ECG-RSI: Towards Recursive Self-Improvement via Feedback-Driven Data Synthesis

Anonymous authors
Paper under double-blind review

ABSTRACT

ECG multimodal language models depend on instruction data that connect waveform evidence to interpretation, while GAN- and diffusion-based generators provide complementary ways to expand signal coverage. A central challenge is to construct examples with specified ECG changes, consistent questions, and checkable evidence. We propose ECG-RSI, centered on reverse waveform perturbation and bidirectional signal–instruction synthesis. The forward route derives questions and explanations from real training-side waveforms; the reverse route translates target features or training-question distractors into constrained waveform edits, remeasures the resulting signal, and constructs supervision from the achieved evidence. Target attainment and non-target preservation jointly constrain acceptance. Two feedback loops connect this data engine to learning: construction feedback guides operator and prompt repair, while learner feedback specifies which capabilities, difficulty levels, and evidence representations to construct next. This process adjusts training data to the model's current errors and capabilities. On ECGBench, ECG-RSI raises the aggregate S7 score from 30.15 to 46.61 across four rounds, a gain of 16.46 points over the base model.

1 INTRODUCTION

ECG interpretation requires connecting visible waveform evidence to measurements and clinical conclusions. Multimodal ECG models learn these connections from instruction data (Liu et al., 2024; Lan et al., 2025; Jin et al., 2026), while conditional signal generators expand the available waveform distribution (Chen et al., 2022; Lai et al., 2025). For targeted training, the challenge is to turn a missing capability into a new example whose relevant ECG factors are controlled and whose supervision agrees with the resulting signal.

Our central data-construction idea is **reverse waveform perturbation within bidirectional signal–instruction synthesis**. Forward synthesis starts from an observed ECG and constructs evidence-supported questions. Reverse synthesis starts from a desired feature contrast, including an eligible distractor from a training question, and edits a training waveform to realize that contrast before regenerating the question and explanation. This provides a way to construct missing examples rather than only rephrase existing reports.

A complementary question is what should happen *after* a learner has been trained: **how should its remaining errors change the data and the procedure used to train the next learner?** The useful training distribution need not remain fixed as model capabilities change. Some examples address persistent diagnostic confusions, some expose newly forgotten capabilities, and others add little beyond what the model already knows.

The second challenge is learning utility. A concise teacher rationale may omit needed intermediate evidence, while a longer one may add little useful supervision. This motivates choosing difficulty and evidence representation alongside capability coverage, then checking the resulting training benefit.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

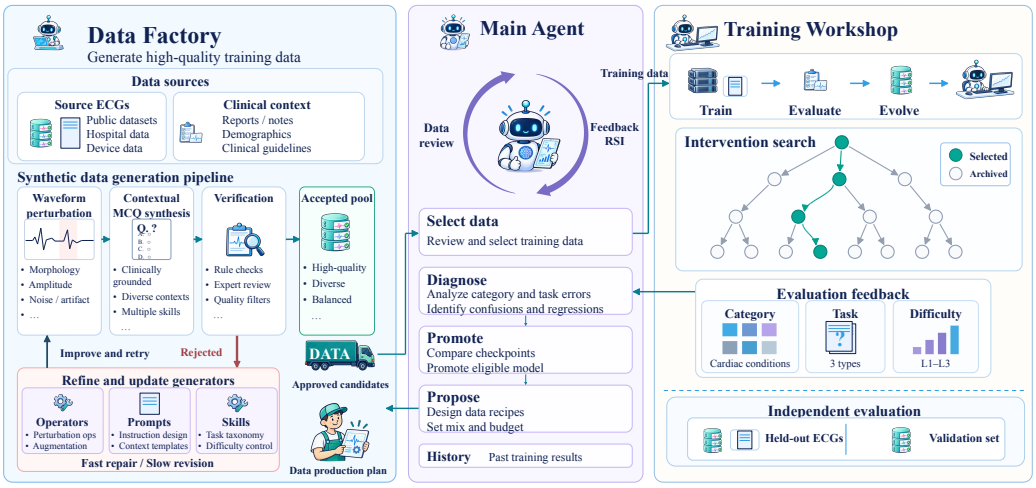

Figure 1: **ECG-RSI: feedback-driven data and model evolution.** The data factory constructs and verifies ECG questions, with rejected candidates driving local revisions of operators, prompts, and skills. The main agent diagnoses feedback, proposes interventions, and selects accepted data for training. The training workshop evaluates candidate checkpoints and supplies category-, task-, and difficulty-level feedback. Independent evaluation is separated from the update loop. The search tree and highlighted path illustrate candidate comparison and model selection.

1.1 From Data Construction to Data Evolution

We frame these challenges as data-centric recursive self-improvement (RSI): a construction loop tests whether the requested example is valid, and a learning loop tests whether accepted data improve the learner. Their connection makes data construction responsive to learner needs.

1.2 ECG-RSI: Improving Training Data from Model Errors

We propose **ECG-RSI (ECG Recursive Self-Improvement)**: identify ECG questions the model struggles with, generate targeted training examples, check their correctness, and compare whether training on them improves the model. Figure 1 summarizes the workflow.

Generate questions from observed errors. We first identify ECG categories and question types with frequent errors, then select relevant training waveforms to construct questions. We can also make constrained changes to an existing waveform, such as prolonging the PR interval, and generate a question and answer from the resulting signal.

Check examples and repair generation. We remeasure the waveform to check whether the target change was achieved, whether other features were affected, and whether the question, answer, and explanation agree with the image. Recurring failures guide improvements to waveform-editing methods and question prompts.

Compare training outcomes and repeat. All previous training data are retained. A/B/C each add 10,000 accepted questions, focusing respectively on weak categories, confusing categories, and difficulty or question-type adjustments. The three candidates train from the same model version and are compared on a fixed development set. Only the best qualifying model and its new data are retained; if none qualifies, the original model is kept. Its errors then guide the next round of question generation. Training and selection rules are detailed in Appendix A.4.

The method asks two questions: are the new examples correct, and does training on them improve the model? Evaluation after each training round determines whether the update is retained.

1.3 CONTRIBUTIONS

1. **Reverse waveform perturbation for targeted data construction.** We formulate feature-level requests as constrained edits of training-side ECGs.
2. **Bidirectional signal–instruction synthesis.** We connect a forward waveform-to-question route with a reverse target-to-waveform-to-question route. Training-question distractors can propose new construction targets; accepted outputs are rendered and paired with evidence-supported questions, compatible context, and explanations.
3. **A dual-feedback framework for construction and learning.** Construction feedback guides versioned operator, prompt, and workflow repair; learner feedback guides category–difficulty–representation recipes.

2 RELATED WORK

2.1 ECG MULTIMODAL LEARNING AND EVIDENCE-GROUNDED SUPERVISION

PULSE constructs ECGInstruct through diverse sources, ECG rendering, and instruction generation, and introduces ECGBench for evaluation (Liu et al., 2024). GEM combines signal and image representations with knowledge-guided supervision linking interpretations to measurable features (Lan et al., 2025). ECG-R1 incorporates medical protocols into instruction generation and uses diagnostic-evidence rewards to supervise intermediate reasoning and final diagnoses (Jin et al., 2026). These methods emphasize task coverage, cross-modal alignment, and evidence quality, respectively. ECG-RSI adds a construction axis: generating a requested feature contrast through a paired waveform intervention and deriving supervision from the achieved signal.

ECG Mirage further distinguishes ECG neglect from ECG confusion by comparing matched, outcome-discordant mismatched, and absent ECG inputs while holding clinical text fixed (Liang et al., 2026). Its results on MDS-ED show that strong clinical prediction does not necessarily imply use of patient-specific ECG evidence; the study mitigates this underutilisation through visual prompt tuning with a frozen VLM backbone. This motivates evaluating dependence on the correct waveform separately from aggregate task performance, complementing ECG-RSI's focus on constructing paired waveform interventions.

2.2 CONDITIONAL ECG SYNTHESIS AND CONTROLLED INTERVENTIONS

ME-GAN uses disease-conditioned adversarial generation (Chen et al., 2022); SSSD-ECG and its extension apply diffusion models to diagnostic conditioning (Lopez Alcaraz & Strodthoff, 2023; Skorik & Avetisyan, 2024). Text-To-ECG generates signals from free-text reports (Chung et al., 2023), while DiffuSETS incorporates report embeddings and patient attributes (Lai et al., 2025). MIDT-ECG and FlowECG further investigate morphological fidelity, inter-lead coherence, and sampling efficiency (Huang et al., 2025; Bondar et al., 2025). These works offer complementary choices of conditioning information and generative architecture.

ECG-RSI focuses on paired edits of an existing training waveform, with explicit target-change and non-target-preservation checks. The two synthesis directions and their implementation are specified in Section 3.1.1.

2.3 FEEDBACK-DRIVEN DATA EVOLUTION AND RECURSIVE SELF-IMPROVEMENT

Feedback methods differ in the object they update. Self-Refine and Reflexion revise responses or persistent textual memory (Madaan et al., 2023; Shinn et al., 2023); Self-Instruct and Evol-Instruct construct instruction corpora for subsequent fine-tuning (Wang et al., 2023; Xu et al., 2024). STaR iteratively generates correct-answer rationales and fine-tunes the learner, providing a precedent for refreshing supervision through learner updates (Zelikman et al., 2022). CoEvolve uses learner behavior to guide validated task synthesis and agent training (Yang et al., 2026), and NeoHorse-1 uses capability and interaction evidence to shape post-training mixtures (NeoHorse Team, 2026). Absolute Zero and VisPlay couple task generation with solver improvement through reinforcement learning, using executable-code verification and image-conditioned task signals, respectively (Zhao et al., 2025; He et al., 2025).

A further line updates the improvement procedure. STOP and DGM modify improver code or agent workflows (Zelikman et al., 2024; Zhang et al., 2025). MetaSkill-Evolve separates task-skill changes from updates to the mechanisms proposing and managing revisions (Wang et al., 2026). Dream-RSI and EvoX investigate adaptive exploration and evolving search strategies (Zheng et al., 2026; Liu et al., 2026b), while RecurSE studies evolving self-evaluation (Liu et al., 2026a). These settings distinguish task, learner, workflow, and evaluator updates.

3 ECG-RSI: METHOD

ECG-RSI consists of two connected components. **Data synthesis** retrieves or edits training-side ECG signals and constructs evidence-supported questions. Construction failures drive local revisions of operators, prompts, and reusable skills. **Feedback-driven RSI** diagnoses the accepted student's errors, proposes direct reinforcement, confusion contrast, and difficulty- and task-aware curriculum interventions.

3.1 REVERSE PERTURBATION FOR WAVEFORM AND QUESTION SYNTHESIS

We expand training data by translating target ECG features into constrained waveform edits, verifying the resulting measurements, and constructing evidence-supported questions and answers. The forward route supplies questions from existing signals; reverse perturbation supplies additional paired feature contrasts.

3.1.1 BIDIRECTIONAL SYNTHESIS WITH REVERSE WAVEFORM PERTURBATION

Forward synthesis measures features of existing training waveforms and generates matching questions, answers, and explanations. Reverse synthesis edits a training waveform to match a target feature, then remeasures and checks it before generating questions and answers (see Appendix A.3 and Figure 3).

