Cardiac Imaging

Performance of 3-Dimensional Echocardiography in Measuring Left Ventricular Volumes and Ejection Fraction

A Systematic Review and Meta-Analysis

Jennifer L. Dorosz, MD,* Dennis C. Lezotte, PHD,† David A. Weitzenkamp, PHD,†
Larry A. Allen, MD, MHS,* Ernesto E. Salcedo, MD*

Aurora, Colorado

Objectives	The primary aim of this systematic review is to objectively evaluate the test performance characteristics of three-dimensional echocardiography (3DE) in measuring left ventricular (LV) volumes and ejection fraction (EF).
Background	Despite its growing use in clinical laboratories, the accuracy of 3DE has not been studied on a large scale. It is unclear if this technology offers an advantage over traditional two-dimensional (2D) methods.
Methods	We searched for studies that compared LV volumes and EF measured by 3DE and cardiac magnetic resonance (CMR) imaging. A subset of those also compared standard 2D methods with CMR. We used meta-analyses to determine the overall bias and limits of agreement of LV end-diastolic volume (EDV), end-systolic volume (ESV), and EF measured by 3DE and 2D echocardiography (2DE).
Results	Twenty-three studies (1,638 echocardiograms) were included. The pooled biases ± 2 SDs for 3DE were −19.1 ± 34.2 ml, −10.1 ± 29.7 ml, and −0.6 ± 11.8% for EDV, ESV, and EF, respectively. Nine studies also included data from 2DE, where the pooled biases were −48.2 ± 55.9 ml, −27.7 ± 45.7 ml, and 0.1 ± 13.9% for EDV, ESV, and EF, respectively. In this subset, the difference in bias between 3DE and 2D volumes was statistically significant (p = 0.01 for both EDV and ESV). The difference in variance was statistically significant (p < 0.001) for all 3 measurements.
Conclusions	Three-dimensional echocardiography underestimates volumes and has wide limits of agreement, but compared with traditional 2D methods in these carefully performed studies, 3DE is more accurate for volumes and more precise in all 3 measurements. (J Am Coll Cardiol 2012;59:1799–808) © 2012 by the American College of Cardiology Foundation

The assessment of left ventricular (LV) volume and ejection fraction (EF) is vital to the practice of cardiology. These measures are used to inform prognosis in most cardiac patient populations, determine treatment decisions for a variety of therapies, and function as eligibility criteria in many clinical trials (1–5). Despite their importance, there is no consensus on the best method to routinely measure EF and volumes.

Two-dimensional echocardiography (2DE) is the most ubiquitous tool for assessing LV size and systolic function. Measurement of EF is the most common reason for referring a patient for an echocardiogram (6). This test is noninvasive, portable, inexpensive, radiation free, and quick. It does not, however, provide reliable, reproducible, and accurate measures of EF or volumes (7). Traditionally, 2DE has used the method of disks to calculate LV volumes based on areas in only 2 imaging planes (6). This method is subject to errors due to foreshortening, poor endocardial definition, narrow echocardiographic windows, and assumptions about LV shape.

Because it is able to image the entire heart in multiple planes and provides excellent endocardial definition, cardiac magnetic resonance (CMR) imaging has been used as the gold standard for measuring LV volumes and EF (8). Nevertheless, the expense, limited availability, and incompatibility with metallic hardware make CMR impractical for widespread clinical use.

Three-dimensional echocardiography (3DE) uses recently developed matrix array echocardiographic probes to image the entire heart in ≤8 beats/min. As opposed to older 3D methods, "live" 3D datasets provide volumes with minimal post-processing. Some echocardiographic labora-

**Abbreviations
and Acronyms**

2D = two-dimensional

2DE = two-dimensional
echocardiography

3D = three-dimensional

3DE = three-dimensional
echocardiography

BA = Bland-Altman

CMR = cardiac magnetic
resonance

EDV = end-diastolic volume

EF = ejection fraction

ESV = end-systolic volume

LV = left ventricular

tories have embraced this technology for routine care and charge for it using a new billable Current Procedural Terminology code for 3DE (9). Yet, it is still unclear if 3DE offers an advantage over 2DE, and, if so, in which patients and with which specific techniques.

In this era when new technology, especially in imaging, has escalated medical costs, advances like 3DE should be thoroughly evaluated before recommending their large-scale use (10). We, therefore, undertook this systematic review of the 3DE published reports to evaluate its performance compared with CMR and its utility over traditional 2D methods in calculating EF and volumes.

Methods

Search selection. To determine the test performance characteristics of 3DE in measuring LV volumes and EF, we included all studies that compared 3DE with CMR in adult patients. Most used Pearson's correlations between CMR and 3DE-derived LV end-diastolic volume (EDV), end-systolic volume (ESV), and EF. As correlation coefficients alone are misleading, we required that acceptable studies contain a Bland-Altman (BA) analysis of agreement for at least 1 of the 3 measurements (11). Although not required for inclusion, many studies also compared 3DE with traditional 2D methods and provided BA analysis on inter- and intraobserver variability. We excluded studies that only investigated freehand (rather than live or "real-time") 3DE, did not include CMR as a gold standard, only included children, only evaluated the right (rather than the left) ventricle, only reported on LV mass (rather than volumes or EF), or were only published in abstract form. In addition, we excluded early studies done with probes or software that are not currently commercially available.

Search strategy. We searched 7 electronic information sources for studies published between January 1, 1990 and September 14, 2011: MEDLINE, Cardiosource Clinical Trials, the Cochrane Central Register of Controlled Trials, the Cochrane Health Technology Assessment Database, the International Standard Randomized Controlled Trial Number Register, National Institutes of Health Clinical-Trials.gov, and UpToDate Online. We used variations and combinations of the following search terms: three-dimensional echocardiography, real-time three-dimensional echocardiography, heart ventricles, and cardiac magnetic resonance imaging. We limited the search to humans, adults, and English language. In addition to database searches, we reviewed the references of included studies and

