

# Learning curve-cumulative summation analysis of visual estimation of left ventricular function in novice practitioners

## A STROBE-compliant article

Yoonje Lee, MD, PhD<sup>a</sup>, Hyungoo Shin, MD<sup>b</sup>, Changsun Kim, MD, PhD<sup>b,\*</sup>, Inhye Lee, MD<sup>b</sup>, Hyuk Joong Choi, MD, PhD<sup>b</sup>

### Abstract

The aim of the present study was to determine the value of cumulative summation (CUSUM) analysis in assessing the proficiency of novice practitioners in estimating the left ventricular ejection fraction (EF).

Seven novice practitioners with no echocardiography experience were recruited in this observational study. Each practitioner assessed EF from echocardiographic video files of 100 cases, one by one, and received feedback and teaching. We obtained a CUSUM score through comparison of the gold standard values of EF and the EF values determined by the practitioners. Then, the practitioners underwent the same test 4 weeks later, except without feedback and teaching, using echocardiographic video files from 100 other cases.

The mean number of visual estimation cases required to pass the learning curve (LC)-CUSUM test was  $56.3 \pm 9.1$  (95% CI 47.8–64.7). The LC-CUSUM average of the 7 novice practitioners showed improvement in visual estimation skill, with an average acceptable level achieved after a mean experience of 55 cases. In the test performed after 4 weeks, 5 of the 7 novice practitioners showed significantly good overall agreement. All novice practitioners had a kappa coefficient greater than .8, and significant and almost perfect agreement was observed. All the participants exhibited a percentage of correct answers greater than 81%.

We found that the novice practitioners could acquire an acceptable level of skill for estimating EF with short-term, self-learning-focused echocardiographic training.

**Abbreviations:** CUSUM = cumulative summation, ED = emergency department, EF = ejection fraction, FADE = Focused Assessment Diagnostic Echocardiography, ICU = intensive care unit, LC = learning curve, LC-CUSUM = learning curve-cumulative summation, LV = left ventricular, LVEF = LV ejection fraction, MSM = Modified Simpson's method.

**Keywords:** Cumulative summation analysis, echocardiography, ejection fraction, learning curve, visual assessment

## 1. Introduction

Echocardiography is a noninvasive method of providing immediate assessment of cardiac function at the bedside and is the most commonly used echocardiographic parameter for evaluating left ventricular (LV) systolic function, known as the LV ejection

fraction (LVEF).<sup>[1]</sup> The reliable assessment of LVEF is essential, especially for evaluating the hemodynamic status of unstable, time-sensitive, critically ill patients.<sup>[2,3]</sup> Although the modified Simpson method (MSM) has been regarded as the standard method for evaluating EF, several critical care physicians have focused on visually estimating EF (eyeballing EF) due to its high reliability and simple application. Thus, this method of estimating EF has been widely used as an alternative method for evaluating cardiac contractility, especially in the emergency department (ED) and intensive care unit (ICU). However, several studies have shown that although eyeballing EF is fairly well correlated with the MSM,<sup>[4]</sup> this method is dependent on the operator's experience. Therefore, we investigated how much experience a novice physician without any experience with echocardiography would require to achieve an acceptable level of performance for eyeballing EF using learning curve-cumulative summation (LC-CUSUM) analysis. LC-CUSUM analysis was developed to signal when an individual's performance had achieved a predefined level of competence. This method has been applied to the analysis of the LC and quality control for monitoring clinical performance.<sup>[5]</sup> Additionally, to the best of our knowledge, LC-CUSUM analysis of competence and quality of visual estimation of LV function through echocardiographic video images in the ED has not yet been studied in the emergency medicine literature.

This study aimed to investigate the process of visual estimation skill acquisition for evaluating EF using 100 echocardiographic

Editor: Salvatore De Rosa.

This study was not supported by any funding sources.

The authors are responsible for the content and composition of this paper.