3.1.2 PARAMETER EXTRACTION AND SELECTION

A waveform delineator extracts candidate heart-rate, interval, and amplitude features. The parameter selector combines extraction reliability and diagnostic utility to choose edit parameters; measurement conversions and selection equations are given in Appendix A.6.

3.1.3 QUESTION CONSTRUCTION, CONTEXT, AND EXPLANATIONS

Stage-wise prompts for construction, review, repair, and learner feedback are documented in Appendix D. Question prompts, waveform-construction examples, and complete responses are provided in Appendix E; the taxonomy of the 23 ECG tasks is shown in Appendix C.

3.2 FEEDBACK SYSTEM AND TREE EVOLUTION

The feedback system connects the accepted learner to the synthesis pipeline in Section 3.1. Learner errors determine data requests, construction failures guide local repair, and same-parent training branches test which candidate recipe should continue the accepted lineage.

3.2.1 SYSTEM STATE AND TWO FEEDBACK INTERFACES

Let θ_t be the accepted learner, D_t the accepted training pool, and ϕ_t the construction policy: source retrieval rules, operator configurations, question prompts, and reusable workflows. We separate the feedback set V_{fb} from a predesignated model-selection set V_{sel} within the development protocol. The following summarizes one accepted candidate path; bounded sibling selection is expanded in

Appendix A.4. Each iteration is

$$\begin{aligned} e_t &= \mathrm{Diagnose}(\theta_t, V_{\mathrm{fb}}), \\ q_t &= \mathrm{Plan}(e_t, D_t; B_t), \\ (A_t, \mathcal{L}_t, \phi_{t+1}) &= \mathrm{ConstructValidate}(q_t, S_{\mathrm{train}}, \phi_t), \\ \widetilde{\theta}_{t+1} &= \mathrm{Train}(\theta_t, A_t \cup R_t), \\ \theta_{t+1} &= \mathrm{Promote}(\theta_t, \widetilde{\theta}_{t+1}; V_{\mathrm{sel}}), \qquad D_{t+1} = D_t \cup A_t. \end{aligned} \tag{1}$$

Here B_t is the budget, q_t specifies the data to generate, R_t is historical training data used for replay, and $\mathcal{L}_t$ records data generation.

The system summarizes the categories, question types, and difficulty levels that the model struggles with on a fixed feedback set, then decides what data to generate and how much.

Generated waveforms and questions are checked against the requirements.

3.2.2 Student-Conditioned Data Recipes

Based on the current learner's error profile, candidate recipe a jointly allocates task category g, learner-relative difficulty d, and supervision representation s:

$$\pi_t^a(g, d, s) = p_t^a(g)\, p_t^a(d \mid g)\, p_t^a(s \mid g, d), \tag{2}$$

Representations include original teacher explanations, evidence-expanded explanations, and error-targeted explanations. Recipes register source constraints, replay, and training budgets; same-parent candidate training tests their utility. Additional settings and ECG examples are detailed in Appendix A.1.

3.2.3 Construction Feedback and Versioned Repair

Construction failures trigger repairs to operators, question/context prompts, or explanation prompts according to their cause. Candidate revisions are validated on the same source signals, requests, and budget with fixed acceptance thresholds. Validated revisions update the operator, prompt, or skill version. Each update records its changes, version, and validation outcome; details are provided in Appendix A.2.

3.2.4 Learner Update and Development-Only Acceptance

For accepted image-question-response triples, the learner update is ordinary supervised fine-tuning,

$$\mathcal{L}_{\mathrm{MCQ}} = -\sum_{\ell} \log p_\theta(y_\ell \mid I, Q, C, y_{<\ell}), \tag{3}$$

where C is optional context and y contains the supervised explanation and answer. Replay and new-data proportions are measured in training tokens as well as examples.

The language model is trainable, while the vision encoder and multimodal aligner are frozen. Sequence, batching, parallelism, and optimizer settings are listed in Appendix A.5, Table 8.

Before training, register a development-only promotion rule: primary score, non-target regression tolerances, valid-output requirements, and stopping rule. Apply it on V_{sel}.

The recipe-search procedure, candidate comparison rules, and training-lineage diagram are provided in Appendix A.4.

Current data protocol. In each round, A/B/C start from the same parent and retain the previous cumulative training set unchanged. Each adds 10,000 previously unused accepted questions, with no additional shared 10,000-question baseline set. The branches use direct reinforcement, confusion contrast, and difficulty/task progression, respectively, with the same historical replay manifest. Only the winner's 10,000 new questions enter the cumulative set retained in the next round.

4 EXPERIMENTS

We report main benchmark results, round-wise evolution, and validation-set performance in this order. Each benchmark retains its accuracy or F1 metric; S7 is the equal-weight mean of four accuracy and three F1 scores on a 0–100 scale.

4.1 MAIN BENCHMARK RESULTS

Table 1 compares the recorded scores of closed-source, open-source, and ECG-specific models with ECG-RSI on seven benchmarks. ECG-RSI R3 has the highest tabulated S7 (46.61), while individual-metric leaders vary across models.

Table 1: Current main benchmark results.

Model / method	**MMMU Acc.**	**CSN Acc.**	**G12 Acc.**	**ECGQA Acc.**	**PTB F1**	**CPSC F1**	**CODE15 F1**	**S7**
Proprietary models								
GPT-5.6-sol	56.00	58.70	60.20	32.10	37.40	24.30	42.80	44.50
Gemini-3.8-Flash	59.10	57.90	59.60	33.40	36.80	23.70	40.90	44.49
Open-source and base models								
LINGSHU-7B	47.60	45.20	46.80	26.30	33.87	18.60	36.50	36.41
LINGSHU-32B	51.30	48.90	50.40	28.70	39.45	20.80	38.20	39.68
HuatuoGPT-V-7B	44.80	41.60	43.20	25.40	31.15	17.30	33.90	33.91
Original Base	42.00	30.91	41.36	23.99	29.50	13.10	30.20	30.15
ECG-R1	38.12	44.87	47.05	28.63	22.91	17.88	34.55	33.43
PULSE	39.24	45.16	43.87	21.55	26.08	14.92	26.82	31.09
ECG-RSI								
ECG-RSI R0	42.00	30.91	41.36	23.99	29.50	13.10	30.20	30.15
ECG-RSI R3	**57.50**	**67.23**	**66.24**	**33.18**	**38.30**	**22.80**	**41.00**	**46.61**

All scores are on a 0–100 scale. Base denotes the initial model. ECG-R1 and PULSE are ECG-specific comparison models.

4.2 EVOLUTION ACROSS ROUNDS

Table 2 and Figure 2 show the R0–R3 trajectory of series R-A. Under the current protocol, the winning branch adds 10,000 previously unused accepted questions to the cumulative training set each round, alongside replay of all previous training data. Round-wise feedback statistics by question type, difficulty, and category remain in Appendix B.

Table 2: Seven benchmark metrics over R0–R3 for series R-A.

R-A round	MMMU Acc.	CSN Acc.	G12 Acc.	ECGQA Acc.	PTB F1	CPSC F1	CODE15 F1	S7
R0	42.00	30.91	41.36	23.99	29.50	13.10	30.20	30.15
R1	48.00	61.68	64.61	33.56	32.90	23.00	38.80	43.22
R2	56.00	64.86	66.39	34.55	36.40	23.90	41.50	46.23
R3	57.50	67.23	66.24	33.18	38.30	22.80	41.00	46.61

R0 is the starting model in R-A.

Appendix B, Figure 5, shows feedback errors by question type, difficulty, and ECG category/capability. Difficulty-level counts remain in Table 9.

4.3 Validation-Set Performance

Table 3 and Figure 2(c) summarize four-round changes on MedX and LUDB. MedX accuracy improves from 51.57% to 57.24%, a gain of 5.67 percentage points, with a small R2 decline. LUDB accuracy increases across the recorded rounds from 53.45% through 61.28% and 63.28% to 67.24%, a total gain of 13.79 percentage points.

Table 3: MedX and LUDB validation trajectories and round-to-round changes.

	Accuracy (%)				Change (pp)			
Dataset	R0	R1	R2	R3	R1−R0	R2−R1	R3−R2	R3−R0
MedX	51.57	55.94	55.86	57.24	+4.37	-0.08	+1.38	+5.67
LUDB	53.45	61.28	63.28	67.24	+7.83	+2.00	+3.96	+13.79

Figure 2 summarizes round-wise benchmark and validation performance together with historical strategy trajectories.

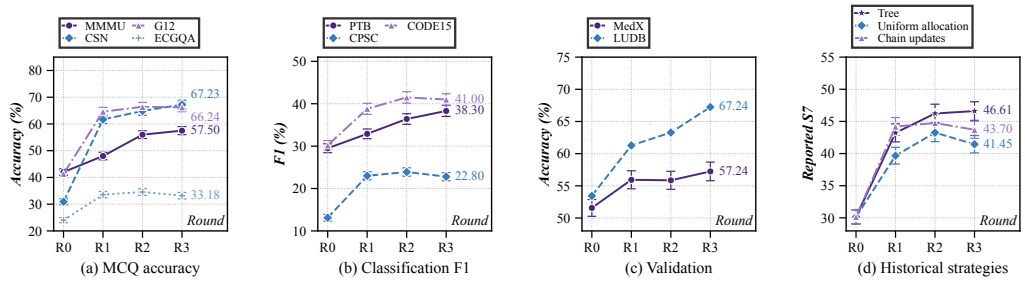

Figure 2: Round-wise performance and historical strategies. (a) Four MCQ accuracy metrics; (b) three classification F1 metrics; (c) MedX and LUDB validation accuracy; (d) reported S7 for tree, uniform-allocation, and chain trajectories. Error bars show the supplied ± 1 SD where available, not confidence intervals. LUDB uncertainty estimates are unavailable, so its curve has no error bars.

5 Ablation Studies

This section evaluates waveform synthesis, compares data construction and feedback selection from a common checkpoint, and studies synthesis-component updates and RSI evolution strategies.

5.1 Effectiveness of the Synthesis Method

Quantitative waveform-feature validation. We report waveform-feature error statistics from a single synthesis run of 10,000 samples in Table 4. For each feature, let $e_i = \hat{x}_i - x_i$ and compute

$$\bar{e} = \frac{1}{n}\sum_i e_i, \qquad \mathrm{MAE} = \frac{1}{n}\sum_i |e_i|,$$
$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_i e_i^2}, \qquad s_e^2 = \frac{1}{n-1}\sum_i (e_i - \bar{e})^2. \tag{4}$$

Table 4 reports errors separately for HR, RR, PR, QRS, QT, and QTc.

Blinded clinical expert review. Two physicians independently review 240 randomly sampled examples, stratified into 40 examples each for HR, RR, PR, QRS, QT, and QTc edits. The review covers waveform morphology, before/after changes, and the generated answers and explanations. Target

Table 4: Waveform-feature error statistics from a single synthesis run of 10,000 samples.

