

Article

An Automated ECG-PCG Coupling Analysis System with LLM-Assisted Semantic Reporting for Community and Home-Based Cardiac Monitoring

Yi Tang ¹ , Fei Cong ², Yi Li ² and Ping Shi ^{1,*}

¹ School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China; tangyi_0719@163.com

² Sino-German College, University of Shanghai for Science and Technology, Shanghai 200093, China; congfeijj@163.com (F.C.); ly499474970@163.com (Y.L.)

* Correspondence: pshi@usst.edu.cn

Abstract

Objective: Cardiac monitoring in community and home environments requires automated operation, cross-state robustness, and interpretable feedback under resource-constrained and uncontrolled conditions. Unlike accuracy-driven ECG–PCG studies focusing on diagnostic performance, this work emphasizes systematic modeling of cardiac electromechanical coupling for long-term monitoring and engineering feasibility validation. **Methods:** An automated ECG–PCG coupling analysis and semantic reporting framework is proposed, covering signal preprocessing, event detection and calibration, multimodal coupling feature construction, and rule-constrained LLM-assisted interpretation. Electrical events from ECG are used as global temporal references, while multi-stage consistency correction mechanisms are introduced to enhance the stability of mechanical event localization under noise and motion interference. A structured electromechanical feature set is constructed to support fully automated processing. **Results:** Experimental results demonstrate that the proposed system maintains coherent event sequences and stable coupling parameter extraction across resting, movement, and emotional stress conditions. The incorporated LLM module integrates precomputed multimodal metrics under strict constraints, improving report readability and consistency without performing autonomous medical interpretation. **Conclusions:** This study demonstrates the methodological feasibility of an ECG–PCG coupling analysis framework for long-term cardiac state monitoring in low-resource environments. By integrating end-to-end automation, electromechanical coupling features, and constrained semantic reporting, the proposed system provides an engineering-oriented reference for continuous cardiac monitoring in community and home settings rather than a clinical diagnostic solution.



Academic Editors: Maryam Ravan and Iman Dehzangi

Received: 5 December 2025

Revised: 27 January 2026

Accepted: 29 January 2026

Published: 2 February 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: ECG–PCG coupling analysis; automated cardiac monitoring; large language model interpretation

1. Introduction

With the acceleration of global population aging and the increasing burden of chronic diseases, the demand for long-term and continuous physiological monitoring in community and home environments has become increasingly prominent [1,2]. In real-world care scenarios, limitations in cognitive function, communication ability, and caregiving resources often hinder timely and accurate reporting of physiological discomfort. Traditional care models

relying on manual observation and subjective judgment are insufficient for continuous and standardized monitoring, potentially amplifying health risks due to delayed responses [3,4]. Driven by the concept of “Active Health,” portable and non-invasive monitoring technologies with automated analysis capabilities that can operate under resource-constrained conditions are increasingly required in primary and community healthcare systems [5].

Among various vital signs, continuous monitoring of cardiac activity plays a central role. Cardiovascular diseases remain one of the leading causes of mortality worldwide [6], and their progression is often accompanied by changes in the coordination between cardiac electrical and mechanical activities [7]. Accordingly, in non-clinical environments, cardiac monitoring systems are expected to provide stable operation, automated analysis, and interpretable feedback, with an emphasis on continuous physiological state perception rather than direct diagnostic decision-making.

The electrocardiogram (ECG) reflects cardiac electrical activity [8,9], while the phonocardiogram (PCG) captures valve motion and related hemodynamic characteristics. Synchronous ECG–PCG acquisition enables comprehensive characterization of cardiac electromechanical processes, allowing quantitative assessment of key mechanical phases such as the pre-ejection period, systole, and diastole. Compared with single-modality analysis, ECG–PCG fusion provides improved information completeness for cardiac state monitoring [10]. However, most existing ECG–PCG studies [11–15] focus on disease classification or diagnostic performance enhancement, whereas systematic frameworks addressing long-term monitoring under nonstationary and motion-corrupted conditions remain limited. In contrast to accuracy-driven fusion paradigms, this study adopts a stability-oriented algorithmic design, emphasizing physiologically consistent coupling and temporal robustness rather than optimizing standalone detection or classification performance.

Distinct from accuracy-driven diagnostic paradigms, the technical contribution of this study lies not in improving the precision of individual detection modules, but in establishing an ECG–PCG electromechanical coupling feature construction paradigm tailored for long-term monitoring. By incorporating rule-driven event consistency constraints and multi-stage correction mechanisms, the proposed approach preserves the temporal logic and cross-state stability of electromechanical events under nonstationary and noisy conditions.

Based on this perspective, we propose an automated ECG–PCG coupling analysis and semantic reporting framework designed for community and home environments under resource constraints. The primary objective is to validate its engineering feasibility and stability rather than clinical diagnostic performance. A structured multimodal feature system is constructed to support a fully automated pipeline, including signal preprocessing, event localization, ECG-guided mechanical event correction, and electromechanical parameter extraction. A large language model (LLM) is further introduced as a strictly constrained semantic interface to transform rule-based structured features into standardized and readable textual reports. This hybrid architecture—combining rule-driven analysis with LLM-assisted expression—provides a deployable and physiologically interpretable ECG–PCG coupling analysis paradigm.

2. Materials and Methods

2.1. Dataset Introduction

This study employs two types of synchronized ECG-PCG datasets—publicly available and self-built—to construct and validate the proposed ECG-PCG coupling analysis system. Data sources include: the publicly accessible EPHNOGRAM database and a self-built multi-state physiological signal database. Public data were used for model development and algorithm performance comparison, while self-built data were used to assess the system’s

robustness and transferability under dynamic physiological and emotional variations. Both datasets comprise synchronously acquired signals, enabling temporal coupling analysis and feature modeling of electro-mechanical activities.

(1) EPHNOGRAM Database

The EPHNOGRAM database was recorded using a portable, low-power acquisition system (hardware version 2.1) developed by an international research team [16], designed to provide high-quality synchronized ECG–PCG data to support cardiac electro-mechanical coupling research. This system simultaneously records three-lead ECG and single-channel PCG. The auscultation position was selected between the tricuspid and mitral valve auscultation areas to simultaneously obtain clear first and second heart sounds.

The database contains ECG-PCG synchronized recordings from 24 healthy adults (aged 23–29 years, mean 25.4 ± 1.9 years). All signals were recorded at an 8 kHz sampling rate with 12-bit quantization precision (approximately 10.5 effective bits). Experimental tasks encompassed multiple typical physiological activity states, including resting, walking, running, and cycling. This database provides a reliable baseline data resource for multimodal cardiac signal research, offering crucial support for algorithm performance evaluation and validation of noise-reduction methods.

(2) Self-Built ECG–PCG Dataset with Physiological Perturbations

To evaluate the robustness of the proposed ECG–PCG coupling analysis algorithm under physiological perturbations and nonstationary conditions, a self-built synchronous ECG–PCG dataset was constructed. The purpose of this dataset was to introduce controlled variations in heart rate, respiratory modulation, and motion-related noise, thereby systematically evaluating algorithmic stability under distribution shifts, rather than to represent specific populations, age groups, or clinical disease conditions, nor to support diagnostic or prognostic clinical conclusions.

Data acquisition was conducted using a PowerLab/16sp system (AD Instruments, Dunedin, New Zealand). Five healthy young volunteers aged 22–26 years were recruited, and all experiments were performed in a quiet environment. The study protocol was approved by the Institutional Review Board of the University of Shanghai for Science and Technology (IRB-AF98-V1.0). Written informed consent was obtained from all participants prior to data acquisition, and all procedures were conducted in accordance with the principles of the Declaration of Helsinki. ECG signals were recorded using a standard Lead I configuration (left arm positive, right arm negative, right leg ground). PCG signals were collected using an MLT209 electronic stethoscope (AD Instruments, Dunedin, New Zealand) positioned between the tricuspid and mitral auscultation areas. Both ECG and PCG signals were sampled at 8 kHz to ensure compatibility with publicly available datasets.

To introduce different types of physiological perturbations, each participant completed three experimental tasks: approximately 5 min of resting recording, 5 min of post-exercise recovery, and approximately 10 min of emotion-induction tasks. The post-exercise recovery task was designed to induce heart rate acceleration, increased respiratory activity, and chest-wall vibration-related noise [17–19]. The emotion-induction task, based on standardized video stimuli, elicited mild and transient emotional responses (e.g., stress or positive affect), resulting in nonstationary heart rate dynamics and occasional motion artifacts [20–22]. These conditions were intentionally incorporated to evaluate algorithmic robustness under non-ideal and dynamically varying signal conditions.

2.2. Overall System Framework

This study proposes an automated ECG-PCG coupling analysis system tailored for community and home settings, aiming to achieve a complete closed-loop process from

multimodal physiological signal acquisition to semantic interpretation of cardiac function. The system adopts a modular pipeline design, sequentially performing signal preprocessing, event detection and calibration, multimodal feature construction, and LLM-driven intelligent report generation.

In its overall design, the system adheres to three core principles:

- (1) Algorithmic robustness: Improving reliability under noisy and motion-intensive conditions through multi-stage processing and redundant verification strategies;
- (2) Cross-modal consistency: Fully leveraging the temporal correlation between ECG and PCG signals, using electrical events as a global reference for mechanical event detection;
- (3) Intelligent interpretability: Achieving a natural transition from quantitative analysis to semantic interpretation through a structured feature system and large language models.

The overall system architecture is shown in Figure 1. The proposed system establishes an end-to-end processing pipeline from low-level signal acquisition to high-level semantic representation, with an architectural design that explicitly considers scalability, deployability, and physiological interpretability at the system and implementation levels. Rather than claiming validated real-world deployment, the proposed framework is intended as a methodologically feasible and engineering-oriented technical reference for cardiac physiological state monitoring under resource-constrained and uncontrolled conditions, such as those commonly encountered in community and home scenarios. Further details are presented in Sections 2.3–2.7.

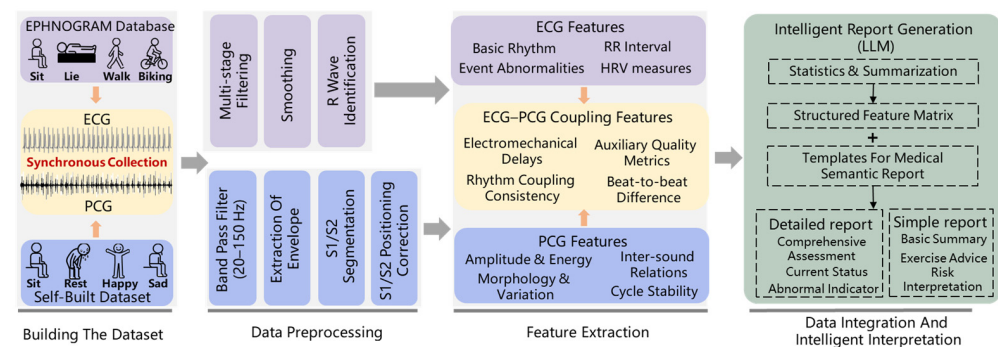


Figure 1. Overall System Framework Diagram.

