JACC: CARDIOVASCULAR IMAGING
VOL. ■, NO. ■, 2021

ISSN 1936-878X/$36.00
PUBLISHED BY ELSEVIER

iMAIL

Systematic Quantification of Sources of Variation in Ejection Fraction Calculation Using Deep Learning

Accurate left ventricular (LV) ejection fraction (LVEF) assessment is essential for diagnosing and managing many medical conditions, including heart failure, myocardial infarction, valvular disease, and even cancer (1,2). Echocardiography is the most frequently used modality to assess LVEF because of its lack of ionizing radiation, widespread availability, and high temporal resolution. However, echocardiographic assessment is also prone to significant intraprovider variability because of its reliance on expert view acquisition and measurements (1). Potential sources of error in tracings and view acquisition are known (3). However, the degree to which small variations affect downstream calculations of LVEF has not been well studied.

In this study, we used deep learning to simulate common variations of echocardiogram tracing and view acquisition across many heart geometries and assessed their individual effects on LVEF quantification. We quantified ventricular volumes and LVEF from 3,906 apical-4-chamber videos randomly sampled from a published database of consecutive echocardiograms from Stanford Healthcare in Palo Alto, California (4). Patients were on average 69 years of age, 45.3% were female, and these patients had cardiovascular comorbidities: 46.4% had coronary artery disease, 68.8% had hypertension, 57.6% had heart failure, 33.4% had diabetes, and 39.0% had chronic kidney disease. The average LVEF was 50.9% $\pm$ 11.6%, and LV diastolic volume was 91.4 mL (SD = 46.5 mL); 32.5% had reduced LVEF, 12.3% had LV dilation, and 25.4% had LV hypertrophy.

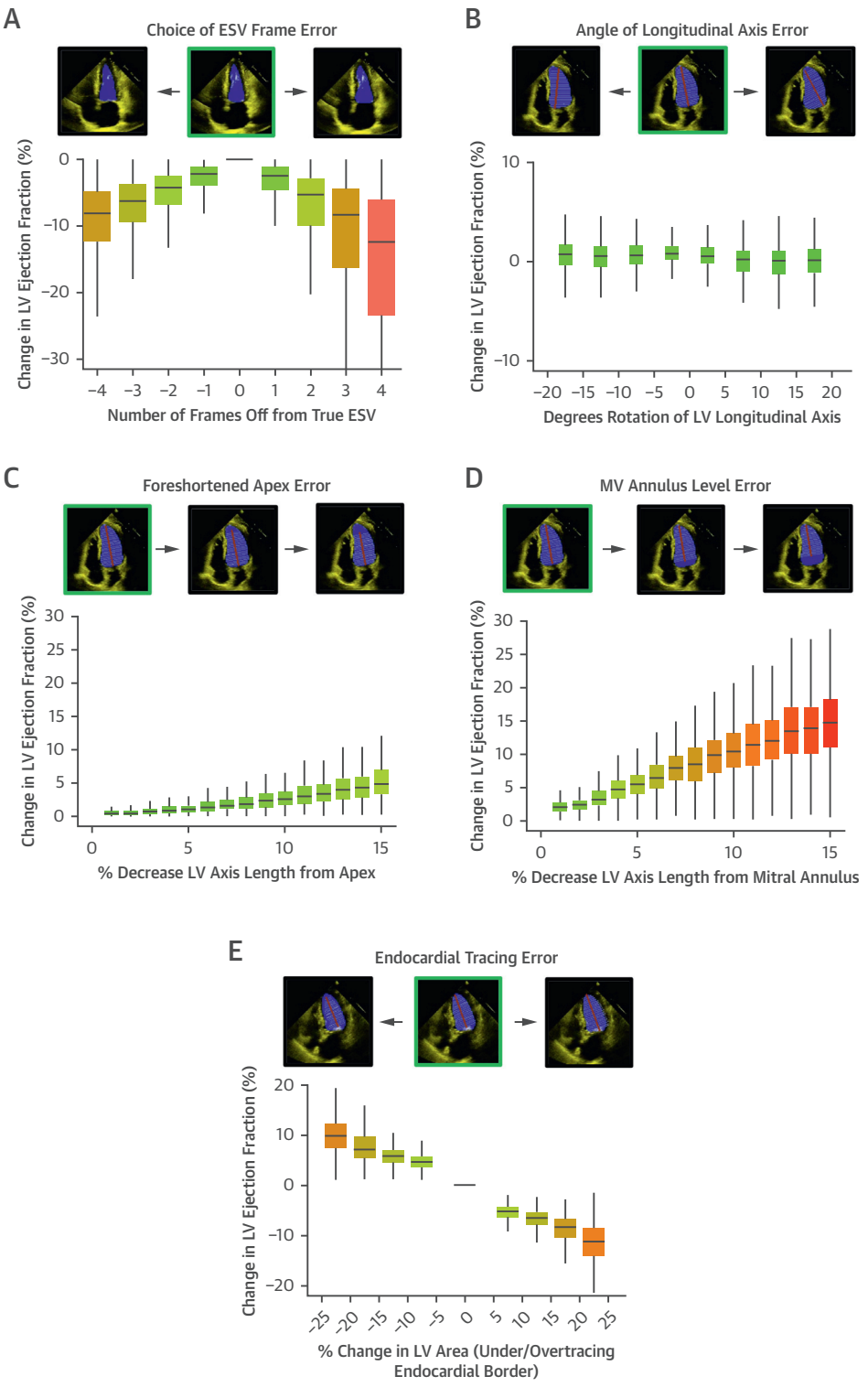
Leveraging a semantic segmentation deep learning architecture, we simulated variations by generating 976,500 LV tracings with varying degrees of individual tracing variations introduced at end-systole (4). Variations included LV apex foreshortening, mitral annular level mismeasurement, overtracing or undertracing of the endocardium, and angle misalignment of the LV longitudinal axis. From the tracings, we calculated the reference LVEF and quantified the impact of introduced variations on single-plane, method-of-disks LVEF assessment. We performed linear regression for change in LVEF vs amount of introduced error. Echocardiograms and code are available at https://echonet.github.io/dynamic/, https://github.com/echonet/variance. The Cedars-Sinai Institutional Review Board approved all research.

