PEDECG: ADULT-TO-PEDIATRIC ECG TRANSFER THROUGH REPRESENTATION REUSE, TARGET REINTERPRETATION, AND EVIDENCE VALIDATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Pediatric ECG interpretation is inherently development-dependent: the diagnostic significance of heart rate, conduction intervals, voltage, axis, and repolarization changes throughout childhood. Although large adult ECG datasets provide rich cardiac knowledge, directly transferring adult models to children can also transfer diagnostic rules calibrated to adult physiology. To address this challenge, we propose PedECG, a decomposable framework for adult-to-pediatric ECG transfer that separates representation reuse, target-side reinterpretation, and record-specific evidence validation. PedECG preserves adult pretrained waveform representations with a frozen encoder, reinterprets them using quantitative ECG measurements and age-indexed pediatric reference context, and verifies candidate diagnoses against evidence from the current ECG record. Experiments show that adult cardiac representations remain highly reusable in the pediatric domain, while structured target-side reinterpretation and record-specific verification further improve how this transferred knowledge is converted into pediatric predictions; the learned interpretation also adapts to changes in the pediatric physiological reference frame. By separating reusable physiological knowledge from population-specific diagnostic interpretation, PedECG offers a general strategy for extending medical AI from data-rich source populations to clinically distinct target populations, helping make large-scale learned medical knowledge usable beyond the cohorts on which it was originally developed.

Keywords: Transfer Learning; Domain Adaptation; Representation Learning; Pediatric ECG; Multimodal Learning; Developmental Adaptation

1 INTRODUCTION

Pediatric ECG diagnosis provides a demanding test of cross-population transfer because the physiological reference frame itself changes with development. From infancy to adolescence, normal heart rate, PR interval, QRS duration, frontal axis, precordial voltage, and repolarization patterns shift substantially [1, 2, 3, 4]. The same waveform morphology can therefore carry different diagnostic meaning at different ages. At the same time, large and consistently annotated pediatric ECG datasets remain far scarcer than adult resources, creating a setting in which abundant source-domain knowledge must be reused under a changing target-domain interpretation.

Large adult ECG cohorts already support models that capture rich structure in rhythm, conduction, chamber-related voltage, and repolarization [5, 6, 7, 8, 9], and recent ECG language models extend this knowledge to multimodal interpretation and reporting [10, 11, 12]. This creates an important distinction for transfer: the cardiac evidence learned from adults may remain useful even when the diagnostic rules attached to that evidence no longer remain valid in children. The central problem is therefore to preserve what is transferable while reconstructing how that evidence should be interpreted under pediatric physiology.

Most transfer pipelines adapt representations and predictors together and report a single target-domain endpoint. Such evaluation can obscure where adaptation is actually required: in the source representation, in the target decision rule, or in the final use of record-specific evidence. We study

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.

adult-to-pediatric ECG transfer by separating these possibilities explicitly. In particular, we ask three questions: what adult cardiac knowledge survives the population shift, how should that knowledge be reinterpreted in the pediatric diagnostic space, and when is a transferred candidate supported by the current ECG record?

PedECG answers these questions through three corresponding stages: preserve, reinterpret, and validate. A frozen adult waveform encoder preserves source-domain cardiac representations; a structured pediatric readout reexpresses this evidence using age, quantitative ECG measurements, and age-indexed pediatric reference context; and a record-specific verifier evaluates candidate diagnoses against the current ECG before final decision integration. Beyond label prediction, PedECG also extends the same record-level evidence pathway to structured ECG reporting, producing organized diagnostic summaries of rhythm, intervals, axis and voltage, conduction, repolarization, and supporting evidence.

Controlled analyses reveal a consistent transfer pattern. Adult pretrained representations retain nearly the same pediatric probe utility as a matched pediatric model trained from scratch, showing that substantial cardiac representation utility survives the population shift without pediatric encoder updates. Structured target-side reinterpretation improves over direct feature concatenation, and record-specific verification provides additional discrimination after candidate generation; blank and mismatched image interventions further show that this verification depends on the current ECG. Developmental-context substitution additionally demonstrates that the learned pediatric readout changes when the physiological reference frame changes.

Together, these findings identify a stage-resolved transfer mechanism in which reusable source representations can be preserved while target-side interpretation and record-level validation are reconstructed for the new population. PedECG operationalizes this mechanism as an explicit learning framework, turning adult-to-pediatric transfer from a single adaptation score into a sequence of testable decisions about what to preserve, what to reinterpret, and what to validate.

In summary, our contributions are threefold:

- **We formulate adult-to-pediatric ECG diagnosis as a stage-resolved cross-population transfer problem.** The formulation separates representation transferability from target-side decision formation, exposing adaptation requirements that are hidden by a single end-to-end score.
- **We introduce PedECG, a preserve–reinterpret–validate framework for pediatric transfer.** The model retains a frozen adult waveform interface, reconstructs its diagnostic use under pediatric physiological context, and grounds final candidate selection in record-specific evidence while also supporting structured ECG reporting.
- **We provide controlled evidence for where adult knowledge transfers and where target adaptation contributes.** Adult waveform representations remain highly reusable after the population shift, while structured target-side reinterpretation and record-level validation improve how preserved source knowledge is converted into pediatric predictions.

