RESEARCH

Impact of image quality on reliability of the measurements of left ventricular systolic function and global longitudinal strain in 2D echocardiography

Yasufumi Nagata MD[1], Yuichiro Kado MD[1], Takeshi Onoue MD[1], Kyoko Otani MD[2], Akemi Nakazono RDCS[2], Yutaka Otsuji MD[1] and Masaaki Takeuchi MD[2]

[1]Second Department of Internal Medicine, University of Occupational and Environmental Health, School of Medicine, Kitakyushu, Japan
[2]Department of Laboratory and Transfusion Medicine, University of Occupational and Environmental Health, School of Medicine, Kitakyushu, Japan

Correspondence should be addressed to Y Nagata: nyasufumi1979@yahoo.co.jp

Abstract

Background: Left ventricular ejection fraction (LVEF) and global longitudinal strain (GLS) play important roles in diagnosis and management of cardiac diseases. However, the issue of the accuracy and reliability of LVEF and GLS remains to be solved. Image quality is one of the most important factors affecting measurement variability. The aim of this study was to investigate whether improved image quality could reduce observer variability.
Methods: Two sets of three apical images were acquired using relatively old- and new-generation ultrasound imaging systems (Vivid 7 and Vivid E95) in 308 subjects. Image quality was assessed by endocardial border delineation index (EBDI) using a 3-point scoring system. Three observers measured the LVEF and GLS, and these values and inter-observer variability were investigated.
Results: Image quality was significantly better with Vivid E95 (EBDI: 26.8±5.9) than that with Vivid 7 (22.8±6.3, $P<0.0001$). Regarding the inter-observer variability of LVEF, the *r*-value, bias, 95% limit of agreement and intra-class correlation coefficient for Vivid 7 were comparable to those for Vivid E95. The % variabilities were significantly lower for Vivid E95 (5.3–6.5%) than those for Vivid 7 (6.5–7.5%). Regarding GLS, all observer variability parameters were better for Vivid E95 than for Vivid 7. Improvements in image quality yielded benefits to both LVEF and GLS measurement reliability. Multivariate analysis showed that image quality was indeed an important factor of observer variability in the measurement of LVEF and GLS.
Conclusions: The new-generation ultrasound imaging system offers improved image quality and reduces inter-observer variability in the measurement of LVEF and GLS.

Key Words

- 2D transthoracic echocardiography
- reliability
- image quality
- endocardial border delineation

Introduction

Two-dimensional (2D) echocardiography is a versatile modality for assessing cardiac function in daily clinical practice because of its portability, usability and high cost performance (1). Indications for specific treatment of congestive heart failure or valvular heart disease (2, 3) and the determination to withdraw cancer therapy

www.echorespract.com
https://doi.org/10.1530/ERP-17-0047

(4, 5) are usually determined by the specific cut-off values of left ventricular ejection fraction (LVEF) assessed by 2D echocardiography. Therefore, it is crucial that the LVEF measurement should be accurate and have high reliability. LV global longitudinal strain (GLS), as assessed by 2D speckle-tracking echocardiography (STE), is another robust parameter for detecting latent LV dysfunction and predicting future prognosis (4, 5, 6). Some previous studies have revealed that GLS is superior to LVEF in terms of reliability and the detectability of subtle functional abnormalities (7, 8, 9). Considering reliability and accuracy in the 2D echocardiography assessment of LV function, intra- and inter-observer variabilities are major concerns (5, 10, 11). As expected, inadequate visualization of the LV endocardial border is associated with high observer variabilities (10, 12). It has been reported that approximately 10% to 15% of routine echocardiograms have poor image quality (10). Several attempts have been made to improve image quality (13, 14) and to reduce observer variability (10, 15).

Ultrasound companies are continuously working to improve their echocardiographic image quality. Thus, the image quality obtained by the latest advanced ultrasound system might be better than that obtained using a previous generation ultrasound system, even from the same ultrasound vendor, which provides a chance to examine the effect of image quality on measurement reliability.

Accordingly, the aims of this study were as follows: (1) to compare image quality between the latest generation and relatively older-generation ultrasound machines from the same company; (2) to examine whether examiner with different levels of the experience could obtain the same amount of image quality improvement between two examinations and (3) to assess whether image quality would affect the interpretation of regional wall motion and the reliability of LVEF and GLS measurements.

Methods

Study subjects

We prospectively enrolled 324 subjects undergoing transthoracic echocardiography in one echocardiography room in the echocardiographic laboratory. The exclusion criteria were age <20 years, frequent premature ventricular contraction and extremely poor image quality, such that at least one of three observers could not measure the LVEF or GLS with confidence. The Ethics Committee of the University of Occupational and Environmental Health approved the study protocol, and informed consent was obtained from all study subjects.

Study protocol and two-dimensional echocardiography

The study protocol is summarized in Fig. 1. 2D harmonic echocardiography images were obtained using two different ultrasound imaging machines (Vivid 7 with the M4S probe and Vivid E95 with the M5Sc probe, GE Vingmed Ultrasound AS, Horten Norway) in the same room. Routine comprehensive echocardiographic examinations were performed using the Vivid 7. Then, three LV apical images (apical four-, two- and three-chamber views) were acquired using the Vivid E95. During the acquisition of images with the Vivid E95, the examiner checked the images just previously acquired using the Vivid 7 and tried to obtain the same 2D cut-plane images. All grayscale images during two or three consecutive cardiac cycles were digitally stored on a hard disk. The overall gain and compression of both machines were adjusted for each patient to minimize the dropout of the LV endocardial borders. The depth and the sector angle were adjusted to include the left ventricle with a higher frame rate. The focus level was set at the middle point between the apex and the mitral annulus on the Vivid 7. Focus setting was not required for the Vivid E95 due to its multifocus capability.

Examiners and observers

The 13 examiners (7 sonographers and 6 physicians) who acquired the echocardiographic images were classified into 3 groups according to their level of experience with echocardiography. The beginner group had 1–5 years of

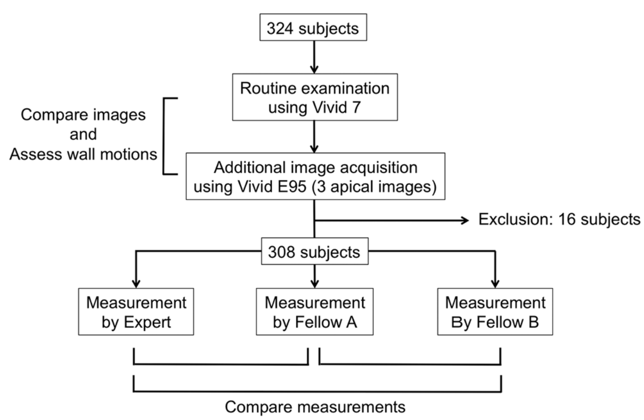


Figure 1
Flow chart of the study protocol.

experience, the senior examiner group had 6–15 years of experience, and the expert group had more than 15 years of experience.