Mistiming end-systole by 3 frames (60 ms), foreshortening by 8% from the mitral annulus, and mistracing the endocardial border by 20% to 25% each resulted in ~10% change in measured LVEF (**Figures 1A to 1E**). In contrast, LV apex foreshortening and LV longitudinal axis misalignment had relatively little effect on LVEF. Mistiming end-systole by 3 frames and mistracing the endocardial border by 20% to 25% reclassified 13.6% (95% CI: 12.5%-14.8%) and 6.0% (3.7%-8.6%) of patients with LVEF >50% to <40%, respectively. Foreshortening by 5% to 10% from the mitral annulus reclassified 31.6% (20.0%-43.8%) of patients with LVEF <40% to LVEF >50%. By linear approximation, choosing the wrong end-systolic frame resulted in a 3.8% (95% CI: 3.7%-3.8%) decrease in LVEF for each frame after true end-systole. Shortening the LV axis from the mitral annular level caused a 9.6% (9.4%-9.8%) increase in LVEF per 10% decrease in LV axis length. There was a 4.4% (4.3%-4.5%) decrease in LVEF for each 10% area increase as a result of overtracing the endocardium.

Taken together, small variation in the selection of the end systolic frame, measurement of the mitral annular level, and tracing of the endocardial border caused large changes in measured LVEF. Although studies have identified variability in LVEF assessment across observers and imaging modalities, none to date have systematically quantified the impact of individual measurement errors, possibly because of the difficulty of reproducing errors manually (1,3).