other relevant review articles to obtain a comprehensive list of included studies (Table 1). The database searches were performed twice and reviewed by 2 authors (J.L.D., E.E.S.).
Data abstraction. The quantitative data collected were the differences in EDV, ESV, and EF calculations determined by 3DE and CMR and expressed as the BA bias and limits of agreement. Where reported, we also recorded the BA differences for 2DE with CMR as well as intra- and interobserver analyses. To assess the applicability of 3DE in a general population of patients referred for echocardiography, we were particularly interested in the types of patients included in each study. Thus, we collected the distributions of disease types, sex, average ages, and the average sizes of the ventricles studied. To assess the general feasibility of 3DE, we noted in each study the method of patient selection and number of patients excluded due to poor image quality. We also logged the type of platform and methods used to obtain volumes.
Statistical analysis. Baseline patient characteristics for each study were re-weighted according to the number of patients in the sample and used to compute a pooled estimate of average age and EDV (expressed as mean ± SD). Our analysis included summary estimates of BA statistics to assess agreement of the 2 competing procedures. We used BA analytical methods because they are more informative than Pearson's correlation coefficients in comparing methods where the main goal is to assess accuracy rather than to determine if 2 competing procedures have associated outcomes (11). BA methods only provide estimates in lieu of statistical tests and level of significance like more traditional clinical trials. Like the BA methodology, our synthesis strategy included weighted summary statistics. Thus, for pooled estimates of 3DE and 2DE performance compared with CMR, we simply computed appropriate summary statistics (weighted sums and sums of squares) across individual studies, which were combined to produce pooled estimates of BA biases and limits of agreement for each of the 3 measurements. We found that the reported BA limits of agreement were inconsistently computed across the studies (which may have used 1, 1.96, or 2 SDs of the reported bias). For the purposes of our meta-analysis, we converted each study's reported estimates and expressed our limits of agreement as bias ±2 SDs. For intraobserver and interobserver variability, we included studies that calculated the mean differences between the observations divided by their average and reported them as mean difference with or without limits of agreement. Again, these average differences were pooled to compute an overall measure of the observer difference and limits of agreement (also expressed as 2 SDs). Significant differences between 3DE and 2DE biases were tested using the paired *t* test. To determine significant differences in the variances, we used Levene's test for the homogeneity of variances (35). A p value <0.05 was considered statistically significant. SAS version 9.2 (SAS Institute, Cary, North Carolina) was used for all calculations and statistical testing.

Table 1	Included Studies and Patient Characteristics						
First Author (Ref. #)	N	Age (yrs)	Mean EDV by CMR (ml)	MRI to Echo Time	Patient Selection and Number Excluded Due to Poor Images		Diseases Studied
Jenkins et al. (12)	50	63 ± 10	207 ± 79	N/A	Consecutive	0	All with WMAs
Soliman et al. (13)	24	51 ± 12	213 ± 63	2 h	"Good to excellent images"	N/A	29% normal, 42% ICM, 29% NICM
Pouleur et al. (14)	83	54 ± 19	187 ± 70	1 day	Consecutive	0	24% normal, 40% CAD, 36% valve disease
Mor-Avi et al. (15)	92	57 ± 16	N/A	1 day	Excluded cardiac surgery patients	0	N/A
Chukwu et al. (16)	69	65 ± 5.5	171 ± 47	24 h	Excluded if >2 unseen segments	1	51% normal, 49% ICM
Bicudo et al. (17)	20	32 ± 15	135 ± 28	N/A	Consecutive	0	100% HCM all with normal EF
Qi et al. (18)	58	59	139 ± 59	1 day	Consecutive	0	24% normal, 40% CAD, 12% NICM, 5% valve disease, 5% ASD/VSD
Jenkins et al. (19)	30	66 ± 7	168 ± 54	1 week	Excluded for poor images or acquisitions	7	All with WMAs
van den Bosch et al. (20)	29	31 ± 9	155 ± 38	2 h	Excluded for poor images and respiratory artifact	3	Congenital heart disease with septal flattening
Sugeng et al. (21)	31	60 ± 15	N/A	1 day	"Adequate endocardial visualization"	N/A	29% normal, 45% CAD, 23% NICM, 1% HCM
Jenkins et al. (22)	110	63 ± 11	180 ± 55	N/A	Excluded due to inability to measure 3D	7	89% WMA, 11% with global dysfunction
Jacobs et al. (23)	50	58 ± 19	N/A	N/A	"Adequate endocardial visualization"	8	48% CAD, 36% NICM, 8% valve disease, 4% coarctation, 4% right atrial mass
Chan et al. (24)	30	62 ± 9	179 ± 56	12 days	Consecutive	0	All with prior MI, altered LV shape, or WMA
Gutiérrez-Chico et al. (25)	35	61 ± 17	N/A	1 day	Consecutive	0	23% ICM, 26% HCM, 17% valve disease, 11% myocarditis, 17% NICM, 6% congenital
Jenkins et al. (26)	50	64 ± 8	172 ± 53	N/A	Excluded if "technically inadequate"	7	14% normal, 82% with WMA, 4% HTN
Nikitin et al. (27)	64	65 ± 12	195 ± 72	N/A	"Good acoustic windows"	16	16% normal, 63% ICM or NICM, 22% HT /HFN EF
Soliman et al. (28)	41	56 ± 11	182 ± 54	1 day	Excluded if >2 unseen segments	N/A	15% ICM, 22% NICM, 22% noncompaction, 41% HCM
Soliman et al. (29)	53	57 ± 11	175 ± 51	1 day	Excluded if >2 unseen segments	N/A	17% ICM, 24% NICM, 21% noncompaction, 38% HCM
Kuhl et al. (30)	24	65 ± 12	N/A	1 day	Excluded if >2 unseen segments	N/A	42% normal, 2% NICM, 42% ICM
Caiani et al. (31)	20	58 ± 17	164 ± 64	1 day	Included only poor images	N/A	5% normal, 35% CAD, 45% NICM, 10% HCM, 5% PAH
Marsan et al. (32)	52	62 ± 10	288 ± 87	1 day	Excluded for >2 unseen segments	5	ICM with aneurysm referred for surgery
Macron et al. (33)	50	59 ± 18	161 ± 59	1 day	Excluded for poor images	16	42% CAD, 10% valve disease, 6% NICM, 6% HCM, 36% other
Chang et al. (34)	109	53 ± 14	137	1 day	Excluded for poor endocardial definition	18	38% normal, 32% CAD, 19% HCM, 7% NICM, 6% other

3D = three-dimensional; ASD = atrial septal defect; CAD = coronary artery disease; CMR = cardiac magnetic resonance; HCM = hypertrophic cardiomyopathy; HTN = hypertension; ICM = ischemic cardiomyopathy; LV = left ventricular; N/A = not available; NICM = nonischemic cardiomyopathy; PAH = pulmonary arterial hypertension; VSD = ventricular septal defect; WMA = wall motion abnormality.

Results

Search results and patient population. Electronic and manual searches from reference lists produced 189 citations, of which 88 were excluded as irrelevant based on title and keywords. The remaining 101 were reviewed in detail by reading the abstract or paper in full. Of these, 78 were excluded for a priori defined criteria (Fig. 1), leaving 23 included articles, comprising 1,174 patients (Table 1) (12–34). Nine of these reported results on separate patient groups or did the 3DE analysis twice using different methods, resulting in 34 distinct protocols with 1,638 separate 3DE analyses. Each of these substudies was treated as a separate data source in this meta-analysis (Table 2). These patients had a wide spectrum of

cardiac abnormalities. Across studies, the following disease states were represented: 10% normal controls, 45% coronary artery disease, 9% dilated cardiomyopathy, 4% systolic dysfunction (not otherwise characterized), 4% valve disease, 8% hypertrophic cardiomyopathy, 3% congenital heart disease, 1.4% hypertension, and 16% other or not reported. The majority (72%) were men. The weighted age was 58.0 ± 13.7 years, and EDV by CMR was 184.9 ± 61.8 ml.

