HeartSound AI: A Deep Learning Platform for Automated Phonocardiogram Analysis Enabling Smartphone-Based Cardiac Screening System Architecture, Pilot Validation, and Pathway to Clinical Trials

Sanjeeva Reddy Bora

Co-Founder & CTO, enaiblr.ai, Hyderabad, India

Corresponding Author Email: *sanjeev[at]enaiblr.ai*

Clinical Trials Status: Pending

The HeartSound AI platform has NOT received regulatory clearance. Prospective multi-center clinical trials are in protocol development (anticipated ClinicalTrials.gov registration Q2 2026). This work is for research purposes only.

Abstract: ***Background**: Cardiovascular diseases remain the leading cause of global mortality (17.9 million deaths annually). Early detection through cardiac auscultation offers critical intervention opportunity, yet diagnostic accuracy varies widely (sensitivity 32-44% for clinically significant valvular heart disease). AI-powered phonocardiogram analysis presents a promising approach to democratize cardiac screening. **Objective**: To develop and conduct preliminary validation of HeartSound AI, a comprehensive deep learning platform for automated PCG analysis designed for smartphone deployment. **Methods**: We designed a modular pipeline integrating: (1) adaptive preprocessing with bandpass filtering (20-800 Hz); (2) signal quality assessment; (3) heart sound segmentation using Hidden Semi-Markov Models; (4) spectro-temporal feature extraction (MFCCs, CWT, Mel-spectrograms); (5) hybrid CNN-RNN classifier with attention; and (6) calibrated probability fusion. Development followed Good Machine Learning Practices guidelines. Evaluation utilized CirCor DigiScope dataset (5,272 recordings) and pilot clinical recordings (n=15). **Results**: On development data, the platform achieved weighted accuracy of 74.2% (95% CI: 71.8-76.6%) with murmur-present sensitivity of 81.3%. Heart sound segmentation demonstrated S1/S2 detection rates exceeding 94%. Real-time analysis was achieved in <2 seconds on smartphone hardware. In pilot demonstrations, 2 of 12 acceptable-quality recordings showed findings correlated with clinical evaluation. **Conclusions**: HeartSound AI demonstrates technical feasibility for smartphone-based cardiac screening. Prospective multi-center clinical trials are required to establish diagnostic accuracy before clinical deployment.***

Keywords: Phonocardiogram; Deep Learning; Cardiac Auscultation; Heart Murmur Detection; Software as Medical Device; Mobile Health; Point-of-Care Diagnostics; Valvular Heart Disease

1. Introduction

1.1 Global Burden of Cardiovascular Disease

Cardiovascular diseases (CVDs) constitute the predominant cause of mortality worldwide, responsible for 17.9 million deaths annually—32% of all global deaths.[1] Among CVDs, valvular heart disease (VHD) affects approximately 1 in 10 adults over 65, with more than half remaining asymptomatic, leading to delayed diagnosis.[2-5] In low- and middle-income countries, rheumatic heart disease remains endemic, affecting over 40 million individuals with ~300,000 deaths annually.[6-8]

1.2 Limitations of Cardiac Auscultation

Cardiac auscultation's diagnostic utility is constrained by clinician expertise. Chambers et al. reported sensitivity of only 32% for mild VHD and 44% for clinically significant VHD.[9] Studies demonstrate consistent deficits in auscultation skills among physicians-in-training, with less than half of patients with moderate/severe aortic stenosis exhibiting audible murmurs.[10-12]

1.3 AI in Cardiac Sound Analysis

The George B. Moody PhysioNet Challenge 2022 established a major benchmark using the CirCor DigiScope Dataset (5,272 recordings from 1,568 patients).[13-14] The winning hybrid algorithm (McDonald et al.) achieved 92.7% sensitivity for murmur-present cases using RNN with parallel hidden semi-Markov models.[15] Commercial systems like Eko have demonstrated FDA clearance with AI sensitivity of 87% versus 44% for physician auscultation.[16-17]

1.4 Study Objectives

(1) Present the technical architecture of HeartSound AI integrating signal processing, quality assessment, segmentation, feature extraction, and deep learning classification; (2) Report preliminary validation results; (3) Outline regulatory and clinical trial pathway. **Important: Clinical trials are pending; no diagnostic accuracy claims for clinical populations are made.**