Feature	Unit	Bias $\bar{e}$	MAE	RMSE	Variance s_e^2
HR	bpm	0.63	2.14	2.22	4.71
RR	ms	3.11	15.28	15.36	238.62
PR	ms	0.57	7.71	7.78	63.84
QRS	ms	4.12	4.58	4.61	6.37
QT	ms	0.94	9.71	9.85	101.43
QTc	ms	0.86	10.22	10.34	114.08

Error is measured minus target. Bias, MAE, and RMSE use the listed unit; sample error variance uses bpm^2 or ms^2 and denominator $n - 1$. Statistics are reported to two decimals.

attainment is 0.82 (95% CI: 0.76–0.87), and non-target preservation is 0.79 (0.73–0.84). Morphological and lead consistency pass rates are 0.86 and 0.83, respectively. Answer accuracy reaches 0.88 (0.83–0.92), while evidence-supported explanations reach 0.76 (0.70–0.81); the expected-parameter reference rate is 0.09. Inter-physician Cohen's κ is 0.78 (0.71–0.85). Across edit types, target attainment ranges from 0.78 to 0.85 and explanation support from 0.72 to 0.78. The complete review procedure, overall metrics, and stratified results appear in Appendix B.4.

Five-model question-quality scores are provided in Appendix B.6, Figure 6.

5.2 Data Construction and Feedback Selection

We compare four data configurations starting from the same checkpoint with S7 = 43.22. Each configuration adds 10,000 questions for continued SFT, using ordinary training-side questions, forward-only synthesis, forward and reverse synthesis with fixed selection proportions, or forward and reverse synthesis with feedback-guided selection.

Table 5: Data-construction and selection ablation from the same checkpoint (S7 = 43.22), with 10,000 added questions per group. ΔS7 is relative to the starting checkpoint.

Group	Added data	S7	ΔS7
A	Ordinary training-side questions	44.15	+0.93
B	Forward-only synthesis	44.27	+1.05
C	Forward + reverse, fixed selection proportions	45.56	+2.34
D	Forward + reverse, feedback-guided selection	**46.23**	**+3.01**

Table 5 shows that ordinary questions and forward-only synthesis reach S7 scores of 44.15 and 44.27, respectively. Adding reverse synthesis with fixed selection proportions raises S7 to 45.56, 1.29 points above forward-only synthesis; feedback-guided selection further raises it to 46.23, an additional 0.67 points.

5.3 Synthesis-Component Ablation

Table 6 compares synthesis updates at a fixed target of 10,000 accepted questions. The full method requires 16,193 candidates to reach this target, versus 36,723 with operator updates frozen and 22,398 with skill updates frozen (2.27 and 1.38 times as many, respectively). Freezing updates reduces question-quality scores and downstream R1 S7.

Table 6: R1 synthesis cost and training outcomes for a fixed target of 10,000 accepted questions.

Method	Candidates required	Pass (%)	Question score	S7	ΔS7
Full method	16,193	61.76	82.30	43.22	0.00
Freeze operator updates	36,723	27.23	71.90	39.29	-3.93
Freeze skill updates	22,398	44.65	79.20	40.15	-3.07

Each method targets 10,000 accepted questions and uses them for training. Candidate counts include rejected questions; ΔS7 is relative to the full method.

5.4 Comparison of RSI Evolution Strategies

Table 7 and Appendix B.7, Figure 7(c) summarize the historical tree, uniform-allocation, and chain trajectories in R-A. The tree trajectory continues improving, finishing 5.16 and 2.91 points above the two controls and 1.86 points above the chain's best round.

Table 7: R-A strategy comparison: S7 trajectories over R0–R3 and saved changes.

Strategy	R0	R1	R2	R3	R3−R0	R3−R2	R3 vs. tree
Tree (R-A)	30.15	43.22	46.23	46.61	+16.46	+0.38	0.00
Uniform allocation (R-A)	30.15	39.67	43.25	41.45	+11.30	-1.80	-5.16
Chain updates (R-A)	30.15	44.20	44.75	43.70	+13.55	-1.05	-2.91

All three rows belong to the R-A source summary. Scores and changes are transcribed from the source.

5.5 Results Analysis

Data validity and learning value. Remeasurement after waveform perturbation, provenance checks, and multi-agent review aim to ensure consistency among each candidate question's signal, context, answer, and explanation.

Multi-round training and strategy comparison. The tree strategy continues to improve in later rounds, while uniform allocation and chain updates decline in R3, showing that data organization and update strategies affect subsequent performance.

6 Conclusion

We presented ECG-RSI, a feedback-driven framework for ECG instruction-data synthesis and iterative model improvement. Its core contribution is a bidirectional signal–instruction construction pipeline: forward synthesis derives questions from waveform evidence, while inverse perturbation constructs targeted feature contrasts and regenerates questions and explanations after remeasurement. Construction checks connect waveform validity to supervision consistency, and the feedback interface translates learner errors into data recipes and candidate branches for subsequent training.

The seven-benchmark aggregate S7 rises from 30.15 to 46.61, and MedXpert ECG validation accuracy increases from 51.57% to 57.24%. Synthesis-component ablations further associate operator and skill updates with lower candidate-generation cost and better downstream performance.

Declaration of AI Assistance. AI tools assisted with language polishing, formatting, and figure preparation.

Reproducibility Statement. The main text and supplementary materials document the proposed method and experimental setup.

REFERENCES

Vitalii Bondar, Serhii Semenov, Vira Babenko, and Dmytro Holovniak. FlowECG: Using flow matching to create a more efficient ECG signal generator. *arXiv preprint arXiv:2509.10491*, 2025.

Jintai Chen, Kuanlun Liao, Kun Wei, Haochao Ying, Danny Z. Chen, and Jian Wu. ME-GAN: Learning panoptic electrocardio representations for multi-view ECG synthesis conditioned on heart diseases. *arXiv preprint arXiv:2207.10670*, 2022.

Hyunseung Chung, Jiho Kim, Joon-myoung Kwon, Ki-Hyun Jeon, Min Sung Lee, and Edward Choi. Text-To-ECG: 12-lead electrocardiogram synthesis conditioned on clinical text reports. *arXiv preprint arXiv:2303.09395*, 2023.

Yicheng He, Chengsong Huang, Zongxia Li, Jiaxin Huang, and Yonghui Yang. VisPlay: Self-Evolving Vision-Language Models from Images. *arXiv preprint arXiv:2511.15661*, 2025. URL https://arxiv.org/abs/2511.15661.

Zhuoyi Huang, Nutan Sahoo, Anamika Kumari, Girish Kumar, Kexuan Cai, Shixing Cao, Yue Kang, Tian Xia, Somya Chatterjee, Nicholas Hausman, Aidan Jay, Eric S. Rosenthal, Soundar Srinivasan, Sadid Hasan, Alex Fedorov, and Sulaiman Vesal. High-fidelity synthetic ECG generation via mel-spectrogram informed diffusion training. *arXiv preprint arXiv:2510.05492*, 2025.

Jiarui Jin, Haoyu Wang, Xingliang Wu, Xiaocheng Fang, Xiang Lan, Zihan Wang, Deyun Zhang, Bo Liu, Yingying Zhang, Xian Wu, Hongyan Li, and Shenda Hong. ECG-R1: Protocol-guided and modality-agnostic MLLM for reliable ECG interpretation. *arXiv preprint arXiv:2602.04279*, 2026.

Yongfan Lai, Jiabo Chen, Deyun Zhang, Yue Wang, Shijia Geng, Hongyan Li, and Shenda Hong. DiffuSETS: 12-lead ECG generation conditioned on clinical text reports and patient-specific information. *arXiv preprint arXiv:2501.05932*, 2025.

Xiang Lan, Feng Wu, Kai He, Qinghao Zhao, Shenda Hong, and Mengling Feng. GEM: Empowering MLLM for grounded ECG understanding with time series and images. *arXiv preprint arXiv:2503.06073*, 2025.

Jinning Liang, Mingcheng Zhu, and Tingting Zhu. ECG Mirage: Revealing and mitigating the underutilisation of ECGs in vision-language models for clinical prediction. *arXiv preprint arXiv:2609.21755*, 2026. URL https://arxiv.org/abs/2609.21755.

Kaiyuan Liu, Ziyuan Zhuang, Rongxiang Weng, and Jieping Ye. RecurSE: Bounded Recursive Self-Evaluation for LLM Rubric Judges. *arXiv preprint arXiv:2608.24231*, 2026a. URL https://arxiv.org/abs/2608.24231.

Ruoqi Liu, Yuelin Bai, Xiang Yue, and Ping Zhang. Teach multimodal LLMs to comprehend electrocardiographic images. *arXiv preprint arXiv:2410.19008*, 2024.

Shu Liu, Shubham Agarwal, Monishwaran Maheswaran, Mert Cemri, Zhifei Li, Qiuyang Mang, Ashwin Naren, Ethan Boneh, Audrey Cheng, Melissa Z. Pan, Alexander Du, Kurt Keutzer, Alvin Cheung, Alexandros G. Dimakis, Koushik Sen, Matei Zaharia, and Ion Stoica. EvoX: Meta-Evolution for Automated Discovery. *arXiv preprint arXiv:2602.23413*, 2026b. URL https://arxiv.org/abs/2602.23413.

Juan Miguel Lopez Alcaraz and Nils Strodthoff. Diffusion-based conditional ECG generation with structured state space models. *arXiv preprint arXiv:2301.08227*, 2023.

Aman Madaan, Niket Tandon, Prakhar Gupta, Skyler Hallinan, Luyu Gao, Sarah Wiegreffe, Uri Alon, Nouha Dziri, Shrimai Prabhumoye, Yiming Yang, Shashank Gupta, Bodhisattwa Prasad Majumder, Katherine Hermann, Sean Welleck, Amir Yazdanbakhsh, and Peter Clark. Self-Refine: Iterative Refinement with Self-Feedback. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2023. URL https://arxiv.org/abs/2303.17651.

NeoHorse Team. NeoHorse-1: Towards Recursive Self-Improvement via Agentic Post-Training with Routing Harness. *arXiv preprint arXiv:2609.08183*, 2026. URL https://arxiv.org/abs/2609.08183.

Noah Shinn, Federico Cassano, Edward Berman, Ashwin Gopinath, Karthik Narasimhan, and Shunyu Yao. Reflexion: Language Agents with Verbal Reinforcement Learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2023. URL https://arxiv.org/abs/2303.11366.

Sergey Skorik and Aram Avetisyan. SSSD-ECG-nle: New label embeddings with structured state-space models for ECG generation. *arXiv preprint arXiv:2407.11108*, 2024.

Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A. Smith, Daniel Khashabi, and Hannaneh Hajishirzi. Self-Instruct: Aligning Language Models with Self-Generated Instructions. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL)*, 2023. URL https://arxiv.org/abs/2212.10560.

Zefeng Wang, Minxi Yan, Jinhe Bi, Sikuan Yan, Volker Tresp, and Yunpu Ma. MetaSkill-Evolve: Recursive Self-Improvement of LLM Agents via Two-Timescale Meta-Skill Evolution. *arXiv preprint arXiv:2607.05297*, 2026. URL https://arxiv.org/abs/2607.05297.

Can Xu, Qingfeng Sun, Kai Zheng, Xiubo Geng, Pu Zhao, Jiazhan Feng, Chongyang Tao, Qingwei Lin, and Daxin Jiang. WizardLM: Empowering large pre-trained language models to follow complex instructions. In *International Conference on Learning Representations (ICLR)*, 2024. URL https://arxiv.org/abs/2304.12244.

Shidong Yang, Ziyu Ma, Tongwen Huang, Yiming Hu, Yong Wang, and Xiangxiang Chu. CoEvolve: Training LLM Agents via Agent-Data Mutual Evolution. *arXiv preprint arXiv:2604.15840*, 2026. URL https://arxiv.org/abs/2604.15840.