Measurements of LV volume, LVEF and GLS were independently performed by three observers, including one expert with more than 10 years of experience ,and two fellows with 6 years (Fellow A) and 2 years of experience in 2D speckle-tracking analysis (Fellow B), to evaluate observer variability. All examiners were blinded to each other's results and the patient's data.

Image interpretation

Image quality was assessed by one observer (Fellow A) who was blinded to the clinical information. The LV was assessed using the 18-segment model (1). Image quality was evaluated in two ways according to the number of visible segments and the degree of endocardial border delineation. Image quality was classified into 3 grades (good: 0–2 segments were poorly visible, fair: 3–5 segments were poorly visible and poor: >5 segments were poorly visible) in the LV 18-segment model (16). Endocardial border delineation was classified using a 3-point scoring system (14, 17) (0 = not visible, 1 = fairly visible, and 2 = clearly visible during the entire cardiac cycle) in each set of 18 segments. The endocardial border delineation index (EBDI) was defined as the summation of scores in all segments (with a maximum score of 36 and a minimum score of 0).

Image analysis

LV parameters were measured using vendor-dependent software (EchoPAC PC version 201, GE Vingmed Ultrasound AS, Horten Norway). For the determination of LV volumes and LVEF, we used the specific software (Auto EF, GE Vingmed Ultrasound). After the manual tracing of the LV endocardial border in the apical 4- and 2-chamber views at end-systole, the software automatically performed speckle-tracking analysis on the LV endocardial border during one cardiac cycle to generate LV volumes and LVEF.

GLS was measured from three apical views with commercially available speckle-tracking software (2D strain, GE Vingmed Ultrasound). Manual tracing of the LV endocardial border at end-systole was performed. Subsequently, the software generated the region of interest (ROI), which was manually adjusted to encompass the entire thickness of the myocardium. The software performed a speckle-tracking analysis on the LV myocardium on a frame-by-frame basis during one cardiac cycle. Finally, the software automatically created time–domain strain curves for each view and bull's eye maps illustrating the GLS, which was defined as the average longitudinal strain at end-systole measured from all segments. The adequacy of the tracking was verified visually, and if the tracking was deemed to be suboptimal, the ROI was manually re-adjusted. If the subsequent tracking was still not satisfactory, the subjects were excluded from the analysis. Although longitudinal strain is generally expressed as a negative value, we used the absolute values of GLS in this manuscript to avoid confusion.

Visual estimations of regional wall motion in each image acquired by the Vivid 7 and the Vivid E95 were separately performed by one observer (Fellow A) at least one week apart to evaluate the differences in interpretation between the two imaging systems. Wall motion was assessed as normal, abnormal or uninterpretable for each segment.

Furthermore, patients were classified into three groups according to the severity of LV systolic dysfunction (normal: LVEF ≥50%, mild LV dysfunction: 30% ≤LVEF <50%, severe LV dysfunction: LVEF <30%) (10), and concordant and discordant rates between the two imaging systems were evaluated. The LVEF values determined by the expert were used for the analysis.

Reproducibility

Measurement reproducibility was determined by comparisons between two of the three observers. Inter-observer variability was determined to be the percentage of variability, which was defined as the absolute differences between the two measurements divided by the values (%variability), the 95% limit of agreements (LOAs), the correlation coefficient and intra-class correlation coefficients (ICC).

Statistical analysis

Continuous data are expressed as the mean ± S.D. or as the median and interquartile range according the data distribution. Normality was evaluated using the Shapiro–Wilk test. Categorical data are presented as a number or percentage. A *t*-test or Wilcoxon rank-sum test was used to evaluate the differences in the continuous variables between the two groups according to data distributions. The Friedman test was performed to compare results among the three observers, with *post hoc* analysis to compare the results between each pair (Expert vs Fellow A, Expert vs Fellow B,

https://doi.org/10.1530/ERP-17-0047

and Fellow A vs Fellow B). The *r* value was analyzed as a Pearson's product–moment correlation coefficient or as a Spearman's rank correlation coefficient. A Bland–Altman analysis was performed to determine the bias and the 95% limit of agreement (LOA) between the two measurements. Multiple regression analysis was performed to test the effects of anthropometric parameters and image quality on observer variability after univariate analysis. A two-sided P value <0.05 was considered significant. All statistical analyses were performed using commercial software (JMP, version11, SAS Institute, Cary, NC, USA).

Results

Study subjects and examiners

Sixteen subjects were excluded from the analysis due to frequent premature contraction ($n=2$), because the software did not provide GLS values ($n=2$), or due to extremely poor image quality that at least one of the three observes judged to be inappropriate for analysis ($n=12$). Thus, the final study cohort consisted of 308 subjects. The clinical characteristics of the study population are summarized in Table 1. Among the 13 examiners that participated in this study, 5 were classified into the beginner group, 4 in the senior group and 4 in the expert group. The number of echo examinations performed by the beginner, senior and expert groups was 128, 96, and 84, respectively.

Table 1 Clinical characteristics in study subjects ($n=308$).

Variables	Mean ± s.d., *n* (%)
Anthropometric parameters	
Age (years old)	65 ± 14
Men (*n* (%))	150 (49%)
Body surface area (m²)	1.60 ± 0.21
Body mass index (kg/m²)	23.1 ± 4.4
Heart rate (bpm)	70 ± 13
Systolic blood pressure (mmHg)	139 ± 22
Diastolic blood pressure (mmHg)	79 ± 13
Medical diagnosis	
Ischemic heart disease	65 (20%)
Valvular heart disease	13 (4%)
Cardiomyopathy	8 (3%)
Arrhythmia	27 (9%)
Malignancy	44 (14%)
Collagen disease	32 (10%)
Pre-operative assessment	47 (15%)
Others	72 (24%)

s.d., standard deviation.

Image interpretation

Among the 308 subjects, image quality was determined to be poor in 101 (33%), fair in 137 (44%) and good in 70 (23%) with the use of the Vivid 7; the corresponding values for the Vivid E95 were 46 (15%), 114 (37%) and 148 (48%), respectively (Fig. 2A). The image quality of the Vivid E95 was significantly better than that of the Vivid 7. The EBDI was also significantly higher for the Vivid E95 than for the Vivid 7 (Fig. 2B). A comparison of EBDI between the Vivid E95 and the Vivid 7 in each