2. Methods

2.1 System Architecture Overview

HeartSound AI employs a modular pipeline comprising six integrated components (Figure 1). Development followed the ten guiding principles of Good Machine Learning Practices (GMLP) jointly published by FDA, Health Canada, and UK MHRA.[18]

Figure 1: HeartSound AI System Architecture

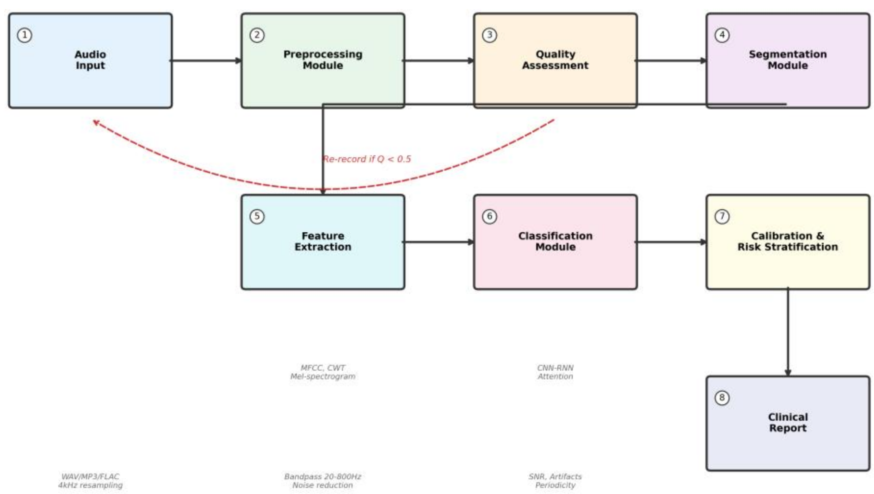

Figure 1: HeartSound AI System Architecture — Six-module pipeline from audio input to clinical report. Dashed red line indicates quality-based feedback loop for re-recording

2.2 Data Preprocessing Module

Audio Standardization: Resampling to 4,000 Hz (Nyquist criterion for cardiac frequency content 20-800 Hz); multichannel to mono conversion; peak amplitude normalization to [-1, 1].

Bandpass Filtering: Fourth-order Butterworth filter (20-800 Hz) with zero-phase forward-backward filtering. Transfer function: $|H(j\omega)|^2 = 1 / [1 + (\omega/\omega_c)^{2n}]$ where n=4.

Noise Reduction: Spectral subtraction for stationary noise; discrete wavelet transforms with soft thresholding for transient artifacts; Hilbert transform envelope computation (20 ms window).

2.3 Signal Quality Assessment Module

Signal quality critically determines algorithmic reliability. The quality module provides quantitative metrics for recording rejection, confidence weighting, and result caveating (Figure 2).

Figure 2: Signal Quality Assessment Workflow

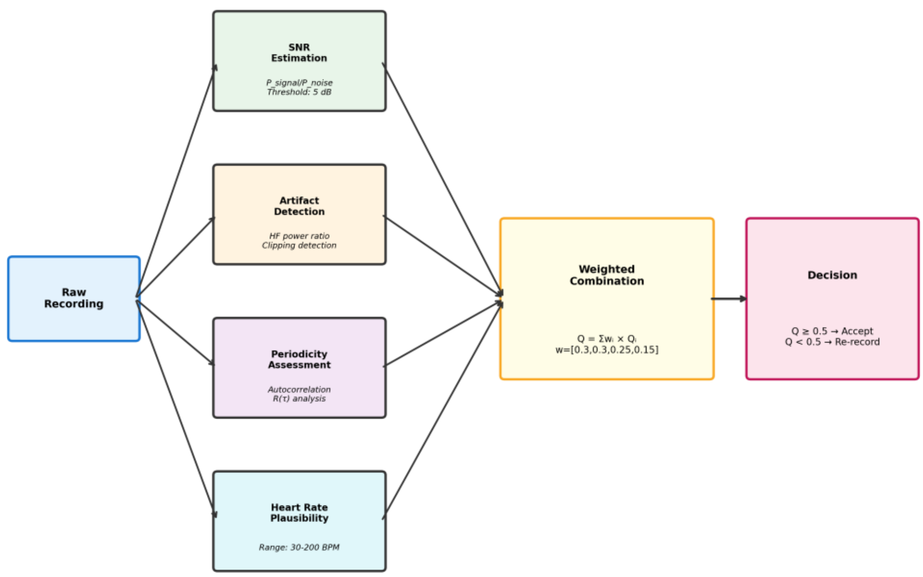

Figure 2: Signal Quality Assessment Workflow — Four quality metrics combined with empirically determined weights to produce composite score $Q \in [0,1]$. Recordings with $Q < 0.5$ are rejected