The authors have no conflicts of interest to disclose.

<sup>a</sup> Department of Emergency Medicine, College of Medicine, Korea University Anam Hospital, Seoul, <sup>b</sup> Department of Emergency Medicine, College of Medicine, Hanyang University Guri Hospital, Guri, Republic of Korea.

\* Correspondence: Changsun Kim, Department of Emergency Medicine, College of Medicine, Hanyang University Guri Hospital, 153, Gyeongchun-ro(st), Guri-si, Gyeonggi-do 11923, Korea (e-mail: flyes98@naver.com).

Copyright © 2019 the Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Medicine (2019) 98:14(e15191)

Received: 27 December 2018 / Received in final form: 14 March 2019 /

Accepted: 18 March 2019

<http://dx.doi.org/10.1097/MD.0000000000015191>

**Table 1****Baseline characteristics of patients of video echocardiographic files.**

Patients (n=100)	Primary learning	Follow-up test
Sex (F: M) (%)	61: 39	63: 37
Age (F: M) (mean±SD)	74.3±11.5: 67.5±13.8	76.7±9.8: 69.1±15.7

F=female, M=male, SD=standard deviation.

video files by 7 novice practitioners, and a follow-up investigation was conducted after 4 weeks using the LC-CUSUM test.

## 2. Methods and materials

### 2.1. Study design and setting

This study was conducted to analyze the consecutive attempts of novice practitioners of eyeballing EF between July 2015 and November 2015 at an academic ED in a tertiary urban hospital. The institutional review board of our institution approved this study (HYI-15-149-1). We registered the study protocol at ClinicalTrials.gov before study initiation (Clinicaltrials.gov: NCT02866318)

### 2.2. Study participants

Echocardiographic video files, without personal information, from 200 patients (aged > 18 years) who visited the ED with a complaint of chest pain or dyspnea between July 2014 and June 2015 were collected in this study. The standard EF values were measured by a cardiologist using the MSM. Two echocardiographic specialists reviewed the videos, and values that had been incorrectly measured were excluded. Patients who were intubated and pregnant were excluded, and those who had arrhythmia, a previous valvular disease or a history of cardiac surgery were also excluded.

Seven medical students who had no experience with echocardiography prior to the study were recruited as novice practitioners to analyze the performance of eyeballing EF. They received a 1 hour lecture including echocardiographic videos that had been classified into 4 categories of wall motion according to the reduced degree of EF (Table 1). A total of 20 patient video files (5 samples of each category) were provided during the training session, and the participants were instructed on how to estimate EF by observing the LV contractility. The movement of the anterior leaflet of the mitral valve (E-point septal separation), percentage of myocardial thickening, fractional shortening, and level of ascent of the base of the heart in the apical 4 chamber view were mainly used when eyeballing EF.<sup>[6]</sup>

### 2.3. Equipment and materials

We use a Logiq 7 ultrasonographic system (M7R model) and a 3.6-MHz microconvex transducer (GE Medical Systems, Milwaukee, USA) to evaluate cardiac function in our ED. The video files of the parasternal long axis, parasternal short axis, and apical 4-chamber view of each patient were transferred to a desktop computer. We used 100 of 200 patient video files, and each case was randomly assigned a number from 1 to 100 for the learning course, regardless of their EF values. We created a PowerPoint presentation that showed each case in order of the assigned number. The reference slide consisted of the correct

**Table 2****Four categories of wall motion according to the reduced degree of ejection fraction.**

Categories of wall motion	Values (%)
1. Normal	> 50
2. Mild reduced	40–49
3. Moderated reduced	30–39
4. Severe reduced	< 30

answer (appropriate classification of 4 predesigned categories), and an explanation followed each case. The last slide of each case compared the echocardiographic videos of that case and the sample videos for each of the 4 categories. Using another 100 patient files, we prepared a follow-up test after 4 weeks. We arranged the PowerPoint file with cases in a random order and distributed them randomly to each practitioner, minimizing the impact of case order on the practitioner's education.