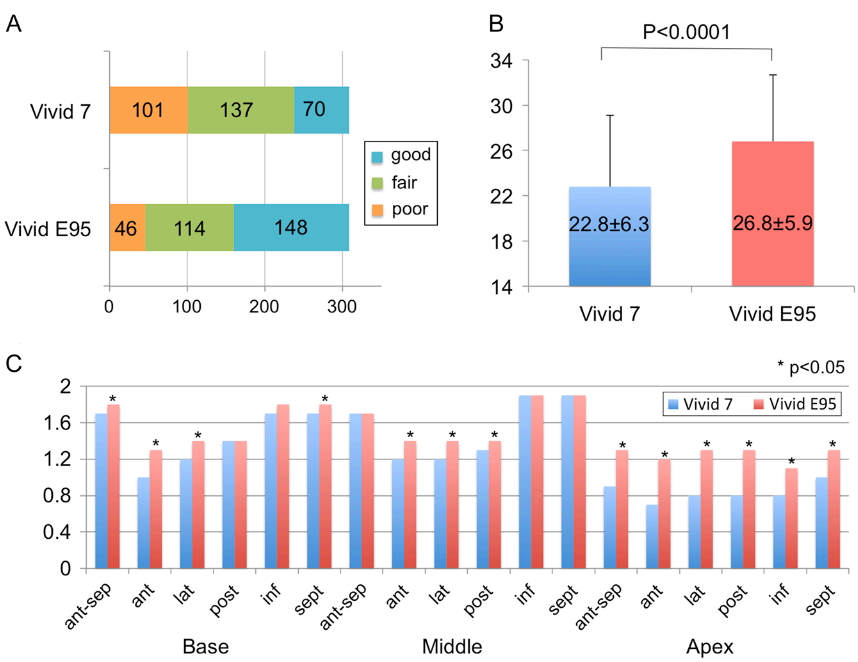


Figure 2
Comparison of segmental image quality and endocardial border delineation. Segmental image quality is shown in the upper left panel (A). The totals of the endocardial border delineation index in whole segments (B) and each segmental endocardial border delineation index (C) are shown in the upper right panel and the bottom panel, respectively.

https://doi.org/10.1530/ERP-17-0047

LV segment is shown in Fig. 2C. The EBDI of the Vivid E95 was significantly higher in the free wall and the apex than that of the Vivid 7. A representative case is shown in Fig. 3.

Figure 4 depicts the effect of experience on image quality improvement between the Vivid 7 and the Vivid E95. Irrespective of the degree of scanning experience, the three groups showed nearly the same trend of image improvement from the Vivid 7 to the Vivid E95.

Image analysis: inter-system variability and inter-observer variability

The mean frame rates of the images acquired using the Vivid 7 and the Vivid E95 were 69 ± 7 and 61 ± 2 frames/s, respectively. A summary of the LV parameters analyzed by the three observers using the images captured by two different ultrasound machines is presented in Table 2. The LV end-diastolic volume measured with the use of the Vivid E95 was slightly but significantly larger for two observers than that measured with the use of the Vivid 7. In contrast, the LV end-systolic volume of the Vivid E95 was significantly smaller in all three observers, resulting in the determination of a larger LVEF using the Vivid E95 for all observers. Although good correlations of LV volumes between the two ultrasound systems (r=0.93–0.94) were noted, the correlation of LVEF was modest (r=0.78–0.80). The GLS determined by 2 observers were slightly but significantly larger using the Vivid E95 than using the Vivid 7. There were good correlations of GLS between the two systems for all observers (r=0.87–0.89). Regarding inter-observer differences, all parameters showed significant differences among the three observers.

The results of regional wall motion assessment and LV systolic function grade are shown in Fig. 5 and Supplementary Fig. 1 and Supplementary Table 1 (see section on supplementary data given at the end of this article). Among the 5544 segments in the regional wall motion analysis, the percentage of uninterpretable segments decreased from 18% (980 segments) for the Vivid 7–7% (393 segments) in for the Vivid E95. In 67 (1%) segments, for which wall motion assessment was possible with the Vivid 7, the assessments were uninterpretable using the Vivid E95; 654 (12%) segments were

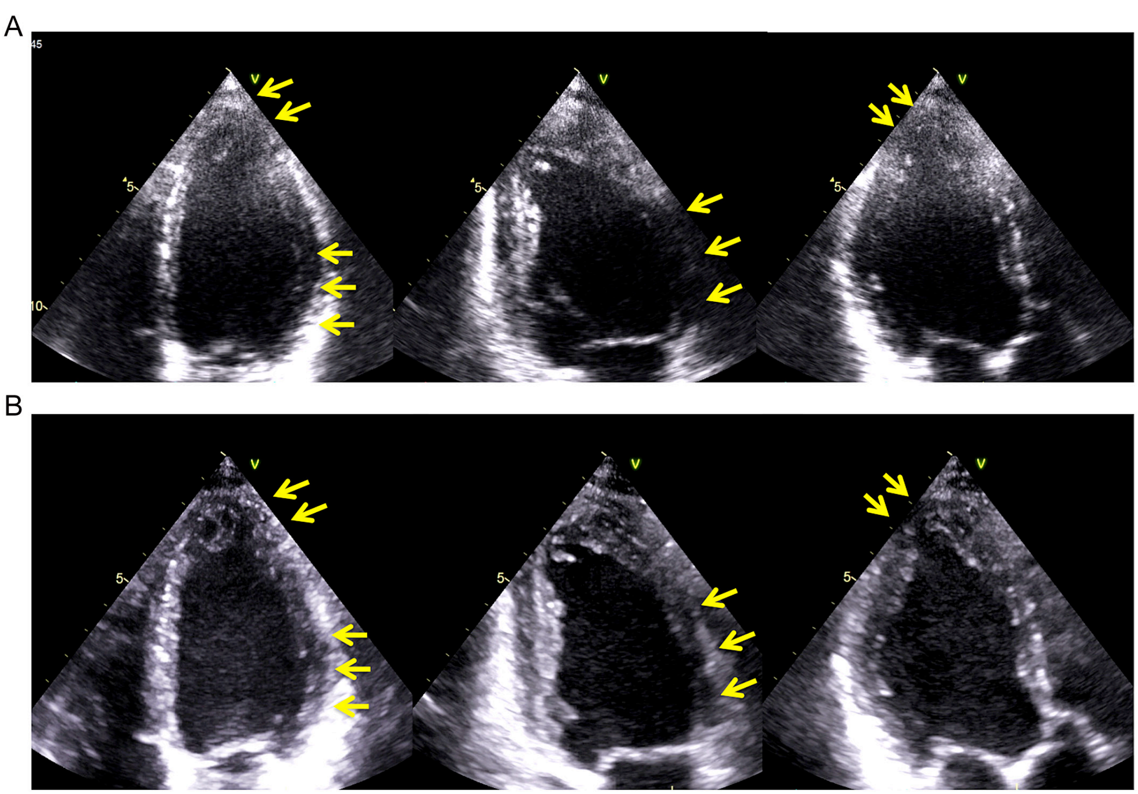


Figure 3
Representative images acquired using the Vivid 7 (A) and the Vivid E95 (B). Improvement in visualization of the endocardial border delineation from the Vivid 7 to the Vivid E95 at segments of the free wall and apex are shown (allows). The endocardial border delineation index was improved from 27 using the Vivid 7–34 using the Vivid E95.