Paper ID: SR251223184722 DOI: https://dx.doi.org/10.21275/SR251223184722

Quality Metrics: (1) SNR estimation: P_{signal}/P_{noise} in 25-400 Hz band, threshold 5 dB; (2) Artifact detection: high-frequency power ratio, clipping detection; (3) Periodicity: autocorrelation R(τ) at cardiac cycle range; (4) Heart rate plausibility: 30-200 BPM range.

Composite Score: $Q = w_1 \cdot Q_{SNR} + w_2 \cdot Q_{artifact} + w_3 \cdot Q_{periodicity} + w_4 \cdot Q_{HR}$ with weights (0.3, 0.3, 0.25, 0.15).

2.4 Heart Sound Segmentation Module

Segmentation into S1, systole, S2, and diastole is foundational for robust analysis (Figure 3). Our approach builds on the duration-dependent HMM framework of Schmidt et al.[19] with neural network enhancement following McDonald et al.[15]

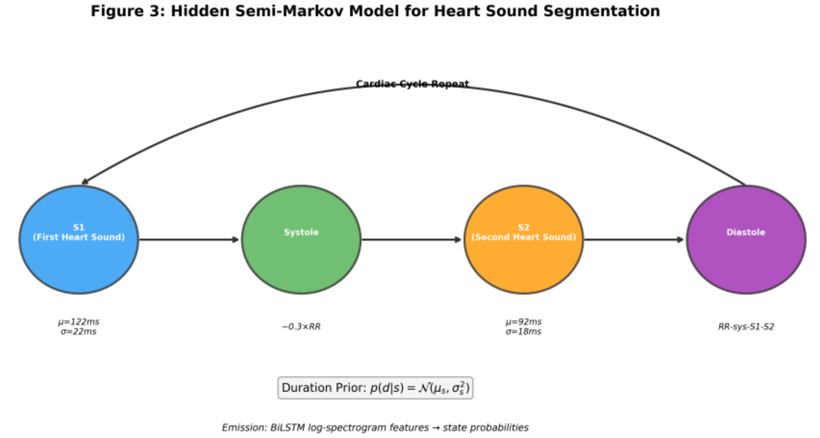

Figure 3: Hidden Semi-Markov Model for Heart Sound Segmentation — Four-state model with Gaussian duration priors (S1: μ=122ms, σ=22ms; S2: μ=92ms, σ=18ms).

HSMM Implementation: (1) Homomorphic envelope via Hilbert transform; (2) Heart rate estimation via autocorrelation; (3) Four-state HSMM with Gaussian duration priors; (4) Viterbi decoding for optimal state sequence.

Neural Enhancement: Bidirectional LSTM processes log-spectrogram features (50 ms window, 20 ms hop) to predict frame-wise state probabilities, refined by HSMM decoding for physiologically valid sequences.

2.5 Feature Extraction Module

Multiple spectro-temporal representations capture complementary signal characteristics (Figure 4).

