⭐ Spotlight on Special Topics

AI-GUIDED ECHOCARDIOGRAPHY SYSTEM MATCHES THE IMAGE QUALITY ASSESSMENT ABILITY OF CARDIAC SONOGRAPHERS

Poster Contributions
Sunday, May 16, 2021, 9:45 a.m.-10:30 a.m.

Session Title: Spotlight on Special Topics: Innovation, Digital Health, and Technology 4
Abstract Category: 60. Spotlight on Special Topics: Innovation, Digital Health, and Technology

Authors: _Ha Hong_, Samuel Surette, Ali Khalid Chaudhry, Nripesh Parajuli, Charles Cadieu, Randolph Martin, James Thomas, Caption Health, Inc., Brisbane, CA, USA, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

Background: The FDA recently authorized artificial intelligence (AI)-guidance for echocardiography (echo). It predicts image quality (IQ) in realtime, which is displayed to the operator as acquisition guidance. We sought to determine whether IQ matched the image quality assessment ability of trained sonographers (S).

Methods: 12 subjects (6 females, 3 w/ BMI 30+, 3 s/p cardiac surgery) and 4 novice operators (RNs w/o prior echo experience) were recruited. Guided by AI, RNs scanned each subject for 10 views (total 480 acquisitions). Echo frames were continuously recorded throughout the entire exam from initial suboptimal images to diagnostic ones (mean duration: 3.6 min/acquisition) with IQ grading. From each view, 1,260 pairs of 1-heartbeat clips were sampled and shown to 5 sonographers (S), who independently and blindly chose the better clip in a given pair (total 63,000 S judgements). Each S's judgements were compared to the other 4 Ss', which were aggregated into a panel read (P) (considered the reference standard), against which the IQ judgements on the better clip were also compared. We then tested whether the agreement between the IQ and P was noninferior to S vs P.

Results: On average IQ agreed with P for 82.4±2.3% of pairs vs 82.2±3.6% for S and P, p=NS for difference, p<0.001 for noninferiority.

Conclusion: AI-adjudicated relative IQ throughout acquisition is noninferior to experts. This suggests that the AI can recognize nuances of widely varying images during scanning, capturing the human expertise.
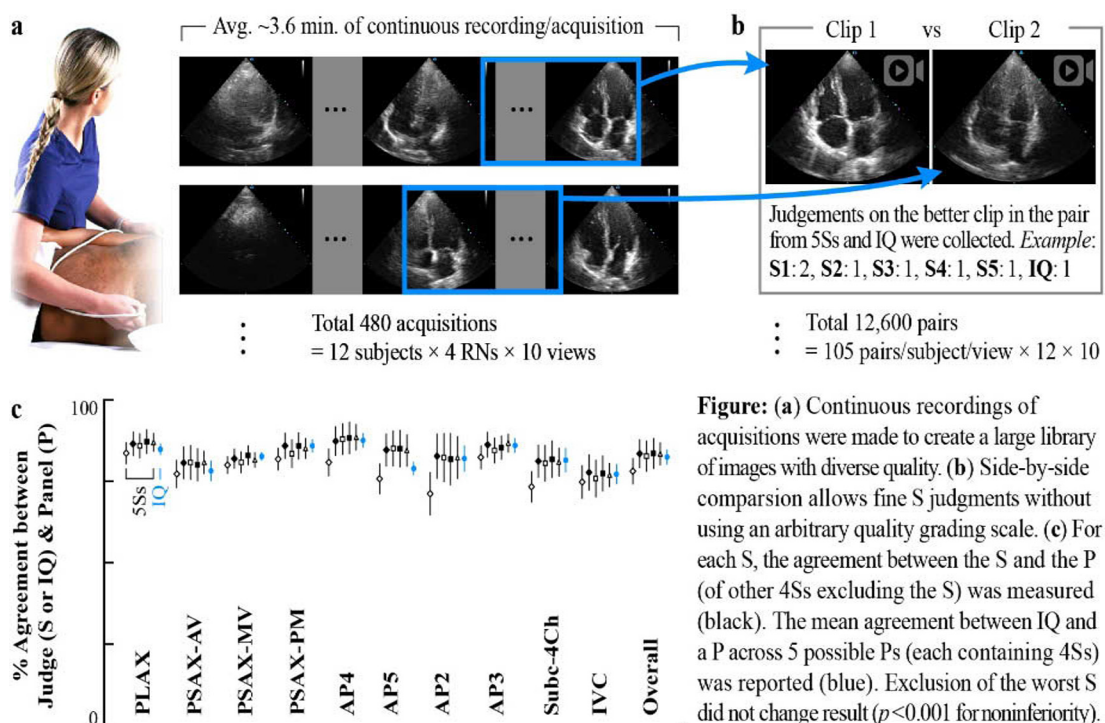
Figure: **(a)** Continuous recordings of acquisitions were made to create a large library of images with diverse quality. **(b)** Side-by-side comparison allows fine S judgments without using an arbitrary quality grading scale. **(c)** For each S, the agreement between the S and the P (of other 4Ss excluding the S) was measured (black). The mean agreement between IQ and a P across 5 possible Ps (each containing 4Ss) was reported (blue). Exclusion of the worst S did not change result (_p_<0.001 for noninferiority).