

**Systematic quantification of sources of variation in ejection fraction calculation
using deep learning**

Neal Yuan, M.D.¹, Ishan Jain¹, Neeraj Rattehalli¹, Bryan He, B.S.², Charles Pollick, M.D.
¹, David Liang M.D., Ph.D.³, Paul Heidenreich, M.D.³, James Zou, Ph.D.^{2,4}, Susan
Cheng, M.D., M.P.H., M.Sc.^{1,6}, David Ouyang, M.D. ^{1,5,6}

1. Smidt Heart Institute, Cedars-Sinai Medical Center, Los Angeles, CA
2. Department of Computer Science, Stanford University, Palo Alto, CA
3. Department of Medicine, Stanford University, Palo Alto, CA
4. Department of Biomedical Data Science, Stanford University, Palo Alto, CA
5. Department of Medicine, Division of Artificial Intelligence in Medicine, Cedars-Sinai Medical Center, Los Angeles, CA
6. Co-Senior Author

Corresponding Author

Neal Yuan, M.D.
8700 Beverly Blvd, Davis 1015
Los Angeles CA 90048
(310) 423-3843
Neal.Yuan@cshs.org

Accurate left ventricular 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 intra-provider variability due to 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 impact 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.⁴ Patients were on average 69 years old, 45.3% female, and had cardiovascular comorbidities: 46.4% coronary artery disease, 68.8% hypertension, 57.6% heart failure, 33.4% diabetes, 39.0% chronic kidney disease. The average LVEF was 50.9% (SD 11.6) and LV diastolic volume was 91.4mL (SD 46.5); 32.5% had reduced LVEF, 12.3% had LV dilation, 25.4% had LV hypertrophy.

Leveraging a semantic segmentation deep learning architecture, we simulated variations by generating 976,500 left ventricular tracings with varying degrees of

individual tracing variations introduced at end systole.⁴ Variations included LV apex foreshortening, mitral annular level mismeasurement, over/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-discs LVEF assessment. We performed linear regression for change in LVEF versus 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 (60ms), foreshortening by 8% from the mitral annulus, and mistracing the endocardial border by 20-25% each resulted in ~10% change in measured LVEF (**Figure 1**). 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 20-25% 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 5-10% 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 due to 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. While 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 due to the difficulty of reproducing errors manually.^{1,3}

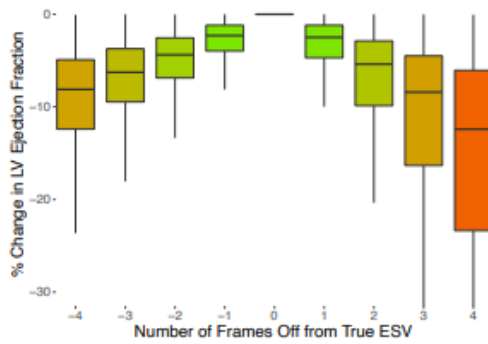
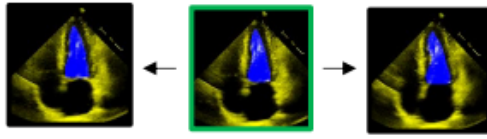
Our findings are consistent with known properties of ventricular timing and geometry. Assuming a 16-frame systolic cycle (E.g. an echocardiogram at 30 frames/second and a heart rate of 90 beats/min), for an LVEF of 50%, at least 3-4% of the ejection fraction occurs every frame, consistent with our simulated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

Regarding limitations, since echocardiograms were from a single academic center, referral bias may affect the generalizability of findings to other centers. Since our segmentation model was previously validated using apical-4-chamber videos only, LVEFs in this study were obtained by single plane method of disks. Quantifying the effects of mistiming end-systole by linear regression remains an approximation given the non-linearity of contraction from isovolumic periods. LVEF errors due to simulated

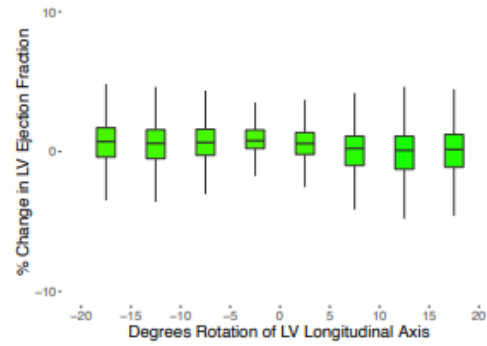
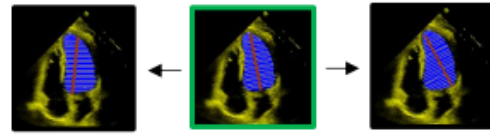
apical foreshortening may be underestimated since foreshortening also affects endocardial border tracings.