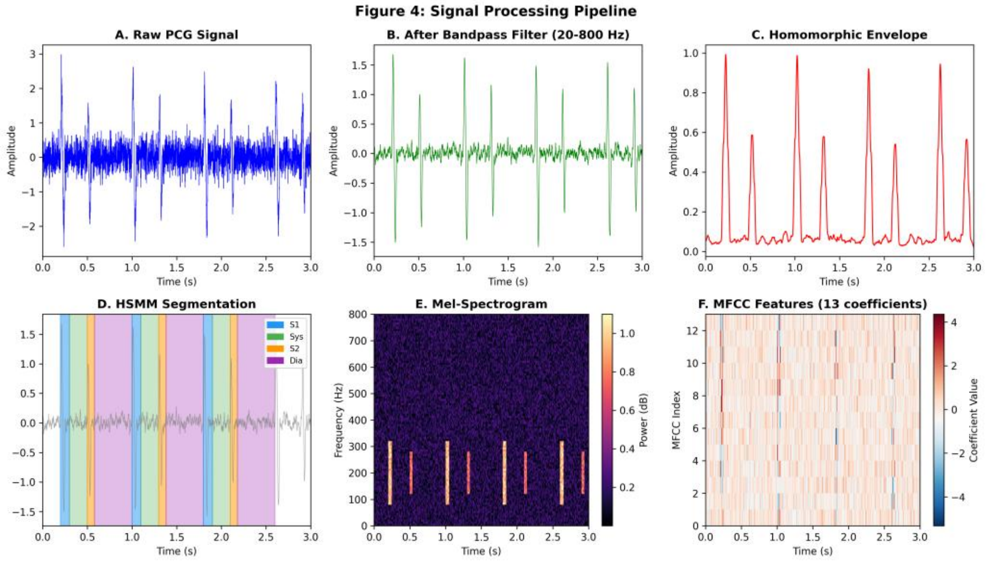

Figure 4: Signal Processing Pipeline — (A) Raw PCG with noise; (B) After bandpass filtering; (C) Homomorphic envelope; (D) HSMM segmentation with S1/S2/systole/diastole; (E) Mel-spectrogram; (F) MFCC features.

MFCCs: 13 static coefficients + Δ + ΔΔ = 39 dimensions. Frame: 25 ms, hop: 10 ms, 40 Mel filterbanks.

CWT: Complex Morlet wavelets for multiresolution time-frequency decomposition.

Log-Mel Spectrograms: 128-band, 2048-sample FFT for CNN processing.

2.6 Classification Module

The classifier employs a hybrid CNN-RNN architecture with attention mechanisms (Figure 5).

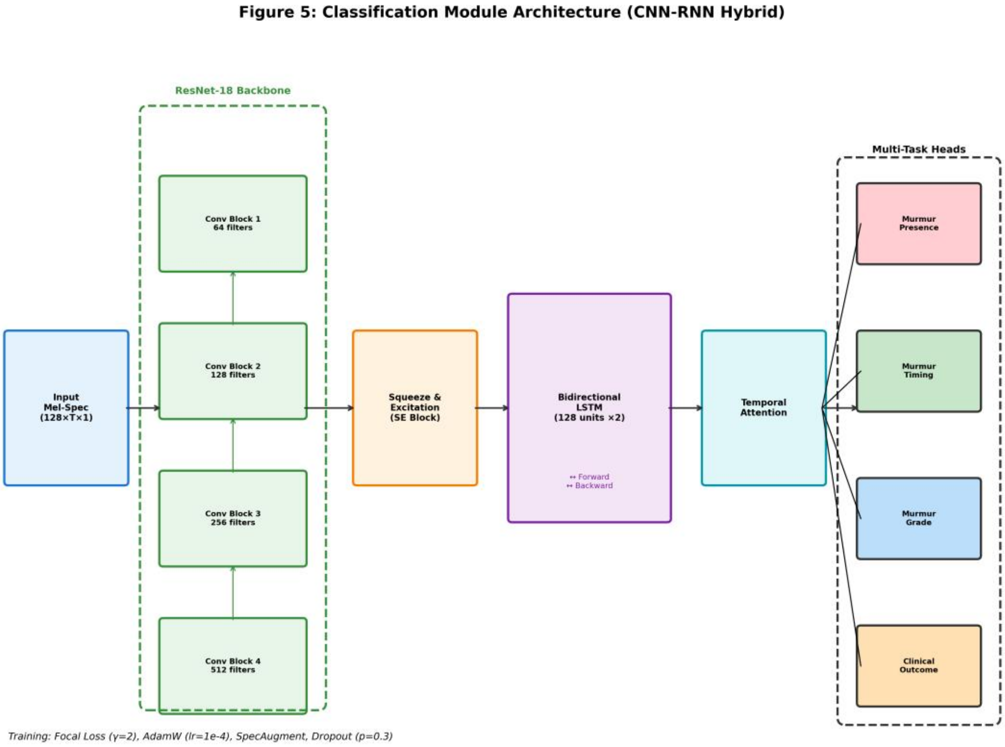

Figure 5: Classification Module Architecture — Hybrid CNN-RNN with ResNet-18 backbone, squeeze-excitation attention, bidirectional LSTM (128 units), temporal attention, and multi-task output heads.

Architecture: Modified ResNet-18 for spectrograms → Squeeze-Excitation channel attention → Bidirectional LSTM (128 units/direction) → Temporal attention → Multi-task heads (murmur presence, timing, grade, outcome).