https://doi.org/10.1530/ERP-17-0047

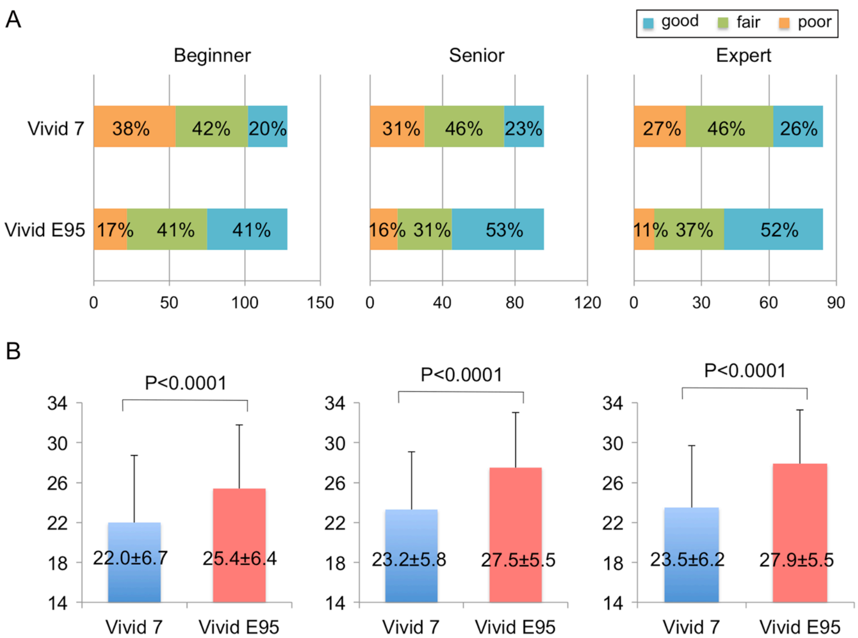


Figure 4
Segmental image quality and endocardial border delineation indexes in three groups classified according to examiners' levels of experience. The upper panel shows the segmental image quality (A). The bottom panel shows the endocardial border delineation index (B).

uninterpretable using with the Vivid 7 but interpretable using the Vivid E95. With the use of the Vivid E95, the percentage of interpretable segments was remarkably increased (442/1848 segments, 24%), especially at the LV apical level. Overall, concordant results were obtained in 4401 segments (79%) between the Vivid 7 and the Vivid E95. In contrast, discordant results were observed in 96 segments (2%). A total of 72 segments were identified as new regional abnormal wall motion only with the Vivid E95. The Vivid E95 also had a significant impact on the assessment of LVEF (Fig. 5B). Discordant LV systolic function grades were observed in 84 (27%) subjects. Compared with the results from the Vivid 7, the Vivid E95 provided better LV function grades in 62 out of 84 subjects.

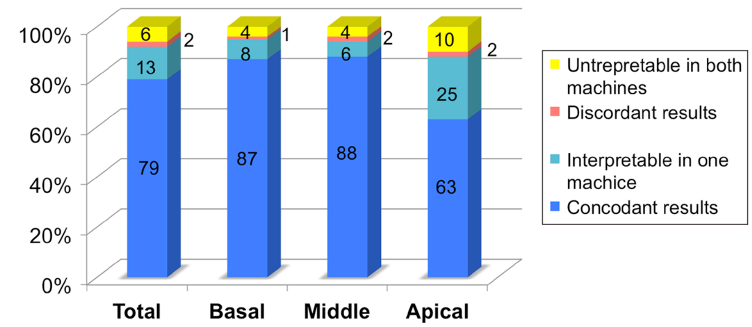
The inter-observer variability of LVEF and GLS according to the different ultrasound systems is shown in Table 3. Regarding observer variability of LVEF, the values

Table 2 Inter-system differences of left ventricular parameters.

	Vivid 7	Vivid E95	Bias	*P* value	*r*
Left ventricular end-diastolic volume (mL)					
Expert	96.5±31.5‡	98.2±32.6‡	1.6	0.0171	0.93
Fellow A	98.3±31.7‡	98.6±32.2‡	0.4	0.5322	0.94
Fellow B	101.3±31.3*,†	102.7±33.3*,†	1.5	0.0367	0.93
ANOVA	<0.0001	<0.0001			
Left ventricular end-systolic volume (mL)					
Expert	50.4±21.4‡	49.1±21.0‡	−1.4	0.0018	0.94
Fellow A	51.7±21.7‡	49.0±20.9‡	−2.6	<0.0001	0.94
Fellow B	54.4±22.0*,†	52.6±23.0*,†	−1.9	<0.0001	0.94
ANOVA	<0.0001	<0.0001			
Left ventricular ejection fraction (%)					
Expert	48.4±8.0‡	50.8±7.6‡	2.4	<0.0001	0.78
Fellow A	48.1±8.1‡	51.0±7.6‡	2.9	<0.0001	0.80
Fellow B	46.9±8.3*,†	49.7±7.8*,†	2.9	<0.0001	0.78
ANOVA	<0.0001	<0.0001			
Global longitudinal strain (%)					
Expert	15.9±3.5†,‡	15.8±3.3†,‡	−0.1	0.1969	0.89
Fellow A	15.1±3.2*	15.4±3.1*	0.4	<0.0001	0.89
Fellow B	15.0±3.4*	15.3±3.2*	0.3	0.0059	0.87
ANOVA	<0.0001	<0.0001			

**P*<0.05 vs expert; †*P*<0.05 vs Fellow A; ‡*P*<0.05 vs Fellow B.
ANOVA, analysis of variance.

A



B

Vivid 7	Vivid E95			
		Normal	Mild LV dysfunction	Severe LV dysfunction
	Normal	5 (2%)	5 (2%)	0 (0%)
	Mild LV dysfunction	1 (0%)	91 (30%)	57 (19%)
	Severe LV dysfunction	0 (0%)	21 (7%)	128 (42%)

Figure 5
Impacts of different generation imaging systems on the interpretation of wall motion abnormality (A) and LV systolic function grade (B). The numbers in the upper graphs represent percentages.

of bias, 95% LOA and ICC with use of the Vivid 7 and the Vivid E95 did not change considerably for two of the three observers. However, the values for %variability between the expert and the two fellows were significantly reduced (Expert vs Fellow A: from 6.5% with the Vivid 7 to 5.3% with the Vivid E95, $P=0.0019$; Expert vs Fellow B: from 7.4% to 6.5%, $P=0.0174$) (Fig. 6A). Although %variability from the Vivid E95 to the Vivid 7 between the two fellows' comparisons tended to be reduced, the difference was not significant (From 7.1% to 6.3%, $P=0.0727$).

Regarding observer variability of GLS, bias and 95% LOA were reduced, and ICC increased from the Vivid 7 to the Vivid E95 for two of the three observers. The changes of %variability of GLS between the Vivid 7 and the Vivid E95 are shown in Fig. 6B. The %variability of GLS was significantly reduced when comparing the Vivid 7 to the Vivid E95 in any pair of comparison.

The relationships between % variability and anthropometric parameters and endocardial border delineation were assessed using multivariate regression analysis. Univariate analysis showed weak but significant correlations between %variability and EBDI in all pairs for both LVEF and GLS (Table 4). Multivariate analysis still showed that EBDI had the largest effect on %variability in almost all analyses except for one pair (Fellow A and Fellow B) in LVEF analysis (Table 5).