2.3. ECG Signal Processing

The ECG signal reflects the electrical activity of the heart during depolarization and repolarization. The R wave, as the most prominent feature within the QRS complex, serves as the primary indicator of ventricular depolarization. Accurate detection of the R wave is critical for heart rate analysis, heart rate variability (HRV) calculation, and multimodal signal synchronization. To achieve robust R-wave identification, this study developed a processing workflow incorporating multi-stage filtering, feature enhancement, and adaptive threshold estimation.

First, the raw ECG signal undergoes preprocessing to suppress noise. Specifically, a 0.5 Hz high-pass filter is first applied to eliminate baseline drift, followed by a 49–51 Hz notch filter to suppress power-line interference, and finally a 40 Hz low-pass filter to remove high-frequency electromyographic noise. Subsequently, amplitude normalization and 20-millisecond moving-average smoothing are applied to standardize the signal's dynamic range and smooth waveform transitions, laying the groundwork for subsequent feature enhancement.

During the R-wave detection phase, this study established a multi-stage recognition framework based on the principles of the Pan–Tompkins algorithm, comprising “frequency domain filtering—feature enhancement—adaptive threshold detection.” First, the preprocessed signal undergoes band-pass filtering in the 5–15 Hz range to enhance the dominant frequency components of the QRS complex. Subsequently, first-order differentiation and the squaring operations are applied to the filtered result to further sharpen the steepness of the R-wave rising edge. Then, a 150-millisecond sliding window is used to compute the signal’s energy envelope, smoothing high-frequency fluctuations and highlighting the overall QRS energy profile.

Based on the statistical properties of this envelope signal, the adaptive threshold is defined as

$$T = \mu + 0.3\sigma \quad (1)$$

Here, μ and σ represent the mean and standard deviation of the signal envelope, respectively. This threshold is used for preliminary detection of candidate peaks, with a minimum inter-peak interval of 300 ms imposed to suppress false detections. Finally, the original smoothed signal is re-examined within ± 100 ms of each candidate peak to identify the local maximum as the definitive R-wave location. This process integrates physiological features with spectral characteristics, achieving an integrated processing pipeline from noise suppression to precise R-wave localization.

2.4. PCG Signal Processing

The PCG signal captures the mechanical-acoustic information generated by valve motion and hemodynamic changes, serving as a key modality for characterizing mechanical event timing in electro-mechanical coupling analysis. The S1/S2 identification strategy adopted in this study follows a rule-based pipeline comprising “frequency-domain enhancement–envelope extraction–heuristic temporal discrimination.” This design targets a computationally efficient and interpretable baseline that remains stable under typical and moderately noisy conditions, particularly for deployment in low-power and low-computation environments.

First, the original PCG signal undergoes the DC-offset correction to eliminate the constant baseline component, facilitating subsequent processing. Considering that the primary spectral energy of PCG signals is concentrated in the low-to mid-frequency range, while simultaneously suppressing low-frequency drift caused by respiration and high-frequency electromyographic and quantization noise, this study employs a 20–150 Hz band-pass filter to preserve key frequency components related to the cardiac cycle. This frequency band has been demonstrated in prior studies to effectively cover the primary acoustic energy bands of S1 and S2, while maintaining good discrimination capability even in dynamic or noisy environments. Following filtering, the signal undergoes mild smoothing using a 10 ms moving-average window to suppress spike noise and enhance the stability of subsequent envelope estimation.

To extract the instantaneous energy profile of the PCG signal, this study employs the Hilbert transform to compute the instantaneous envelope:

$$E(t) = |H\{x(t)\}| \quad (2)$$

Here, $x(t)$ denotes the smoothed bandpass PCG, $H\{\cdot\}$ represents the Hilbert transform, and $E(t)$ is the envelope signal. The Hilbert envelope preserves transient energy details while avoiding the excessive amplification of extreme values caused by squaring operations.

In the envelope domain, a peak-detection algorithm identifies candidate PCG events, setting the minimum peak height to 30% of the envelope maximum and the minimum

inter-peak interval to 0.2 s to filter out adjacent false peaks caused by noise. Furthermore, based on the physiological pattern of PCG cycles—where systole is typically shorter than diastole—candidate peaks are classified as S1 or S2: if the interval between a pair of adjacent peaks is shorter than the following interval, the earlier peak is assigned as S1 and the later peak as S2, and this rule is iteratively applied to reconstruct the complete PCG cycle sequence. This method demonstrates excellent noise robustness while maintaining computational efficiency.

It should be noted that this rule-based approach represents an engineering simplification by design. Its robustness may be limited under extreme heart rate conditions (pronounced tachycardia or bradycardia), complex arrhythmias, pronounced split heart sounds, or recordings with strong motion artifacts. Addressing such scenarios by incorporating adaptive thresholds or lightweight machine learning models as complementary or fallback solutions constitutes an important direction for future work.

2.5. ECG-Based S1/S2 Localization Assistance and Correction

To improve the consistency and physiological plausibility of PCG event localization, this study introduces the ECG R-wave as a cross-modal temporal reference for assisting the correction of preliminarily detected S1/S2 events. This strategy leverages the relatively stable electromechanical coupling between ECG and PCG to explicitly constrain event timing and compensate for the inherent uncertainty of PCG under low signal-to-noise conditions.

In the current implementation, fixed search windows derived from typical physiological ranges (e.g., 30–150 ms after the R-wave for S1 and 250–400 ms for S2) are employed to constrain candidate PCG events. Importantly, these windows are not treated as rigid physiological constants, but rather as initial search constraints that balance algorithmic stability, interpretability, and deterministic execution in embedded systems.

We acknowledge that fixed windows may underfit inter-subject variability and state-dependent changes, such as differences in resting heart rate, physical activity level, or autonomic modulation. A more generalizable solution would involve adaptive windowing strategies, for example, by parameterizing search ranges as functions of individual baseline heart rate. The current system design leaves explicit interfaces for such personalization, which is identified as a key direction for future work.

2.6. Multimodal Coupling Features and Other Feature Extraction

To deeply reveal the associative patterns between cardiac electrical activity and mechanical acoustic responses, this study constructed a multimodal feature system based on synchronously acquired electrocardiogram (ECG) and phonocardiogram (PCG) signals. This system encompasses temporal, biomechanical amplitude, statistical trend, and coupling synchrony features. This framework not only incorporates independent information from ECG and PCG signals but also emphasizes characterizing their cross-modal dynamic interactions—specifically, multimodal electro-mechanical coupling features, which constitute a core focus of this methodology. All features were extracted from the R-wave peaks and S1/S2 heart sound localization results corrected by the aforementioned algorithms, ultimately forming a 60-dimensional structured feature matrix (Table 1). This provides a unified data foundation for subsequent cardiac state assessment and automated interpretation by large language models.

Table 1. Sixty-Dimensional Structured Features Extracted in This Study and Their Descriptions.

Type	Number	Feature Name	Description/Physical Meaning
Temporal dimension	F1	RS1_mean	Mean delay from R-peak to subsequent S1 onset (R-S1 interval)
	F2	RS1_median	Median delay from R-peak to subsequent S1 onset (R-S1 interval)
	F3	RS1_std	Standard deviation of the R-S1 interval
	F4	RS1_rms	Root-mean-square of the R-S1 interval
	F5	S1S2_mean	Mean interval from S1 onset to subsequent S2 onset (S1-S2 interval)
	F6	S1S2_median	Median interval from S1 onset to subsequent S2 onset (S1-S2 interval)
	F7	S1S2_std	Standard deviation of the S1-S2 interval
	F8	S1S2_rms	Root mean square of the S1-S2 interval
	F9	S2S1_mean	Mean interval from S2 onset to subsequent S1 onset (S2-S1 interval)
	F10	S2S1_median	Median interval from S2 onset to subsequent S1 onset (S2-S1 interval)
	F11	S2S1_std	Standard deviation of the S2-S1 interval
	F12	S2S1_rms	Root mean square of the S2-S1 interval
	F13	RR_mean	Mean interval between consecutive R-peaks (RR interval)
	F14	RR_median	Median interval between consecutive R-peaks (RR interval)
	F15	RR_std	Standard deviation of the RR interval
	F16	RR_rms	Root mean square of the RR interval
	F17	S1S1_mean	Mean interval between consecutive S1 onsets (S1-S1 interval)
	F18	S1S1_median	Median interval between consecutive S1 onsets (S1-S1 interval)
	F19	S1S1_std	Standard deviation of the S1-S1 interval
	F20	S1S1_rms	Root mean square of the S1-S1 interval
	F21	RS1_CV	Coefficient of variation of the R-S1 interval (RS1_std/RS1_mean)
	F22	S1S1_CV	Coefficient of variation of the S1-S1 interval (S1S1_std/S1S1_mean)
	F23	HRV_SDNN	Standard deviation of normal-to-normal RR intervals (SDNN)
	F24	HRV_RMSSD	Root mean square of successive differences between normal RR intervals (RMSSD)
	F25	HRV_pNN50	Percentage of successive RR intervals that differ by more than 50 ms (pNN50)

Table 1. Cont.

Type	Number	Feature Name	Description/Physical Meaning
Magnitude and energy dimensions	F26	S1_peaks	Mean peak amplitude of S1 heart sounds
	F27	S1_rms	Mean root mean square acoustic energy of S1 heart sounds within a 100-ms window
	F28	S1_energy	Mean energy of S1 heart sounds within a 100-ms window
	F29	S2_peaks	Mean peak amplitude of S2 heart sounds
	F30	S2_rms	Mean root mean square acoustic energy of S2 heart sounds within a 100-ms window
	F31	S2_energy	Mean energy of S2 heart sounds within a 100-ms window
	F32	S1_rms_var	Variance of the S1 RMS energy across beats
	F33	S1_rms_cv	Coefficient of variation of the S1 RMS energy across beats
	F34	S2_rms_var	Variance of the S2 RMS energy across beats
	F35	S2_rms_cv	Coefficient of variation of the S2 RMS energy across beats
	F36	S1_skew	Skewness of S1 amplitude distribution
	F37	S1_kurt	Kurtosis of S1 amplitude distribution
	F38	S2_skew	Skewness of S2 amplitude distribution
	F39	S2_kurt	Kurtosis of S2 amplitude distribution
Statistical trend dimension	F40	S1_trend	Slope of linear regression on S1 RMS across consecutive beats, quantifying the trend of S1 intensity
	F41	S1S2_ratio	Ratio of average S1 peak amplitude to S2 peak amplitude, indicating systolic-diastolic mechanical coordination
	F42	S1S2_ratio_var	Variance of the S1/S2 ratio across beats, quantifying the variability of coordination
Multimodal coupling	F43	RR_S1S1_corr	Pearson correlation coefficient between RR intervals and S1-S1 intervals, serving as a consistency measure between electrical and mechanical rhythm detections
	F44	RR_S1S1_pval	<i>p</i> -value of the above correlation
	F45	diff_mean	Mean difference between mechanical (S1-S1) and electrical (RR) cycles
	F46	diff_median	Median difference between mechanical and electrical cycles
	F47	diff_std	Standard deviation of the difference between mechanical and electrical cycles
	F48	diff_rms	Root mean square of the difference between mechanical and electrical cycles

Table 1. *Cont.*

Type	Number	Feature Name	Description/Physical Meaning
Others (frequency domain/anomaly)	F49	LF_power	Low-frequency power of HRV (0.04–0.15 Hz)
	F50	HF_power	High-frequency power of HRV (0.15–0.4 Hz)
	F51	LF_HF_ratio	Ratio of LF to HF power
	F52	bradycardia	Indicator of bradycardia (RR interval > 1.2 s)
	F53	tachycardia	Indicator of tachycardia (RR interval < 0.6 s)
	F54	HRV_abnormal	Indicator of abnormal HRV (SDNN > 0.15 s)
	F55	long_RS1	Indicator of prolonged R-S1 interval (R-S1 > 0.2 s)
	F56	S1S2_abnormal	Indicator of abnormal S1-S2 interval (<0.1 s or >0.4 s)
	F57	S1S2_ratio_abnormal	Indicator of abnormal S1/S2 ratio (<0.5 or >2.0)
	F58	estimated_HR	Estimated mean heart rate (beats per minute)
	F59	missing_data_percent	Percentage of missing or invalid signal segments
	F60	signal_quality	Composite signal quality score/flag

First, in the temporal dimension, the study examined the relative temporal relationship between ECG events (R waves) and PCG events (S1/S2), extracting multiple indicators characterizing electro-mechanical conduction efficiency. To enhance the quantification of temporal stability, this study further calculated the RS1 delay coefficient of variation, S1–S1 period variability, and time-domain heart rate variability (HRV) metrics such as SDNN, RMSSD, and pNN50. This approach provides a multi-layered characterization of autonomic nervous system regulation and periodic stability.