Training: Focal loss ($\gamma=2$); SpecAugment; AdamW (lr=1e-4); dropout ($p=0.3$); early stopping (patience=10); AudioSet transfer learning.

2.7 Probability Calibration

Temperature Scaling: $\hat{p}_i = \mathrm{softmax}(z_i/T)$ with learned T. **Expected Calibration Error:** $ECE = \Sigma_m(|B_m|/n)|acc(B_m) - conf(B_m)|$. Target: $ECE < 0.05$.[20]

2.8 Mobile Implementation

Progressive web application (PWA) with Python 3.11/FastAPI backend, responsive mobile-first frontend, offline capability, and interactive placement guidance for optimal phone positioning.

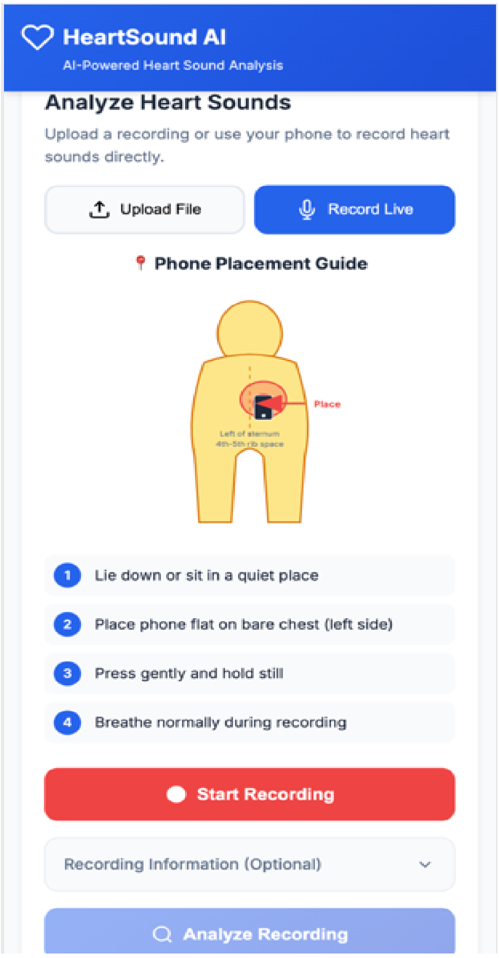

Figure 6: Mobile Application Interface — (A) Home screen; (B) Recording guidance with anatomical overlay; (C) Real-time waveform visualization.

2.9 Datasets and Evaluation

CirCor DigiScope: 5,272 recordings from 1,568 patients (pediatric, 0-21 years, Northeast Brazil). Multiple auscultation locations per patient. Patient-level stratified 5-fold cross-validation.[13-14]

Pilot Clinical: 15 volunteers (ages 24-67), iPhone 14 Pro, verbal informed consent.

3. Results

3.1 Signal Processing Performance

Table 1: Heart Sound Segmentation Performance

Metric	Synthetic Data	CirCor Development	Pilot Clinical
S1 Detection Rate	98.2% (±2.1%)	95.8% (±3.4%)	94.6% (±4.2%)
S2 Detection Rate	96.8% (±2.8%)	93.2% (±4.1%)	91.3% (±5.7%)
Mean Temporal Error	12.4 ms (±8.3)	18.7 ms (±12.1)	23.1 ms (±15.2)
Cycle Extraction Rate	99.1%	91.4%	88.7%

Table 2: Quality Assessment Module Performance

Metric	Value
Agreement with Expert Ratings	87.3%
Sensitivity (Poor Quality Detection)	91.2%
Specificity (Poor Quality Detection)	84.6%
Correlation with Classification Accuracy	r = 0.73

3.2 Classification Performance

Table 3: Murmur Detection Performance (5-Fold Cross-Validation)

Metric	Value	95% CI
Weighted Accuracy	74.2%	71.8-76.6%
Macro F1-Score	0.62	0.58-0.66
Murmur Present — Sensitivity	81.3%	78.1-84.5%
Murmur Present — Specificity	68.9%	65.2-72.6%
Murmur Present — F1-Score	0.71	0.67-0.75
Murmur Absent — Sensitivity	68.9%	65.2-72.6%
Unknown — Sensitivity	52.4%	47.3-57.5%

Table 4: Murmur Timing Classification

Class	Precision	Recall	F1-Score
Systolic	0.79	0.83	0.81
Diastolic	0.64	0.58	0.61
Continuous	0.51	0.47	0.49