Variation in LVEF measurement can be clinically significant, potentially determining whether a patient is started on heart failure medications, receives an ICD, has a valve procedure, or stops chemotherapy. Clinicians should therefore consider paying special attention to the highlighted pitfalls in LVEF assessment. While 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}

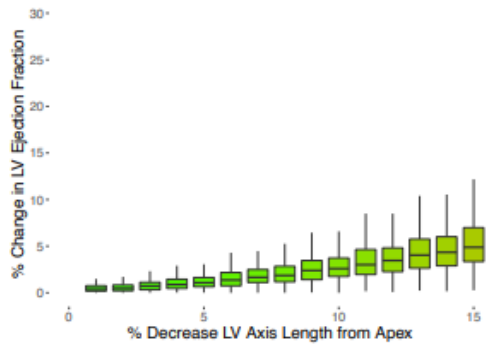
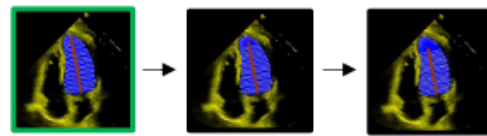
A. Choice of ESV Frame Error



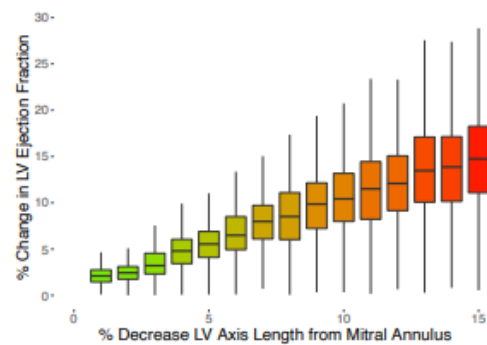
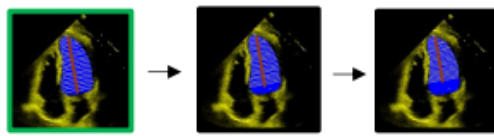
B. Angle of Longitudinal Axis Error



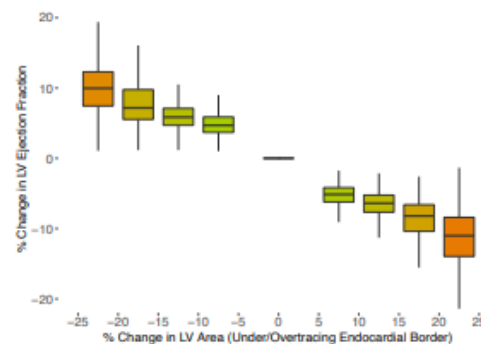
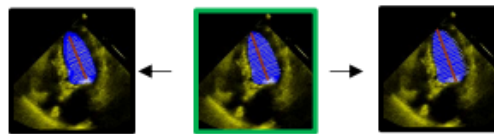
C. Foreshortened Apex Error



D. MV Annulus Level Error



E. Endocardial Tracing Error



106 **Figure 1. Effects of manual tracing and view acquisition errors on measured left**
107 **ventricular ejection fraction**
108 Absolute changes in calculated left ventricular (LV) ejection fraction after systematic
109 introduction of varying degrees of common tracing and view acquisition errors, including
110 errors in **(A)** choice of end systolic volume (ESV) frame, **(B)** angle of left ventricular
111 longitudinal axis, **(C)** foreshortening of the left ventricular apex, **(D)** mitral valve (MV)
112 annulus level, and **(E)** tracing of the endocardial borders.
113

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. doi: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. doi: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. doi: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. doi: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. doi:10.1161/CIRCULATIONAHA.118.034338