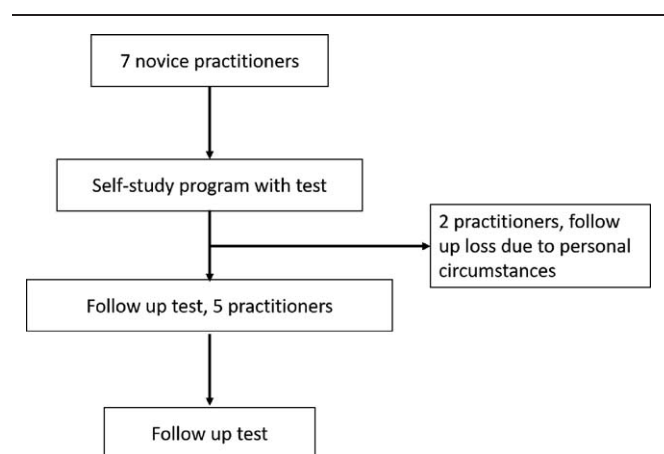
### 2.4. Intervention

The novice participants attempted eyeballing EF by watching the predesigned PowerPoint slide on a desktop monitor. They determined the EF according to the preclassified 4 categories of wall motion (Table 2). After rating the EF, the participants viewed the next slide, which contained the correct answer and an explanation. They could verify whether their rating was correct. Lastly, they compared the wall motion of each case with those of the preselected reference videos representing the 4 categories of wall motion. They received feedback after determining the EF of each case and in this manner could learn how to correctly eyeball EF via self-learning.

The second procedure was performed after 4 weeks to evaluate the same participants' ability to eyeball EF. The participants eyeballed the EF while watching a randomly assigned predesigned slide containing only echocardiographic videos of each case. The participants did not receive any feedback in this phase (Fig. 1).

### 2.5. Outcomes and statistical analysis

The LC-CUSUM test has a holding barrier of zero, which prevents the score from drifting too far from the decision limit

