Bidirectional Fiducial Matching of Electrocardiography and Phonocardiography for Multimodal Signal Quality Assessment

Daniel Proaño-Guevara[1,2,3], André Lobo[4], Cristina Oliveira[5], Cátia Isabel Costa[4], Ricardo Fontes-Carvalho[5], Hugo Plácido da Silva[2,6], Francesco Renna[1,7]

[1] INESC TEC, Porto, Portugal
[2] Instituto de Telecomunicações, Lisboa, Portugal
[3] Faculdade de Engenharia da Universidade do Porto, Porto, Portugal
[4] Unidade Local de Saúde de Gaia e Espinho (ULSGE), Porto, Portugal
[5] Unidade de Investigação e Desenvolvimento Cardiovascular (UnIC@RISE), Porto, Portugal
[6] Instituto Superior Técnico da Universidade de Lisboa, Lisboa, Portugal
[7] Faculdade de Ciências da Universidade do Porto, Porto, Portugal

Abstract

We introduce a multimodal Signal Quality Indicator (SQI) for assessing fidelity of synchronous electrocardiogram (ECG) and phonocardiogram (PCG) signals recorded in ambulatory, non-standardized settings. The method uses a bidirectional fiducial-matching algorithm to test the temporal alignment of QRS complexes and T waves (ECG) with S1 and S2 sounds (PCG) respectively. Validation employed 564 synchronous ECG–PCG pairs collected with the FDA-cleared Rijuven Cardiosleeve at the aortic, pulmonary, tricuspid, and mitral valves sites. Expert annotations served as ground truth. In a three-class task, the SQI reached an area under the ROC curve greater than 79%, showing strong discriminative power. This physiology-based metric supports batch-online monitoring and reliable quality control of opportunistic cardiac recordings.

1. Introduction

Cardiovascular diseases (CVDs) are the leading cause of death globally, accounting for an estimated 17.9 million deaths in 2019 [1]. Early detection remains critical, and effective monitoring depends on the availability of accurate and interpretable physiological signals.

Electrocardiography (ECG) and phonocardiography (PCG) offer complementary views of cardiac function. ECG captures the heart's electrical activity, while PCG reflects mechanical events. When combined, these modalities provide a more comprehensive representation of cardiac activity [2]. However, signal quality is often compromised in real-world settings, particularly when data are acquired using multimodal stethoscopes. These devices prioritize PCG for auscultation, with ECG captured opportunistically, leading to increased susceptibility to noise, motion artifacts, and non-standard sensor placement [3].

Signal quality analysis (SQA) seeks to assess the reliability of acquired signals before diagnostic interpretation or algorithmic processing. Most existing SQA methods are unimodal, evaluating ECG and PCG independently or using one high-fidelity signal to assess the other [4, 5]. Furthermore, many rely on standard lead placements and are unsuitable for opportunistic acquisition environments [6]. PCG-specific SQA remains underdeveloped, often relying on statistical heuristics with limited interpretability [7]. Currently, no standardized, physiologically grounded framework exists for assessing multimodal cardiac signal quality, particularly for signals recorded with non-standard lead configurations and outside controlled clinical environments.

This study addresses the literature gap by proposing a novel multimodal Signal Quality Indicator (SQI) based on bidirectional matching of ECG and PCG fiducial events—specifically, QRS complexes and T waves in ECG, and S1 and S2 sounds in PCG. The method relies on energy-based segmentation, which improves robustness to noise and supports applications in opportunistic settings. Unlike previous approaches, our method does not assume one modality as the reference. Instead, it performs quality assessment in both ECG-to-PCG (E2P) and PCG-to-ECG (P2E) directions, verifying physiologically expected alignments (e.g., QRS with S1, T with S2) within defined temporal windows. The resulting SQI variants are validated against expert manual annotations and show strong correlation with human-assigned quality levels.

2. Materials

Data were collected in an ambulatory auscultation setting in collaboration with the Cardiology Department of the *Unidade Local de Saúde Gaia e Espinho* (ULSGE), using the FDA-approved Cardiosleeve device (Rijuven). Cardiosleeve integrates synchronous ECG and PCG acquisition via a conventional stethoscope. PCG signals were sampled at 3 kHz with 16-bit resolution, and ECG at 500 Hz with 12-bit resolution.

Recordings were acquired from 190 participants (167 valid), aged 2–93 years (60% male), at four auscultation points: aortic, pulmonary, tricuspid, and mitral valves. Each session lasted approximately 30 seconds per point, with ECG recorded opportunistically. Multiple sessions were collected for some participants.