Our findings are consistent with known properties of ventricular timing and geometry. Assuming a 16-frame systolic cycle (eg, an echocardiogram at 30 frames/s and a heart rate of 90 beats/min), for an LVEF of 50%, at least 3% to 4% of the ejection fraction occurs every frame, consistent with our simulated 3.8% absolute LVEF decrease for each frame after true end-systole. Because the left ventricle is wider at the mitral annulus than the apex, mistracing the annular level produced more LVEF error compared with foreshortening from the apex. Overtracing or undertracing the endocardium grows or shrinks the disk diameters used by the Simpson method; LVEF errors therefore grow quickly, being proportional to the diameter squared.

FIGURE 1 **Effects of Manual Tracing and View Acquisition Errors on Measured Left Ventricular Ejection Fraction**



Absolute changes in calculated left ventricular (LV) ejection fraction after systematic introduction of varying degrees of common tracing and view acquisition errors, including errors in **(A)** choice of end systolic volume (ESV) frame, **(B)** angle of left ventricular longitudinal axis, **(C)** foreshortening of the left ventricular apex, **(D)** mitral valve (MV) annulus level, and **(E)** tracing of the endocardial borders.

Regarding limitations, given that the echocardiograms were from 1 academic center, referral bias may affect the generalizability of findings to other centers. Because our segmentation model was previously validated using apical 4-chamber videos only, LVEFs were not calculated using apical 2-chamber views. Quantifying the effects of mistiming end-systole by linear regression remains an approximation given the nonlinearity of contraction during isovolumic periods. LVEF errors resulting from simulated apical foreshortening may be underestimated because foreshortening also affects endocardial border tracings.

Variation in LVEF measurement can be clinically significant, potentially determining whether a patient is started on heart failure medications, receives an implantable cardioverter-defibrillator, has a valve procedure, or stops chemotherapy. Clinicians should therefore consider paying special attention to the highlighted pitfalls in LVEF assessment. Although our study was limited to echocardiograms, manually tracing chambers and selecting frames affect all imaging modalities. Methods that reduce this variation, such as using automated artificial intelligence methods, may help improve the precision of care that relies on LVEF (4,5).

Neal Yuan, MD*
Ishan Jain
Neeraj Rattehalli
Bryan He, BS
Charles Pollick, MD
David Liang, MD, PhD
Paul Heidenreich, MD
James Zou, PhD
Susan Cheng, MD, MPH, MS†
David Ouyang, MD†

*Smidt Heart Institute
Cedars-Sinai Medical Center
8700 Beverly Blvd, Davis 1015
Los Angeles, California 90048, USA
Email: neal.yuan@cshs.org

https://doi.org/10.1016/j.jcmg.2021.06.018

†Drs Cheng and Ouyang are joint senior authors of this work.

The authors have reported that they have no relationships relevant to the contents of this paper to disclose.

The authors attest they are in compliance with human studies committees and animal welfare regulations of the authors' institutions and Food and Drug Administration guidelines, including patient consent where appropriate. For more information, visit the Author Center.

REFERENCES

1. Pellikka PA, She L, Holly TA, et al. Variability in ejection fraction measured by echocardiography, gated single-photon emission computed tomography, and cardiac magnetic resonance in patients with coronary artery disease and left ventricular dysfunction. *JAMA Netw Open*. 2018;1(4):e181456. https://doi.org/10.1001/jamanetworkopen.2018.1456

2. Plana JC, Galderisi M, Barac A, 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. *J Am Soc Echocardiogr*. 2014;27(9):911–939. https://doi.org/10.1016/j.echo.2014.07.012

3. Lang RM, Badano LP, Mor-Avi V, et al. Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *J Am Soc Echocardiogr*. 2015;28(1):1–39.e14. https://doi.org/10.1016/j.echo.2014.10.003

4. Ouyang D, He B, Ghorbani A, et al. Video-based AI for beat-to-beat assessment of cardiac function. *Nature*. 2020;580(7802):252–256. https://doi.org/10.1038/s41586-020-2145-8

5. Zhang J, Gajjala S, Agrawal P, et al. Fully automated echocardiogram interpretation in clinical practice. *Circulation*. 2018;138(16):1623–1635. https://doi.org/10.1161/CIRCULATIONAHA.118.034338