Discussion

The main findings of this study can be summarized as follows: (1) the latest-generation of ultrasound imaging system provided better image quality than that of a relatively older-generation system, irrespective to the

Table 3 Inter-observer differences of LVEF and GLS assessed with Vivid 7 and Vivid E95.

	Vivid 7						**Vivid E95**					
	LVEF						LVEF					
	r	Bias	*P* value	LOA	% variability	ICC	*r*	Bias	*P* value	LOA	% variability	ICC
Expert vs Fellow A	0.89	−0.3	.1867	±7.4	6.5%	0.888	0.89	0.2	.3118	±6.9	5.3%	0.893
Expert vs Fellow B	0.87	−1.5	<.0001	±8.0	7.4%	0.859	0.87	−1.1	<.0001	±7.4	6.5%	0.859
Fellow A vs Fellow B	0.88	−1.2	<.0001	±7.8	7.1%	0.871	0.86	−1.3	<.0001	±8.0	6.3%	0.848
	GLS						GLS					
	r	Bias	*P* value	LOA	% variability	ICC	*r*	Bias	*P* value	LOA	% variability	ICC
Expert vs Fellow A	0.93	−0.8	<.0001	±2.5	8.1%	0.900	0.95	−0.3	<.0001	±2.0	5.5%	0.945
Expert vs Fellow B	0.91	−0.9	<.0001	±2.9	8.8%	0.882	0.93	−0.5	<.0001	±2.5	6.7%	0.916
Fellow A vs Fellow B	0.92	−0.1	.5426	±2.6	7.2%	0.900	0.94	−0.1	.0234	±2.2	5.9%	0.936

GLS, global longitudinal strain; ICC, intra-class correlation coefficient; LOA, limits of agreement; LVEF, left ventricular ejection fraction.

www.echorespract.com
https://doi.org/10.1530/ERP-17-0047

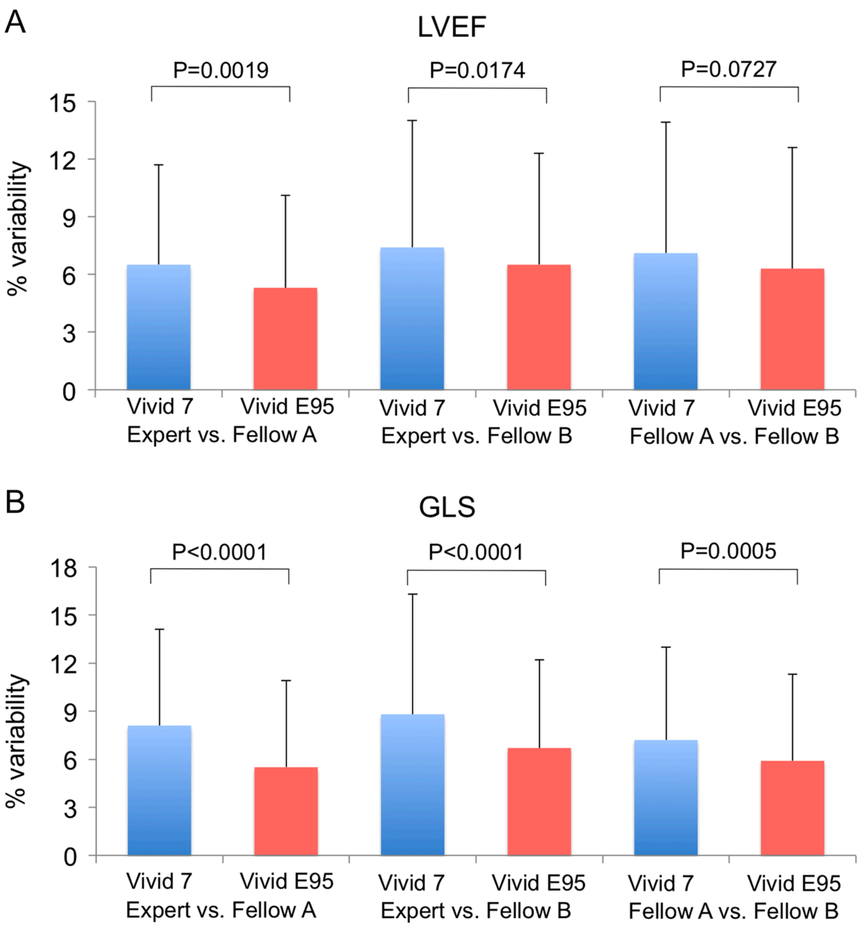


Figure 6
The %variability of LVEF (A) and GLS (B) between two of the three different experienced observers using the Vivid 7 and the Vivid E95.

acquisition skills of the examiner; (2) the change in ultrasound imaging systems had significant impacts on the assessment of LV systolic function grade and the diagnosis of regional wall motion abnormalities and (3) improved image quality was associated with a reduction of inter-observer variability in the measurement of LVEF and GLS.

LVEF is not only a decision-making parameter but also a robust predictor of cardiac outcomes (1, 2, 3, 5). However, LVEF is not sufficiently sensitive to detect subtle abnormalities at an early stage of LV dysfunction (7, 8). GLS measurements derived by STE have been reported to be superior to LVEF in their capacity to detect subtle changes in LV dysfunction and to predict adverse outcomes in some clinical situations (5, 7, 8, 18). However, the concern regarding observer variability cannot be ignored. Several factors contribute to observer variability in echocardiography, including level of experience, differences in tracing of the endocardial border and quality of the acquired images (19). Among these factors, the quality of the acquired images depends on the skill (i.e., experience) of the examiners and the performance of the ultrasound imaging system itself. New-generation ultrasound imaging systems produce better image quality due to their new imaging technology. Although the effect of improved image quality with the use of a contrast agent or tissue harmonic imaging on measurements has been reported (12, 17, 20), the effect of improving image quality on observer variability has not been fully investigated. This study is unique in that it focuses on these points.

Comparison of image quality between new- and old-generation ultrasound imaging systems

The results of our study showed that the image quality was obviously improved in a newer-generation ultrasound system. Every examiner benefited from the image quality improvements with use of the new-generation ultrasound imaging system. In particular, the prevalence of poor image quality was remarkably reduced in the expert group. These results may be due to the improved partial and contrast resolution in the new technology, which renders segments that were invisible by the old-generation system visible by the new-generation system. The EBDI, based

Table 4 Relationship between %variability and physical parameters and endocardial border delineation index.