Excluded patients. All studies excluded patients who were not in sinus rhythm at the time of their echocardiograms. Due to incompatibility with CMR, patients with internal pacemakers or defibrillators or severe claustrophobia were also excluded. Most studies assessed echocardiographic
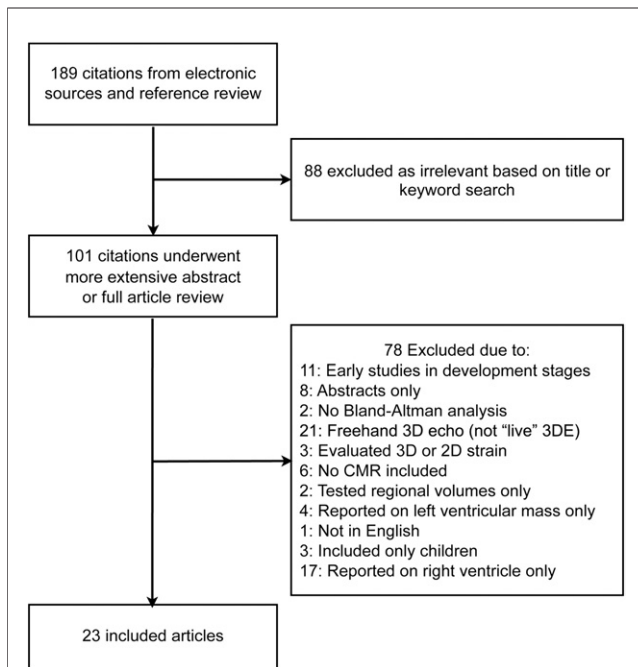
Figure 1   **Published Report Selection**

This flow sheet describes how the 23 included studies were selected for the meta-analysis. 2D= two-dimensional; 3D = three-dimensional; CMR = cardiac magnetic resonance.

image quality and excluded patients with suboptimal images (Table 1). Among studies that reported the number of exclusions, a total of 88 patients (7.5%) were excluded; 6 studies (representing 123 patients) did not report on the number of excluded patients. One study used contrast imaging and included only those with poor images (31). Five studies recruited a variety of patients consecutively without exclusions for image quality (12,14,15,18,25). Of these, Jenkins et al. (12) reported that 20% of patients had technically difficult images; Pouleur et al. (14) noted that 16%, 35%, 32%, and 17% had excellent, good, moderate, and fair images, respectively.

3DE methods used in the selected studies. All studies except 2 used a Philips Medical Systems 2 to 4 MHz 3D matrix array transducer (Andover, Massachusetts). One study used the General Electric 2.5 MHz transducer (Horton, Norway) (33), and one used the Siemens Medical Systems Acuson SC2000 transducer (Mountain View, California) (34). For analysis of the 3D datasets, there were 2 general methods used to calculate volumes (Fig. 2). The first was manual tracing of equally spaced individual long- or short-axis slices at end-systole and end-diastole. To correspond with CMR standards, papillary muscles and trabeculations were treated as part of the LV cavity. The number of slices was chosen by the operator and ranged from 2 to 12; however, according to studies that specifically assessed the quality of data based on number of planes, a minimum of 8 planes were needed to achieve the best results (16,25). The

second method required the user to identify 3 to 5 points at the apex and mitral annuls in the 2- and 4-chamber end-diastolic and end-systolic views. From these points, the software used automated border-detection to create a 3D endocardial shell of the entire ventricle from which volumes were calculated. There were 4 principal software platforms (QLab, Philips Medical Systems, Andover, Massachusetts; TomTec Imaging systems, Munich, Germany; Echopak, General Electric Vingmed Ultrasound, Horton, Norway; and Argus, Siemens Medical Solutions, Mountain View, California). Our selection of studies used a variety of slices versus mesh techniques on all 4 platforms (Table 2). Four studies were done to specifically compare the slices versus mesh techniques (16,22,29,30). Only 2 studies used contrast for LV opacification (12,31). As reported by Jenkins et al. (22), the average time to trace 12 slices was 10.5 ± 1 min, whereas 4 ± 0.3 minutes were required to identify the points for the mesh technique. This compared with 1.5 ± 0.45 min to trace 2D images for the method of discs.

3DE and CMR agreement results. Table 2 and Figure 3 show the BA bias and limits of agreement for EDV, ESV, and EF from the 34 studies. The overall pooled bias $\pm$ 2 SDs were -19.1 ± 34.2 ml, -10.1 ± 29.7 ml, and $-0.6 \pm 11.8\%$ for EDV, ESV, and EF, respectively. To better approximate a typical population seen in an echocardiographic laboratory, we performed a subanalysis that included the 6 substudies (368 echocardiograms) that did not exclude patients for image quality. In this case, the BA biases ± 2 SDs were -28.8 ± 38 ml, -17.9 ± 34 ml, and $0.3 \pm 15.5\%$ for EDV, ESV, and EF, respectively. As the quicker mesh-based technique is also more likely to be used routinely in echocardiographic laboratories, we also performed a subanalysis with the 4 studies (512 echocardiograms) that compared the mesh versus slices methods (16,19,29,30). The pooled BA biases ± 2 SDs for the mesh technique compared with CMR were -22.7 ± 29 ml, -12.1 ± 25 ml, and $-0.4 \pm 9.1\%$; for slices, they were -12.6 ± 25 ml, -8.1 ± 21 ml, and $1.0 \pm 9.0\%$ for EDV, ESV, and EF, respectively.

3DE and 2DE agreement results. Nine articles (14 substudies) reported comparisons of both 3DE and 2DE with CMR (12,16,17,19,22,23,25,26,32). Table 3 and Figure 4 show the BA estimates for these studies. For 2DE, the pooled BA biases $\pm$ 2 SDs were -48.2 ± 55.9 ml, -27.7 ± 45.7 ml, and $0.1 \pm 13.9\%$ for EDV, ESV, and EF, respectively. The 3DE pooled BA biases ± 2 SDs in those same studies were -15.7 ± 31.0 ml, -9.6 ± 25.8 ml, and $0.0 \pm 9.2\%$ for EDV, ESV, and EF, respectively. The differences in biases between 3DE and 2DE were statistically significant for volumetric measurements (p = 0.01 for both EDV and ESV), but not for EF (p = 0.42). The difference in width of the variances was statistically significant for all 3 measurements (p < 0.001).