Second, in the amplitude and energy dimensions, a fixed 100-millisecond window was designed for each PCG event to extract key acoustic parameters. The coefficient of variation (CV) of S1 and S2, along with morphological features such as skewness and kurtosis, was also extracted to characterize stability and morphological structural differences within PCG cycles.

Third, in the statistical trend dimension, this study focused on the dynamic changes in PCG intensity over time. By performing linear regression on the S1 RMS sequence of consecutive heartbeats, the slope of the S1 intensity trend was obtained to quantify whether myocardial contraction exhibits an enhancing or attenuating acoustic trend. Additionally, the S1/S2 amplitude ratio and its variability were extracted. This ratio serves as a crucial indicator of the coordination between ventricular systolic and diastolic mechanical responses, reflecting the impact of valve state changes or load alterations on PCG.

Finally, this study focuses on constructing a set of multimodal coupling features to quantify the electro-mechanical synchrony between ECG and PCG. For instance, calculating the Pearson correlation coefficient between RR intervals and S1–S1 periods can assess synchrony between cardiac electrical and mechanical rhythms. Metrics such as cycle-difference distributions, RS1 delay variability, and S1/S2 dynamic ratios can reveal potential electro-mechanical decoupling, valve delay, or abnormal myocardial mechanical responses. Furthermore, this study incorporates simplified frequency domain features based on RR intervals (LF, HF, and their ratio) alongside a series of clinical threshold-based anomaly markers—such as bradycardia, tachycardia, abnormal RS1 delay, and abnormal S1/S2 ratio—to enable fundamental automated screening capabilities.

Overall, the 60 multimodal features extracted in this section form the foundational data structure for subsequent large language model interpretation, automated summarization,

and personalized cardiac state assessment. Among these, cross-modal electro-mechanical coupling features are particularly crucial, as they effectively integrate complementary information from ECG and PCG signals, thereby establishing the core technological foundation for deploying intelligent cardiac health monitoring systems in everyday settings such as communities and homes.

2.7. LLM-Based Automated Data Integration and Intelligent Interpretation System

Building upon the aforementioned multimodal ECG–PCG coupling features, this study developed an LLM-driven automated interpretation module for ECG–PCG signal analysis. This module enables efficient bridging from structured features to clinical semantic outputs. Operating under unified input specifications, it achieves hierarchical semantic mapping of multidimensional physiological indicators through three steps: feature summarization, prompt construction, and structured text generation.

First, the system automatically computes and compresses key temporal, energy, and electromechanical coupling features, generating structured feature entries containing range information and anomaly markers. This ensures the large language model receives inputs with clear boundaries, stability, and medical significance. Subsequently, prompt templates constructed based on cardiac physiology and clinical interpretation logic impose explicit constraints on the output format, restricting the generated content to three sections: “Comprehensive Assessment,” “Current Status,” and “Abnormal Indicators.” Minimal physiological background information is embedded within the prompts to guarantee semantic consistency and medical controllability of the generated results. During inference, the LLM generates standardized report text based on structured features. A hierarchical parsing module then extracts the structure and standardizes the format to support subsequent clinical review and system-level invocation. Additionally, the module automatically generates a simplified summary for non-specialist users, enhancing readability and practicality in home and community health monitoring scenarios.

This interpretation module achieves an automated closed-loop process from “feature computation” to “semantic interpretation” by deeply integrating the language generation capabilities of LLMs with structured physiological features. Its design emphasizes structural constraints, semantic consistency, and medical controllability, providing a scalable and deployable technical pathway for the intelligent interpretation of complex multimodal physiological signals.

3. Experiments and Results

3.1. Core Algorithm Performance Evaluation

3.1.1. Ablation Experiments and Results of PCG Localization Assisted by R-Wave Prior Correction Strategy

To quantitatively assess the contribution of R-wave temporal a priori information to PCG localization accuracy, ablation experiments were designed in this study. Experiments were conducted under four typical physiological activity scenarios—sitting, lying down, walking, and cycling—to perform a comparative analysis of long-term synchronously collected PCG and ECG data. The system systematically compared the detection performance of S1/S2 heart sound peaks under two conditions: “without R-wave correction” (baseline method) and “with R-wave-assisted correction.”

The experimental results (Table 2) clearly demonstrate the effectiveness and necessity of the R-wave correction strategy. Specifically: (1) Correction of missed detections in low-motion scenarios: In certain seated recordings (e.g., ECGPCG0011), the baseline method exhibited severe missed detections, yielding calculated heart rates significantly below the normal physiological range (60–90 beats per minute). After introducing R-wave

timing constraints, missing PCG events were effectively supplemented, restoring heart rate estimates to the normal range. (2) Removal of false detections in supine scenarios: In supine recordings (e.g., ECGPCG0016), R-wave correction not only aided localization but also successfully filtered erroneous peaks caused by noise or physiological interference, enhancing detection specificity. (3) Improved robustness in dynamic scenarios: During walking and cycling, the difference in detection counts between the two methods was significantly greater than in static scenarios. This indicates that when motion artifacts intensify, the reliable electrophysiological timing reference provided by R waves becomes crucial, substantially enhancing the robustness of PCG detection algorithms in dynamic environments.

Table 2. Ablation Experiment Results for R-Wave-Assisted PCG Correction Strategy.

Document ID	Status	Number of S1	Number of S1 After R-Wave Correction	Difference
ECGPCG0040	Sit	45	45	0
ECGPCG0011	Sit	38	66	28
ECGPCG0010	Sit	58	63	5
ECGPCG0009	Sit	30	30	0
ECGPCG0008	Sit	32	36	4
ECGPCG0007	Sit	31	36	5
ECGPCG0006	Sit	31	34	3
ECGPCG0003	Sit	45	45	0
ECGPCG0013	Rest	54	56	2
ECGPCG0014	Rest	39	59	20
ECGPCG0015	Rest	66	67	1
ECGPCG0016	Rest	68	63	−5
ECGPCG0021	Rest	79	80	1
ECGPCG0022	Rest	28	69	41
ECGPCG0023	Rest	8	68	60
ECGPCG0029	Bike	59	103	44
ECGPCG0033	Bike	54	87	33
ECGPCG0050	Walk	45	91	46
ECGPCG0053	Walk	30	84	54

Overall, across all four experimental conditions, quantifiable and significant differences were observed between the “uncorrected” and “corrected” detection results. The specific manifestations of these differences—such as reduced missed detections and eliminated false positives—exhibited systematic variations in magnitude depending on exercise intensity and interference levels. This fully validates the cross-scenario effectiveness of the proposed R-wave-assisted correction strategy.

3.1.2. Comparison with the OSET Benchmark Algorithm

To establish a reliable reference for performance evaluation, the ground truth for R-wave, S1, and S2 positions was generated through a semi-automatic, correction-based protocol. First, initial event positions were obtained using the OSET benchmark algorithm [23]. Subsequently, these preliminary detections were meticulously reviewed and manually corrected by two experienced researchers. This correction process involved visual inspection and auditory verification of synchronously recorded ECG and PCG signals to ensure physiological plausibility and temporal precision. Any discrepancies between reviewers were resolved through discussion, and when necessary, adjudicated by a third senior researcher, resulting in a consensus-based, corrected reference dataset. This approach leverages the efficiency of an established algorithm while ensuring expert-level accuracy through human oversight.

All detection accuracies (for R, S1, and S2) and mean timing errors (RR, S1-S1, S2-S2, and R-S1 intervals) reported in Table 3 were computed by comparing the outputs of the evaluated algorithms against this corrected ground truth. It is important to note that for the OSET algorithm, this represents a comparison between its original output and its manually corrected version. For the proposed method, the comparison is against the finalized reference standard. This framework allows for a direct assessment of how each algorithm's detections deviate from the expert-verified benchmark.

Table 3. Quantitative comparison between the proposed method and OSET benchmark using manually curated ground truth.

Document ID	Statue	Acc of R (OSET)	Acc of R (Prop)	Acc of S1 (OSET)	Acc of S1 (Prop)	Acc of S2 (OSET)	Acc of S2 (Prop)	Mean RR Error (ms)	Mean S1S1 Error (ms)	Mean S2S2 Error (ms)	Mean RS1 Error (ms)
ECGPCG0040	Sit	100.0	97.9	100.0	100.0	100.0	100.0	5.4	63.5	59.3	24.9
ECGPCG0011	Sit	98.8	98.6	98.3	98.2	96.7	97.4	35.1	104.4	109.2	3.6
ECGPCG0010	Sit	99.6	99.1	97.8	96.6	97.5	96.8	12.2	81.9	84.9	17.9
ECGPCG0009	Sit	100.0	100.0	94.0	95.5	92.0	96.9	0.2	73.2	69.4	17.2
ECGPCG0008	Sit	100.0	100.0	94.7	94.7	90.2	91.3	2.5	80.4	159.6	42.4
ECGPCG0007	Sit	99.0	97.4	94.7	92.1	85.7	100.0	44.0	102.8	212.5	7.3
ECGPCG0006	Sit	100.0	100.0	97.1	92.8	86.8	89.2	0.5	77.3	170.6	22.9
ECGPCG0003	Sit	100.0	100.0	100.0	100.0	100.0	97.8	0.7	49.4	57.6	22.5
ECGPCG0013	Rest	100.0	99.0	98.3	92.0	97.3	92.6	32.9	114.1	119.9	18.0
ECGPCG0014	Rest	99.6	98.5	92.3	90.4	90.8	89.8	36.5	100.9	114.1	26.0
ECGPCG0015	Rest	99.4	98.1	98.8	98.6	95.7	97.9	31.9	81.0	82.4	14.2
ECGPCG0016	Rest	100.0	98.9	94.3	86.5	96.2	92.6	9.5	99.9	115.1	13.9
ECGPCG0021	Rest	100.0	98.5	97.3	97.4	94.0	97.6	23.9	97.9	105.6	12.2
ECGPCG0022	Rest	98.5	98.5	92.6	94.52	90.1	98.1	17.0	106.2	180.2	14.2
ECGPCG0023	Rest	100.0	98.6	98.3	98.1	85.6	92.7	22.9	75.2	110.1	6.9
ECGPCG0029	Bike	79.2	87.4	83.8	93.3	84.2	88.9	117.2	101.2	173.5	4.3
ECGPCG0033	Bike	70.5	74.8	97.9	98.1	94.9	99.2	21.6	124.1	164.2	31.7
ECGPCG0050	Walk	71.0	73.2	88.5	94.0	85.7	100.0	53.1	94.0	168.2	28.4
ECGPCG0053	Walk	68.9	73.7	80.6	92.9	73.9	91.6	194.6	319.2	412.8	23.2