	Left ventricular ejection fraction: observer variability (% variability)					
	Vivid 7			Vivid E95		
	Expert vs Fellow A	Expert vs Fellow B	Fellow A vs Fellow B	Expert vs Fellow A	Expert vs Fellow B	Fellow A vs Fellow B
Age	$r<0.01$, $P=0.0987$	$r=0.08$, $P=0.1811$	$r=-0.05$, $P=0.3872$	$r=0.10$, $P=0.0738$	$r=-0.01$, $P=0.8414$	$r=0.04$, $P=0.4654$
Sex	$r=0.03$, $P=0.6332$	$r=0.03$, $P=0.5621$	$r=0.06$, $P=0.2997$	$r=0.02$, $P=0.7891$	$r=0.04$, $P=0.4666$	$r=0.02$, $P=0.6911$
Body mass index	$r=0.06$, $P=0.2838$	$r=0.17$, $P=0.0029$	$r=0.14$, $P=0.0124$	$r=0.13$, $P=0.0215$	$r=0.16$, $P=0.0039$	$r=0.15$, $P=0.0072$
Heart rate	$r=0.07$, $P=0.2495$	$r=0.07$, $P=0.2181$	$r=0.09$, $P=0.1240$	$r=0.02$, $P=07268$	$r=-0.02$, $P=0.7422$	r<0.01, $P=0.9687$
EBDI Vivid 7	$r=-0.16$, $P=0.0042$	$r=-0.21$, $P=0.0002$	$r=-0.14$, $P=0.0128$			
EBDI Vivid E95				$r=-0.20$, $P=0.0004$	$r=-0.23$, $P<0.0001$	$r=-0.11$, $P=0.0422$

	Global longitudinal strain: observer variability (% variability)					
	Vivid 7			Vivid E95		
	Expert vs Fellow A	Expert vs Fellow B	Fellow A vs Fellow B	Expert vs Fellow A	Expert vs Fellow B	Fellow A vs Fellow B
Age	$r=0.12$, $P=0.0358$	$r=0.06$, $P=0.2579$	$r=-0.04$, $P=0.4802$	$r=0.01$, $P=0.8591$	$r=0.07$, $P=0.2162$	$r=0.17$, $P=0.0022$
Sex	$r=0.08$, $P=0.1527$	$r=0.02$, $P=0.7301$	$r=0.05$, $P=0.4129$	$r=0.09$, $P=0.1265$	$r=0.03$, $P=0.5734$	$r=0.05$, $P=0.4249$
Body mass index	$r=-0.05$, $P=0.4254$	$r=0.10$, $P=0.0784$	$r=0.11$, $P=0.0458$	$r=-0.07$, $P=0.2119$	$r=0.03$, $P=0.5547$	$r=-0.05$, $P=0.4001$
Heart rate	$r=0.06$, $P=0.3058$	$r=0.06$, $P=0.2707$	$r=0.07$, $P=0.2228$	$r=0.09$, $P=0.1304$	$r=0.05$ $P=0.4253$	$r=0.04$, $P=0.4967$
EBDI Vivid 7	$r=-0.22$, $P<0.0001$	$r=-0.17$, $P=0.0034$	$r=-0.18$, $P=0.0020$			
EBDI Vivid E95				$r=-0.17$, $P=0.0026$	$r=-0.15$, $P=0.0083$	$r=-0.14$, $P=0.0120$

EDBI, endocardial border delineation index.

on the visibility of the endocardial border delineation, was significantly increased in the new-generation system. Interestingly, the EBDIs of the free wall and apical segments showed marked increases. This finding is clinically important, particularly in patients with ischemic heart disease, including acute coronary syndrome (1), because the visualization of these segments has been a weak point of echocardiography in the detection of abnormal wall motion and aneurysm (14, 17).

Although the new-generation ultrasound imaging systems indeed yielded improvements in image quality in a large number of subjects, approximately 4% of subjects had still poor image quality that precluded appropriate analysis in this study. These findings suggest that there may be a certain number of subjects who still have poor image quality, even when the latest version of ultrasound machine is used. The application of ultrasound contrast agents could resolve some of these problems (10, 20). The guideline by the American Society of Echocardiography and the European Association of Cardiovascular imaging also recommend the use of contrast agents in patients with poor image quality (1).

Although every examiner was benefited from the new-generation ultrasound imaging system (Fig. 4), the percentage of patients who were categorized as having good image quality and the differences in EBDI scores between the two systems were larger in the expert group than in the beginner group. These findings imply that daily training and skill improvement are still important to obtain good-quality images.

Comparison of LV parameters and wall motion between new- and old-generation ultrasound imaging systems

In this study, small but significant differences in LV volume measurements were observed between the latest version of the ultrasound machine and the older version. This finding is inconsistent with previous publications using contrast agent that demonstrated that LV volumes determined from contrast-enhanced images were larger

https://doi.org/10.1530/ERP-17-0047

Table 5 Multivariate regression analysis of association with % variability and endocardial border delineation index.

	Left ventricular ejection fraction: observer variability (% variability)					
	Standardized β coefficient, *P* value					
	Vivid 7			Vivid E95		
	Expert vs Fellow A	Expert vs Fellow B	Fellow A vs Fellow B	Expert vs Fellow A	Expert vs Fellow B	Fellow A vs Fellow B
Age	0.05, $P=0.3872$	0.06, $P=0.3039$		0.09, $P=0.1297$		
Sex						
Body mass index		0.12, $P=0.0480$	0.11, $P=0.0696$	0.08, $P=0.1553$	0.10, $P=0.0791$	0.13, $P=0.0292$
Heart rate						
EBDI Vivid 7	−0.20, $P=0.0003$	−0.17, $P=0.0046$	−0.11, $P=0.0722$			
EBDI Vivid E95				−0.17, $P=0.0062$	−0.20, $P=0.0006$	−0.08, $P=0.2053$
	Global longitudinal strain: observer variability (% variability)					
	Standardized β coefficient, *P* value					
	Vivid 7			Vivid E95		
	Expert vs Fellow A	Expert vs Fellow B	Fellow A vs Fellow B	Expert vs Fellow A	Expert vs Fellow B	Fellow A vs Fellow B
Age	0.10, $P=0.0815$					0.16, $P=0.0049$
Sex						
Body mass index		0.03, $P=0.5783$	0.06, $P=0.2751$			
Heart rate						
EBDI Vivid 7	−0.16, $P=0.0066$	−0.21, $P=0.0003$	−0.15, $P=0.0097$			
EBDI Vivid E95				$r=-0.17$, $P=0.0026$	$r=-0.15$, $P=0.0083$	−0.12, $P=0.0294$

EDBI, endocardial border delineation index.

than those determined from unenhanced images (10, 20). A potential cause of this discrepancy may be the difference in the way the endocardial border was determined. In this study, the new-generation ultrasound system allowed for visualization of the cardiac structure itself, while contrast agents enable an observer to detect the border between the compacted myocardium and the LV cavity (1, 20). The LVEF values measured from the images acquired using the new-generation system were significantly higher than those obtained using the old system for all three observers. As a result, approximately 20% of patients in this study were reclassified into the better LV systolic function grades using the new-generation system. These results are consistent with previous publications on image quality (10, 12, 20). Poor image quality could cause the underestimation of LVEF. Furthermore, in this study, LVEF was semi-automatically measured using the speckle-tracking algorithm; therefore, the tracking quality may directly affect endocardial border visualization. Regarding GLS, the measurement differences between the two systems were small compared with those for LVEF. One possible explanation for this discrepancy is the size of the ROI. The ROI of GLS was transmural, while the ROI of LVEF included only the endocardium. Therefore, LVEF can be more easily influenced by image quality, particularly that of the endocardium, than the GLS.