Eric Zelikman, Yuhuai Wu, Jesse Mu, and Noah D. Goodman. STaR: Bootstrapping Reasoning With Reasoning. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2022. URL https://arxiv.org/abs/2203.14465.

Eric Zelikman, Eliana Lorch, Lester Mackey, and Adam Tauman Kalai. Self-Taught Optimizer (STOP): Recursively Self-Improving Code Generation. In *Conference on Language Modeling (COLM)*, 2024. URL https://arxiv.org/abs/2310.02304.

Jenny Zhang, Shengran Hu, Cong Lu, Robert Lange, and Jeff Clune. Darwin Gödel Machine: Open-Ended Evolution of Self-Improving Agents. *arXiv preprint arXiv:2505.22954*, 2025. URL https://arxiv.org/abs/2505.22954. Version 3, revised March 2026.

Andrew Zhao, Yiran Wu, Yang Yue, et al. Absolute Zero: Reinforced Self-play Reasoning with Zero Data. *arXiv preprint arXiv:2505.03335*, 2025. URL https://arxiv.org/abs/2505.03335.

Tong Zheng, Xidong Wu, Zheng Zhang, Zhankui He, Chaoyi Zhang, Benjamin Coleman, Ruoqiao Wei, Di Bai, Haolin Liu, Rui Liu, Xue Wang, Yue Zhuan, Wang-Cheng Kang, Renkai Xiang, Heng Huang, Xinwu Cheng, and Yunsong Guo. Dream-RSI: Recursive Self-Improvement through Evolving Worlds. Technical report, Google and collaborating institutions, 2026. URL https://dream-rsi.com/assets/dream-rsi.pdf.

A Supplementary Method Details

A.1 Student-Conditioned Data Recipes

Equation (2) defines the joint recipe over category, difficulty, and supervision representation.

Here g is a task/capability category, d is learner-relative difficulty, and s is a supervision representation. These are task categories rather than additional input modalities. Each recipe also registers retrieval/edit constraints, replay, and budgets. The operator and prompt library ϕ determines which requests can be constructed. Selecting among fixed templates changes the recipe; editing a template or operator changes the library version. These interventions must be evaluated separately.

Difficulty estimation. With the current learner and decoding settings fixed, answer each training-side candidate independently K_i times and estimate difficulty from its fraction of correct responses:

$$\widehat{p}_{it} = \frac{1}{K_i} \sum_{r=1}^{K_i} \mathbf{1}[\widehat{y}_{ir} = y_i]. \tag{5}$$

Group candidates into reliably solved, mixed-success, and unresolved questions.

Supervision representations. A verified question can have (i) the original teacher rationale, (ii) an evidence-expanded rationale making required intermediate observations explicit, and (iii) an error-targeted rationale adding checks motivated by the student's observed failure pattern. "CoT" denotes teacher-generated explanations used for supervision. Store token count under the student's tokenizer, required evidence-step coverage, and the checks added; concise reasoning can be complete, and long reasoning can be redundant or unsupported. Evidence expansion must preserve the image, question, options, answer, and clinical facts in a paired comparison.

Difficulty belongs to the current learner. Register the rollout budget, concept-level scoring, position-counterbalancing policy, minimum valid-output rate, and binning thresholds before branch comparison. Approximate bins are reliably solved, mixed success, and unresolved; $0/K_i$ means unresolved under this probe, not intrinsically impossible. Few rollouts yield uncertain bins, and sparse or unstable estimates require review, more probes, or conservative pooling. Parsing/truncation remain separate flags: report end-to-end correctness and correctness conditional on valid output.

ECG-specific supervision repair. For a QRS-width error, a candidate expanded target identifies QRS onset/offset, converts grid distance to milliseconds, and checks the threshold, using independently verified fiducials. If measurements are already correct but the label is wrong, add the relevant cross-lead differential evidence rather than a longer measurement explanation. These interventions are chosen from the observed error type and waveform evidence. Contrastive probes, scored outputs, and signal references supply observable support.

Figure 4 shows how aggregate feedback becomes candidate selection and generation requests for subsequent branch validation.

A.2 Construction Feedback and Versioned Repair

Acceptance also checks leakage and diversity.

Repair by failure type. Acceptance records cover target attainment, non-target drift, waveform continuity, lead consistency, morphology, context, answer uniqueness, and generation cost. Target errors or boundary artifacts trigger operator-parameter or implementation repairs; missing morphology triggers applicability or source-retrieval changes; question ambiguity triggers question and context template repairs; unsupported explanations trigger rationale-prompt revisions.

Reusable rules and version records. Each skill specifies applicability, permitted sources, operator order, parameter ranges, failure handling, and question rules. Updates record the triggering feedback, changes, proposer, old and new versions, validation results, and next-round scope. Compare the original and revised versions on the same source signals, requests, random seeds, and attempt budgets, and adopt revisions after independent operator and question regression checks.

A.3 Waveform Construction and Validation Details

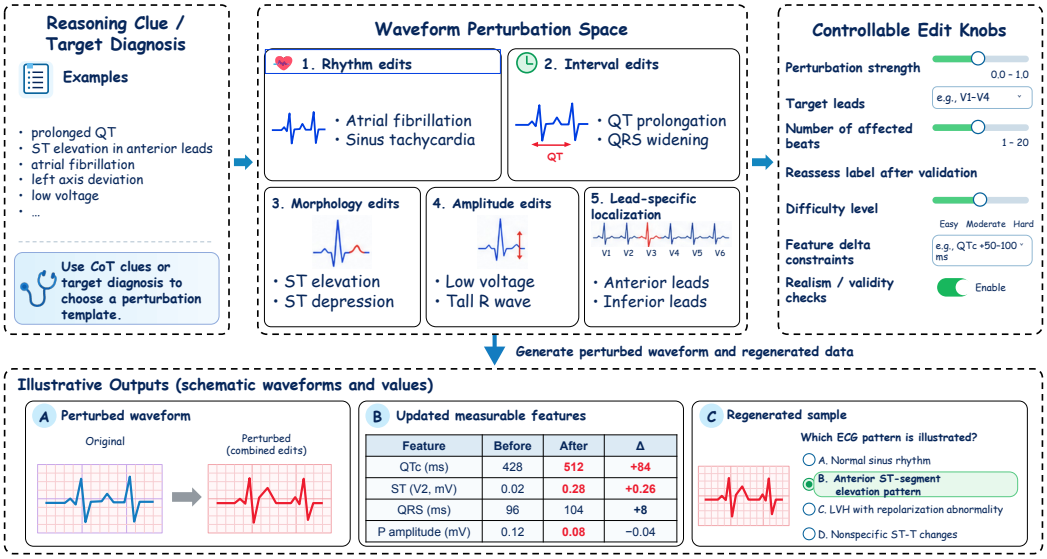

Figure 3: **Controllable waveform perturbations.** Editing options and output examples for rhythm, intervals, morphology, amplitude, and individual leads.

Bidirectional synthesis and reverse perturbation. The **forward route** starts from a real training waveform, extracts reliable measurements and supported labels, renders the ECG, and constructs a question with compatible context and an evidence-linked answer. The **reverse route** starts from a target feature or contrast, selects an eligible training waveform, applies a constrained operator sequence, and re-extracts the resulting evidence before rendering and question construction.

For example, a PR-interval contrast can request a longer PR only when P onset and QRS onset are measurable. The edit must meet the requested interval range while preserving protected heart rate and QRS morphology within fixed tolerances. More complex disease patterns require their own verified morphology and clinical evidence. Both routes then use the same question, context, and explanation checks. Bidirectional synthesis determines how examples can be created; dual feedback determines how this construction is repaired and which examples to request next.

More data can come from (i) newly retrieved real waveforms in the permitted training pool, (ii) reuse or reweighting of accepted training questions, (iii) translating *training-question* distractors into proposed target-feature requests, (iv) controlled edits, and (v) audited training-side donors for complex morphology. Feasibility, applicability, and the output evidence must be checked before creating supervision.

For a source S and permitted donor set, construct $S' = T_\delta(S)$ and re-extract $u' = f(S')$. Store target error $|[f_k(S') - f_k(S)] - \Delta_k^{\text{request}}|$ and protected drift $|f_j(S') - f_j(S)|$ separately. Check continuity at edit boundaries, behavior outside the edit mask, lead alignment, and class-specific morphology. A source and every donor must be patient-disjoint from feedback, selection, and independent evaluation data.

A.4 Bounded Recipe Search within the Learning Loop

The two-loop structure is retained. Construction feedback decides whether examples are valid and how to repair their generation. Learner feedback supplies both the demand profile and a delayed utility signal for choosing a recipe. Each round follows five registered steps:

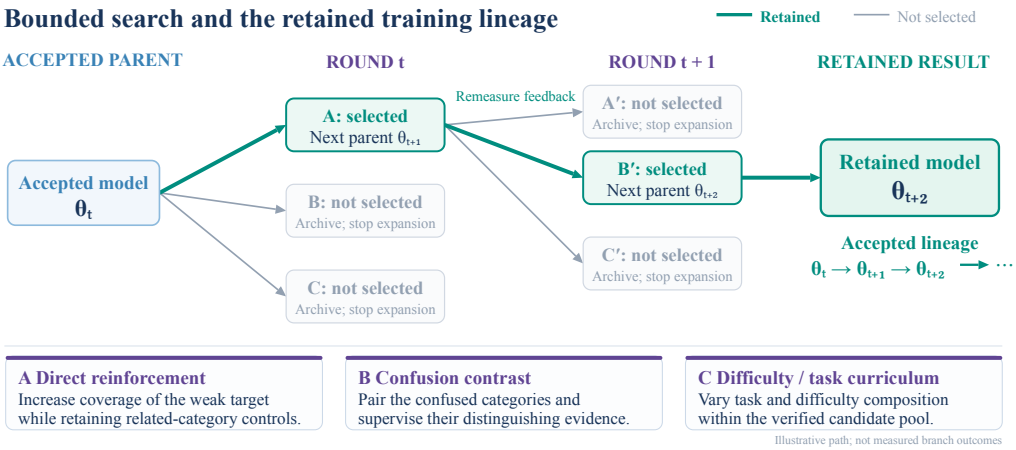

Figure 4: **Bounded search and the retained training lineage.** Each round trains sibling candidates from the same accepted parent using direct reinforcement, confusion contrast, or a difficulty/task curriculum, with matched replay and training budgets. Fixed development evaluation promotes an eligible winner; only that checkpoint is expanded in the next round, while unselected branches are archived. If no candidate is eligible, the parent is retained. The highlighted path illustrates the retained model updates.

1. Score the accepted learner on V_{fb} and probe eligible training candidates to update category, difficulty, and observable-error profiles.
2. Propose a bounded set $\mathcal{A}_t$ of recipes using those profiles and prior branch outcomes; save the proposer identity and exact recipe.
3. Construct and validate each recipe's training data under fixed gates. From the *same parent checkpoint*, run matched small-budget SFT branches with fixed replay and optimizer settings.
4. Compare candidate learners on the predesignated selection-development protocol, including forgetting and output validity.
5. Promote the accepted candidate and its associated recipe/policy versions, then remeasure the learner for the next round. If a pilot winner is retrained at full budget, restart from the registered parent and validate that full-budget checkpoint before promotion.