Interobserver and intraobserver variability. The pooled results for EDV from the 21 substudies (541 echocardio-

| Table 2 | Comparison of Three-Dimensional Echocardiography Versus Magnetic Resonance Imaging |

First Author (Ref. #) Substudy		R Values			Bland-Altman (Mean ± 2 SDs)*			Software/ No. of Slices vs. Mesh
		EDV	ESV	EF	EDV (m)	ESV (m)	EF (%)	
Jenkins et al. (12)								
Noncontrast	1a	N/A	N/A	N/A	−15 ± 36	−9 ± 24	0 ± 6	TomTec/12
Contrast	1b	N/A	N/A	N/A	−6 ± 28	−3 ± 20	0 ± 6	TomTec/12
Soliman et al. (13)	2	0.98	0.98	0.97	−7 ± 20	−4 ± 8	0 ± 6	Qlab/mesh
Pouleur et al. (14)	3	0.97	0.98	0.94	−20 ± 31	−12 ± 31	1 ± 11	Qlab/mesh
Mor-Avi et al. (15)	4	0.91	0.93	0.81	−67 ± 46	−41 ± 46	−3 ± 22	Qlab/mesh
Chukwu et al. (16)								
Normals 8-plane	5a	0.99	0.95	0.79	−1 ± 10	−4 ± 11	2 ± 8	TomTec/8
Normals contours	5b	0.99	0.95	0.84	1 ± 6	−2 ± 10	2 ± 7	TomTec/mesh
MI 8-plane	5c	0.97	0.97	0.90	−7 ± 35	−8 ± 33	1 ± 14	TomTec/8
MI contours	5d	0.97	0.97	0.90	−5 ± 33	−4 ± 31	0 ± 14	TomTec/mesh
Bicudo et al. (17)	6	0.94	0.95	0.83	−4 ± 16	0 ± 10	−2 ± 11	Qlab/8
Qi et al. (18)	7	0.92	0.94	0.92	−22 ± 46	−15 ± 40	5 ± 20	TomTec/8
Jenkins et al. (19)	8	0.89	0.92	0.92	−15 ± 31	−8 ± 26	−1 ± 7	TomTec/12
Van den Bosch et al. (20)	9	0.97	0.98	0.94	−3 ± 12	1 ± 10	−1 ± 7	TomTec/8
Sugeng et al. (21)	10	0.94	0.85	0.87	−5 ± 52	−6 ± 52	0 ± 8	TomTec/mesh
Jenkins et al. (22)								
On-line	11a	0.78	0.86	0.64	−44 ± 35	−21 ± 28	−2 ± 10	Qlab/mesh
Off-line	11b	0.86	0.91	0.81	−15 ± 28	−10 ± 22	1 ± 8	TomTec/12
Jacobs et al. (23)	12	0.96	0.97	0.93	−14 ± 34	−7 ± 32	−1 ± 13	Qlab/mesh
Chan et al. (24)	13	0.90	0.94	N/A	−10 ± 53	−1 ± 38	N/A	TomTec/12
Gutierrez-Chico et al. (25)	14	0.99	0.99	0.98	−13 ± 34	−10 ± 25	N/A	TomTec/8
Jenkins et al. (26)	15	N/A	N/A	N/A	−4 ± 29	−3 ± 18	0 ± 7	TomTec/mesh
Nikitin et al. (27)	16	0.98	0.98	0.94	7 ± 28	3 ± 22	−1 ± 10	TomTec/mesh
Soliman et al. (28)								
TomTec	17a	0.99	0.99	0.98	−9 ± 9	−5 ± 10	0 ± 5	TomTec/mesh
Qlab	17b	0.99	0.98	0.97	−16 ± 13	−9 ± 14	1 ± 6	Qlab/mesh
Soliman et al. (29)								
Contours	18a	0.99	0.99	0.98	−10 ± 9	−5 ± 10	1 ± 5	TomTec/mesh
Multiplane	18b	0.96	0.98	0.95	−24 ± 19	−11 ± 17	1 ± 7	TomTec/8
Kuhl et al. (30)								
Manual	19a	0.99	0.99	0.99	−1 ± 17	0 ± 11	0 ± 5	TomTec/8
Contours	19b	0.98	0.98	0.98	−14 ± 38	−13 ± 41	1 ± 9	TomTec/mesh
Caiani et al. (31)								
Dual-triggering	20a	0.95	0.96	0.77	−14 ± 41	−5 ± 33	−4 ± 22	Qlab/mesh
Continuous imaging	20b	0.89	0.93	0.76	−23 ± 59	−15 ± 41	0 ± 22	
Marsan et al. (32)	21	0.97	0.98	0.97	−22 ± 40	−20 ± 41	1 ± 5	Qlab/mesh
Macron et al. (33)								
1-beat	22a	0.92	0.93	0.85	−21 ± 50	−3 ± 44	−5 ± 16	GE/Mesh
2-beat	22b	0.94	0.96	0.92	−17 ± 42	−9 ± 32	1 ± 12	
4-beat	22c	0.93	0.95	0.88	−15 ± 46	−12 ± 34	3 ± 14	
Chang et al. (34)	23	0.91	0.94	0.91	−41 ± 37	−8 ± 33	−8 ± 13	Siemens/Mesh

*Reported values are rounded to nearest milliliter.

EDV = end-diastolic volume; EF = ejection fraction; ESV = end-systolic volume; MI = myocardial infarction; N/A = not available.

grams) that reported BA analysis for 3DE interobserver and intraobserver variability are shown in Figure 5 (12,13,15,19,21,23,24,26–30,33,34). Of these, 4 studies (110 echocardiograms) also reported results for 2DE (12,19,23,26). For 3DE, mean percent differences (mean ± 2 SDs) were 5.8 ± 12.54 and 3.9 ± 8.5, for interobserver and intraobserver variability, respectively. For 2DE, the mean percent differences were 4.8 ± 21.1 and 0.2 ± 19.6. The difference in width of the variances between 3DE and 2DE was statistically significant

(p < 0.0001) for both interobserver and intraobserver variability.