Under static conditions, this algorithm achieves an R-wave detection accuracy ranging from 97.5% to 100%, with an average RR interval error consistently below 50 ms. For S1 and S2 detection, accuracies were consistently high, predominantly above 90% and reaching up to 100% in multiple recordings. Meanwhile, the R–S1 delay error was predominantly under 25 ms, indicating precise electro-mechanical interval estimation. The proposed R-wave-assisted correction strategy ensures that S1 and S2 detection strictly follows their physiological sequence, effectively avoiding physiologically implausible misdetections such as “one S1 corresponding to multiple S2s” occasionally observed in the OSET algorithm (Figure 2).

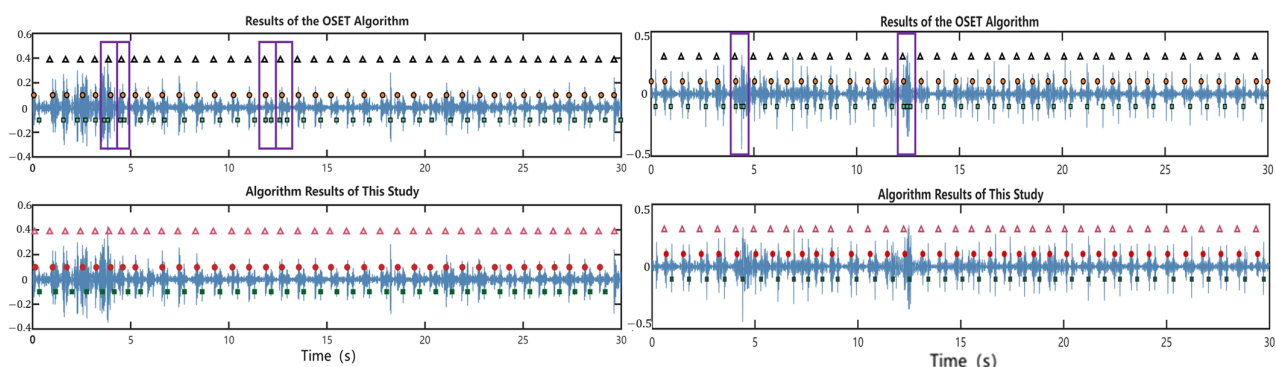


Figure 2. Comparison of PCG Segmentation Results Between the OSET Algorithm and the Algorithm Developed in This Study. In the phonocardiogram (PCG) tracings, triangle markers denote R-waves of the electrocardiogram (ECG) reference. Red markers indicate the first heart sound (S1), and green markers indicate the second heart sound (S2). Purple boxes highlight instances of physiologically implausible misdetections (e.g., one S1 corresponding to multiple S2s) occasionally produced by the OSET algorithm.

Under exercise conditions, R-wave detection accuracy decreased to approximately 73%, with average RR interval errors increasing to 53.1 ms and 194.6 ms, respectively. Notably, despite the decline in R-wave detection performance, the detection accuracy of S1 and S2 remained above 90%. This robustness may stem from the distinct interference patterns affecting PCG signals compared with ECG signals during motion. However, because R-wave detection forms the basis of PCG correction, its degraded performance still induces cascading effects on overall timing accuracy, manifested as increased S1–S1 and S2–S2 interval errors.

Overall, the comparison indicates that the proposed method maintains high detection accuracy while producing physiologically consistent PCG event sequences. Although overall performance degrades under dynamic conditions, it remains comparable to the traditional benchmark method, supporting the feasibility of adapting established ECG-guided PCG segmentation principles into a lightweight pipeline suitable for low-resource deployment.

3.1.3. Baseline Performance on the Self-Built Physiological Perturbation Dataset

Using the self-built ECG–PCG dataset, the baseline performance of the proposed system was evaluated under four representative perturbation conditions: resting, post-exercise recovery, positive emotion (happiness), and negative emotion (anger). Manual inspection confirmed that, after signal enhancement, ECG recordings across all conditions exhibited clear R-wave morphology with no evident missed detections. Accordingly, R-wave counts were used as a reference for cardiac cycle estimation, and the consistency between the number of detected S1/S2 events and R-wave counts was employed to quantify heart sound localization stability. Typical recognition results for the four class states are shown in Figure 3a–d.

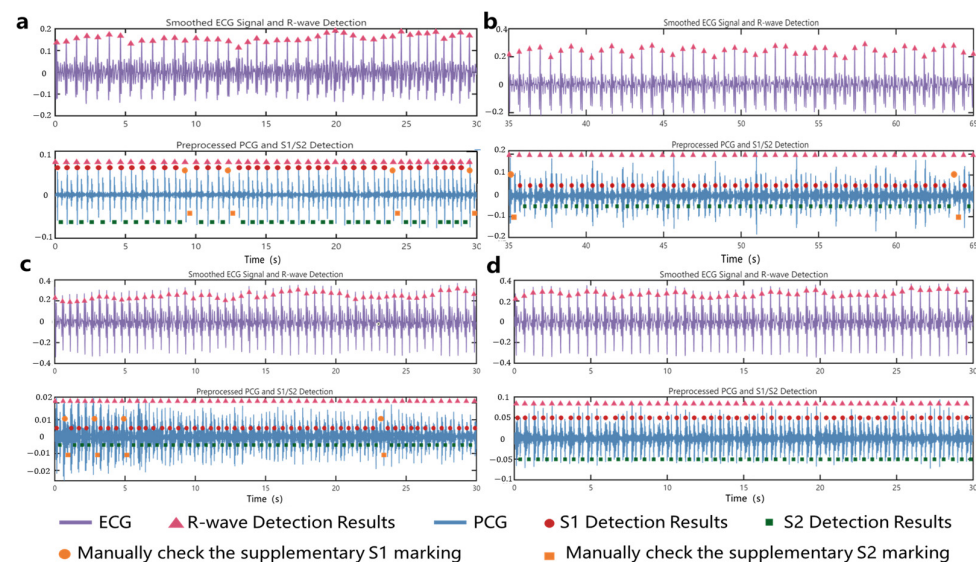


Figure 3. Demonstration of the basic performance of the data processing algorithm on the self-built dataset: (a) Preprocessed ECG and R-wave recognition results in the resting state, along with preprocessed PCG signals and S1/S2 segmentation results corrected based on R-waves. (b) Preprocessed ECG and R-wave recognition results in the post-exercise recovery state, along with preprocessed PCG signals and S1/S2 segmentation results corrected based on R-wave. (c) Preprocessed ECG and R-wave recognition results during the positive emotional state, along with preprocessed PCG signals and R-wave-corrected S1/S2 segmentation results. (d) Preprocessed ECG and R-wave recognition results during the negative emotional state, along with preprocessed PCG signals and R-wave-corrected S1/S2 segmentation results.

PCG signals exhibited distinct noise characteristics and morphological variations across perturbation conditions. During resting, heart sound structures were clear, and envelope features were stable. In post-exercise recovery, increased respiratory activity, chest-wall vibration, and muscle noise led to pronounced envelope fluctuations. During emotion-induction tasks, heart rate dynamics became more nonstationary, and occasional transient disturbances were introduced by motion or mild speech.

Despite these perturbations, the proposed system consistently maintained stable alignment between S1/S2 detections and R-wave-defined cardiac cycles. Quantitative results indicated that S1/S2 detection was most stable during resting, with detection counts matching R-wave counts or deviating only slightly in most recordings, yielding accuracies above 95%. In post-exercise recovery, although PCG signals were subject to the strongest noise contamination, the system maintained reasonable localization performance, with S1/S2 detection accuracies exceeding 90% in most samples; occasional missed detections occurred under conditions of intense respiration or motion.

Under emotion-induction conditions, system performance fell between resting and post-exercise states. During negative emotion, S1/S2–R count consistency was maintained at 93–97%, while during positive emotion, mild speech-related interference caused a slight reduction in accuracy (approximately 88–95%). Overall, these results demonstrate that the proposed system exhibits robust performance under nonstationary, noise-contaminated conditions with pronounced heart rate variability.

3.2. Feature Stability and Physiological Consistency

Following validation of the core algorithm's performance, the stability and physiological consistency of the system's multimodal coupled features were evaluated. This section analyzes high-quality signal recordings from seated, supine, cycling, and walking states within public databases. Relevant results are presented in Figure 4.

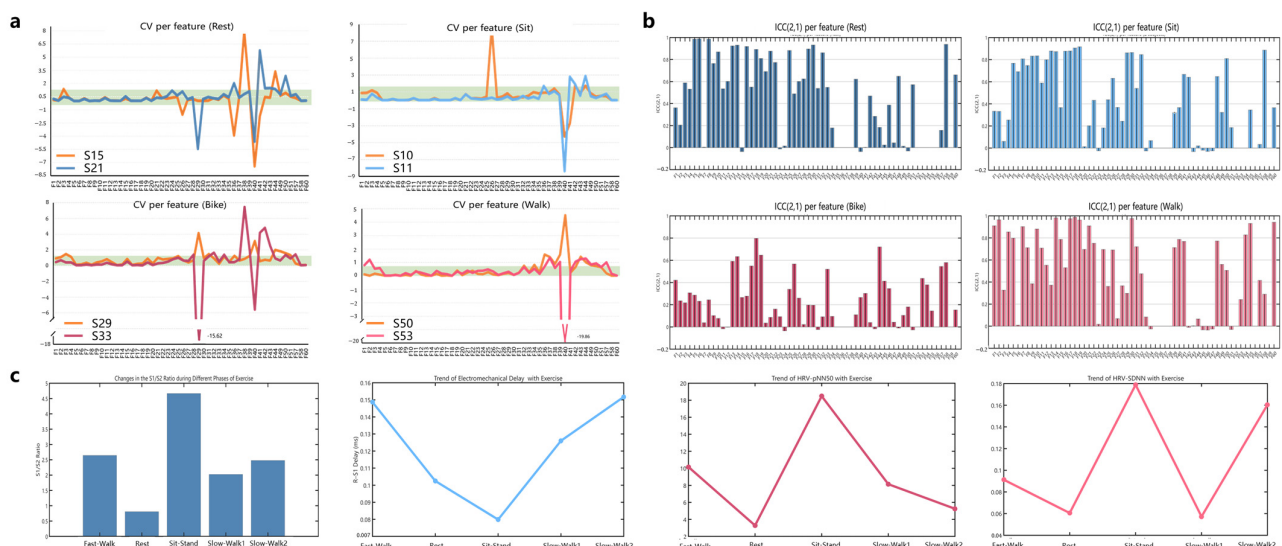


Figure 4. Results of Feature Stability and Physiological Consistency: (a) Intra-segment stability assessment of 60-dimensional multimodal coupling features across Rest, Sit, Bike, and Walk states. Bars highlighted in green ($0 \leq \text{mean CV} \leq 1$) indicate features with acceptable temporal stability, whereas features with $\text{mean CV} \leq 0$ or >1 (not highlighted) exhibit low signal strength or high variability, respectively. (b) Inter-segment repeatability analysis of 60-dimensional multimodal coupling features across Rest, Sit, Bike, and Walk states. (c) Changes in R–S1, S1–S2, HRV_SDNN, and HRV_pNN50 across multiple states for a single subject.

