Stratifying the performance of AI models for echocardiography by image quality

Pere Lopez-Gutierrez[1,2], J. Garrido-Oliver[1,2], M. Elbatel[3], M. Boonstra[4], N. Cauwenberghs[5], A. Morales-Galan[1], JP. Betancur-Rengifo[1], S. van der Zwaard[4], Y. Qin[3], FW. Asselbergs[4], T. Kuznetsova[5], X. Li[3], J. Rodríguez-Palomares[6], L. Dux-Santoy[1], A. Guala[1]

(1) Vall d'Hebron Research Institute. Barcelona, Spain (2) Universitat Autònoma de Barcelona. Spain (3) Hong Kong University of Science and Technology, Hong Kong, China (4) Amsterdam UMC, Amsterdam, Netherlands (5) KU Leuven, Leuven, Belgium, (6) University Hospital Vall d'Hebron, Barcelona, Spain

Background

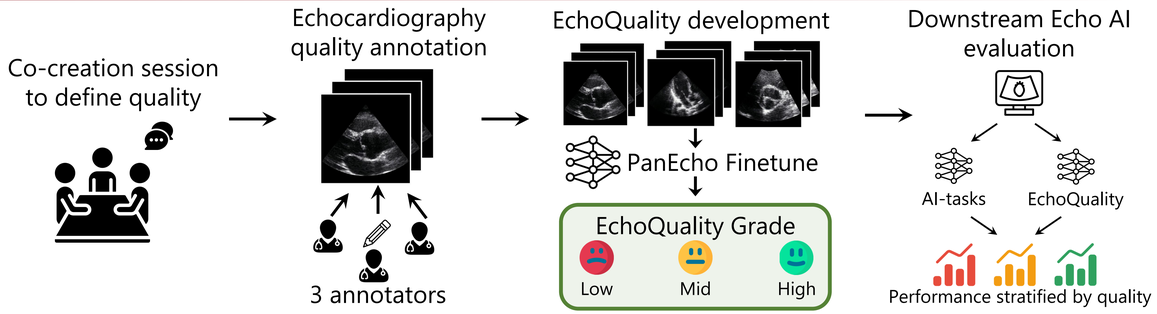
Image quality is crucial for echocardiography assessment, yet grading is...
SUBJECTIVE

Purpose

Develop a **deep learning quality assessment** model.

Understand **quality impact** in AI and human tasks.

Methodology



Model Development

Training
5,056 videos
374 studies
HUVH, Spain

Internal validation
1,418 videos
89 studies
HUVH, Spain

External validation
1,000 videos
500 studies
CAMUS, France[5]

Evaluation Cohorts

440,090 videos
14,500 studies
HUVH, Spain

279,660 videos
7,059 studies
MIMIC-IV-ECHO, USA

34,310 videos
1,004 studies
AUMC, Netherlands

112,883 videos
1,310 studies
PWH, Hong Kong

59,431 videos
1,332 studies
FLEMENGHO, Belgium

1,000 videos
500 studies
CAMUS, France

Echo AI assessments

Clinical tasks [1, 2]

12 **measurements**

5 **severity dysfunction detections**

6 **diseases classification**

Pre-processing tasks

View identification [3]

LV Segmentation [4]

Downstream Echo AI evaluation

EchoQuality Validation

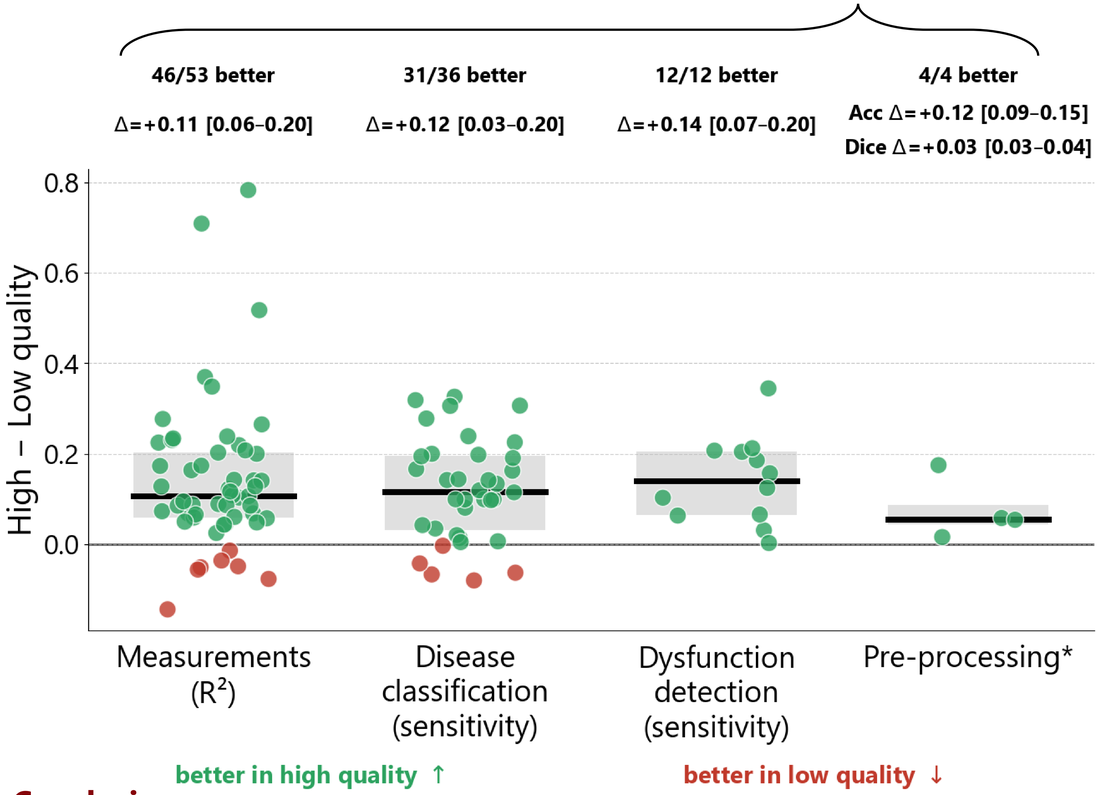
Int.
Balanced Accuracy: **63.2 %**
Inter-observer Agg.: **47.9%**

Ext.
Balanced Accuracy: **50.9 %**
Inter-observer Agg.: **21.0%**

Downstream Echo AI evaluation

High-quality studies achieved the best performance on:
93/105 (88%)
evaluations.

Scan me for detailed results!



*Accuracy is reported for View identification, and Dice Score for LV segmentation.

*A minimum of 50 measurements or 20 positives cases were needed to report performance.

Conclusions

EchoQuality provides 2D echocardiography video **quality grading.**

Image quality **impacts the performance** of **AI** models, and highlights the need for considering it in the evaluation.

1. Holste G, JAMA. 2025
2. Vukadinovic M, Nature. 2025
3. Zhang J. Circulation. 2018
4. Van de Vyver G. Sci Rep. 2025
5. Leclerc S. IEEE. 2019

Funded by the European Commission (DVPS, GA: 101213369), ISCIII (CP24/00121), Spanish Ministry of Science, Innovation and Universities (FORT23/00034), Generalitat de Catalunya (AGAUR-FI 2023 FI-1 00322 Joan Oró) and the Biomedical Research Networking Center on Cardiovascular Diseases (CIBERCV).