A total of 564 ECG–PCG signal pairs were manually annotated by one external reviewer using a 6-point unified quality scale, adapted from the International Telecommunication Union (ITU) [8] grading system (Table 1). A multimodal score was derived as the minimum of both modalities. Table 2 summarizes the distribution of quality scores.

Table 1. Manual SQA scoring system. Indicates the minimum features each signal must show to be labeled with that score. Each * indicates the impairness severity.

Score	0	1	2	3	4	5
ECG	None	R only	QRS	QRS+T**	QRS+T*	QRS+T
PCG	None	S1 or S2*	S1 or S2	S1+S2**	S1+S2*	S1+S2

Table 2. Quality annotation distribution.

Score	0	1	2	3	4	5
ECG	121	125	59	116	89	54
PCG	141	152	82	109	58	22
Multimodal	218	155	72	85	29	5

3. Methods

We employed a U-Net-based segmentation architecture, following the methodology described in [9], to extract fiducial events from ECG and PCG signals. While the architecture was shared, separate training weights and label sets were used for each modality.

Recordings from the ULSGE dataset were preprocessed before segmentation. Due to differences in acquisition resolution, ECG and PCG signals were first standardized using z-score normalization. ECG preprocessing included zero-phase filtering with a sixth-order Butterworth bandpass filter (0.5–150 Hz), 50 Hz IIR notch filtering, and detrending via median subtraction. Given the non-standard ECG lead configuration, features were derived from signal energy rather than morphology [3], and downsampled to 50 Hz. Extracted ECG features included the Hilbert, Shannon, homomorphic (median window = 21 samples), and smoothed (Hamming window = 21 samples) energy envelopes.

PCG signals were downsampled from 3 kHz to 1 kHz using a polyphase antialias FIR filter. The despiking method in [10] was applied, followed by zero-phase Butterworth bandpass filtering (15–450 Hz). PCG features were computed using the same energy-based approach, and also included two continuous wavelet scalograms (20–200 Hz), using Morlet and Mexican Hat wavelets.

ECG segmentation was trained using the LUDB dataset [11], while the PhysioNet 2016 dataset [12] was used for PCG. Feature vectors were segmented into 64-sample patches (1.28 s) with a stride of 8 samples. Labels followed standard event sequences: 'baseline', 'P-wave', 'QRS-complex', and 'T-wave' for ECG; and 'S1', 'systole', 'S2', and 'diastole' for PCG.

We developed a bidirectional multimodal signal quality scoring algorithm based on the alignment of fiducial intervals. Each QRS complex and T wave was expected to precede an S1 and S2, respectively, within a tolerance window. The SQI quantifies the number of matches across modalities, using ECG-to-PCG (E2P) and PCG-to-ECG (P2E) directions.

A penalized recall formulation, presented in Equation (1) was used to define the SQI.

$$\text{Score} = \frac{M - \lambda(N - M)}{N} = (1 + \lambda)\frac{M}{N} - \lambda \quad (1)$$

where N is the number of expected events, M is the number of correctly matched events, and $\lambda = 0.1$ is the mismatch penalty. The selected λ value is the result of a grid-test with the values 0.1, 0.5, and 1.0, where 0.1 showed the greatest correlation with the manual annotations.

Events were represented as intervals, extended by 8 samples (160 ms). An overlap was considered valid if the intersection of reference and test intervals was non-zero. Figure 1 illustrates the evaluation of ECG quality using PCG segmentation as the reference. In this example, several detected S1 and S2 sounds lack corresponding QRS complexes or T waves, indicating a lower signal fidelity in the PCG channel.

We implemented two scoring strategies:

1. **Total signal:** Fiducials from the same modality (e.g., QRS + T wave) were pooled and matched jointly.
2. **Individual events:** Matching was computed separately per fiducial type (QRS, T, S1, S2), and the minimum used as a modality-wise score.

Bidirectional scores were computed for both E2P and P2E. These were combined using two methods:

- **avg_lin:** average of total-signal E2P and P2E
- **min_lin:** minimum of total-signal E2P and P2E
- **avg_min:** average of individual-event E2P and P2E

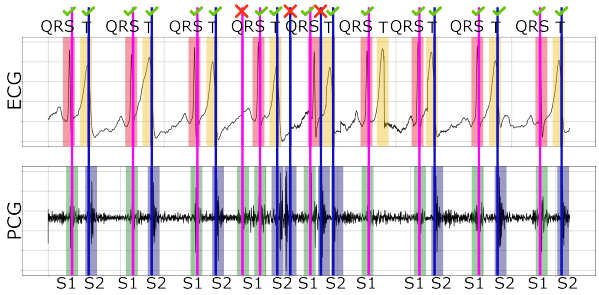


Figure 1. PCG to ECG multimodal quality assessment.