3.2.1. Reproducibility Analysis of Feature Extraction

To evaluate the measurement consistency of ECG-PCG multimodal coupling features at short-term scales, this study conducted reproducibility analyses at two levels: intra-segment (within the same subject) and inter-segment (across subjects).

In the intrasegment stability assessment, 30 one-minute segments were extracted from 30 min of synchronized recordings from two subjects for each of four typical states: sitting, lying down, cycling, and walking. The coefficient of variation (CV) for 60 features and their 95% confidence intervals based on 1000 bootstrap samples were calculated. Results (Figure 3) indicate that most features exhibit good to excellent stability ($CV < 0.3$) within the same state. Time-series metrics such as RR_mean and RR_rms maintain extremely low variability ($CV < 0.1$) across all states with narrow confidence intervals. The S1/S2 amplitude ratio exhibited moderate stability ($CV 0.24\text{--}0.58$) during dynamic states. PCG energy-based features remained stable at rest but showed greater variability during exercise. Certain features, like S1_peaks and S2_peaks, exhibited higher CV and wider confidence intervals. Morphological features (S1_skew, S2_skew) and trend-based features (S1_trend) also showed significant variability across multiple states.

In the inter-segment repeatability analysis, ICC(2,1) was used to quantify the cross-subject consistency of features across states. Results showed that multiple core features exhibited ICC values exceeding 0.75. During the seated state, features such as S1_rms, S1_energy, and RR_mean all achieved ICC values surpassing 0.86. In the supine state, S1-S2_mean and RR_median both achieved ICC values above 0.9. During walking, features like RS1_median, S1_peaks, S1_rms, and RR_median reached ICC values between 0.96 and 0.99. In contrast, most features showed reduced ICC during cycling, with only a few, like S1_rms and RR_median, maintaining moderate consistency. Morphometric statistics and trend-based features (e.g., S2_peaks, S1_trend) consistently exhibited low ICC across all states.

3.2.2. Physiological Consistency Validation of Electromechanical Coupling Parameters

Intergroup differences in electromechanical coupling parameters were tested using multi-state recordings from a single subject in the public dataset (Slow Walk 1—Fast Walk—Sitting—Standing Repeated—Slow Walk 2—Rest). The tested indicators included R-S1, S1-S2, HRV_SDNN, and HRV_pNN50. Results showed: Intergroup comparison of R-S1 did not reach statistical significance ($p = 0.3826$); S1-S2 exhibited significant intergroup differences ($p = 0.0111$); Intergroup differences for HRV_SDNN and HRV_pNN50 were $p = 0.0158$ and $p = 0.0025$, respectively. The multimodal coupling indicators extracted in this study effectively captured changes in cardiac electro-mechanical dynamics and autonomic regulation characteristics triggered by different conditions. Their evolutionary trends were highly consistent with existing studies in the literature [24–26], validating the system's reliability in physiological correlation and dynamic tracking capabilities.

3.3. LLM-Based Intelligent Interpretation System

Following feature extraction and multimodal coupling computation, this study further evaluated the system's performance in automated text interpretation. This module generates structured cardiac health reports tailored to diverse clinical scenarios using LLMs as input, thereby validating the system's capability to translate complex ECG-PCG metrics into structured semantic outputs.

The system employs a dual-layer output framework: one layer delivers detailed reports for professionals, covering RR interval variability, HRV metrics, autonomic regulation status, S1/S2 mechanical contraction dynamics, and interpretation of electro-mechanical coupling relationships; the other layer provides summary reports for general users, present-

ing heart rate levels, rhythm stability, baseline risk alerts, and lifestyle recommendations through templated summarization. This dual-layer mechanism enables multi-tiered information presentation, bridging medical interpretability to public comprehensibility.

3.3.1. Comparative Performance Evaluation of Different LLMs and the Necessity of Template-Based Constraint Mechanisms

To enhance the reliability of automatically generated reports, this study systematically compared two lightweight open-source models (DeepSeek-R1:1.5b and Qwen2.5:1.5b). The evaluation is based on a set of real physiological parameters (derived from an ECG-PCG recording during the recovery phase after exercise in a 26-year-old healthy female). By programmatically constructing 20 standardized test cases, the model's four key dimensions are systematically tested: (1) medical knowledge comprehension, (2) numerical logical reasoning, (3) report generation capability, and (4) error detection capability. This design enables us to conduct deep, multidimensional capability assessments based on limited real-world data.

Each evaluation dimension consisted of five standardized test items. For each item, the model response was scored according to predefined quantitative rules: 1 point for fully correct answers or correct identification of errors, 0.5 points for partially correct responses containing core relevant elements, and 0 points for incorrect answers, failures to identify errors, or generation failures.

The score of each dimension was calculated as the mean value of its five sub-test scores and converted to a percentage scale. The overall score reported in Table 4 was computed as the unweighted arithmetic mean of the four dimensions' scores, since these dimensions represent complementary and equally critical aspects for safe and reliable medical report generation:

$$Score_{overall} = \frac{1}{4} \sum_{i=1}^4 Score_i \quad (3)$$

Table 4. Performance comparison of two lightweight LLMs evaluated on 20 standardized test cases across four dimensions. Each dimension score is the average of five sub-tests and is reported on a 0–100 scale. The overall score is the unweighted arithmetic mean of the four dimensions' scores.

Evaluation Dimensions	DeepSeek-R1: 1.5b	Qwen2.5:1.5b	Key Findings
Overall Score	72.5/100	70.0/100	DeepSeek-R1 shows slight superiority
Medical Knowledge Accuracy	60.0	60.0	Both models exhibit significant terminology confusion
Numerical Logical Reasoning	80.0	80.0	Insufficient unit conversion capability
Report Generation Capability	80.0	60.0	DeepSeek-R1 demonstrates superior data integration
Error Identification Capability	70.0	80.0	Qwen2.5 exhibits stronger logical detection

Detailed descriptions of the test case construction, prompt design, automation pipeline, and scoring criteria are provided in Appendix A.

Test results indicate that the models exhibit certain limitations in understanding medical terminology and numerical reasoning. Common issues include terminology confusion (e.g., misinterpreting LF/HF as financial metrics or misjudging S1/S2 as non-medical quantities) and misreading numerical values for RR intervals and HRV ranges. Both models can generate basic structured descriptions, but their performance differs:

Qwen2.5 scored higher in logical consistency and error identification, while DeepSeek-R1 demonstrated greater consistency in overall report structuring and feature integration.

Based on the above findings, this study employs a template-constrained hybrid generation strategy. Within this mechanism, the conclusion section is first determined by the rule-based control module based on structured feature calculation results; the language model is solely responsible for converting the predetermined conclusion into natural language expressions tailored for different user groups. This design effectively reduces potential biases arising from medical terminology misuse, numerical misinterpretation, and free generation while maintaining linguistic flexibility.

3.3.2. Classic Case Showcase

Figure 5 and Appendix B illustrate two tiers of system-generated reports: (a) simplified reports based on universal templates, and (b) personalized reports triggered by specific statuses (neither of which has been manually edited). This demonstrates the system’s full capability, from foundational structured generation to context-aware personalized adaptation.

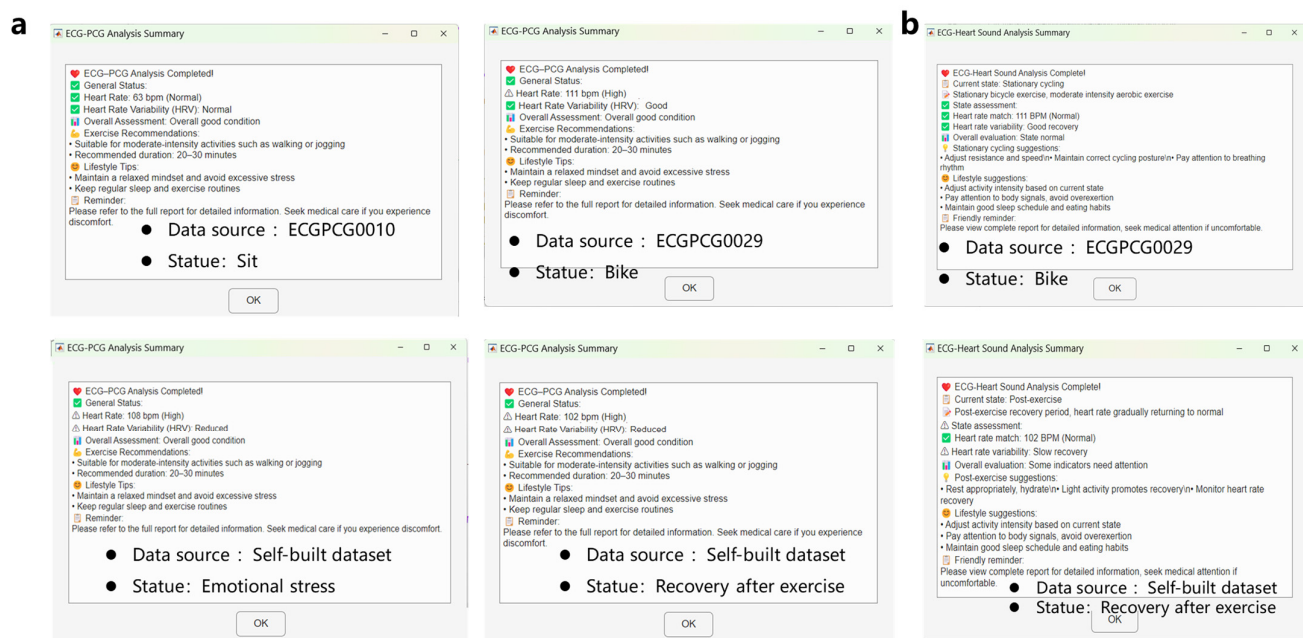


Figure 5. Examples of Automated Report Generation: (a) Four simplified reports generated using a common template, demonstrating the system’s basic capability to produce structured output across different states. (b) Personalized reports for two specific states, generated when the system’s rule engine is explicitly triggered by the contextual state label (“Bike” and “Recovery after exercise”). These examples demonstrate the system’s capacity to produce tailored recommendations.

