Validation of machine learning model for automated estimation of left ventricular ejection fraction on cardiac point of care ultrasound

Christina L Luong*, Mohammad H. Jafari*, Delaram Behnami*, Yaksh R Shah*, Lynn Straatman, Nathan Van Woudenberg, Leah Christoff, Nancy Gwadry, Nathaniel Hawkins, Eric C. Sayre, Darwin Yeung, Michael Tsang, Ken Gin, John Jue, Parvathy Nair, Purang Abolmaesumi#, Teresa Tsang#

Background: Machine learning algorithms (ML) have been shown to estimate the left ventricular ejection fraction (LVEF) from echo with a high degree of accuracy. However, few studies validate ML models on point-of-care ultrasound (POCUS). We aimed to test our LVEF ML model on clinician-scanned cardiac POCUS compared with Level III echocardiographers' interpretation.

Methods: Clinicians at a tertiary care Cardiac Function Clinic prospectively scanned 138 participants using hand-carried ultrasound devices to acquire the apical 2 and apical 4 chamber views. The videos were analyzed offline by an ML model for automated estimation of LVEF. Reference standards were established by a Level III echocardiographer assigning an LVEF on the acquired 1257 cardiac POCUS videos either in randomized order or in the context of all videos for a given participant. The Level III echocardiographer estimated LVEF visually and using Simpson biplane method of discs (segmentation) when feasible.

Results: Of 138 participants scanned, 120 had videos sufficient for Level III echocardiographer estimation of LVEF (on at least one video file). The ML model generated LVEF predictions on 91 subjects. The mean age was 66.2±14.3 years, 86% were male, 29% were overweight (BMI ≥ 30), and 39% were in atrial fibrillation at the time of the scan. When examining the ML model performance based on individual video files, we observed good intraclass correlation (ICC) compared with reference standards. The ICC was used to compare the ML model estimates and level III echocardiographer estimates for each sample. When comparing LVEF estimates for randomized single POCUS videos, the level III echocardiographer's visual estimates and ML model produced an ICC of 0.772 (good correlation) and 0.778 for videos that could be segmented. When the level III echocardiographer was able to review all videos for a participant, the ICC was 0.794, which improved to 0.843 for studies that could be segmented.

Conclusion: We conclude that our ML model can produce automated LVEF estimates on clinician-driven cardiac POCUS videos that correlate well with expert interpretation (ICC = 0.77 to 0.84). ML models for cardiac ultrasound may provide a tool for clinical support in settings without ready access to expert ultrasound interpretation.

*Joint first authors
#Joint senior authors