Fewer segments were uninterpretable when we used the new-generation ultrasound system. Interestingly, the observed beneficial effect of the new system over the older system was more obvious at the LV apex, where wall motion assessment was particularly important, especially in patients with ischemic heart disease (17). A small number of segments (1%) showed the opposite findings. The main reason for this theoretically paradoxical finding might be related to the examiners' expertise or the subjects' posture. Subtle changes in the transducer location and/or the patient's position could sometimes have a large impact on image quality.

There were also some segments showing discordant results in wall motion interpretation between the older system and the new system. Since we did not have a reference standard for wall motion analysis, we could not determine which assessment was accurate. However, the judgment of wall motion in images acquired by the new system could be more reliable compared with those of the older system because its endocardial border delineation visualization capability was superior to that of the older system.

Effect of image quality on observer variability

In this study, inter-observer variability was determined by comparing two of three observers with different levels

www.echorespract.com
https://doi.org/10.1530/ERP-17-0047

of expertise. All parameters regarding inter-observer variability in all pairs were unchanged or improved with the new-generation ultrasound imaging system. In particularly, %variability was significantly reduced except for one paired comparison (Fellow A vs Fellow B) of LVEF. Thus, the new-generation system had desirable effects on the reliability of LVEF and GLS measurements. The reduction of %variability in GLS was larger than that in LVEF in this study. The reason for these differences can be explained by the size of the ROI. The LVEF measurements depend on an endocardial border, whereas speckle-tracking analysis for GLS was performed not on the endocardium but on the entire myocardium. Moreover, the amount of information contained in the images may affect tracking quality. The images acquired using the new-generation system most likely had more information, which could be more difficult to visually recognize than those acquired using the old-generation system. Thus, although the improvement of image quality obtained by the new-generation ultrasound imaging system offered preferable effects on both LVEF and GLS measurement reliability, the benefit was greater for GLS than for LVEF.

Multivariate analysis to assess the influence of anthropometric parameters and image quality on observer variability showed that image quality (EBDI) was the most significant independent factor of %variability in the majority of cases (Table 4). These results indicate that image quality indeed has an effect on measurement reliability, in addition to fundamental observer variability itself (15). Notably, EBDI was not a significant independent factor of %variability in LVEF between the two fellows. The probable reason for this finding might be that the measurement reliability and accuracy of both fellows were fundamentally suboptimal due to their lower levels experience compared with the expert. The relatively small value of the standardized β coefficient in the comparative analysis between the two fellows might support this indication.

Expected impact of improved image quality on the reproducibility of 3D echocardiographic measurements

Theoretically, three-dimensional (3D) echocardiography is more accurate and reliable for the assessment of LV volumes, LVEF and GLS than 2D echocardiography, because 3D images encompass the entire left ventricle and LV chamber, and functional assessment with 3D echocardiography does not require any geometric assumptions (1, 21, 22). However, the lower temporal and spatial resolutions of current 3D echocardiography compared with those of 2D echocardiography are associated with some degradation of image quality, and this drawback may nullify the advantage (1, 19, 23). Therefore, improvement in 3D image quality with the advancement of imaging technology is quite important for the further improvement of measurement reproducibility. Further studies are required to investigate the impact of image quality improvement on measurement reproducibility using 3D echocardiography.

Study limitations

There are several limitations to be addressed in this study. First, there was no reference standard to validate which measurements were more accurate. We used the measurement values determined by the expert as a reference, which was close to the practical situation of the echocardiographic laboratory. Second, observers were not blinded to which systems were used for the analysis. Third, this study demonstrated the improvement of image quality and reliability in one vendor: these findings may not apply to the other vendors. Fourth, the order of image acquisition with the older system and the new system was not randomized due to the reimbursement of the cost. It is not permitted to use a trial instrument for medical services under health insurance in Japan. Finally, this study was observational. Therefore, further studies will be necessary to determine whether the improvements in image quality and measurement reliability using an updated imaging system would favorably affect the diagnosis and the decision of treatment strategy in patients with cardiac disease.

Conclusions

The new-generation ultrasound imaging system produced obviously improved image quality compared with that produced by the relatively old-generation system. Furthermore, the new-generation ultrasound system provides enhanced image quality for all examiners, who had varying levels of experience. This finding could be associated with beneficial impacts on the assessment of regional wall motion and LV systolic function grading and reduced inter-observer variability. The improvement of image quality yielded benefits to both LVEF and GLS measurement reliability, with the benefit being more obvious for GLS than for LVEF.

www.echorespract.com
https://doi.org/10.1530/ERP-17-0047

Supplementary data
This is linked to the online version of the paper at https://doi.org/10.1530/ERP-17-0047.

Declaration of interest
The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding
This work did not receive any specific grant from any funding agency in the public, commercial, or not-for-profit sector.

References

1 Lang RM, Badano LP, Mor-Avi V, Afilalo J, Armstrong A, Ernande L, Flachskampf FA, Foster E, Goldstein SA, Kuznetsova T, *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. *European Heart Journal Cardiovascular Imaging* 2015 **16** 233–270. (https://doi.org/10.1093/ehjci/jev014)

2 Moss AJ, Zareba W, Hall WJ, Klein H, Wilber DJ, Cannom DS, Daubert JP, Higgins SL, Brown MW & Andrews ML. Prophylactic implantation of defibrillator in patients with myocardial infarction and reduced ejection fraction. *New England Journal of Medicine* 2002 **346** 877–883. (https://doi.org/10.1056/NEJMoa013474)

3 Nishimura RA, Otto CM, Bonow RO, Carabello BA, Erwin JP 3rd, Guyton RA, O'Gara PT, Ruiz CE, Skubas NJ, Sorajja P, *et al*. 2014 AHA/ACC guideline for the management of patients with valvular heart disease: a report of the American College of Cardiology/American Heart Association task force on practice guidelines. *Circulation* 2014 **129** e521–e643. (https://doi.org/10.1161/CIR.0000000000000031)

4 Negishi K, Negishi T, Haluska BA, Hare JL, Plana JC & Marwick TH. Use of speckle strain to assess left ventricular responses to cardiotoxic chemotherapy and cardioprotection. *European Heart Journal Cardiovascular Imaging* 2014 **15** 324–331. (https://doi.org/10.1093/ehjci/jet159)

5 Plana JC, Galderisi M, Barac A, Ewer MS, Ky B, Scherrer-Crosbie M, Ganame J, Sebag IA, Agler DA, Badano LP, *et al*. Expert consensus for multimodality imaging evaluation of adult patients during and after cancer therapy: a report from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *European Heart Journal Cardiovascular Imaging* 2014 **15** 1063–1093. (https://doi.org/10.1093/ehjci/jeu192)