**Figure 1.** Flow diagram.

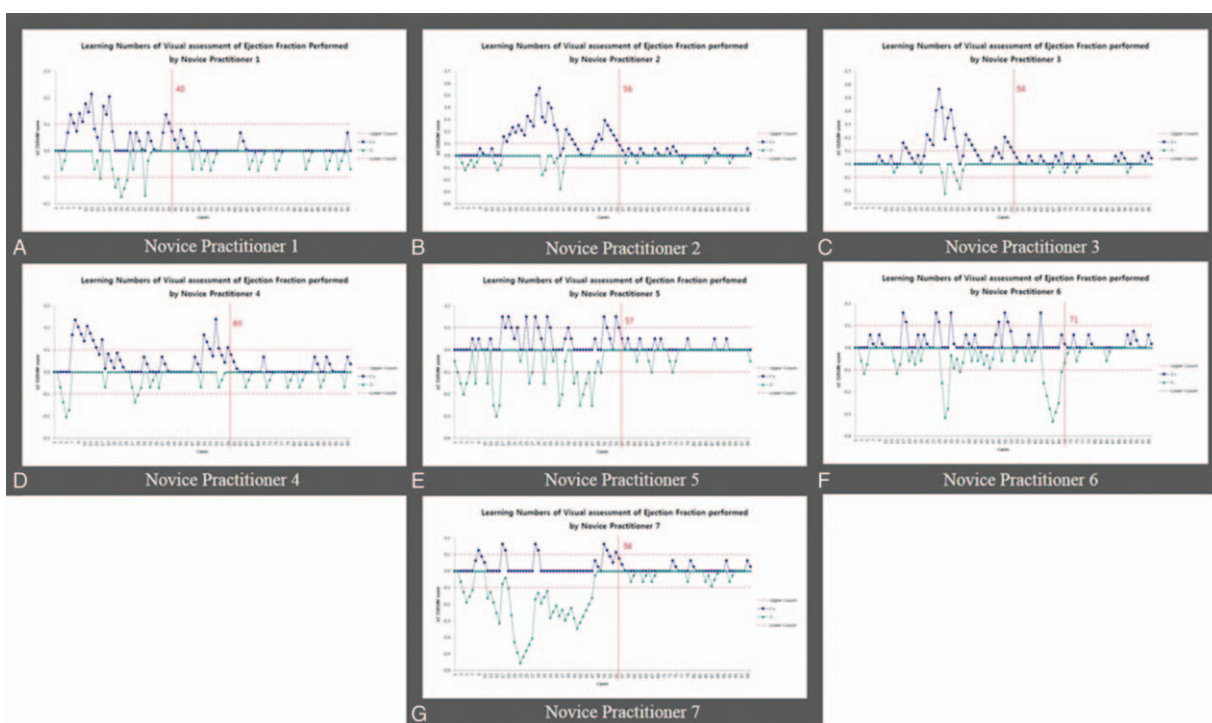


Figure 2. Each novice practitioner's Learning Curve-CUSUM charts.

range. When the graph approaches this limit, it stays in the limit area, which includes zero, and monitoring continues. Therefore, the participants do not have to compensate for all the accumulated failures to show acceptable performance. The CUSUM is designed to detect a shift from an adequate to an inadequate performance level. The concordance outcome is recorded above the x-axis. The CUSUM score increased with discordance and decreased with concordance between the gold standard and the EF value estimated by the participant. The CUSUM can be used after the LC-CUSUM has shown that the novice practitioner reached proficiency to ensure that the performance is maintained within an acceptable range.

To analyze the LC-CUSUM test, we assigned points to each case by comparing the categories determined by the novice practitioners and those determined by the MSM (standard values). If the category range of the 2 methods was concordant, we assigned zero points. If the category according to the 2 methods was discordant by as much as 1 category range, we assigned  $\pm 1$  point. Using this rule, we also assigned  $\pm 2$  and  $\pm 3$  points for each case. When the score reached a certain predefined limit, the LC-CUSUM test indicated that the visual estimation skill of the participant had reached an adequate level. For example, 0 points were assigned when the EF range of the gold standard result matched the category rating of the participant.

LC-CUSUM calculations were performed using SAS University Edition (SAS Institute Inc, Cary, NC). Continuous variables are expressed as the means with 95% CIs, while categorical variations are expressed in numbers and percentages.

The primary outcomes of 7 practitioners were the LC-CUSUM graph and the number of cases of visual estimation required to achieve an acceptable EF visual estimation performance level. After 4 weeks, 2 practitioners were not participated because of personal circumstances. The secondary outcome of 5

practitioners was the weighted kappa index of each novice practitioner's evaluation performance level after 4 weeks.

### 3. Results

A total of 100 echocardiographic video files were analyzed by 7 novice practitioners for visual assessment self-learning. Among the 100 cases, 61 patients were female with a mean age of  $74.3 \pm 11.5$  years old, and 39 patients were male with a mean age of  $67.5 \pm 13.8$  years. After 4 weeks, another 100 echocardiographic video files were analyzed by 5 novice practitioners among the 7 participants. Among 100 cases, 63 patients were female, 37 patients were male with average ages of  $76.7 \pm 9.8$  and  $69.1 \pm 15.7$  years, respectively (Table 2).

#### 3.1. LC-CUSUM analysis of visual assessment of LV systolic function by novice practitioners

The mean number of visual estimation cases required to reach a proficient level according to the LC-CUSUM test was  $56.3 \pm 9.1$  (95% CI 47.8–64.7). The LC-CUSUM graph of visual estimation by 7 novice practitioners is presented in Figure 2. This figure shows that the novice practitioners reached an acceptable level at a minimum of 40 cases and a maximum of 71 cases.