For branch a, the simplest candidate utility is

$$\Delta U_t(a) = U_{\text{sel}}(\text{SFT}(\theta_t, D_t^a \cup R_t; b)) - U_{\text{sel}}(\theta_t), \tag{6}$$

where b is the common pilot training budget and U_{sel} is a preregistered capability-macro score on a fixed eligible task set.

Each round trains candidate branches from the same parent, compares them on a fixed development set, and retains the best qualifying branch. The next round starts from that model, while unselected branches are archived. The search tree records candidates; the retained path records the selected model updates.

Figure 4 illustrates the candidate exploration tree and the training lineage retained after validation.

A.5 Training Configuration

Table 8: Training hyperparameters and distributed configuration.

Parameter	Setting
Sequence and batching	
Maximum sequence length	28,672 tokens
Packing	Enabled
Padding-free batching	Enabled
Micro-batch per GPU	1
Global batch size	32
Data parallelism (DP)	32
Gradient accumulation	1
Distributed topology	
Total GPUs	128 (16 nodes $\times$ 8 GPUs)
Tensor / pipeline / context parallelism	TP = 4, PP = 1, CP = 1
Expert / expert tensor parallelism	EP = 8, ETP = 1
Sequence parallelism	Enabled
Optimizer and learning rate	
Optimizer	Distributed Adam
Peak learning rate	10^{-5}
Learning-rate schedule	Cosine decay
Warmup fraction	5%
Minimum learning rate	10^{-7}
Weight decay	0.1 (constant)
Gradient clipping	1.0
Adam $\beta_1/\beta_2/\epsilon$	0.9 / 0.95 / 10^{-8}
Compute precision	bfloat16
Master parameters / gradients / Adam states	float32
Optimizer CPU offload	Enabled; fraction = 1.0
Freezing and recomputation	
Language model	Trainable
Vision encoder	Frozen
Multimodal aligner	Frozen
Selective recomputation	`core_attn`, `moe`, `mlp`
Attention backend	FlashAttention
Mixture-of-experts settings	
Grouped GEMM	Enabled
Permutation fusion	Disabled
Shared-expert overlap	Enabled
Auxiliary-loss coefficient	10^{-6}
Router dtype	float32

A.6 Parameter Extraction and Selection Details

A delineator supplies candidate measurements

$$u = f(S), \quad u \in \mathbb{R}^M. \tag{7}$$

At 25 mm/s, a 1 mm horizontal square is 40 ms and a 5 mm square is 200 ms; at 10 mm/mV, 1 mm vertically is 0.1 mV (a 5 mm square is 0.5 mV). PR is measured from P onset to QRS onset, and QT from QRS onset to T offset.

The parameter selector uses reliability and diagnostic utility,

$$r_k = \frac{1/\hat{\sigma}_k}{\sum_j 1/\hat{\sigma}_j}, \tag{8}$$

$$s_{ik} = a_{ik} r_k d_k, \tag{9}$$

$$\mathcal{C}_i = \{k : s_{ik} > \tau_{\text{param}}\}. \tag{10}$$

Here a_{ik} is definability, $\hat{\sigma}_k$ an extraction-error scale, and d_k a probe-based utility.

B SUPPLEMENTARY EXPERIMENTS AND DATA

This appendix summarizes round-wise feedback statistics, branch-evaluation results, and waveform data.

B.1 FEEDBACK STATISTICS ACROSS ROUNDS

All three partitions are recorded as being measured on a fixed feedback list across R0–R3. The source labels the ECG-class partition modality/capability.

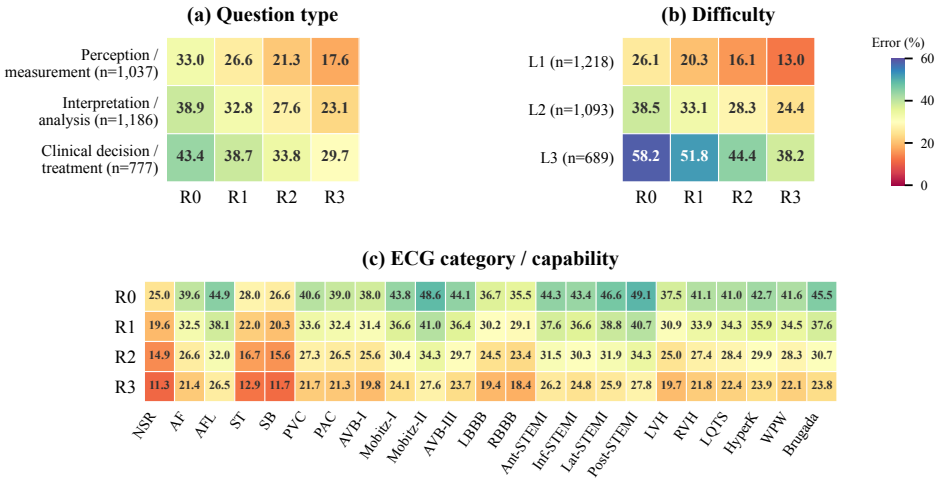

Figure 5: Feedback error rates across R0–R3: (a) question type, (b) difficulty, and (c) ECG category/capability, in the same category order as the appendix. Each cell is the group error count divided by its question count.

Table 9: Feedback statistics by difficulty.

		Error count				Error rate (%)			
Group	N	R0	R1	R2	R3	R0	R1	R2	R3
L1	1,218	318	247	196	158	26.11	20.28	16.09	12.97
L2	1,093	421	362	309	267	38.52	33.12	28.27	24.43
L3	689	401	357	306	263	58.20	51.81	44.41	38.17

Table 10: Feedback statistics for all 23 ECG categories.

Group	N	Error count				Error rate (%)			
		R0	R1	R2	R3	R0	R1	R2	R3
Normal Sinus Rhythm	168	42	33	25	19	25.00	19.64	14.88	11.31
Atrial Fibrillation	154	61	50	41	33	39.61	32.47	26.62	21.43
Atrial Flutter	147	66	56	47	39	44.90	38.10	31.97	26.53
Sinus Tachycardia	132	37	29	22	17	28.03	21.97	16.67	12.88
Sinus Bradycardia	128	34	26	20	15	26.56	20.31	15.63	11.72
Premature Ventricular Contraction	143	58	48	39	31	40.56	33.57	27.27	21.68
Premature Atrial Contraction	136	53	44	36	29	38.97	32.35	26.47	21.32
First-Degree AV Block	121	46	38	31	24	38.02	31.40	25.62	19.83
Mobitz I AV Block	112	49	41	34	27	43.75	36.61	30.36	24.11
Mobitz II AV Block	105	51	43	36	29	48.57	40.95	34.29	27.62
Complete AV Block	118	52	43	35	28	44.07	36.44	29.66	23.73
Left Bundle Branch Block	139	51	42	34	27	36.69	30.22	24.46	19.42
Right Bundle Branch Block	141	50	41	33	26	35.46	29.08	23.40	18.44
Anterior STEMI	149	66	56	47	39	44.30	37.58	31.54	26.17
Inferior STEMI	145	63	53	44	36	43.45	36.55	30.34	24.83
Lateral STEMI	116	54	45	37	30	46.55	38.79	31.90	25.86
Posterior STEMI	108	53	44	37	30	49.07	40.74	34.26	27.78
Left Ventricular Hypertrophy	152	57	47	38	30	37.50	30.92	25.00	19.74
Right Ventricular Hypertrophy	124	51	42	34	27	41.13	33.87	27.42	21.77
Long QT Syndrome	134	55	46	38	30	41.04	34.33	28.36	22.39
Hyperkalemia	117	50	42	35	28	42.74	35.90	29.91	23.93
Wolff-Parkinson-White Syndrome	113	47	39	32	25	41.59	34.51	28.32	22.12
Brugada Syndrome	101	46	38	31	24	45.54	37.62	30.69	23.76
Original total	3,000	1,140	966	811	688	38.00	32.20	27.03	22.93

B.2 Supplementary Branch-Evaluation Records

The following tables compare the parent and A/B/C candidates in overall performance, group scores, and selection outcomes.

Table 11: Evaluation and selection results for the parent and A/B/C branches.

Model	Overall Acc. (%)	Five-group macro (%)	Task-type macro (%)	Difficulty macro (%)	S	ΔS
Parent	30.15	29.80	30.05	29.60	29.93	–
A	43.22	42.10	42.85	41.90	42.66	+12.73
B	41.09	40.25	40.80	39.95	40.51	+10.58
C	38.78	38.10	38.45	37.70	38.27	+8.34

Model	Worst group change (pp)	Score / decline gate	Selection note
Parent	–	–	–
A	-1.20	Pass	Promoted
B	-1.80	Pass	Eligible, below A
C	-2.10	Pass	Eligible, below A

B.3 Supplementary Waveform Feature Data

Table 12: Supplementary waveform-feature statistics.

Feature	Unit	Sample count	MAE	Mean bias	Target attainment (%)	Non-target drift
HR	bpm	HR	bpm	2,486	5.83	-0.47
RR	ms	RR	ms	2,147	12.46	1.38
PR	ms	PR	ms	1,324	14.72	-2.19
QRS	ms	QRS	ms	1,289	9.35	0.86
QT	ms	QT	ms	1,568	17.24	-1.73
QTc	ms	QTc	ms	1,186	19.48	2.05

B.4 Blinded Clinical Expert Review

We randomly sample 240 examples stratified by six edit types (HR, RR, PR, QRS, QT, and QTc), with 40 examples per type; the sample size is determined by estimation precision. Two physicians score the examples independently. In stage 1, target labels, generation parameters, and teacher responses are hidden while waveform morphology and lead consistency are assessed. In stage 2, the physicians inspect the before/after waveforms and editing targets to assess target attainment and non-target preservation. In stage 3, they review answers and explanations and record references to expected parameters.

Table 13: Overall metrics and acceptance outcomes of the blinded clinical expert review.

Metric	Value	95% CI	n	Source	Status
Target attainment rate	0.82	[0.76, 0.87]	240	Stage 2	Pass
Non-target feature preservation rate	0.79	[0.73, 0.84]	240	Stage 2	Pass
Morphological consistency pass rate	0.86	[0.81, 0.90]	240	Stage 1	Pass
Lead consistency pass rate	0.83	[0.78, 0.88]	240	Stage 1	Pass
Answer accuracy	0.88	[0.83, 0.92]	240	Stage 3	Pass
Evidence-supported explanation rate	0.76	[0.70, 0.81]	240	Stage 3	Pass
Expected-parameter reference rate	0.09	[0.06, 0.13]	240	Stage 3	Pass
Inter-physician κ	0.78	[0.71, 0.85]	240	Two physicians	Pass

CIs use Wilson intervals or bootstrap; κ denotes Cohen's κ. Lower expected-parameter reference rates are better; higher values are better for the other proportion metrics.

Table 14: Clinical review results stratified by edit type, with 40 randomly sampled examples per type and 240 in total.