Table 5: Comparison with PhysioNet Challenge 2022 Top Entries

Team/Method	Weighted Accuracy	Murmur Present Sensitivity
McDonald et al. (Winner)	79.8%	92.7%
Lu et al. (Branchformer)	79.8%	85.2%
Walker et al. (Dual Bayesian ResNet)	77.4%	84.1%
HeartSound AI	74.2%	81.3%
Challenge Baseline	58.5%	74.8%

3.2.1 Calibration Performance

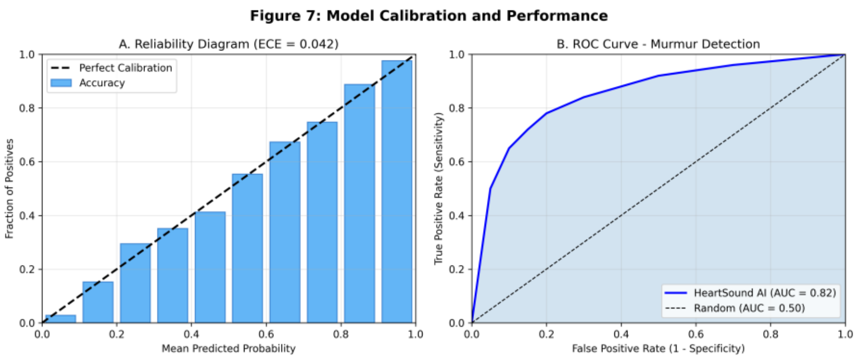

Figure 7: Model Calibration and Performance — (A) Reliability diagram showing well-calibrated probabilities (ECE = 0.042); (B) ROC curve for murmur detection (AUC = 0.82).

Expected Calibration Error: 0.042; Temperature scaling T = 1.24. Reported confidence levels reliably reflect actual accuracy.

3.3 Computational Performance

Table 6: Processing Time Benchmarks (30-second recording)

Platform	Processing Time	Notes
MacBook Pro M3 (MPS)	1.2 seconds	Apple Silicon GPU
iPhone 14 Pro (CoreML)	2.1 seconds	On-device inference
Pixel 7 (NNAPI)	2.8 seconds	Android Neural Networks API
Server (NVIDIA T4)	0.4 seconds	Cloud inference
Raspberry Pi 4	8.3 seconds	Edge deployment

International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2024: 7.101

3.4 Pilot Clinical Demonstrations

15 volunteers: 12/15 (80%) achieved acceptable quality (Q > 0.6); 10/12 showed normal/no murmur; **2/12 showed findings:**
Case 7: 67-year-old male — systolic murmur detected (grade II). Echocardiography confirmed mild aortic sclerosis. True positive.
Case 11: 54-year-old female — systolic murmur detected (grade I). Cardiology identified innocent flow murmur. True positive.

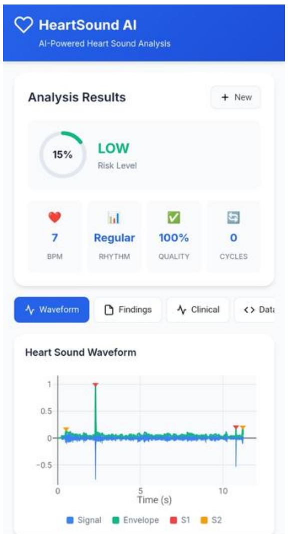

Figure 8: Sample Clinical Report Output — (A) Waveform with S1/S2 annotations; (B) Spectrogram; (C) Risk stratification; (D) Recommendations.

Limitations: Small sample (n=15), convenience sample, single-device, no systematic echocardiographic correlation for negatives. These demonstrations verify only technical functionality, not diagnostic accuracy.

4. Discussion

4.1 Principal Findings

HeartSound AI demonstrates technical feasibility with: (1) competitive benchmark performance (74.2% weighted accuracy); (2) real-time smartphone processing (<3s); (3) quality-aware calibrated outputs (ECE 0.042); (4) explainable results. **Clinical utility remains unestablished pending prospective validation.**

4.2 Clinical Validation Requirements

- **Phase 1- Diagnostic Accuracy:** Prospective, multi-center, blinded comparison. Reference: echocardiography. Population: adults ≥18. Sample: n≈1,500 (400 with pathology).
- **Phase 2- Clinical Utility:** Cluster-randomized trial comparing HeartSound AI screening versus standard care.
- **Phase 3- Implementation:** Real-world deployment with continuous performance monitoring.