- **min_min:** minimum of individual-event E2P and P2E

We evaluated the correlation between each SQI variant and the manual annotation vector. The correlation metrics included:

- Pearson's ρ_P: linear association
- Spearman's ρ_S: monotonic trends
- Kendall's τ: ordinal agreement

Significance was assessed using two-tailed tests at $\alpha = 0.05$. Pairwise Welch's t-tests with Bonferroni correction were applied to test group separability.

To reduce class imbalance and improve interpretability, we quantized the annotations into a 3-class scale: *low quality* (0–1), *uncertain* (2–3), and *high quality* (4–5). SQIs were used as input to multinomial logistic regression with 5-fold cross-validation and class-weighted loss. Predictive performance was evaluated using the area under the ROC curve (AUC) under a one-vs-rest scheme. Confusion matrices were computed to assess class-level accuracy.

Feature relevance was assessed using permutation importance, calculated as the average decrease in AUC across shuffled inputs, using the Gini index as a scoring metric.

4. Results

We evaluated the statistical association between the proposed multimodal SQIs and manual annotations. Table 3 reports the Pearson, Spearman, and Kendall correlation coefficients computed between the manual labels and each SQI variant.

Table 3. Correlation coefficients between SQI variants and manual annotations.

	min_lin	avg_lin	min_min	avg_min
Pearson's ρ_P	0.21	0.19	0.19	0.18
Spearman's ρ_S	0.24	0.22	0.22	0.21
Kendall's τ	0.19	0.18	0.18	0.17

Among all variants, `min_lin` exhibited the highest correlation with manual labels, always with $p < 10^{-50}$.

The resulting `min_lin` scores spanned across the low (0.44), uncertain (0.78), and high (0.90) quality classes. Figure 2 presents the violin plot of the `min_lin` score.

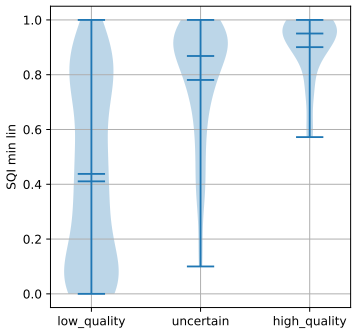


Figure 2. Distribution of `min_lin` SQI across quantized quality levels.

Statistical analysis using the Kruskal–Wallis test confirmed significant group separation ($H = 153.72$, $p = 4.17 \times 10^{-34}$), with all three pairwise Welch's t-tests remaining significant after Bonferroni correction.

To assess predictive utility, we trained multinomial logistic regression models using each SQI variant to predict the labels "low quality", "uncertain", and "high quality". The confusion matrix for this task is shown in Figure 3. Among all features, `min_lin` exhibited the highest permutation importance (0.1122 ± 0.0178), and achieved an AUC of 0.797 when used as the sole predictor. In contrast, other variants showed negative importance values.

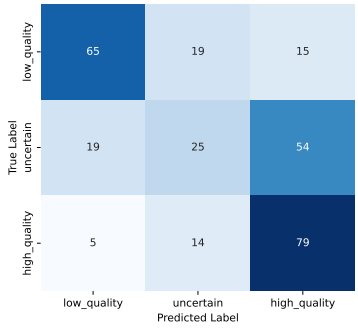


Figure 3. Confusion matrix for 3-class prediction using SQI variants. (% per true label.)

These results demonstrate that `min_lin` is the most effective and interpretable variant, showing strong agreement with manual labels and robust classification performance. It provides a reliable basis for multimodal cardiac signal quality assessment in real-world acquisition settings.

5. Discussion and Conclusion

This study presents a multimodal Signal Quality Indicator that scores synchronous ECG and PCG recordings

via bidirectional fiducial matching. The indicator quantifies fidelity by aligning electrical landmarks(the QRS complex and T wave) with mechanical landmarks (S1 and S2), and evaluates both ECG-to-PCG (E2P) and PCG-to-ECG (P2E) directions, so neither modality is assumed as the baseline.

Of four scoring schemes, the minimum-linear variant (`min_lin`), derived from a penalized-recall metric applied to the entire trace, showed the strongest link to expert labels ($p < 10^{-50}$).

On the three-level quality scale, Kruskal–Wallis and pairwise Welch t-tests confirmed significant group separability. A multinomial logistic-regression model that used `min_lin` achieved an AUC of 0.797, underscoring the metric's predictive utility.

