound technique for subxiphoid visualization of the inferior vena cava. *Crit Ultrasound J.* 2016;8:4. doi:10.1186/s13089-016-0040-1
6. Zeiler FA, Ziesmann MT, Goeres P, et al. A unique method for estimating the reliability learning curve of optic nerve sheath diameter ultrasound measurement. *Crit Ultrasound J.* 2016;8:9. doi:10.1186/s13089-016-0044-x
7. Deacon AJ, Melhuishi NS, Terblanche NC. CUSUM method for construction of trainee spinal ultrasound learning curves following standardised teaching. *Anaesth Intensive Care.* 2014;42:480-486. doi:10.1177/0310057X1404200409
8. Wright J, Jarman R, Connolly J, Dissmann P. Echocardiography in the emergency department. *Emerg Med J.* 2009;26:82-86. doi:10.1136/emj.2008.058560
9. Muse ED, Topol EJ. Guiding ultrasound image capture with artificial intelligence. *Lancet.* 2020;396:749. doi:10.1016/S0140-6736(20)31875-4
10. Schneider M, Bartko P, Geller W, et al. A machine learning algorithm supports ultrasound-naïve novices in the acquisition of diagnostic echocardiography loops and provides accurate estimation of LVEF. *Int J Cardiovasc Imaging.* 2020;1-10. doi:10.1007/s10554-020-02046-6
11. Hope MD, de la Pena E, Yang PC, Liang DH, McConnell MV, Rosenthal DN. A visual approach for the accurate determination of echocardiographic left ventricular ejection fraction by medical students. *J Am Soc Echocardiogr.* 2003;16(8):824-831. doi:10.1067/S0894-7317(03)00400-0
12. Grant K, McParland A, Mehta S, Ackery AD. Artificial intelligence in emergency medicine: Surmountable barriers with revolutionary potential. *Ann Emerg Med.* 2020;75(6):721-726. doi:10.1016/j.annemergmed.2019.12.024
13. Ehrlich H, McKenney M, Elkbuli A. The niche of artificial intelligence in trauma and emergency medicine. *Am J Emerg Med.* 2021;45:669-670. doi:10.1016/j.ajem.2020.10.050
14. Chan KS, Zary N. Applications and challenges of implementing artificial intelligence in medical education: Integrative review. *JMIR Med Educ.* 2019;5:e13930. doi:10.2196/13930
15. Cheng CT, Chen CC, Fu CY, et al. Artificial intelligence-based education assists medical students' interpretation of hip fracture. *Insights Imaging.* 2020;11:119. doi:10.1186/s13244-020-00932-0
16. Simpson SA, Cook TS. Artificial Intelligence and the Trainee Experience in Radiology. *J Am Coll Radiol.* 2020;17:1388-1393. doi:10.1016/j.jacr.2020.09.028
17. Wu D, Xiang Y, Wu X, et al. Artificial intelligence-tutoring problem-based learning in ophthalmology clerkship. *Ann Transl Med.* 2020;8:700. doi:10.21037/atm.2019.12.15
18. Mirchi N, Bissonnette V, Yilmaz R, Ledwos N, Winkler-Schwartz A, Del Maestro RF. The Virtual Operative Assistant: An explainable artificial intelligence tool for simulation-based training in surgery and medicine. *PLoS One.* 2020;15:e0229596. doi:10.1371/journal.pone.0229596
19. Sugahara M, Masuyama T. Echocardiography tips in the emergency room. *Heart Fail Clin.* 2020;16:167-175. doi:10.1016/j.hfc.2019.12.003
20. Hindricks G, Potpara T, Dagres N, et al. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association of Cardio-Thoracic Surgery (EACTS). *Eur Heart J.* 2021;42:373-498. doi:10.1093/eurheartj/ehaa612
21. Moazed F, Cohen ER, Furiasse N, et al. Retention of critical care skills after simulation-based mastery learning. *J Grad Med Educ.* 2013;5:458-463. doi:10.4300/JGME-D-13-00033.1
22. McGraw R, Chaplin T, Rocca N, et al. Cognitive load theory as a framework for simulation-based, ultrasound-guided internal jugular catheterization training: Once is not enough. *CJEM.* 2019;21:141-148. doi:10.1017/cem.2018.456
23. Bernard A, Chemaly P, Dion F, et al. Evaluation of the efficacy of a self-training programme in focus cardiac ultrasound with simulator. *Arch Cardiovasc Dis.* 2019;112:576-584. doi:10.1016/j.acvd.2019.06.001