In the CUSUM chart of novice practitioner 1, the first 4 cases were judged to be lower than the standard result of the MSM; after feedback and self-learning, the results were judged to be higher than the standard result of the MSM. As this pattern repeated, it gradually reached the acceptable range, and from the 40th case onward, the results exhibited an acceptable level.

In the CUSUM chart of practitioner 2, at first, the ratings were lower than the standard results. After feedback and self-learning, the ratings increased. After another feedback and self-learning

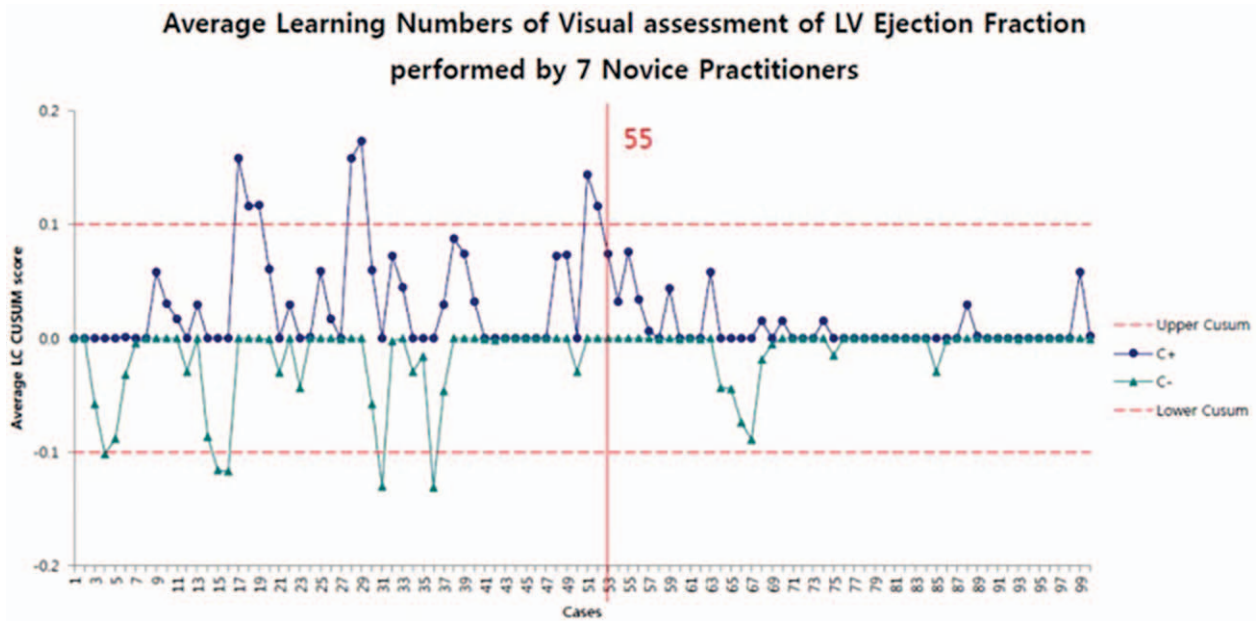


Figure 3. Average of 7 novice practitioner’s Learning Curve-CUSUM charts.

phase, the ratings that had decreased below the standard result were more highly rated. The rating, which was now higher, decreased again with the feedback and self-learning system, and finally the ratings reached an acceptable range after 56 cases.

In the CUSUM chart of practitioner 3, the acceptable range was maintained beginning at the 15th case, but the ratings did not exceed the limit of the novice practitioner, and eventually, the ratings started to deviate from the acceptable range after the 15th case. The practitioner repeatedly judged the function of LV to be higher and then lower than the standard, and finally returned to an acceptable range at case 54.

In the CUSUM chart of novice practitioner 4, from the beginning, LV function was judged to be lower than the standard results. The ratings were corrected through the feedback and self-learning system, but then the ratings were higher than the standard results. Through this process, the practitioner reached the acceptable range at the 20th case, which seemed to be well maintained, but the 50th case was again out of range. Eventually, from the 60th case onward, the ratings were maintained in the acceptable range.