Edit	n	Target attainment	Non-target preservation	Morphological consistency	Answer accuracy	Explanation support
HR	40	0.85	0.80	0.88	0.90	0.78
RR	40	0.78	0.75	0.82	0.85	0.72
PR	40	0.80	0.78	0.85	0.88	0.75
QRS	40	0.83	0.81	0.87	0.89	0.77
QT	40	0.82	0.79	0.86	0.87	0.76
QTc	40	0.84	0.80	0.88	0.89	0.78
Total	**240**	**0.82**	**0.79**	**0.86**	**0.88**	**0.76**

Overall target attainment is 0.82, non-target preservation is 0.79, answer accuracy is 0.88, and evidence-supported explanation rate is 0.76; inter-physician Cohen's κ is 0.78. Across edit types, target attainment ranges from 0.78 to 0.85 and explanation support ranges from 0.72 to 0.78.

B.5 Multi-Seed Results on ECG Tasks

Table 15: Three-seed results on ECG tasks. Means and sample standard deviations are computed across the three results.

Model	Seed 1	Seed 2	Seed 3	Mean ± SD	Source
Base	30.15	30.15	30.15	30.15 ± 0.00	Fixed initialization
ECG-RSI R1	33.42	33.18	33.61	33.40 ± 0.22	3 independent runs
ECG-RSI R2	35.27	35.04	35.48	35.26 ± 0.22	3 independent runs
ECG-RSI R3	37.86	37.52	38.09	37.82 ± 0.29	3 independent runs

B.6 Multi-Model Question-Quality Scores

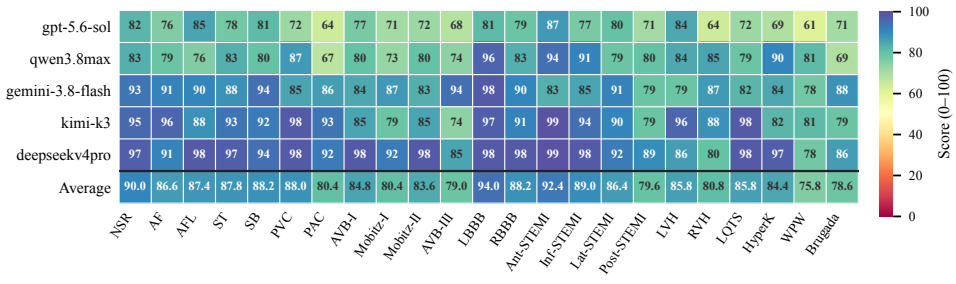

Figure 6: Recorded scores from five models for synthetic ECG questions across 23 categories (0–100 points). The bottom row is the per-category arithmetic mean across models; the color scale runs from red (low) to blue-purple (high). These scores are not diagnostic accuracy.

Figure 6 presents recorded scores from five models for synthetic ECG questions across 23 categories.

Question quality scoring and acceptance gates. Figure 6 summarizes quality scores from five models across 23 synthetic-question categories; the bottom row reports per-category means. The first-round record contains 16,193 candidate questions and a training setting of 10,000.

B.7 Synthesis Updates and Strategy Trajectories

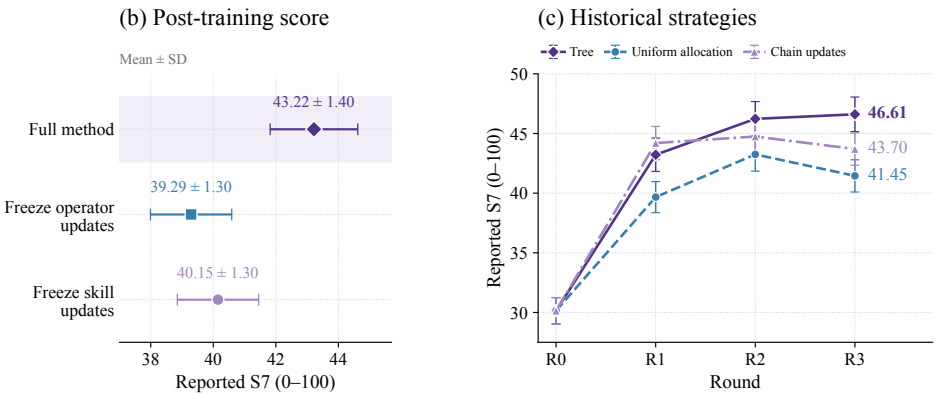

Figure 7: Synthesis-update and strategy results. (b) R1 S7 for the full method and frozen operator/skill updates from Table 6; (c) R-A tree, uniform-allocation, and chain trajectories from Table 7.

C ECG Task Taxonomy

Figure 8 shows the 23 ECG tasks currently used in this paper. Tasks are organized by medical category. Sector area represents the number of categories, not training-sample counts or model scores.

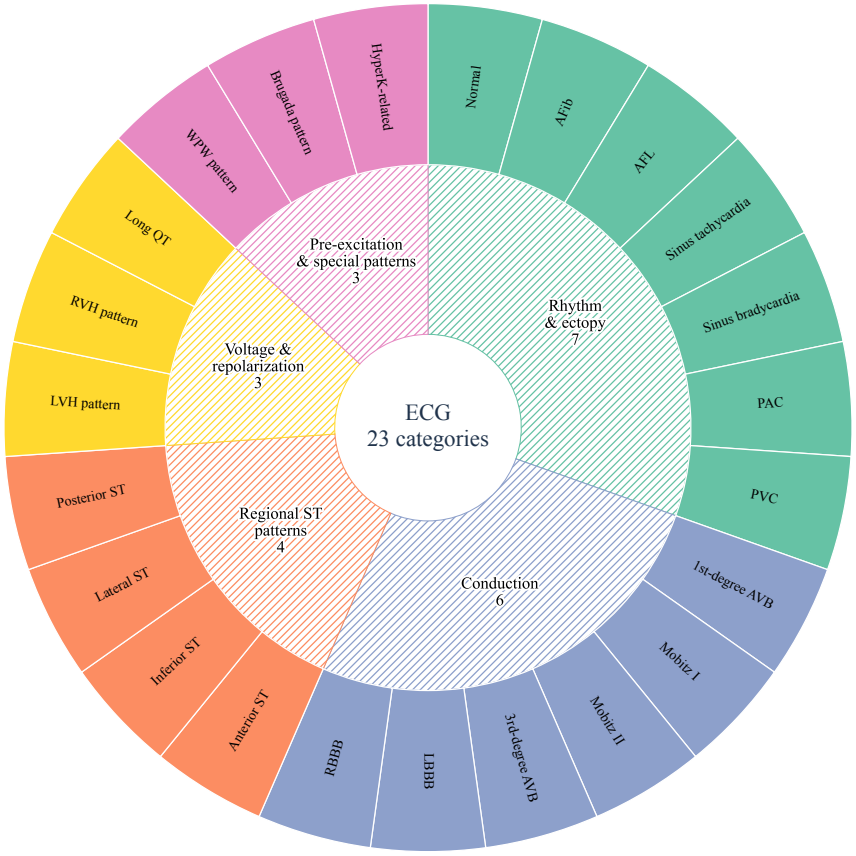

Figure 8: Medical taxonomy of the 23 ECG tasks. Each outer sector represents one category; inner-sector areas and counts indicate the number of categories in each group.

D Stage-wise Prompt Templates

This appendix documents reusable prompt templates for the method interfaces. Bracketed fields are filled with request-specific inputs and tool results. D.3 reproduces the teacher-answer instruction; the other templates are organized by function. The following cases retain their original questions and saved responses.

D.1 Constrained waveform construction

Prompt P1

System role. You are an ECG waveform-construction agent. Convert a target feature request into an auditable edit plan.

User inputs. [request ID], [training-source waveform and sampling rate], [lead names and calibration], [measured landmarks and uncertainty], [target ranges], [protected features and tolerances], [allowed operators and donors], [retry budget].

Instructions. Check source eligibility and landmark reliability. For the forward route, extract supported features without changing the signal. For the reverse route, specify operator order, parameters, and edit

mask. Invoke the supplied editing and measurement tools, then compare observed targets and protected drift with the fixed tolerances. Check boundary continuity, lead alignment, and required morphology. If tools or required landmarks are unavailable, return blocked; if verification fails, repair within budget or reject.

Output fields. Status; source identities; operator plan; requested and observed measurements with units; protected drift; validity checks; failure codes; edited-signal reference. Report executed results separately from proposed operations.

D.2 Evidence-grounded question and context generation

Prompt P2

System role. You construct ECG multiple-choice questions from verified signal evidence. Keep internal labels and generator intent out of the student-visible question.

User inputs. [verified ECG image], [measurement evidence], [supported label level], [target category], [question type and difficulty], [option count], [context policy], [permitted source context].

Instructions. Write one question with exactly the requested number of options and one evidence-supported answer. Use plausible distractors that remain incorrect for this signal. Ask only at the supported level: measurement, ECG pattern, or clinically supported diagnosis. Default to image-only or traceable demographics. Mark any synthetic teaching context explicitly and check it against the achieved waveform.

Output fields. Student-visible question, options, and allowed context; evaluator-only answer, evidence mapping, context provenance, category, and difficulty; ambiguity or insufficient-evidence flag.

D.3 Teacher answer and evidence explanation

Prompt P3

Please reason step by step, and put your final answer within `\boxed{}`.

D.4 Independent quality and acceptance review

Prompt P4

System role. You audit a candidate ECG training example.

User inputs. [source and edited evidence], [target/protected-feature checks], [question and options], [context provenance], [proposed answer and teacher response], [fixed rubric and aggregation rule].

Instructions. Independently solve the question, then audit target attainment, protected drift, morphology, context consistency, answer uniqueness, leakage, and explanation support. Return separate decisions for each gate and cite the relevant observation. Score only with the supplied rubric. The aggregate-score threshold is 80/100, but every mandatory gate must also pass. Route unresolved disagreement to evidence review.

Output fields. Independent answer; gate-wise pass/fail/unresolved decisions and evidence; rubric scores; aggregate if defined; accept/review/reject status; failure codes and repair target.

D.5 Construction failure and versioned repair

Prompt P5

System role. You propose a minimal repair to a failed data-construction component while keeping the acceptance criteria fixed.

User inputs. [failed attempt and evidence], [failure codes], [current operator/prompt/skill version], [allowed repair scope], [fixed regression suite], [attempt budget].

Instructions. Map numerical or boundary failures to operator repair; missing morphology to applicability or donor retrieval; ambiguity/leakage to question or context repair; unsupported explanation to response-

prompt repair. Specify the exact change and why it addresses the observed failure. Compare old and new versions on the same sources, requests, seeds, and budget. Quarantine or reject when the budget is exhausted.

Output fields. Triggering failure; component; exact revision; old/new version; proposer; validation results or pending status; acceptance rationale; next-round scope.

D.6 Learner-feedback diagnosis and recipe proposal

Prompt P6

System role. You propose training-data interventions from development feedback.

User inputs. [accepted parent and feedback version], [category/type/difficulty counts and errors], [confusions and regressions], [training-pool coverage], [construction yield/cost], [history], [token and attempt budgets].

Instructions. Prioritize deficits with adequate support; report sparse or uncertain groups explicitly. Propose distinct candidates for direct weakness coverage, confusion contrasts, and difficulty or evidence-granularity progression. For each, give normalized category, difficulty, and supervision-style allocations; requested feature ranges; protected attributes; allowed routes/operators; source filters; and cost limits. A/B/C retain the same previous cumulative training set unchanged and each add 10,000 first-use accepted questions, without an extra shared baseline set; only the winning additions enter the next cumulative set.