In the resting scenario example (corresponding to the “Sit” state in Figure 5a), the system generates qualitative descriptions of rhythm stability, autonomic regulation levels, and structural PCG characteristics based on the low-to-high frequency ratio, SDNN, RR distribution, and S1–S2 dynamics parameters. This includes a low-risk assessment and foundational health management recommendations. In the exercise scenario example (corresponding to the “Bike” state in Figure 5a and the personalized report in Figure 5b), the system identifies physiological characteristics such as increased heart rate, elevated LF/HF ratio, and structural changes in S1/S2. When provided with an explicit state label, it generates personalized recommendation text centered on rhythm recovery and exercise load adjustment, as shown in Figure 5b.

As demonstrated in the examples, the LLM interpretation module generates context-sensitive textual descriptions based on multidimensional physiological features such as rhythm variability, autonomic regulation, and electro-mechanical coupling. It can adjust the structure and linguistic style of the reports based on the specific state input. This module constitutes the system's key interaction layer for end-users, enabling long-term physiological monitoring and health management applications in community and home settings.

4. Discussion

4.1. Overall System Performance and Methodological Positioning

This study presents an automated ECG–PCG coupling analysis system organized along a unified processing pipeline comprising signal-level denoising, event-level calibration, feature-level fusion, and semantic-level expression. It should be explicitly noted that the primary objective of this work is to validate the methodological feasibility and engineering robustness of such a system under resource-constrained and uncontrolled community or home environments, rather than to assess its effectiveness for disease diagnosis or clinical decision-making. In contrast to existing ECG–PCG studies that predominantly pursue improved diagnostic accuracy or disease classification performance, this work deliberately adopts a monitoring-oriented paradigm, focusing on stable physiological state perception, cross-condition robustness, and interpretable feedback suitable for long-term deployment.

The system design leverages the complementary characteristics of ECG and PCG signals, combining the high temporal precision of ECG with the mechanical information captured by PCG. By introducing cross-modal temporal constraints, the proposed framework improves the stability and consistency of heart sound event localization. Based on this foundation, a structured multimodal feature set is constructed to describe cardiac rhythm characteristics, acoustic energy patterns, and electro-mechanical coupling relationships, providing reliable inputs for subsequent state-level analysis and semantic representation.

The robustness of the overall system is reflected at multiple levels: (1) the preprocessing module effectively mitigates common nonstationary noise encountered in home and community settings; (2) the R-wave-guided heart sound calibration strategy substantially enhances S1/S2 temporal alignment under conditions of weak heart sounds or motion artifacts, shifting the system from acoustics-dependent detection to a cross-modal constraint-driven approach; and (3) the multimodal feature set spans rhythm, energy, and electro-mechanical timing dimensions, enabling the characterization of global physiological state variations across different conditions. (4) The system integrates real-time evaluation and failure protection mechanisms based on signal quality. When signal quality is too low, it prompts users to retake measurements rather than outputting unreliable results, thereby ensuring safety and practicality in actual deployment.

On top of the structured feature computation, a semantic expression module is incorporated solely as an interface for translating predefined features into standardized textual descriptions, thereby improving output interpretability across different user groups. This module does not perform medical reasoning or diagnosis; instead, it serves to reduce the cognitive burden associated with understanding complex multimodal physiological information.

In summary, this work demonstrates a multi-level collaborative analysis framework tailored for community and home environments, with its main contribution lying in the validation of ECG–PCG coupling analysis and structured expression under constrained conditions. The proposed system is intended as a physiological state awareness and decision-support tool, providing early and understandable indications of potential changes

and encouraging users to seek professional medical evaluation, when necessary, rather than replacing clinical diagnosis or treatment.

4.2. Contribution of R-Wave-Assisted Correction to PCG Analysis

Introducing the R-wave as a temporal prior for auxiliary correction of heart sound events (S1/S2) is a critical step in achieving stable PCG analysis under complex conditions. Under pure PCG segmentation, heart sound localization is significantly impacted by reduced signal-to-noise ratio, envelope distortion, and motion artifacts—whether at rest or during dynamic movement—resulting in PCG counts that deviate markedly from actual beat counts. Particularly in high-interference scenarios like walking or cycling, baseline methods become nearly ineffective, with some recordings yielding single-digit or significantly underestimated PCG counts. This further highlights the structural limitations of unimodal acoustic strategies in non-stationary environments.

In contrast, the R-wave-assisted correction strategy leverages the high temporal precision of ECG signals and the stability of electro-mechanical coupling to provide reliable cross-modal temporal anchors for PCG events, thereby mechanistically addressing these shortcomings. Based on ablation experiment results (Section 3.1.1), R-wave correction effectively compensates for missed detections caused by weak PCG, restoring counting accuracy to levels consistent with physiological heart rate, or achieving false detection suppression. This demonstrates that introducing the R-wave prior establishes a robust “structural supervision channel,” significantly enhancing system stability in non-resting environments.

From a system workflow perspective, this correction strategy not only enhances the accuracy of PCG localization but also provides high-quality foundational data support for subsequent PCG classification, temporal feature calculation, and multimodal coupling index analysis. The stability of S1/S2 locking directly impacts the reliability of key temporal metrics such as RR–S1, S1–S2, and mechanical delay. These metrics serve as crucial inputs for automated interpretation and ECG–PCG coupling assessment in this study. Therefore, R-wave-assisted correction functions not merely as a localized optimization module but as a foundational safeguard throughout the entire system analysis chain.

At the methodological level, the strategy employed in this study demonstrates distinct advantages in robustness and scene adaptability. Traditional PCG segmentation methods predominantly rely on envelope detection [27], Hilbert transform [28], wavelet analysis [29], or statistical learning models [30], which perform well under quiet conditions but exhibit a sharp decline in performance in dynamic scenarios such as low signal-to-noise ratios or motion. While prior studies [31–33] have demonstrated the auxiliary value of ECG for PCG segmentation, most follow a data-driven paradigm. Their performance heavily depends on the scale and quality of training data, limiting their generalization capability and interpretability in unseen noisy scenarios. To address this, this study proposes a mechanism-first regularization framework centered on a systematic cross-modal temporal coupling mechanism: “window constraint—candidate matching—physiological tolerance—missed detection compensation.” This approach drives the segmentation process by explicitly encoding the physiological relationship between PCG and ECG, rather than implicitly learning it from data. This design significantly reduces reliance on large-scale annotated datasets, ultimately achieving stable and consistent PCG localization performance under complex conditions such as motion and long-term recordings—a feat previously difficult for conventional methods.

It should be emphasized that the R-wave-assisted correction strategy is not intended to rigidly enforce fixed physiological windows across all subjects and states. Rather, it functions as a structural constraint layer that mediates between rule-based PCG segmentation

and complex physiological variability. Its core value lies in explicitly encoding the temporal relationship between electrical and mechanical events, thereby substantially mitigating the disruptive impact of false positives and false negatives without relying on large-scale training data. When signal quality deteriorates substantially or the rule-based pipeline fails, incorporating modern PCG segmentation approaches—such as lightweight learning models or template-based methods—as secondary checks or fallback modules represents a reasonable and extensible system evolution path. Within this framework, the lightweight rule-based pipeline proposed in this study should be regarded as a stable and interpretable baseline layer rather than an exclusive final solution.

In summary, R-wave-assisted correction effectively mitigates the acoustic sensitivity issues of PCG in dynamic environments by introducing cross-modal temporal constraints, providing a fundamental mechanism for reliable localization of core cardiac sound events. Results demonstrate that this strategy constitutes an effective approach for achieving stable cardiac sound analysis across diverse scenarios. Its rule-based, low-data-dependency design also lays a crucial foundation for constructing more robust and interpretable multi-modal physiological signal analysis systems.

4.3. Physiological Interpretability and Application Potential of Multimodal Coupling Features

The ECG–PCG multimodal coupling feature framework developed in this study is designed to provide a structured representation of physiological state variations by integrating cardiac electrical activity, mechanical acoustic responses, and their temporal relationships. It should be emphasized that the following discussion focuses on the potential of these features to reflect physiological state changes and regulatory trends, rather than their use for diagnosing specific pathological conditions.

Systematic evaluation results indicate that multiple coupling features exhibit good short-term stability and cross-subject consistency, supporting their methodological reliability as descriptors of physiological states. For instance, RR interval-related temporal features demonstrate low variability and high agreement across different experimental conditions, suggesting their robustness in capturing heart rate regulation and rhythm dynamics. Such indicators have been widely used in prior studies to characterize autonomic modulation and load responses [34], and their stability forms a necessary basis for continuous monitoring in community and home environments.

In addition, heart sound energy- and morphology-related features show favorable inter-subject consistency in resting and certain dynamic conditions, indicating their ability to consistently reflect variations in cardiac mechanical activity [35]. In the context of this study, these features are primarily interpreted as physiological descriptors of mechanical response changes, rather than indicators of specific valvular abnormalities or disease states.

More notably, cross-modal coupling features (e.g., RS1, S1–S2 intervals) integrate the temporal relationship between electrical and mechanical cardiac events, offering a quantitative means to analyze electro-mechanical synchrony under different physiological conditions. Experimental results demonstrate that these features maintain good measurement stability across multiple states and exhibit consistent trends in response to perturbations such as exercise, respiration, and emotion induction, highlighting their capacity to reflect dynamic cardiac adaptations to physiological load.

Further analysis reveals that feature stability is significantly modulated by physiological state. Resting and low-interference conditions are more suitable for extracting baseline-related characteristics, whereas rhythmic movement conditions may enhance the consistency of certain temporal and coupling features. These findings suggest that feature selection should be adapted to specific usage scenarios to maximize reliability in practical deployments. At this stage, the feature set was intentionally kept comprehensive to enable

a systematic evaluation of ECG–PCG coupling information under diverse physiological conditions. Feature selection or dimensionality reduction was therefore not performed in this study, as the primary goal was methodological feasibility assessment rather than application-specific model optimization. Future work will focus on task-oriented feature selection to reduce redundancy and improve computational efficiency.

Compared with previous studies relying primarily on single-modality ECG or PCG features [36,37], this work systematically evaluates the stability and state responsiveness of multimodal coupling features, demonstrating their potential advantages in capturing physiological state variations. Importantly, these features are not intended to directly indicate specific diseases but rather to provide richer information for long-term monitoring, physiological trend analysis, and individual state awareness.

In conclusion, the proposed multimodal coupling feature framework exhibits favorable stability and physiological consistency at the methodological level, making it suitable as a foundational tool for physiological state monitoring in community and home settings. Its value lies in supporting longitudinal observation of cardiac regulatory dynamics, rather than replacing standard clinical assessment or disease diagnosis.