Discussion

Under controlled study settings in patients with adequate image quality, 3DE offers better accuracy and precision in measuring LV volumes and better precision in measuring LV EF compared with 2DE. Although 3DE shows promise in providing the accessibility of echocardiography and the multi-
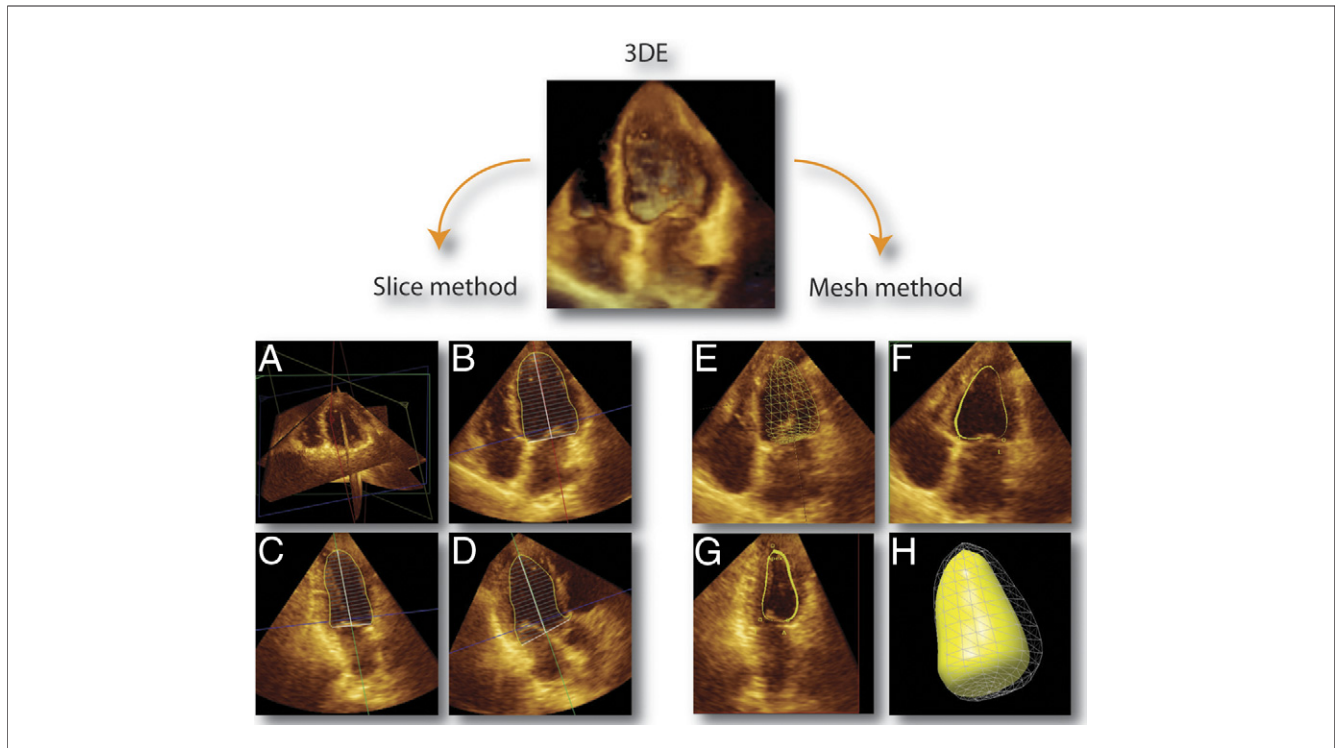
Figure 2   **Examples of Three-Dimensional Echocardiographic Analysis Using Mesh and Slices**

Volumes can be obtained from three-dimensional echocardiographic (3DE) datasets by 1 of 2 methods. With the first, equally spaced slices are traced in end-systole and end-diastole and combined to obtain volumes (3 traced slices are shown in **A to D**). The second method uses automated software to generate a 3D mesh from points identified on the 4- and 2-chamber views **(E to H)**.

planar imaging of CMR, this nascent technology still has limited spatial and temporal resolution compared with CMR as evidenced by clinically significant biases and limits of agreement.

How good is 3DE? From individual studies and summary results, it is clear that 3DE, like 2DE, consistently underestimates LV volumes (but not EF). More importantly, there is substantial variability in 3DE calculations, such that to achieve 95% confidence intervals of true values, one would have to allow for ±34 ml for EDV, ±30 ml for ESV, and ±12% for EF, numbers that would change management for many patients. Furthermore, these results represent images selected for higher quality, and consequently underestimate the expected variability of routine practice. For example, an analysis of those studies that accepted all 3D datasets, regardless of image quality, increased the 95% confidence interval to ±38 ml for EDV, ±34 ml for ESV, and ±15% for EF.

Which patient populations or methods alter the performance of 3DE? Although 3DE does not make assumptions about the LV shape, it fares worse in sicker patients with large ventricles, as seen in the study by Chukwo et al. (16), who reported results on controls separately from those with myocardial infarcts. Very large ventricles may not fit within the scanning sector allotted by the probe (32,34). With an average EDV of 195 ml, our pooled patient population is one with significant LV remodeling, which may account, in part, for the wide confidence interval we found. LV contrast may improve these results; however, in the study by Caiani et al. (who only studied patients with poor images), there was still a large confidence interval (31), and Jenkins et al. showed only a minimal clinical benefit with LV opacification (12). As described in the preceding text, there were 2 very different methods of obtaining volumes from the 3D datasets. As the mesh technique involves substantially less analysis time compared with tracing slices, it is more likely to be used routinely. Although some studies that directly compared the 2 methods reported better results with tracing slices (16,22), others found that the quicker mesh technique was more accurate (29,30). Our pooled analysis of the 4 studies demonstrated little difference in the bias and variances between the mesh and slices methods, which may validate the quicker mesh technique. Only 1 study directly compared 2 separate vendors and found that the volumes (but not EF) were more accurate in one vendor (28). This study was small and limited to a single center, so more information might be needed before concluding that one platform is superior to the others.

Does 3DE offer an advantage over 2DE? Despite its limitations, 3DE may be superior to 2D techniques. For volumes, 3DE under-represents true values about 50% less
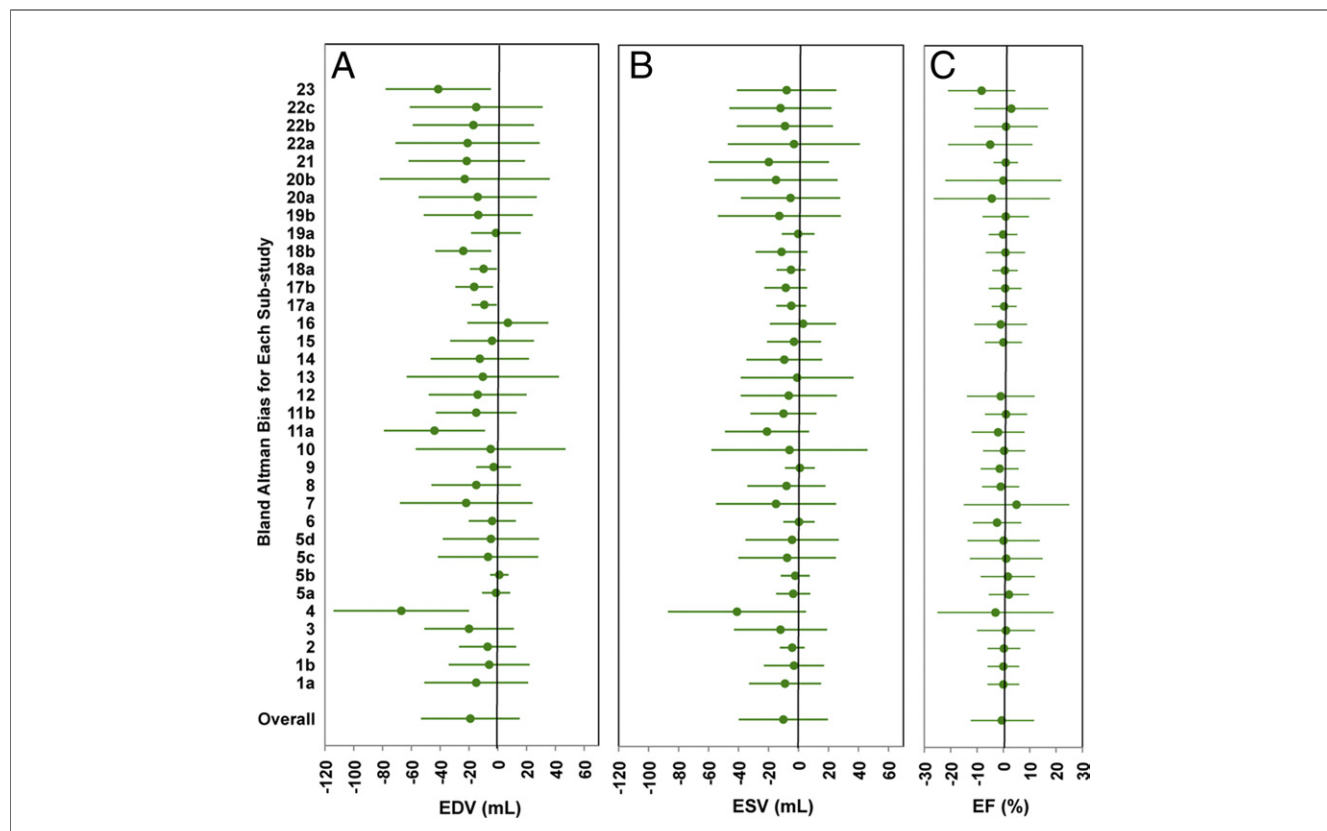
Figure 3   **Three-Dimensional Echocardiography Versus Cardiac Magnetic Resonance**