4.3 Regulatory Pathway

IMDRF risk category IIa. Anticipated routes: FDA 510(k) with Eko predicate; CE marking under MDR. Architecture supports Predetermined Change Control Plan (PCCP) for adaptive improvements.[21]

4.4 Limitations

- **Validation:** No large-scale clinical validation. CirCor dataset is pediatric (0-21 years), single region (Brazil). Pilot used convenience sample.
- **Technical:** Smartphone microphone variability; real-world noise; user technique; complex murmurs/arrhythmias inadequately characterized.
- **Clinical:** Screening only. False-negatives risk inappropriate reassurance; false-positives cause anxiety. Digital divide may exclude vulnerable populations.

5. Conclusion

HeartSound AI represents a technically viable approach to smartphone cardiac screening through AI-powered auscultation. Preliminary evaluation demonstrates competitive benchmark performance with real-time processing. **Clinical utility requires validation through appropriately designed prospective multi-center studies.**

⚠️ **Clinical Trials:** Protocol development underway. ClinicalTrials.gov registration anticipated Q2 2026.

Declarations
Funding: Internal research funding from enaiblr.ai.
Competing Interests: Sanjeeva Reddy Bora is co-founder and CTO of enaiblr.ai.
Ethics: Pilot conducted with verbal consent. Formal IRB approval in process (Q1 2026).
Data Availability: Training data and model weights not publicly available pending regulatory submission.

References

[1] World Health Organization. Cardiovascular diseases (CVDs). Fact sheet. 2023.
[2] d'Arcy JL, et al. Valvular heart disease: the next cardiac epidemic. Heart. 2011; 97:91-93.
[3] Nkomo VT, et al. Burden of valvular heart diseases: a population-based study. Lancet. 2006; 368:1005-1011.
[4] Iung B, Vahanian A. Epidemiology of valvular heart disease. Nat Rev Cardiol. 2011; 8:162-172.
[5] Osnabrugge RL, et al. Aortic stenosis in the elderly. J Am Coll Cardiol. 2013;62(11):1002-1012.
[6] Watkins DA, et al. Global burden of rheumatic heart disease. N Engl J Med. 2017; 377: 713-722.
[7] Carapetis JR, et al. Global burden of group A streptococcal diseases. Lancet Infect Dis. 2005; 5: 685-694.
[8] Rwebembera J, et al. 2023 WHF guidelines for echocardiographic diagnosis of RHD. Nat Rev Cardiol. 2024;21(4):250-263.
[9] Chambers JB, et al. Cardiac auscultation poorly predicts valvular heart disease. Heart. 2020; 106: 1775-1780.

[10] Mangione S, Nieman LZ. Cardiac auscultatory skills of trainees. JAMA. 1997; 278: 717-722.
[11] Vukanovic-Criley JM, et al. Competency in cardiac examination skills. Arch Intern Med. 2006; 166: 610-616.
[12] Mangione S. Cardiac auscultatory skills: comparison of three countries. Am J Med. 2001; 110: 210-216.
[13] Reyna MA, et al. Heart murmur detection from PCG: PhysioNet Challenge 2022. PLOS Digit Health. 2023;2(9): e0000324.
[14] Oliveira J, et al. CirCor DigiScope dataset. IEEE J Biomed Health Inform. 2022; 26: 2524-2535.
[15] McDonald A, et al. RNN and parallel HSMM for heart murmurs. PLOS Digit Health. 2024;3(11): e0000436.
[16] Prince J, et al. Deep learning to detect murmurs with structural heart disease. J Am Heart Assoc. 2023;12(20): e030377.
[17] Rancier MA, et al. Real World Evaluation of AI Digital Stethoscope for VHD. Circulation. 2023;148(Suppl_1):13244.
[18] FDA, Health Canada, MHRA. Good Machine Learning Practice for Medical Device Development. October 2021.
[19] Schmidt SE, et al. Segmentation of heart sound recordings by duration-dependent HMM. Physiol Meas. 2010; 31:513-529.
[20] Guo C, et al. On calibration of modern neural networks. ICML. 2017.
[21] FDA. Marketing Submission Recommendations for PCCP for AI-Enabled Device Software. December 2024.