4.4. Potential and Areas for Improvement of LLMs in Heart Health Monitoring

Based on a multimodal ECG–PCG coupling feature framework, this study developed an embedded lightweight LLM-based intelligent interpretation module, achieving a fully automated closed-loop pipeline from feature extraction to structured semantic report generation. It should be clarified that the primary objective of this study is not to validate independent medical reasoning by LLMs, but rather to investigate whether lightweight models can still accomplish structured feature integration and dual-mode semantic expression under extreme resource constraints. Experimental results demonstrate that the module effectively integrates multidimensional ECG and PCG indicators across resting and exercise scenarios, generating context-sensitive and semantically consistent dual-layer explanatory texts for different user groups. This data-driven semantic abstraction approach significantly lowers the expertise barrier for interpreting complex physiological signals and provides a feasible technical pathway for long-term cardiac monitoring in community and home settings. Sample reports show that the system delivers coherent interpretations of rhythm stability, heart rate variability, and autonomic balance, while distinguishing physiological load responses from potential abnormal fluctuations during exercise—bridging the gap between “metric output” and “understandable health interpretation.”

The effectiveness of the LLM module stems not only from its language generation capability but also from the highly stable ECG–PCG electromechanical coupling feature system developed in this study. Compared with prior approaches relying primarily on single heart rate or HRV metrics [38], the proposed coupling features exhibit improved physiological consistency across states, providing reliable and verifiable structured inputs for semantic integration. By extending monitoring from the “rhythmic level” to the “electromechanical dynamics level,” the system enhances the informational depth of home-based monitoring, approaching that of clinical equipment. Leveraging the organizational and expressive capabilities of LLMs, complex cardiac dynamics can thus be transformed into medically coherent, follow-up-oriented textual descriptions that support long-term monitoring, load adjustment analysis, and individualized risk awareness.

At the same time, systematic evaluation reveals that lightweight LLMs struggle to independently generate medical reports without stringent constraints. Comparative experiments with DeepSeek-R1:1.5B and Qwen2.5:1.5B demonstrate notable limitations in medical terminology comprehension and numerical reasoning. Accordingly, this study adopts a rule-driven and template-constrained hybrid generation architecture, in which

all numerical computations and state determinations are completed and validated by rule-based modules, while the LLM is restricted to generating natural language expressions within predefined semantic boundaries. This strategy enhances interpretive safety and consistency while preserving readability. Notably, the selection of 1.5B-scale models was intentional, aiming to evaluate feasibility under extreme constraints on computation, memory, and energy consumption.

Although this study adopts a cautious stance regarding independent medical reasoning by lightweight LLMs, it does not preclude the potential of larger models under richer resource conditions [39]. The structured feature tables, interpretation templates, and semantic interfaces produced by the system are designed to be interoperable with external interpretation tools, allowing users to submit results to trusted, networked large-scale models or professional intelligent systems for further analysis. Recent advances in traceable LLM-driven medical reasoning agents demonstrate that, with appropriate model scale and system design, language models may play a more central role in future decision-support paradigms.

Overall, rather than positioning LLMs as medical reasoning engines, this study validates a practical pathway for embedding lightweight LLMs as constrained semantic integration and expression interfaces within cardiac monitoring systems under resource-limited conditions. This work provides an engineering reference for flexible deployment of language models across varying resource environments and for the gradual integration of higher-level intelligent interpretation capabilities.

5. Conclusions

This study proposes an automated ECG–PCG coupling analysis system for community and home use, implementing a full workflow including signal preprocessing, event detection and calibration, multimodal feature construction, and rule-constrained LLM-assisted semantic report generation. Results show that the system exhibits robust performance and physiological consistency across multiple states, outputs logically coherent PCG sequences, and maintains accurate coupling parameter extraction. The introduced LLM module integrates pre-computed multimodal metrics under rule-based constraints, focusing on improving report readability and consistency rather than autonomous medical interpretation or diagnosis. Thus, it serves as an interpretation-friendly interface for non-professional settings, not as a clinical decision-support tool.

Leveraging its automation, lightweight design, and deployability, the system shows potential for cardiac monitoring in community and home environments. It supports long-term monitoring without specialized operation, aiding in the observation and tracking of heart rate variations, load fluctuations, and changes in electro-mechanical coupling abnormalities in primary care contexts. However, it must be emphasized that the system remains a research prototype, lacking the clinical validation and regulatory certification required for medical devices, and it should not be used for medical decision-making. Future product development would require substantial work in clinical validation, quality management, data security, and privacy protection.

This study proposes a systematic methodology for validation and engineering feasibility demonstration. Its core value lies in providing a proof-of-concept for subsequent long-term scenario research, rather than direct application in clinical or daily health decision-making. From an algorithmic perspective, this work exemplifies a stability-oriented multimodal system design, where physiologically constrained coupling is used to enhance temporal robustness and system consistency under nonstationary, low-quality signal conditions, rather than pursuing accuracy-driven standalone detection or classification optimization. The currently constructed ECG-PCG coupled feature set focuses on systematic

exploration and has not yet been optimized for specific applications. Future research may target specific monitoring tasks (e.g., home follow-up for heart failure patients) by conducting feature selection and lightweight model design to enhance computational efficiency, system stability, and practical deployment feasibility. Overall, this study provides a methodologically viable technical framework for daily cardiac monitoring in low-resource, uncontrolled environments, laying the foundation for further engineering and standardized development of multimodal cardiac assessment systems.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/a19020117/s1>, Detailed Report File S1: “In_the_resting_scenario_example.txt”; Detailed Report File S2: “In_the_exercise_scenario_example.txt”.

Author Contributions: Conceptualization, Y.T. and P.S.; methodology, Y.T. and P.S.; software, Y.T. and F.C.; validation, Y.T., Y.L. and F.C.; data curation, Y.T., Y.L. and F.C.; writing—original draft preparation, Y.T.; writing—review and editing, Y.T., Y.L. and P.S.; visualization, Y.T.; supervision, P.S. and Y.L.; project administration, P.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Restrictions apply to the datasets. The dataset presented in this article cannot be directly obtained, as it pertains to an ongoing project at the University of Shanghai for Science and Technology. Requests to access the datasets should be directed to the first author, Yi Tang.

Acknowledgments: The authors would like to acknowledge the use of the EPHNOGRAM Database and the OSET Benchmark Algorithm in the preparation of this research. The EPHNOGRAM Database is referenced from [16], and the OSET Benchmark Algorithm, developed by [23], was instrumental for the algorithmic analysis. The authors sincerely thank all the participants for their time and contribution to this study.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

LLM	Large Language Model
ECG	Electrocardiography
PCG	Phonocardiography
S1	First Heart Sound
S2	Second Heart Sound

Appendix A

To systematically evaluate the capability boundaries of lightweight open-source large language models in cardiac health semantic generation tasks, this study designed and implemented a comprehensive performance test. This test aims to quantify model performance across four key dimensions: medical knowledge comprehension, numerical logical reasoning, structured report generation, and error detection, thereby providing empirical evidence for model selection and application strategies.

The testing environment is based on the MATLAB R2025b platform, deploying two open-source models with 1.5 billion parameters locally via the Ollama framework: DeepSeek-R1:1.5b and Qwen2.5:1.5b. The evaluation was based on physiological parameters from a real ECG-PCG recording during the recovery phase after exercise in a 26-year-old healthy female. Through programmatic construction and transformation, 20 standardized test cases covering four evaluation dimensions were generated. Core physiological pa-

rameters included mean heart rate (101.6 BPM), heart rate variability SDNN (0.0159 s), autonomic balance LF/HF ratio (2.80), and RR interval (0.596 s).

The evaluation system encompasses four structured dimensions, each comprising five specific test items. The medical knowledge comprehension dimension assesses understanding of fundamental concept definitions (e.g., heart rate), interpretation of specialized terminology (e.g., LF/HF ratio), recognition of normal physiological parameter ranges, and knowledge of exercise physiology fundamentals. The numerical logical reasoning dimension tests the model's ability to interpret physiological values appropriately in specific scenarios, explain the physiological significance of ratios, perform accurate unit conversions, and determine normal value ranges. The Report Generation Capability dimension evaluates the model's ability to summarize information, semantically integrate multiple indicators, generate practical recommendations, correctly use medical terminology, and adhere to specified output formats. The Error Detection Capability dimension assesses the model's ability to identify obvious numerical anomalies (e.g., resting heart rate of 300 BPM), contradictory data (e.g., abnormal LF/HF values post-exercise), unit errors, logical paradoxes, and domain knowledge errors.

The testing employs procedural automation. For each test item, the system sends standardized prompts to the model and receives and parses its natural language responses. Scoring is based on predefined quantitative rules: answers that fully meet expectations or correctly identify errors receive 1 point; partially compliant responses or those containing relevant core keywords receive 0.5 points; and responses that fail to meet expectations, fail to identify errors, or generate failures receive 0 points. The final score for each dimension is the average of its five sub-test scores, converted to a percentage. The composite score is the arithmetic mean of the four-dimensional scores. The testing process incorporates inter-request delays to ensure system stability.

Evaluation results indicate that both models performed poorly in medical knowledge comprehension (scoring 60/100), revealing significant confusion with domain terminology—such as incorrectly associating LF/HF ratios with machine learning or financial concepts. In numerical logical reasoning, while models could perform basic judgments (80/100), they frequently erred in precise calculations like unit conversions. Report generation capabilities diverged, with DeepSeek-R1 (80/100) outperforming Qwen2.5 (60/100) in data integration and structured output. For error detection, Qwen2.5 (80/100) demonstrated slightly superior logical inconsistency identification compared to DeepSeek-R1 (70/100). Both models scored 72.5 and 70.0, respectively, overall, placing them within a range requiring external constraints for cautious use.

This evaluation method systematically reveals the primary limitations of general-purpose language models with 1.5B parameters when directly applied to professional medical report generation, particularly the issues of insufficient domain knowledge and unreliable numerical reasoning. This empirical finding directly supports the core argument presented in Section 3.3 of the main text: a hybrid generation architecture—where templates govern core logic and models perform limited refinement—is essential to maintain linguistic flexibility while rigorously ensuring the safety and reliability of medical interpretation. This evaluation framework is scalable and provides methodological guidance for assessing the task applicability of lightweight models in other vertical domains.

Appendix B

Example 1: See the Supplementary Material “In the resting scenario example.txt”.

Example 2: See the Supplementary Material “In the exercise scenario example.txt”.