The absolute Bland-Altman difference between 3DE and CMR, expressed as bias ±2 SDs for each substudy. The overall pooled results are shown at the bottom. End-diastolic volume (EDV), end-systolic volume (ESV), and ejection fraction (EF) are shown in **A, B, and C,** respectively. Abbreviations as in Figures 1 and 2.

and has one-half the 95% confidence interval compared with 2DE. However, more clinical decisions are based on EF. With EF, there is no difference in the bias between 3DE and 2DE, and the difference in the variance is modest (±4.7%). The benefit of 3DE can be appreciated by evaluating the intraobserver and interobserver variability, where 3DE demonstrates much lower variance on both tests. Low observer variability is particularly important in a real-world echocardiographic laboratories, with a variety of readers and sonographers and for patients undergoing serial examinations to determine clinical worsening.

Strengths and limitations of this meta-analysis. Our pooled patient sample included a large number of patients with a wide variety of cardiac diseases. Although we did include a few studies that were limited to specific patient groups (17,20,32), most recruited consecutive patients representing the spectrum of patients seen in everyday practice. To ascertain the overall bias of 3DE as seen in Figure 3, we combined all the included studies, regardless of technique and platform. We did this to demonstrate the overall bias and variability of the technology as it performs with the variety of methods used among different clinical and research laboratories. We recognize, however, that technique and software may matter, which is why we also conducted a

subanalysis on those studies that included both the mesh and slices methods. Many studies reported the analysis separately for the different techniques, and we chose to consider these substudies as individual data sources. Although this strategy means that some, but not all, patients were represented twice, it allows the variety of 3D methods to be included. The disadvantage of this approach is that it gives more weight to some patient groups; however, there were subgroup analysis in enough studies, most of which had a variety of patients, that this approach did not significantly change the patient mix. As the aim of this meta-analysis was to evaluate 3DE performance and not patient characteristics, it was more important to include all the different 3DE analyses. We also demanded that 3DE be compared with CMR for determination of "true" values. We realize that CMR itself is not perfect, because it also has errors related to border-detection and controversy on the inclusion of basal LV planes. Despite these potential sources of error, CMR is most often used as the gold standard, both clinically and in research studies. As with any meta-analysis, ours could also have reported bias, because only those studies showing positive results for 3DE might have been published. However, funnel plot analysis of our citations demonstrated no publication bias (36).

Table 3	Comparison of Three- Versus Two-Dimensional Echocardiography														

			3D Echocardiography						2D Echocardiography						
			R -Values			Bland-Altman*			R-Values			Bland-Altman*			
First Author (Ref. #) Substudy		N	EDV	EDS	EF	EDV (ml)	ESV (ml)	EF (%)	EDV	ESV	EF	EDV (ml)	ESV (ml)	EF (%)
Jenkins et al. (12)														
Non-contrast	1a	50	N/A	N/A	N/A	−15 ± 36	−9 ± 24	0 ± 6	N/A	N/A	N/A	−41 ± 42	−22 ± 36	−2 ± 8
Contrast	1b	50	N/A	N/A	N/A	−6 ± 28	−3 ± 20	0 ± 6	N/A	N/A	N/A	−18 ± 38	−8 ± 32	−2 ± 8
Chukwu et al. (16)														
Normals 8-plane	5a	35	0.99	0.95	0.79	−1 ± 10	−4 ± 11	2 ± 8	0.50	0.86	0.23	−38 ± 51	−25 ± 26	9 ± 16
Normals contours	5b	35	0.99	0.95	0.84	1 ± 6	−2 ± 10	2 ± 7	0.50	0.86	0.23	−38 ± 51	−25 ± 26	9 ± 16
MI 8-plane	5c	34	0.97	0.97	0.90	−7 ± 35	−8 ± 33	1 ± 14	0.79	0.86	0.86	−58 ± 81	−46 ± 73	5 ± 18
MI contours	5d	34	0.97	0.97	0.90	−5 ± 33	−4 ± 31	0 ± 14	0.79	0.86	0.86	−58 ± 81	−46 ± 73	5 ± 18
Bicudo et al. (17)	6	20	0.94	0.95	0.83	−4 ± 16	0 ± 10	−2 ± 11	0.89	0.88	0.75	1 ± 23	−0 ± 14	−1 ± 14
Jenkins et al. (19)	8	30	0.89	0.92	0.92	−15 ± 31	−8 ± 26	−1 ± 7	0.72	0.79	0.65	−57 ± 40	−26 ± 33	−5 ± 12
Jenkins et al. (22)														
On-line	11a	110	0.78	0.86	0.64	−44 ± 35	−21 ± 28	−2 ± 10	0.71	0.8	0.61	−70 ± 39	−34 ± 33	−3 ± 11
Off-line	11b	110	0.86	0.91	0.81	−15 ± 28	−10 ± 22	−1 ± 8	0.68	0.78	0.70	−70 ± 39	−34 ± 33	−3 ± 11
Jacobs et al. (23)	12	50	0.96	0.97	0.93	−14 ± 34	−7 ± 32	−1 ± 13	0.89	0.92	0.86	−23 ± 58	−15 ± 48	0.8 ± 19
Gutierrez-Chico et al. (25)	14	35	0.99	0.99	0.98	−13 ± 34	−10 ± 25	N/A	0.78	0.86	0.92	−25 ± 132	−21 ± 107	3 ± 18
Jenkins et al. (26)	15	50	N/A	N/A	N/A	−4 ± 29	−3 ± 18	0 ± 7	N/A	N/A	N/A	−54 ± 33	−28 ± 28	−1 ± 13
Marsan et al. (32)	21	52	0.97	0.98	0.97	−22 ± 40	−20 ± 40	1 ± 5	0.94	0.95	0.96	−41 ± 49	−33 ± 45	1 ± 5

*Rounded to nearest milliliter.