In the CUSUM chart of novice practitioner 5, the participant mainly under-estimated LV function, and after receiving the feedback and self-learning, repeatedly slightly over-estimated and under-estimated LV function until reaching the acceptable range after 57 cases.

In the CUSUM chart of novice practitioner 6, initially the practitioner varied in and out of the acceptable range but did not maintain a pattern until the 70th case. From the 71st case onward, the ratings were maintained within the acceptable range.

In the CUSUM chart of novice practitioner 7, this practitioner under-estimated LV function from the beginning. However, through the feedback and self-learning system, this practitioner entered the acceptable range after 56 cases.

The LC of the novice practitioners showed improvement in visual estimation skill (Fig. 3), with an average acceptable level after 55 cases according to the CUSUM charts.

### 3.2. Effect of the feedback and self-learning program after 4 weeks

In the test 4 weeks later, 5 practitioners among the 7 novice practitioners exhibited significantly good overall agreement (Table 3). All the novice practitioners had kappa coefficients higher than 0.8, and significant, almost perfect agreement was observed. All the participants provided at least 81% correct answers (Table 3).

## 4. Discussion

Echocardiography is an essential study for evaluating patients with symptoms related to cardiac problems. Its use has been widely adopted in various areas such as EDs and ICUs. Considering that cardiology experts are not always available, we believe that emergency physicians, intensivists, and critical care physicians should be able to perform emergency echocardiography, including determining cardiac function. Eyeballing EF is known to be a relatively accurate and easily applicable method for determining cardiac function;<sup>[1]</sup> thus, noncardiologists have focused on eyeballing EF instead of using the standard MSM in emergency situations. We were interested in a teaching method to help beginners achieve an acceptable level of skill for eyeballing EF. In this study, as a first step, we determined the number of cases with feedback that would be required for beginners to reach an acceptable rating level. We are confident that this study

**Table 3**  
Effect of feedback and self-learning program.

In test (after 4 wks)	Weighted Kappa (P value)
Novice Practitioner 1	0.8272 (<.0001)
Novice Practitioner 2	0.9207 (<.0001)
Novice Practitioner 3	0.8102 (<.0001)
Novice Practitioner 5	0.8544 (<.0001)
Novice Practitioner 6	0.8821 (<.0001)

provides useful information to educators in the field of emergency echocardiography.

We obtained several interesting results through the self-learning course. On average, approximately 50 cases with feedback were required for novice practitioners to reach an acceptable level of skill for eyeballing EF in this study. Individual differences were noted in acquiring the skill among the 7 novice practitioners. We cannot explain why practitioner 1 required 40 cases and practitioner 6 required 71 cases to reach an acceptable performance level. The reason is likely related to differences in learning ability, attitude, and other characteristics among the learners. Practitioner 1 showed excellent ability and attitude for learning in other fields, such as intubation skill acquisition (unpublished data). Although individual differences were noted in learning the skill of eyeballing EF, the skills for all practitioners improved as the learning course progressed. The error range of eyeballing EF compared with the standard value in each case was wide in the early stage, and this range narrowed in the process of the learning course. Eventually, the error was maintained within an acceptable range, which indicates that our teaching method for eyeballing EF introduced in this study might be effective for acquiring the skill. Although an average of approximately 50 cases was needed to reach an acceptable level of skill, a minimum of 80 cases with feedback should be provided in the self-learning with feedback course, considering the variation in learning among the practitioners. We believe 100 cases would be sufficient for beginners.