Output fields. Feedback summary with uncertainty; candidate recipes and hypotheses; normalized allocations and budgets; construction requests; provenance restrictions; expected tests, not claimed gains.

D.7 Candidate comparison and model promotion

Prompt P7

System role. You apply a preregistered development-only promotion rule to comparable candidate learners.

User inputs. [parent and candidate checkpoint IDs], [new/replay manifests], [supervised-token budgets], [fixed selection-set results], [primary score], [regression tolerances], [tie/stopping rules].

Instructions. Check that candidates share the required parent, training budget, and evaluation protocol. Mark missing or incomparable records unresolved. Apply the supplied score, regression, and valid-output gates exactly; do not invent thresholds. Promote only an eligible candidate, otherwise retain the parent. Count only previously unused accepted items from the promoted branch toward cumulative additions, preserving replay provenance. Archive other branches and log the decision.

Output fields. Comparability checks; gate results; chosen checkpoint or retain-parent/unresolved decision; supporting scores; admitted item manifest; replay manifest; stopping status.

E Prompts and Case Studies

E.1 Case 1: PR prolongation

Case 1: PR prolongation

Waveform construction

Copy the P wave and QRS–T segment, lengthen the PQ segment, and shorten the TP segment to preserve the RR interval. The target PR is 300 ms; remeasurement yields PR 298 ms and HR 74.6 bpm.

Case 1: PR prolongation (continued)

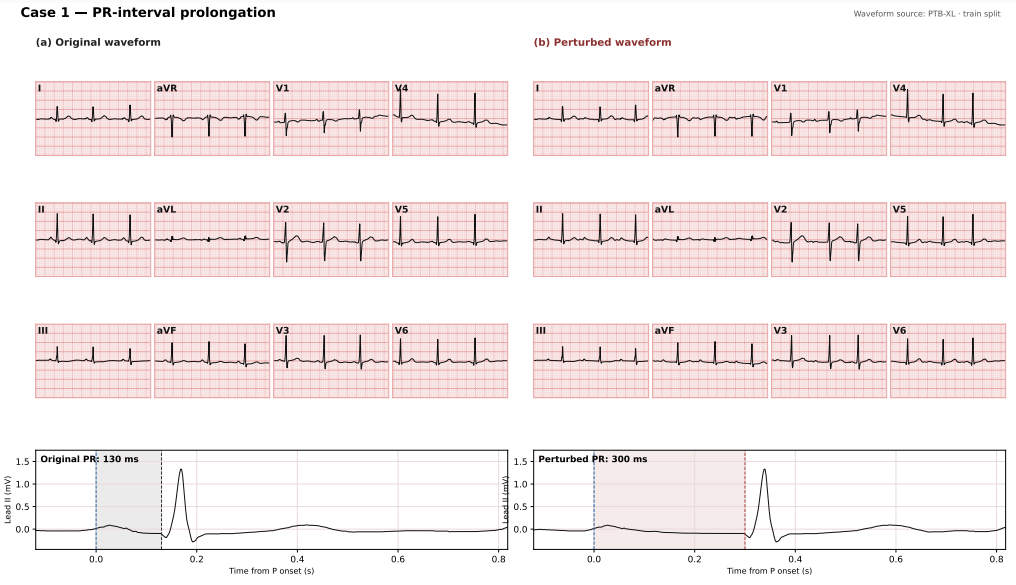

Figure 9: Original and edited ECG with lead II details.

Question

Question: Under the automatic closed-loop measurement convention of this data-construction pipeline, which parameter analysis and conclusion are correct for this 12-lead ECG?

Options:

A: HR 74.6 bpm, PR 130.0 ms, QRS 96.0 ms, QTc 393.7 ms -> Conclusion: no first-degree atrioventricular block

B: HR 48.0 bpm, PR 298.0 ms, QRS 96.0 ms, QTc 393.7 ms -> Conclusion: sinus bradycardia with first-degree atrioventricular block

C: HR 74.6 bpm, PR 298.0 ms, QRS 96.0 ms, QTc 393.7 ms -> Conclusion: first-degree atrioventricular block

D: HR 74.6 bpm, PR 298.0 ms, QRS 130.0 ms, QTc 393.7 ms -> Conclusion: first-degree atrioventricular block with intraventricular conduction delay

Please reason step by step, and put your final answer within \boxed{}.

Qwen3.8 response

Let me analyze this ECG question.

The question is asking (in Chinese): "According to the automatic closed-loop measurement criteria of this data construction pipeline, looking at this 12-lead ECG, which of the following parameter analyses and conclusions is correct?"

Options:

A: HR 74.6 bpm, PR 130.0 ms, QRS 96.0 ms, QTc 393.7 ms → Conclusion: No first-degree AV block observed

B: HR 48.0 bpm, PR 298.0 ms, QRS 96.0 ms, QTc 393.7 ms → Conclusion: Sinus bradycardia with first-degree AV block

C: HR 74.6 bpm, PR 298.0 ms, QRS 96.0 ms, QTc 393.7 ms → Conclusion: First-degree AV block

D: HR 74.6 bpm, PR 298.0 ms, QRS 130.0 ms, QTc 393.7 ms → Conclusion: First-degree AV block with intraventricular conduction delay

I need to look at the ECG and determine the parameters.

Case 1: PR prolongation (continued)

Looking at the ECG: the rhythm strip (II rhythm) shows about 12 beats over some length. Let me estimate the heart rate. The rhythm strip is probably 10 seconds? Let me count: I see 12 QRS complexes in the II rhythm strip. If the strip is 10 seconds, HR ≈ 12 * 6 = 72 bpm... actually 12 beats in 10 seconds = 72 bpm. Hmm, but the option says 74.6 bpm. That's plausible.

Looking at the PR interval: in lead II, there's a P wave before each QRS. The PR interval looks... let me see. In the rhythm strip, the P wave appears just before the QRS? Actually, there's a small bump (P wave) before the QRS, and a T wave after the QRS. Let me look more carefully.

In lead II (second row first column), there's a small P wave, then QRS, then T wave. The distance from P to QRS looks short? Hmm.

But the question says "perturbed waveform"—this is a synthetic case with perturbations added. The question is about the "automatic closed-loop measurement criteria" of the data construction pipeline. Probably the ground truth parameters are: HR 74.6, PR 298 ms (prolonged → first-degree AV block), QRS 96 ms (normal), QTc 393.7 ms.

Looking at the waveform: in lead II, before each QRS there's a P wave... hmm, actually looking at the rhythm strip: after each QRS there's a T wave, and before the next QRS there's a P wave? Let me look: pattern per beat: QRS spike, then T wave bump, then another bump (P wave?), then QRS. Actually, in the rhythm strip I see: spike, then bump, then bump, then spike... spacing: after the spike at ~x=290, bump at ~340, bump at ~390, spike at ~455. Hmm, so two bumps between spikes: T wave and P wave. If the P wave is at ~390 and QRS at ~455, then PR interval is ~65 px? Need scale.

Scale: 25 mm/s. Grid: large box = 5 mm = 0.2 s = 200 ms. Small box = 1 mm = 40 ms.

Let me estimate pixel scale. Each lead panel width is ~500 px? Hmm, need to guess from image coordinates. The image is 2338 px wide? Actually the given image dimensions: width ~2338? Let me think: the figure shows 4 columns of panels. Each panel ~500 px wide representing 10 seconds? No, each panel usually shows ~2.5 seconds? Hmm, but the rhythm strip at the bottom shows 12 beats over full width.

Better approach: use RR interval to determine HR. If HR is 74.6 bpm, RR = 804 ms. If HR is 48, RR = 1250 ms.

Looking at the rhythm strip: 12 beats over full width. If full width is 10 seconds (standard), then average RR = 10/11? Actually 12 beats → 11 intervals over ~10 s → RR ≈ 0.909 s → HR ≈ 66. Hmm.

But maybe the strip is longer? Let me count beats in rhythm strip: spikes at approx x positions: 290, 455, 610, 765, 920, 1075, 1225, 1380, 1530, 1680, 1835, 1990? That's 12 spikes. Spacing ~155 px. If spacing = 804 ms (HR 74.6), then px per ms = 155/804 ≈ 0.1928 px/ms → 25 mm/s = 0.025 mm/ms... hmm, let me compute mm: at 25 mm/s, 804 ms = 20.1 mm. So 155 px = 20.1 mm → 7.7 px/mm. Large box (5 mm) = 38.5 px. Small box = 7.7 px.

Now PR interval: P wave bump before QRS: in rhythm strip, bump at ~390 and spike at ~455: difference 65 px = 65/7.7 ≈ 8.4 mm = 337 ms? Hmm, that's too long. But maybe the bump at 390 is the P wave and the bump at 340 is the T wave. PR = 455-390 = 65 px → at 7.7 px/mm → 8.44 mm → at 25 mm/s → 337 ms. Hmm, option says PR 298 ms. 298 ms = 7.45 mm = 57 px. Close to 65 px? My pixel estimates are rough.

Alternatively, if HR were 48 (RR 1250 ms), spacing 155 px → 155/1250 = 0.124 px/ms → px per mm = 0.124*40 = 4.96 px/mm. Then PR 65 px = 13.1 mm = 524 ms, unlikely.

So HR ~74.6 is consistent with 12 beats over ~9.6 seconds? 11 intervals * 0.804 s = 8.84 s. Rhythm strip width ~1980 px? Hmm, from x=170 to x=2150 ≈ 1980 px. At 7.7 px/mm → 257 mm → at 25 mm/s → 10.3 s. OK, plausible: 10-second rhythm strip with 12 beats → HR ~72-75. Good.

Now PR: is it prolonged (298 ms)? Looking at lead II panel: P wave before QRS: in lead II panel, spike at ~190, preceding bump at ~130? Difference 60 px ≈ 310 ms? Hmm, at 7.7 px/mm: 60 px = 7.8 mm = 312 ms. Yes, consistent with PR ~298 ms.

But wait, option A says PR 130 ms (normal). Which is correct visually? Let me look at lead II more carefully: coordinates within panel: lead II panel spans x from ~70 to ~560. Spikes at ~190, ~350, ~505.

Case 1: PR prolongation (continued)

Preceding bumps: ~130? Actually I see a small bump at ~130 (P wave), then spike at ~190. Then T wave bump at ~230, then another bump at ~290? Hmm.

Actually, typical synthetic ECG with first-degree AV block: P wave clearly separated from QRS with long PR. In the image, in lead II, I see a small bump just before each QRS? Hmm, at ~130 there's a bump, then flat, then spike at 190. Distance 60 px.

Alternatively, the bump at ~130 might be the P wave and PR = 60 px ≈ 300 ms. Yes.

Also, the T wave after QRS at ~230-240, and then maybe a U wave? Hmm, the bump at ~290 might be the P wave of the next beat? No, next spike at 350; bump at ~290 is 60 px before → P wave. Then bump at ~230 is T wave. OK, consistent: T wave ~40-60 px after QRS, P wave ~60 px before next QRS.

So PR ≈ 60 px ≈ 300 ms → first-degree AV block. HR ≈ 74.6. QRS: narrow spike (~96 ms normal). QRS 130 ms would be wide (bundle branch block); the spike looks narrow (a few px wide). So QRS 96 ms.