EDV = end-diastolic volume; EF = ejection fraction; ESV = end-systolic volume; N/A = not available.

Conclusions

Three-dimensional echocardiography underestimates true LV volumes and EF and has a substantial degree of variance, especially in patients with poor images or large ventricles. In these patient groups, or when the results are critical and contradict other clinical data, a degree of
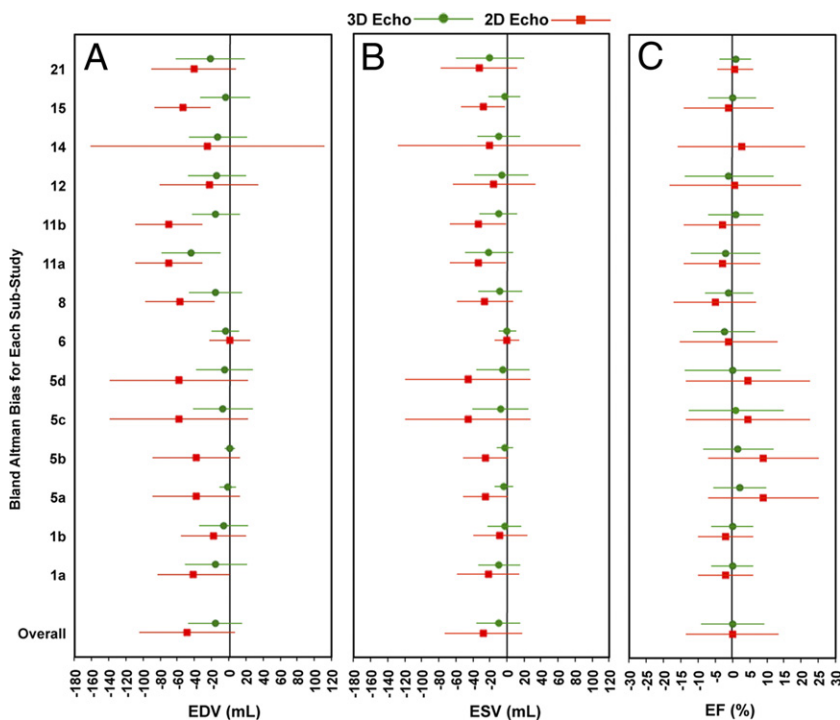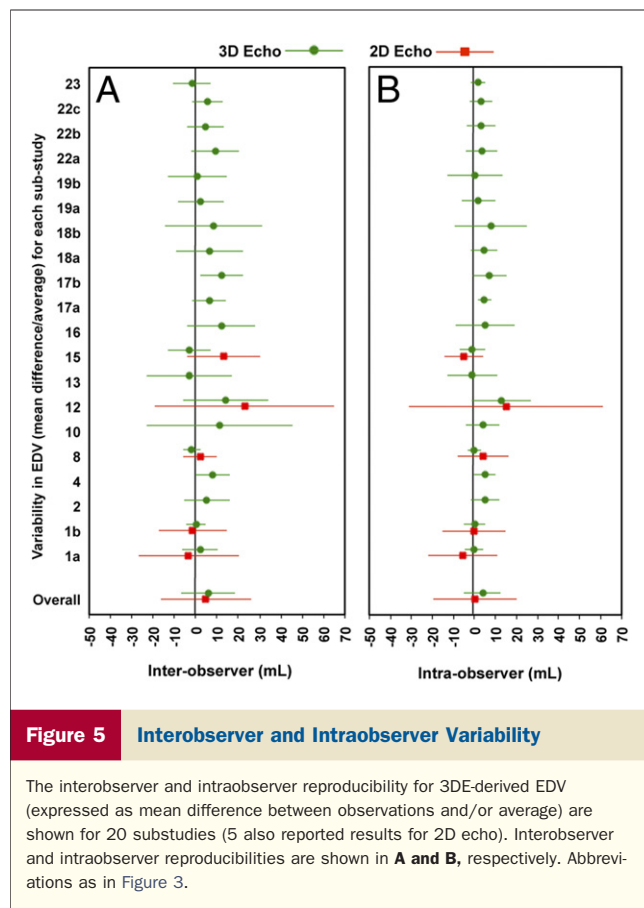

Figure 4	3-Versus 2-Dimensional Echocardiography

Comparison of the Bland-Altman differences (bias ±2 SDs) of 3DE compared with CMR versus that of 2D echocardiography in the 12 substudies that reported on both methods. The overall pooled results are shown at the bottom. EDV, ESV, and EF are shown in **A, B, and C,** respectively. Abbreviations as in Figures 1 and 3.

Figure 5   Interobserver and Intraobserver Variability

The interobserver and intraobserver reproducibility for 3DE-derived EDV (expressed as mean difference between observations and/or average) are shown for 20 substudies (5 also reported results for 2D echo). Interobserver and intraobserver reproducibilities are shown in **A and B,** respectively. Abbreviations as in Figure 3.

skepticism is warranted. With these caveats in mind, 3DE offers an advantage over 2DE in providing better accuracy, precision, and reproducibility for volume measurements. The advantage in measuring EF, however, is limited to a modest increase in precision.

Reprints requests and correspondence: Dr. Jennifer L. Dorosz, University of Colorado Denver, 12631 East 17th Avenue, Mail Stop B130, Aurora, Colorado 80045. E-mail: Jennifer.Dorosz@ucdenver.edu.