We confirmed that the skills for eyeballing EF were highly maintained without any learning or feedback after 4 weeks, which indicates that the skill acquired from repetitive learning in this self-learning course could be effectively maintained for a certain time period (4 weeks in this study). The weighted kappa values of all practitioners were greater than 0.8 (Table 3), which represents almost perfect (very good) agreement between the judgement values by eyeballing EF and standard values obtained using the MSM. The weighted kappa is calculated using a predefined table of weights that measures the degree of agreement between the 2 raters, and greater agreement results in a higher weight.<sup>[7]</sup> The weighted kappa index value is interpreted as follows: 0.01 to 0.2 indicates poor agreement, 0.21 to 0.4 indicates fair agreement, 0.41 to 0.6 indicates moderate agreement, 0.61 to 0.8 indicates good agreement, and 0.81 to 1.0 indicates very good agreement.

Various studies related to the learning effects of focused echocardiographic training for novice practitioners have been conducted. Vignon et al<sup>[8]</sup> reported that effective focused echocardiography training was conducted in an ICU for novice residents. Townsend et al<sup>[9]</sup> provided a Focused Assessment Diagnostic Echocardiography (FADE) course to surgical residents and reported an accuracy of 45% in primary education and improved diagnostic ability through FADE, showing 88% accuracy at the fourth session. These findings indicate that novice practitioners could effectively acquire a certain level of skill via short-term, focused echocardiographic training. To our knowledge, this study is the first to analyze the learning curve for skill acquisition of novices in eyeballing EF using CUSUM analysis. Most previous studies using CUSUM analysis dichotomized the decision parameter, for example, whether trainees correctly diagnosed a condition (acceptable measurement, success or failure); however, we classified the decision parameter into 4 categories (normal, mildly reduced, moderately reduced, and severely reduced).<sup>[5,10–15]</sup> We believe that this detailed

categorization could more accurately reflect the practitioner's skill level than dichotomous classification.

There are several limitations of this study. First, the sample size was very small (7 participants), and 2 of the 7 participants left the study after 4 weeks. Second, this self-learning system alone does not allow beginners to acquire an expert level of eyeballing EF. We classified the decision parameter into 4 categories with an interval of 10%, and the novice practitioners could effectively distinguish a 10% difference after self-learning with feedback, which indicates that they could distinguish normal (EF > 50%), mildly reduced (40–49%), moderately reduced (30–39%), and severely reduced LV function (< 30%) relatively well. However, the practitioners might not be able to distinguish the more detailed differences, such as differentiating moderate to severely reduced LV function (30–34%) from moderately reduced LV function (35–39%). If the decision parameter is subdivided in detail into 7 categories with an interval of 5% (normal: > 55%, borderline: 50–55%, mildly reduced: 45–49%, mild to moderately reduced: 40–44%, moderately reduced: 35–39%, moderate to severely reduced: 30–34%, severely reduced: < 30%), more cases would be needed to achieve an acceptable skill level. However, we suggest that the detection of a difference of approximately 10% in wall motion might be sufficient to evaluate LV function in an emergency situation. Finally, this study did not evaluate the skill of echocardiographic image acquisition. We investigated only a restricted competence in visual assessment of LV function by watching echocardiographic video files that had been acquired in advance. This study determined the number of cases necessary to correctly interpret the EF via eyeballing estimation. However, image acquisition is another important skill involved in echocardiography. Further studies focusing on image acquisition (how to use an ultrasound device, including manipulation of the cardiac probe and how to find an optimal view for the echocardiographic image) are needed.

## 5. Conclusion

We found that novice practitioners could acquire an acceptable skill level for eyeballing EF via short-term, self-learning-focused echocardiographic training.

## Author contributions

YL and CK designed this study. CK supervised the overall data collection process, had full access to all the data in the study, and takes responsibility for the integrity of the data. YL and HS conducted the data analysis. YL wrote the initial draft of the article. All authors provided substantial review and feedback on the final version of the article. YL and HS take responsibility for the paper as a whole. YL and HS contributed equally in this study. All authors have read and approved the submitted manuscript. This manuscript has not been submitted nor published elsewhere in whole or in part.