Thus correct option: C: HR 74.6, PR 298, QRS 96, QTc 393.7 → first-degree AV block.

Check option B: HR 48 wrong. Option D: QRS 130 wrong. Option A: PR 130 wrong (would be normal). So answer C.

But wait: the question says "according to the automatic closed-loop measurement criteria of this data construction pipeline"—probably meaning the ground truth from the synthesis pipeline. The waveform is "perturbed waveform" case 1. Probably the intended answer: C.

Let me double-check HR: 12 beats in 10 s → 72? But option says 74.6; fine.

Also check sinus rhythm: P wave positive in II, negative in aVR? In aVR, after QRS (negative spike) there's a negative bump? Hmm, aVR shows negative QRS and then negative T? P wave in aVR should be negative; hard to see but plausible.

Thus conclusion: first-degree AV block (PR > 200 ms) with normal HR (74.6) and normal QRS (96 ms). So option C.

Therefore answer: \boxed{C}.

Let me reconfirm PR visually: in lead II rhythm strip, between QRS spikes: T wave bump right after spike, then P wave bump before next spike. Distance from P bump to spike: e.g., P at ~390, spike at ~455: 65 px. Large box size: let me estimate grid: panel width ~500 px with ~13 large boxes? Hmm, 500/13 ≈ 38.5 px per large box → consistent with 7.7 px/mm (5 mm = 38.5 px). So 65 px = 1.69 large boxes = 338 ms? Hmm, that's more than 298. But my pixel guesses are rough; maybe P at ~400. Anyway, among options, PR 298 is the prolonged one; 130 would be about half a large box (2.6 small boxes = 26 px). Visually the P-QRS distance is clearly more than 1 large box? Hmm, 298 ms = 7.45 small boxes = 1.49 large boxes. 130 ms = 3.25 small boxes = 0.65 large boxes. Visually the distance from P bump onset to QRS onset looks like ~1.5 large boxes? Plausible.

Actually, PR interval is measured from P onset to QRS onset. The P bump peak at ~390-400, onset maybe ~380; QRS onset ~450. Difference ~70 px ≈ 1.8 large boxes = 360 ms? Hmm, too long. But my pixel readings from memory of the image are unreliable. Given the options, only C is internally consistent: HR 74.6 (not bradycardia), PR prolonged 298 → first-degree AV block, QRS normal 96 → no intraventricular conduction delay. So C.

Thus final answer: C.

Final answer

Looking at the rhythm strip: the R-R intervals correspond to a rate of about 75 bpm (not bradycardia), each P wave is followed by a QRS with a clearly prolonged P-R interval (about 300 ms, i.e., >200 ms), the QRS complexes are narrow (about 96 ms, no intraventricular conduction delay), and QTc is normal (~394 ms). Therefore the correct interpretation is sinus rhythm at ~74.6 bpm with a prolonged PR interval (298 ms) = first-degree AV block, with normal QRS duration.

- A is wrong because the PR is clearly prolonged, not 130 ms.

Case 1: PR prolongation (continued)

- B is wrong because the rate is ~75 bpm, not 48 bpm.
- D is wrong because the QRS is narrow (96 ms), not 130 ms.

\boxed{C}

E.2 Case 2: Heart-rate slowing

Case 2: Heart-rate slowing

Waveform construction

Copy the P–QRS–T segment and extend the TP segment between beats. The target heart rate is 48 bpm, corresponding to an RR interval of 1.25 s. Remeasurement yields HR 47.9 bpm, PR 150 ms, QRS 48 ms, and QTc 348.5 ms.



Figure 10: Original and edited ECG with lead II details.

Question

Question: Under the automatic closed-loop measurement convention of this data-construction pipeline, which parameter analysis and conclusion are correct for this 12-lead ECG?

Options:

A: HR 72.3 bpm, PR 150.0 ms, QRS 48.0 ms, QTc 348.5 ms -> Conclusion: normal heart rate

B: HR 47.9 bpm, PR 150.0 ms, QRS 48.0 ms, QTc 348.5 ms -> Conclusion: sinus bradycardia

C: HR 47.9 bpm, PR 220.0 ms, QRS 48.0 ms, QTc 348.5 ms -> Conclusion: sinus bradycardia with first-degree atrioventricular block

D: HR 47.9 bpm, PR 150.0 ms, QRS 48.0 ms, QTc 490.0 ms -> Conclusion: sinus bradycardia with QT prolongation

Please reason step by step, and put your final answer within \boxed{}.

Qwen3.8 response

Let me analyze this ECG. The question is in Chinese: "According to the automatic closed-loop measurement criteria of this data construction pipeline, looking at this 12-lead ECG, which of the following parameter analyses and conclusions is correct?"

Options:

Case 2: Heart-rate slowing (continued)

A: HR 72.3 bpm, PR 150.0 ms, QRS 48.0 ms, QTc 348.5 ms → Conclusion: normal heart rate

B: HR 47.9 bpm, PR 150.0 ms, QRS 48.0 ms, QTc 348.5 ms → Conclusion: sinus bradycardia

C: HR 47.9 bpm, PR 220.0 ms, QRS 48.0 ms, QTc 348.5 ms → Conclusion: sinus bradycardia with first-degree AV block

D: HR 47.9 bpm, PR 150.0 ms, QRS 48.0 ms, QTc 490.0 ms → Conclusion: sinus bradycardia with prolonged QT interval

I need to look at the ECG and determine the heart rate, PR interval, QRS width, and QTc.

Looking at the rhythm strip (II rhythm): there are about 8 QRS complexes across the entire strip. The strip length is probably 10 seconds? Let me estimate. Paper speed 25 mm/s. Each lead panel shows about 2 beats across roughly 5.5 seconds? Let me look at the grid.

Actually, let me estimate the RR interval from the image. In the II rhythm strip, the QRS complexes appear at approximately x positions: 253, 483, 713, 943, 1200?, ... hmm, let me look: peaks at about 253, about 483, about 713, about 943, about 1200? Actual spacing: from the figure, peaks at x ≈ 253, 483, 713, 943, 1200, 1430, 1660, 1890? Hmm, the spacing looks like about 230 pixels.

Grid: at 25 mm/s, small box = 1 mm = 0.04 s; large box = 5 mm = 0.2 s. The panel width is probably 250 mm? Hmm.

Let me think about the intended answer. This seems to be from a synthetic dataset ("Case 2 · Perturbed waveform"), with ground truth parameters: HR 47.9 bpm (bradycardia), PR 150 ms (normal), QRS 48 ms (narrow? actually 48 ms is quite narrow but plausible as synthetic), QTc 348.5 ms (normal). So the conclusion: sinus bradycardia. Option B.

But let me check the visual RR interval: HR 47.9 bpm → RR ≈ 1.25 s = 31.25 small boxes = 6.25 large boxes. In the rhythm strip, the distance between QRS peaks: if the strip is 10 s with 8 beats, then RR ≈ 1.25 s → 8 beats in 10 s gives HR about 48? Actually 8 beats in 10 s: RR = 10/7 ≈ 1.43 s → HR 42. Hmm, but with 8 beats across the strip: intervals = 7; 7 * 1.25 = 8.75 s, plus margins ≈ 10 s strip. Plausible.

Check PR: in lead II, the P wave appears just before the QRS with a short interval (about 150 ms = 3.75 small boxes). Visually, the P wave is close to the QRS, not far away. So PR 150 ms is plausible; PR 220 would be first-degree AV block, but the P-QRS distance looks short (P wave just before QRS). In the rhythm strip, the P wave appears about 1 small box? Hmm.

QTc: QT interval looks normal: T wave follows QRS soon after. QTc 348.5 is normal (shortish but normal). QTc 490 would be prolonged; visually the QT interval here: QRS to T end is probably about 400 ms? At HR 48, QTc = QT / sqrt(RR). If QT ≈ 400 ms and RR = 1.25, then QTc = 400/1.118 = 357. So QTc 348.5 is consistent.

Therefore the answer is B.

But wait: option A says HR 72.3 with the same other parameters and conclusion "normal heart rate." The rhythm clearly shows bradycardia: RR interval is large (about 6 large boxes). So HR about 48. Thus B.

Let me double-check the grid: each lead panel width is about 540 pixels? The rhythm strip width is about 1965 pixels? Hmm.

Let me estimate: in the rhythm strip, the QRS peaks are at x ≈ 253, 483, 713, 943, 1200? Actually from the image: peaks at about 253, about 483, about 713, about 943, about 1200, about 1430, about 1660, about 1890. Spacing about 230 pixels. Grid large box: need pixel per mm. The rhythm strip spans from x=165 to x=2130 ≈ 1965 pixels. If the strip is 10 s at 25 mm/s = 250 mm, then pixel/mm ≈ 7.86. Then RR = 230 px / 7.86 ≈ 29.3 mm ≈ 1.17 s → HR ≈ 51. Hmm, close to 48.

Alternatively, the panel: each lead panel width about 540 px? For lead I panel from x=65 to x=560 ≈ 495 px, showing 2 beats with RR about 230 px? Same scale.

Given the options, HR 47.9 corresponds to RR = 1.2526 s = 31.3 mm. At 7.86 px/mm that's 246 px. Observed spacing about 230 px → RR 1.17 s → HR 51. Hmm, not exact, but my pixel estimates are rough.

Case 2: Heart-rate slowing (continued)

Check PR: P wave before QRS: in the rhythm strip, the P wave peak is at about 232? QRS at 253. Distance about 20 px ≈ 2.5 mm ≈ 0.1 s? Hmm, PR is from P onset to QRS onset. P onset probably at 225, QRS onset at 250: 25 px ≈ 3.2 mm ≈ 128 ms. Close to 150 ms. Not 220 ms (which would be 5.5 mm = 43 px). Visually the P wave is close to QRS, so PR 150.

QT: QRS at 253 to T end at about 320? T wave peak at about 300, end at about 330. QT ≈ 77 px ≈ 9.8 mm ≈ 392 ms. QTc (Bazett) = 392 / sqrt(1.25) = 350. Matches 348.5. Good.

Therefore the answer is B.

The question states "according to the automatic closed-loop measurement criteria of this data construction pipeline" — meaning the ground truth parameters used to generate the synthetic ECG: HR 47.9, PR 150, QRS 48, QTc 348.5, conclusion sinus bradycardia.

Thus the answer: B.

Let me double-check option C: PR 220 would be first-degree AV block; but the visual PR is short. Option D: QTc 490 would require QT ≈ 548 ms at RR 1.25, which visually is not the case (T wave ends well before the next P). So B.

Therefore final answer: \boxed{B}.

Final answer

Looking at the rhythm strip: the R-R intervals span roughly 6+ large boxes (≈1.25 s), giving a rate of ≈48 bpm → bradycardia, not 72 bpm (excludes A). Each QRS is preceded by a normal sinus P wave with a short, constant P-R distance of about 3–4 small boxes (~150 ms), not the >5 small boxes (220 ms) needed for first-degree AV block (excludes C). The QRS is narrow (~48 ms), and the QT interval ends well before the next P wave (~390–400 ms); corrected for the slow rate (Bazett), QTc ≈ 350 ms, which is normal, not 490 ms (excludes D). Thus the measurements HR ≈ 47.9 bpm, PR 150 ms, QRS 48 ms, QTc 348.5 ms correspond to sinus bradycardia with otherwise normal intervals.

\boxed{B}