References

- Schulz, R.; Beach, S.R.; Czaja, S.J.; Martire, L.M.; Monin, J.K. Family Caregiving for Older Adults. *Annu. Rev. Psychol.* **2020**, *71*, 635–659. [\[CrossRef\]](#) [\[PubMed\]](#)
- Huang, B.; Hu, S.; Liu, Z.; Lin, C.-L.; Su, J.; Zhao, C.; Wang, L.; Wang, W. Challenges and prospects of visual contactless physiological monitoring in clinical study. *Npj Digit. Med.* **2023**, *6*, 231. [\[CrossRef\]](#) [\[PubMed\]](#)
- Moermans, V.R.; Mengelers, A.M.; Bleijlevens, M.H.; Verbeek, H.; de Casterle, B.D.; Milisen, K.; Capezuti, E.; Hamers, J.P. Caregiver decision-making concerning involuntary treatment in dementia care at home. *Nurs. Ethics* **2022**, *29*, 330–343. [\[CrossRef\]](#) [\[PubMed\]](#)
- Lee, C.Y.; Jeon, Y.H.; Fethney, J.; Watson, K.; Low, L.F.; Mowszowski, L.; Woods, R.T. The Association Between Caregiving Context and the Health and Well-Being of Carers and Their Care Recipients Living with Dementia: A Cross-Sectional Study. *J. Adv. Nurs.* **2025**, 1–14. [\[CrossRef\]](#)
- Ruggiano, N.; Brown, E.L.; Shaw, S.; Geldmacher, D.; Clarke, P.; Hristidis, V.; Bertram, J. The Potential of Information Technology to Navigate Caregiving Systems: Perspectives from Dementia Caregivers. *J. Gerontol. Soc. Work* **2019**, *62*, 432–450. [\[CrossRef\]](#)
- Bidwell, J.T.; Fauer, A.J.; Howe, R.J.; Saylor, M.A.; Lee, C.S.; López, J.E.; Godden, M.; Hinton, L. Older Adults with Cardiovascular Disease and Their Care Partners: An Analysis of Care Needs, Care Activities, and Care Partner Stress and Mental Health. *J. Cardiovasc. Nurs.* **2026**, *41*, E1–E10. [\[CrossRef\]](#)
- Grisot, R.; Laurent, P.; Migliaccio, C.; Dauvignac, J.Y.; Brulc, M.; Chiquet, C.; Caruana, J.P. Monitoring of Heart Movements Using an FMCW Radar and Correlation with an ECG. *IEEE Trans. Radar Syst.* **2023**, *1*, 423–434. [\[CrossRef\]](#)
- Weidlich, S.; Mannhart, D.; Serban, T.; Krisai, P.; Knecht, S.; Du Fay De Lavallaz, J.; Schaer, B.; Osswald, S.; Kuehne, M.; Sticherling, C.; et al. Smart-device ECGs and inconclusive results: Impact of ECG anomalies. *Eur. Heart J.* **2023**, *44*, ehad655–314. [\[CrossRef\]](#)
- He, M.; Deng, X.; Niu, W. Physicians' Electrocardiogram Interpretations. *JAMA Intern. Med.* **2021**, *181*, 721–722. [\[CrossRef\]](#)
- Zang, J.; An, Q.; Li, B.; Zhang, Z.; Gao, L.; Xue, C. A novel wearable device integrating ECG and PCG for cardiac health monitoring. *Microsyst. Nanoeng.* **2025**, *11*, 7. [\[CrossRef\]](#)
- Kalatehjari, E.; Hosseini, M.M.; Harimi, A.; Abolghasemi, V. Advanced ensemble learning-based CNN-BiLSTM network for cardiovascular disease classification using ECG and PCG signal. *Biomed. Signal Process. Control* **2025**, *108*, 107846. [\[CrossRef\]](#)
- Sun, C.; Liu, X.; Liu, C.; Wang, X.; Liu, Y.; Zhao, S.; Zhang, M. Enhanced CAD Detection Using Novel Multi-Modal Learning: Integration of ECG, PCG, and Coupling Signals. *Bioengineering* **2024**, *11*, 1093. [\[CrossRef\]](#) [\[PubMed\]](#)
- Chakir, F.; Jilbab, A.; Nacir, C.; Hammouch, A. Recognition of cardiac abnormalities from synchronized ECG and PCG signals. *Phys. Eng. Sci. Med.* **2020**, *43*, 673–677. [\[CrossRef\]](#) [\[PubMed\]](#)
- Li, J.; Ke, L.; Du, Q.; Ding, X.; Chen, X. Research on the Classification of ECG and PCG Signals Based on BiLSTM-GoogLeNet-DS. *Appl. Sci.* **2022**, *12*, 11762. [\[CrossRef\]](#)
- Wang, J.; Zang, J.; An, Q.; Wang, H.; Zhang, Z. A pooling convolution model for multi-classification of ECG and PCG signals. *Comput. Methods Biomech. Biomed. Eng.* **2025**, *28*, 628–641. [\[CrossRef\]](#)
- Kazemnejad, A.; Karimi, S.; Gordany, P.; Clifford, G.D.; Sameni, R. An open-access simultaneous electrocardiogram and phonocardiogram database. *Physiol. Meas.* **2024**, *45*, 055005. [\[CrossRef\]](#)
- Karimi, S.; Karimi, S.; Shah, A.J.; Clifford, G.D.; Sameni, R. Electromechanical Dynamics of the Heart: A Study of Cardiac Hysteresis During Physical Stress Test. *arXiv* **2024**, arXiv:2410.19667. [\[CrossRef\]](#)
- Schmitt, E.E.; McNair, B.D.; Polson, S.M.; Cook, R.F.; Bruns, D.R. Mechanisms of Exercise-Induced Cardiac Remodeling Differ Between Young and Aged Hearts. *Exerc. Sport Sci. Rev.* **2022**, *50*, 137–144. [\[CrossRef\]](#)
- Butterworth, J.B.; Dekerle, J.; Greenhouse-Tucknott, A.; Critchley, H.D.; Smeeton, N.J. Having the Heart to Exercise Control: Cardiac Interoception Influences Self-Paced Exercise Regulation. *Eur. J. Sport Sci.* **2025**, *25*, e12263. [\[CrossRef\]](#)
- Balzarotti, S.; Biassoni, F.; Colombo, B.; Ciceri, M.R. Cardiac vagal control as a marker of emotion regulation in healthy adults: A review. *Biol. Psychol.* **2017**, *130*, 54–66. [\[CrossRef\]](#)
- Dong, H.; He, S.; Wu, W.; Zhang, X.; Li, M.; Millham, R.; Bian, G.; Wu, W. A Novel Approach to Explore Internal Cardiac Electrophysiological Pattern under Emotional Stress. *IEEE J. Biomed. Health Inform.* **2025**, 1–12. [\[CrossRef\]](#)
- Salo, K.I.; Pauw, L.S.; Schubotz, R.I.; Milek, A. At the heart of couple conflict: Emotion regulation and cardiac reactivity. *Int. J. Psychophysiol.* **2025**, *212*, 112581. [\[CrossRef\]](#) [\[PubMed\]](#)
- Sameni, R. The Open-Source Electrophysiological Toolbox (OSET), Version 3.14, 2006–2023. Available online: <https://github.com/alphamericslab/OSET.git> (accessed on 15 March 2024).
- Studinger, P.; Goldstein, R.; Taylor, J.A. Mechanical and neural contributions to hysteresis in the cardiac vagal limb of the arterial baroreflex. *J. Physiol.* **2007**, *583*, 1041–1048. [\[CrossRef\]](#) [\[PubMed\]](#)
- Cooke, W.H.; Hoag, J.B.; Crossman, A.A.; Kuusela, T.A.; Tahvanainen, K.U.; Eckberg, D.L. Human responses to upright tilt: A window on central autonomic integration. *J. Physiol.* **1999**, *517*, 617–628. [\[CrossRef\]](#) [\[PubMed\]](#)
- Liu, Q.; Yan, B.P.; Yu, C.M.; Zhang, Y.T.; Poon, C.C. Attenuation of systolic blood pressure and pulse transit time hysteresis during exercise and recovery in cardiovascular patients. *IEEE Trans. Biomed. Eng.* **2014**, *61*, 346–352. [\[CrossRef\]](#)

27. Chakir, F.; Jilbab, A.; Nacir, C.; Hammouch, A. Phonocardiogram signals processing approach for PASCAL Classifying Heart Sounds Challenge. *Signal Image Video Process.* **2018**, *12*, 1149–1155. [\[CrossRef\]](#)
28. Deng, S.-W.; Han, J.-Q. Towards heart sound classification without segmentation via autocorrelation feature and diffusion maps. *Future Gener. Comput. Syst.* **2016**, *60*, 13–21. [\[CrossRef\]](#)
29. Varghees, V.N.; Ramachandran, K.I. Effective Heart Sound Segmentation and Murmur Classification Using Empirical Wavelet Transform and Instantaneous Phase for Electronic Stethoscope. *IEEE Sens. J.* **2017**, *17*, 3861–3872. [\[CrossRef\]](#)
30. Arora, V.; Leekha, R.; Singh, R.; Chana, I. Heart sound classification using machine learning and phonocardiogram. *Mod. Phys. Lett. B* **2019**, *33*, 1950321. [\[CrossRef\]](#)
31. Silva, A.; Teixeira, R.; Fontes-Carvalho, R.; Coimbra, M.; Renna, F. On the Impact of Synchronous Electrocardiogram Signals for Heart Sounds Segmentation. In *Proceedings of the 2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Sydney, Australia, 24–27 July 2023*; IEEE: New York, NY, USA, 2023; pp. 1–5.
32. Springer, D.B.; Tarassenko, L.; Clifford, G.D. Logistic Regression-HSMM-Based Heart Sound Segmentation. *IEEE Trans. Biomed. Eng.* **2016**, *63*, 822–832. [\[CrossRef\]](#)
33. Schmidt, S.E.; Toft, E.; Holst-Hansen, C.; Graff, C.; Struijk, J.J. Segmentation of heart sound recordings from an electronic stethoscope by a duration dependent Hidden-Markov Model. In *Proceedings of the 2008 Computers in Cardiology, Bologna, Italy, 14–17 September 2008*; IEEE: New York, NY, USA, 2008; pp. 345–348.
34. Han, H.; Xiang, M.; Lian, C.; Liu, D.; Zeng, Z. A Multimodal Deep Neural Network for ECG and PCG Classification with Multimodal Fusion. In *Proceedings of the 2023 13th International Conference on Information Science and Technology (ICIST), Cairo, Egypt, 8–14 December 2023*; IEEE: New York, NY, USA, 2023; pp. 124–128.
35. Singh, S.A.; Meitei, T.G.; Majumder, S. 6—Short PCG classification based on deep learning. In *Deep Learning Techniques for Biomedical and Health Informatics*; Agarwal, B., Balas, V.E., Jain, L.C., Poonia, R.C., Manisha, Eds.; Academic Press: Cambridge, MA, USA, 2020; pp. 141–164.
36. Kiranyaz, S.; Zabihi, M.; Rad, A.B.; Ince, T.; Hamila, R.; Gabbouj, M. Real-time phonocardiogram anomaly detection by adaptive 1D Convolutional Neural Networks. *Neurocomputing* **2020**, *411*, 291–301. [\[CrossRef\]](#)
37. Siontis, K.C.; Noseworthy, P.A.; Attia, Z.I.; Friedman, P.A. Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. *Nat. Rev. Cardiol.* **2021**, *18*, 465–478. [\[CrossRef\]](#)
38. Kim, S.H.; Lim, K.R.; Seo, J.-H.; Ryu, D.R.; Lee, B.-K.; Cho, B.-R.; Chun, K.J. Higher heart rate variability as a predictor of atrial fibrillation in patients with hypertension. *Sci. Rep.* **2022**, *12*, 3702. [\[CrossRef\]](#)
39. Zhao, W.; Wu, C.; Fan, Y.; Zhang, X.; Qiu, P.; Sun, Y.; Zhou, X.; Wang, Y.; Zhang, Y.; Yu, Y.; et al. An Agentic System for Rare Disease Diagnosis with Traceable Reasoning. *arXiv* **2025**, arXiv:2506.20430. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.