6 Stanton T, Leano R & Marwick TH. Prediction of all-cause mortality from global longitudinal speckle strain: comparison with ejection fraction and wall motion scoring. *Circulation Cardiovascular Imaging* 2009 **2** 356–364. (https://doi.org/10.1161/CIRCIMAGING.109.862334)

7 Ng A, Delgado V, Bertini M, van der Meer RW, Rijzewijk LJ, Shanks M, Nucifora G, Smit JW, Diamant M, Romijn JA, *et al*. Findings from left ventricular strain and strain rate imaging in asymptomatic patients with type 2 diabetes mellitus. *American Journal of Cardiology* 2009 **104** 1398–1401. (https://doi.org/10.1016/j.amjcard.2009.06.063)

8 Nakai H, Takeuchi M, Nishikage T, Lang RM & Otsuji Y. Subclinical left ventricular dysfunction in asymptomatic diabetic patients assessed by two-dimensional speckle tracking echocardiography: correlation with diabetic duration. *European Journal of Echocardiography* 2009 **10** 926–932. (https://doi.org/10.1093/ejechocard/jep097)

9 Barbier P, Mirea O, Cefalu C, Maltagliati A, Savioli G & Guglielmo M. Reliability and feasibility of longitudinal AFI global and segmental strain compared with 2D left ventricular volumes and ejection fraction: intra- and inter-operator, test-retest, and inter-cycle reproducibility. *European Heart Journal Cardiovascular Imaging* 2015 **16** 642–652. (https://doi.org/10.1093/ehjci/jeu274)

10 Kurt M, Shaikh KA, Peterson L, Kurrelmeyer KM, Shah G, Nagueh SF, Fromm R, Quinones MA & Zoghbi WA. Impact of contrast echocardiography on evaluation of ventricular function and clinical management in a large prospective cohort. *Journal of the American College of Cardiology* 2009 **53** 802–810. (https://doi.org/10.1016/j.jacc.2009.01.005)

11 Jacobs LD, Salgo IS, Goonewardena S, Weinert L, Coon P, Bardo D, Gerard O, Allain P, Zamorano JL, de Isla LP, *et al*. Rapid online quantification of left ventricular volume from real-time three-dimensional echocardiographic data. *European Heart Journal* 2006 **27** 460–468. (https://doi.org/10.1093/eurheartj/ehi666)

12 Lafitte S, Santos PD, Kerouani A, Robhan T & Roudaut R. Improved reliability for echocardiographic measurement of left ventricular volume using harmonic power imaging mode combined with contrast agent. *American Journal of Cardiology* 2000 **85** 1234–1238. (https://doi.org/10.1016/S0002-9149(00)00734-7)

13 Spencer KT, Bednarz J, Mor-Avi V, Weinert L, Tan J, Godoy I & Lang RM. The role of echocardiographic harmonic imaging and contrast enhancement for improvement of endocardial border delineation. *Journal of the American Society of Echocardiography* 2000 **13** 131–138. (https://doi.org/10.1016/S0894-7317(00)90024-5)

14 Kasprzak JD, Paelinck B, Cate FJT, Vletter WB, Jong Nd, Poldermans D, Elhendy A, Bouakaz A & Roelandt JRTC. Comparison of native and contrast-enhanced harmonic echocardiography for visualization of left ventricular endocardial border. *American Journal of Cardiology* 1999 **83** 211–217. (https://doi.org/10.1016/S0002-9149(98)00826-1)

15 Malm S, Frigstad S, Sagberg E, Larsson H & Skjaerpe T. Accurate and reproducible measurement of left ventricular volume and ejection fraction by contrast echocardiography: a comparison with magnetic resonance imaging. *Journal of the American College of Cardiology* 2004 **44** 1030–1035. (https://doi.org/10.1016/j.jacc.2004.05.068)

16 Takigiku K, Takeuchi M, Izumi C, Yuda S, Sakata K, Ohte N, Tanabe K & Nakatani S. Normal range of left ventricular 2-dimensional strain. *Circulation Journal* 2012 **76** 2623–2632. (https://doi.org/10.1253/circj.CJ-12-0264)

17 Yoshitani H, Takeuchi M, Hirose M, Miyazaki C, Otani S, Sakamoto K & Yoshikawa J. Head-to-head comparison of fundamental, tissue harmonic and contrast harmonic imaging with or without an air-filled contrast agent, levovist, for endocardial border delineation in patients with poor quality image. *Circulation Journal* 2002 **66** 494–498. (https://doi.org/10.1253/circj.66.494)

18 Nagata Y, Takeuchi M, Wu VC, Izumo M, Suzuki K, Sato K, Seo Y, Akashi YJ, Aonuma K & Otsuji Y. Prognostic value of LV deformation parameters using 2D and 3D speckle-tracking echocardiography in asymptomatic patients with severe aortic stenosis and preserved LV ejection fraction. *JACC Cardiovascular Imaging* 2015 **8** 235–245. (https://doi.org/10.1016/j.jcmg.2014.12.009)

19 Mor-Avi V, Jenkins C, Kuhl HP, Nesser HJ, Marwick T, Franke A, Ebner C, Freed BH, Steringer-Mascherbauer R, Pollard H, *et al*. Real-time 3-dimensional echocardiographic quantification of left ventricular volumes: multicenter study for validation with magnetic resonance imaging and investigation of sources of error. *JACC Cardiovascular Imaging* 2008 **1** 413–423. (https://doi.org/10.1016/j.jcmg.2008.02.009)

20 Larsson MK, Da Silva C, Gunyeli E, Ilami AA, Szummer K, Winter R & Bjallmark A. The potential clinical value of contrast-

enhanced echocardiography beyond current recommendations. *Cardiovascular Ultrasound* 2016 **14** 2. (https://doi.org/10.1186/s12947-015-0045-0)

21 Kleijn SA, Aly MF, Terwee CB, van Rossum AC & Kamp O. Reliability of left ventricular volumes and function measurements using three-dimensional speckle tracking echocardiography. *European Heart Journal Cardiovascular Imaging* 2012 **13** 159–168. (https://doi.org/10.1093/ejechocard/jer174)

22 Santoro C, Arpino G, Esposito R, Lembo M, Paciolla I, Cardalesi C, de Simone G, Trimarco B, De Placido S & Galderisi M. 2D and 3D strain for detection of subclinical anthracycline cardiotoxicity in breast cancer patients: a balance with feasibility. *European Heart Journal Cardiovascular Imaging* 2017 **18** 930–936. (https://doi.org/10.1093/ehjci/jex033)

23 Obokata M, Nagata Y, Wu VC, Kado Y, Kurabayashi M, Otsuji Y & Takeuchi M. Direct comparison of cardiac magnetic resonance feature tracking and 2D/3D echocardiography speckle tracking for evaluation of global left ventricular strain. *European Heart Journal Cardiovascular Imaging* 2016 **17** 525–532. (https://doi.org/10.1093/ehjci/jev227)

Received in final form 26 November 2017
Accepted 15 January 2018