Figures 2 and 3 illustrate performance: the model cleanly separates low- and high-quality segments but tends to classify uncertain cases as high quality. This bias arises because the segmentation algorithm detects fiducials more reliably than human reviewers, who may miss subtle cues in noisy signals.

Limitations remain. The SQI has not been tested for fine-grained windows, continuous scoring, or single-beat rejection, and it has not been benchmarked against ensembles of unimodal indicators. Future work will compare multimodal and optimized unimodal metrics and explore real-time windowing.

To our knowledge, this is the first SQI to assess ECG and PCG jointly without a predefined reference signal. Its transparency and batched-online capability suit wearable and telemedicine platforms, where data quality is critical for downstream AI and clinical decision support.

Acknowledgments

This work is co-financed by Component 5 - Capitalization and Business Innovation, integrated in the Resilience Dimension of the Recovery and Resilience Plan within the scope of the Recovery and Resilience Mechanism (MRR) of the European Union (EU), framed in the Next Generation EU, for the period 2021 - 2026, within the project HfPT, with reference 41. This work is also co-financed by FCT/MECI through national funds and when applicable co-funded EU funds under UID/50008: Instituto de Telecomunicações. The work of Daniel Proaño-Guevara was supported by the individual FCT - Fundação para a Ciência e a Tecnologia Grant under Reference 2024.01213.BD.

References

[1] World Health Organization. Cardiovascular diseases (cvds), 2021. Accessed: 2025-04-09.

[2] Monteiro SM, Silva HPd. A novel approach to simultaneous phonocardiography and electrocardiography during auscultation. IEEE Access 2023;11:78224–78236.

[3] Guo L, Pressman GS, Kieu SN, Marrus SB, Mathew G, Prince J, Lastowski E, McDonough RV, Currie C, Maidens JN, Al-Sudani H, Friend E, Padmanabhan D, Kumar P, Kersh E, Venkatraman S, Qamruddin S. Automated detection of reduced ejection fraction using an ecg-enabled digital stethoscope: A large cohort validation. JACC Advances 2025;4(3):101619.

[4] Kuetche F, Alexendre N, Pascal NE, Colince W, Thierry S. Signal quality indices evaluation for robust ecg signal quality assessment systems. Biomedical Physics Engineering Express August 2023;055016.

[5] Macfarlane PW, Kennedy J. Automated ecg interpretation—a brief history from high expectations to deepest networks. Hearts September 2021;2:433–448.

[6] Lueken M, Gramlich M, Leonhardt S, Marx N, Zink MD. Automated signal quality assessment of single-lead ecg recordings for early detection of silent atrial fibrillation. Sensors 2023;23(12).

[7] Springer DB, Brennan T, Ntusi N, Abdelrahman HY, Zühlke LJ, Mayosi BM, Tarassenko L, Clifford GD. Automated signal quality assessment of mobile phone-recorded heart sound signals. Journal of Medical Engineering Technology September 2016;40(7–8):342–355.

[8] ITU - International Telecommunication Union. General methods for the subjective assessment of sound quality, 2019.

[9] Renna F, Oliveira J, Coimbra MT. Deep convolutional neural networks for heart sound segmentation. IEEE Journal of Biomedical and Health Informatics November 2019; 23(6):2435–2445.

[10] Schmidt SE, Holst-Hansen C, Graff C, Toft E, Struijk JJ. Segmentation of heart sound recordings by a duration-dependent hidden markov model. Physiological Measurement March 2010;31(4):513–529.

[11] Kalyakulina AI, Yusipov II, Moskalenko VA, Nikolskiy AV, Kosonogov KA, Osipov GV, Zolotykh NY, Ivanchenko MV. Ludb: A new open-access validation tool for electrocardiogram delineation algorithms. IEEE Access 2020;8:186181–186190.

[12] Liu C, Springer D, Li Q, Moody B, Juan RA, Chorro FJ, Castells F, Roig JM, Silva I, Johnson AEW, Syed Z, Schmidt SE, Papadaniil CD, Hadjileontiadis L, Naseri H, Moukadem A, Dieterlen A, Brandt C, Tang H, Samieinasab M, Samieinasab MR, Sameni R, Mark RG, Clifford GD. An open access database for the evaluation of heart sound algorithms. Physiological Measurement November 2016; 37(12):2181–2213.

Address for correspondence:

Daniel Proaño-Guevara
INESC TEC, Faculdade de Ciencias da Universidade do Porto,
Rua Campo Alegre 1055, 4169-007, Porto, Portugal.
daniel.proano@inesctec.pt