**Conceptualization:** Changsun Kim.

**Data curation:** Yoonje Lee.

**Formal analysis:** Yoonje Lee, Hyungoo Shin.

**Investigation:** Yoonje Lee, Hyungoo Shin, Inhye Lee, Hyuk Joong Choi.

**Methodology:** Hyungoo Shin, Changsun Kim.

**Supervision:** Changsun Kim.

**Writing – original draft:** Yoonje Lee.

**Writing – review & editing:** Yoonje Lee, Hyungoo Shin, Changsun Kim, Inhye Lee, Hyuk Joong Choi. Changsun Kim orcid: 0000-0001-8475-3453.

## References

- [1] Akinboboye O, Sumner J, Gopal A, et al. Visual estimation of ejection fraction by two-dimensional echocardiography: the learning curve. *Clin Cardiol* 1995;18:726–9.
- [2] Bhatnagar SK, Moussa MA, Al-Yusuf AR. The role of prehospital discharge two-dimensional echocardiography in determining the prognosis of survivors of first myocardial infarction. *Am Heart J* 1985;109:472–7.
- [3] Cohn PF, Gorlin R, Cohn LH, et al. Left ventricular ejection fraction as a prognostic guide in surgical treatment of coronary and valvular heart disease. *Am J Cardiol* 1974;34:136–41.
- [4] Hope MD, de la Pena E, Yang PC, et al. A visual approach for the accurate determination of echocardiographic left ventricular ejection fraction by medical students. *J Am Soc Echocardiogr* 2003;16:824–31.
- [5] Je S, Cho Y, Choi HJ, et al. An application of the learning curve-cumulative summation test to evaluate training for endotracheal intubation in emergency medicine. *Emerg Med J* 2015;32:291–4.
- [6] Kim C, Hur J, Kang BS, et al. Can an offsite expert remotely evaluate the visual estimation of ejection fraction via a social network video call? *J Digit Imaging* 2017;30:718–25.
- [7] Warrens MJ. Cohen's linearly weighted kappa is a weighted average. *Adv Data Anal Classif* 2012;6:67–79.
- [8] Vignon P, Dugard A, Abraham J, et al. Focused training for goal-oriented hand-held echocardiography performed by noncardiologist residents in the intensive care unit. *Intensive Care Med* 2007;33:1795–9.
- [9] Townsend NT, Kendall J, Barnett C, et al. An effective curriculum for focused assessment diagnostic echocardiography: establishing the learning curve in surgical residents. *J Surg Educ* 2016;73:190–6.
- [10] Bazot M, Daraï E, Biau DJ, et al. Learning curve of transvaginal ultrasound for the diagnosis of endometriomas assessed by the cumulative summation test (LC-CUSUM). *Fertil Steril* 2011;95:301–3.
- [11] Rodriguez A, Guillén JJ, López MJ, et al. Learning curves in 3-dimensional sonographic follicle monitoring during controlled ovarian stimulation. *J Ultrasound Med* 2014;33:649–55.
- [12] Weerasinghe S, Mirghani H, Revel A, et al. Cumulative sum (CUSUM) analysis in the assessment of trainee competence in fetal biometry measurement. *Ultrasound Obstet Gynecol* 2006;28:199–203.
- [13] Balsyte D, Schäffer L, Burkhardt T, et al. Continuous independent quality control for fetal ultrasound biometry provided by the cumulative summation technique. *Ultrasound Obstet Gynecol* 2010;35:449–455.
- [14] El Hachem L, Momeni M, Friedman K, et al. Safety, feasibility and learning curve of robotic single-site surgery in gynecology. *Int J Med Robot* 2016;12:509–16.
- [15] Alcázar JL, Díaz L, Flórez P, et al. Intensive training program for ultrasound diagnosis of adnexal masses: protocol and preliminary results. *Ultrasound Obstet Gynecol* 2013;42:218–23.