REFERENCES

1. Bonow RO, Carabello BA, Chatterjee K, et al. 2008 focused update incorporated into the ACC/AHA 2006 guidelines 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 (Writing Committee to Revise the 1998 Guidelines for the Management of Patients With Valvular Heart Disease). Endorsed by the Society of Cardiovascular Anesthesiologists, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. J Am Coll Cardiol 2008;52:e1–142.
2. Moss AJ, Zareba W, Hall WJ, et al. Prophylactic implantation of a defibrillator in patients with myocardial infarction and reduced ejection fraction. N Engl J Med 2002;346:877–83.
3. Packer M, Fowler MB, Roecker EB, et al. Effect of carvedilol on the morbidity of patients with severe chronic heart failure: results of the carvedilol prospective randomized cumulative survival (COPERNICUS) study. Circulation 2002;106:2194–9.
4. Young JB, Abraham WT, Smith AL, et al. Combined cardiac resynchronization and implantable cardioversion defibrillation in ad-
vanced chronic heart failure: the MIRACLE ICD Trial. JAMA 2003;289:2685–94.
5. SOLVD Investigators. Effect of enalapril on mortality and the development of heart failure in asymptomatic patients with reduced left ventricular ejection fractions. N Engl J Med 1992;327:685–91.
6. Lang RM, Bierig M, Devereux RB, et al. Recommendations for chamber quantification: a report from the American Society of Echocardiography's Guidelines and Standards Committee and the Chamber Quantification Writing Group, developed in conjunction with the European Association of Echocardiography, a branch of the European Society of Cardiology. J Am Soc Echocardiogr 2005;18:1440–63.
7. Jensen-Urstad K, Bouvier F, Höjer J, et al. Comparison of different echocardiographic methods with radionuclide imaging for measuring left ventricular ejection fraction during acute myocardial infarction treated by thrombolytic therapy. Am J Cardiol 1998;81:538–44.
8. Alfakih K, Reid S, Jones T, Sivananthan M. Assessment of ventricular function and mass by cardiac magnetic resonance imaging. Eur Radiol 2004;14:1813–22.
9. American Society of Echocardiography Coding and Reimbursement Newsletter. December 2009. Available at: http://www.asecho.org/files/public/CodingnewsJan09.pdf. Accessed July 15, 2011.
10. Rosenthal MB. Hard choices—alternatives for reining in Medicare and Medicaid spending. N Engl J Med 2011;364:1887–90.
11. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.
12. Jenkins C, Moir S, Chan J, Rakhit D, Haluska B, Marwick TH. Left ventricular volume measurement with echocardiography: a comparison of left ventricular opacification, three-dimensional echocardiography, or both with magnetic resonance imaging. Eur Heart J 2009;30:98–106.
13. Soliman OII, Kirschbaum SW, van Dalen BM, et al. Accuracy and reproducibility of quantitation of left ventricular function by real-time three-dimensional echocardiography versus cardiac magnetic resonance. Am J Cardiol 2008;102:778–83.
14. Pouleur A-C, le Polain de Waroux J-B, Pasquet A, et al. Assessment of left ventricular mass and volumes by three-dimensional echocardiography in patients with or without wall motion abnormalities: comparison against cine magnetic resonance imaging. Heart 2008;94:1050–7.
15. Mor-Avi V, Jenkins C, Kühl HP, 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. J Am Coll Cardiol Img 2008;1:413–23.
16. Chukwu EO, Barasch E, Mihalatos DG, et al. Relative importance of errors in left ventricular quantitation by two-dimensional echocardiography: insights from three-dimensional echocardiography and cardiac magnetic resonance imaging. J Am Soc Echocardiogr 2008;21:990–7.
17. Bicudo LS, Tsutsui JM, Shiozaki A, et al. Value of real time three-dimensional echocardiography in patients with hypertrophic cardiomyopathy: comparison with two-dimensional echocardiography and magnetic resonance imaging. Echocardiography 2008;25:717–26.
18. Qi X, Cogar B, Hsiung MC, et al. Live/real time three-dimensional transthoracic echocardiographic assessment of left ventricular volumes, ejection fraction, and mass compared with magnetic resonance imaging. Echocardiography 2007;24:166–73.
19. Jenkins C, Leano R, Chan J, Marwick TH. Reconstructed versus real-time 3-dimensional echocardiography: comparison with magnetic resonance imaging. J Am Soc Echocardiogr 2007;20:862–8.
20. van den Bosch AE, Robbers-Visser D, Krenning BJ, et al. Real-time transthoracic three-dimensional echocardiographic assessment of left ventricular volume and ejection fraction in congenital heart disease. J Am Soc Echocardiogr 2006;19:1–6.
21. Sugeng L, Mor-Avi V, Weinert L, et al. Quantitative assessment of left ventricular size and function: side-by-side comparison of real-time three-dimensional echocardiography and computed tomography with magnetic resonance reference. Circulation 2006;114:654–61.
22. Jenkins C, Chan J, Hanekom L, Marwick TH. Accuracy and feasibility of online 3-dimensional echocardiography for measurement of left ventricular parameters. J Am Soc Echocardiogr 2006;19:1119–28.
23. Jacobs LD, Salgo IS, Goonewardena S, et al. Rapid online quantification of left ventricular volume from real-time three-dimensional echocardiographic data. Eur Heart J 2006;27:460–8.

24. Chan J, Jenkins C, Khafagi F, Du L, Marwick TH. What is the optimal clinical technique for measurement of left ventricular volume after myocardial infarction? A comparative study of 3-dimensional echocardiography, single photon emission computed tomography, and cardiac magnetic resonance imaging. J Am Soc Echocardiogr 2006;19:192–201.
25. Gutiérrez-Chico JL, Zamorano JL, Pérez de Isla L, et al. Comparison of left ventricular volumes and ejection fractions measured by three-dimensional echocardiography versus by two-dimensional echocardiography and cardiac magnetic resonance in patients with various cardiomyopathies. Am J Cardiol 2005;95:809–13.
26. Jenkins C, Bricknell K, Hanekom L, Marwick TH. Reproducibility and accuracy of echocardiographic measurements of left ventricular parameters using real-time three-dimensional echocardiography. J Am Coll Cardiol 2004;44:878–86.
27. Nikitin NP, Constantin C, Loh PH, et al. New generation 3-dimensional echocardiography for left ventricular volumetric and functional measurements: comparison with cardiac magnetic resonance. Eur J Echocardiogr 2006;7:365–72.
28. Soliman OII, Krenning BJ, Geleijnse ML, et al. A comparison between QLAB and TomTec full volume reconstruction for real time three-dimensional echocardiographic quantification of left ventricular volumes. Echocardiography 2007;24:967–74.
29. Soliman OII, Krenning BJ, Geleijnse ML, et al. Quantification of left ventricular volumes and function in patients with cardiomyopathies by real-time three-dimensional echocardiography: a head-to-head comparison between two different semiautomated endocardial border detection algorithms. J Am Soc Echocardiogr 2007;20:1042–9.
30. Kühl HP, Schreckenberg M, Rulands D, et al. High-resolution transthoracic real-time three-dimensional echocardiography: quantita-tion of cardiac volumes and function using semi-automatic border detection and comparison with cardiac magnetic resonance imaging. J Am Coll Cardiol 2004;43:2083–90.
31. Caiani EG, Coon P, Corsi C, et al. Dual triggering improves the accuracy of left ventricular volume measurements by contrast-enhanced real-time 3-dimensional echocardiography. J Am Soc Echocardiogr 2005;18:1292–8.
32. Marsan NA, Westenberg JM, Roes SD, et al. Three-dimensional echocardiography for the preoperative assessment of patients with left ventricular aneurysm. Ann Thorac Surg 2011;91:113–21.
33. Macron L, Lim P, Bensaid A, et al. Single-beat versus multibeat real-time 3D echocardiography for assessing left ventricular volumes and ejection fraction: a comparison study with cardiac magnetic resonance. Circ Cardiovasc Imaging 2010;3:450–5.
34. Chang SA, Lee SC, Kim EY, et al. Feasibility of single-beat full-volume capture real-time three-dimensional echocardiography and auto-contouring algorithm for quantification of left ventricular volume: validation with cardiac magnetic resonance imaging. J Am Soc Echocardiogr 2011;24:853–9.
35. Levene H. Robust tests for equality of variances. In: Olkin I, Ghurye S, Hoeffding W, et al. editors. Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling. Palo Alto, CA: Stanford University Press, 1960:278–92.
36. Begg CB. A measure to aid in the interpretation of published clinical trials. Stat Med 2006;4:1–9.