2 Related Work

ECG representation and multimodal modeling. Large adult ECG cohorts have enabled strong supervised models for arrhythmia recognition, multilabel diagnosis, and broader cardiac phenotyping [5, 6, 7, 8, 9, 13, 14]. These results establish adult ECG data as a rich source of reusable cardiac representations. More recent ECG language and multimodal models, including MEIT, ECG-Chat, GEM, and ECG-R1, combine waveforms, rendered images, measurements, and text for interpretation and report generation [10, 12, 15, 11]. Their primary objective, however, is improved inference within a largely fixed diagnostic population or ontology. Our focus is different: whether source representations remain useful when the diagnostic meaning attached to that evidence changes across populations.

Pediatric ECG and developmental variation. Pediatric ECG interpretation is defined by developmental physiology. Reference studies document substantial age-dependent variation in heart rate, conduction intervals, axis, voltage, and repolarization [1, 2, 3, 4]. Pediatric AI work has demonstrated

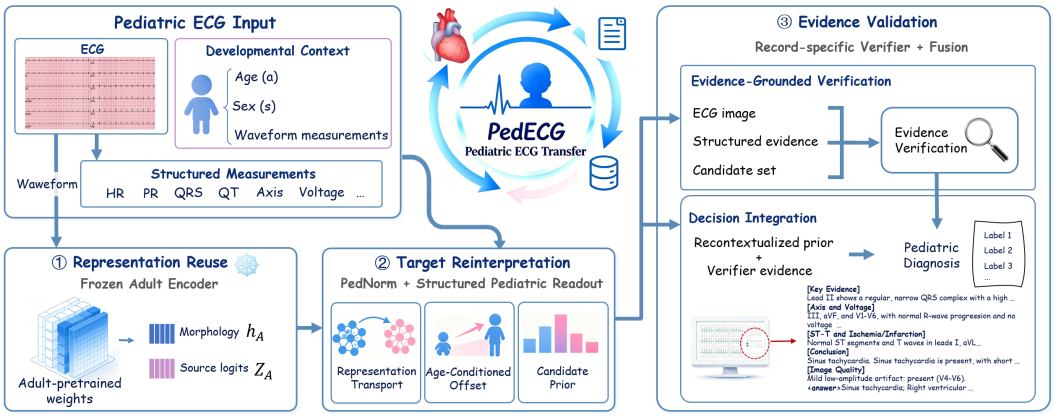


Figure 1: **PedECG: Preserve → Reinterpret → Validate.** A frozen adult encoder preserves reusable cardiac knowledge; pediatric physiological context reconstructs its target-side interpretation; and evidence from the current ECG validates candidate diagnoses before final selection. The same record-level evidence pathway also supports structured diagnostic reporting.

the value of direct pediatric modeling for congenital heart disease detection, ventricular dysfunction screening, and ECG classification [16, 17, 18, 19, 20], while ZZU-pECG provides a relatively large labeled pediatric ECG resource [21]. These studies show that pediatric ECG prediction is feasible when target-domain data are available. PedECG instead studies how knowledge learned from much larger adult cohorts can be reused when the target physiological reference frame itself changes with development.

Transfer under changing decision rules. Classical transfer learning and domain adaptation distinguish reusable source features from target-specific predictors, often through representation alignment or target-classifier adaptation [22, 23, 24]. Work on deep feature transferability further shows that source representations can remain useful even when later prediction layers become task- or domain-specific [25, 26, 27]. Adult-to-pediatric ECG transfer introduces an additional structure: the target interpretation is indexed by developmental context, so similar source evidence may cross different diagnostic boundaries at different ages.

PedECG operationalizes this distinction by keeping the adult waveform interface fixed and separating three questions that are usually entangled in end-to-end adaptation: whether source representations remain reusable, how their evidence should be mapped into the pediatric label space, and whether the resulting candidate diagnosis is supported by the current record. This makes representation reuse, target-side reinterpretation, and evidence validation experimentally separable within one transfer setting.

3 Method

3.1 Problem Setup and Transfer Interface

PedECG treats adult pretraining as a fixed source of cardiac evidence rather than a predictor to be fine-tuned wholesale. This design makes it possible to ask whether the adult representation itself transfers before introducing pediatric decision adaptation. For each pediatric ECG record, let $x_s \in \mathbb{R}^{L\times T}$ denote the 10-second waveform with L leads and T samples, x_i its rendered image, $f \in \mathbb{R}^{D_f}$ the vector of waveform derived quantitative measurements from FeatureDB [28], and $d = (a, \text{sex})$ the demographic context with chronological age a. The diagnostic target is a multilabel set $Y \subseteq \mathcal{Y}_P$, where $\mathcal{Y}_P$ denotes the pediatric label space. Because the adult and pediatric datasets use different diagnostic ontologies, we do not assume a one to one mapping between the adult label space $\mathcal{Y}_A$ and $\mathcal{Y}_P$.

Adult pretraining provides a frozen waveform encoder E_A and source head C_A, producing morphology features $h_A = E_A(x_s) \in \mathbb{R}^{D_h}$ and adult ontology evidence $z_A = C_A(h_A) \in \mathbb{R}^{|\mathcal{Y}_A|}$. Both E_A and C_A remain fixed during pediatric adaptation, so pediatric learning on the candidate pathway is confined to the target-side interpretation of the fixed source interface (h_A, z_A).

We denote the age-indexed pediatric normative reference table as PEDNORM. For each covered measurement j, let $[l_j(a), u_j(a)]$ be its healthy reference interval at age a and f_j the observed value. We define the reference-relative coordinate

$$\delta_j(a) = \mathrm{clip}\left(\frac{f_j - \frac{1}{2}(l_j(a) + u_j(a))}{\max(u_j(a) - l_j(a), \epsilon)}, [-5, 5]\right), \tag{1}$$

where $\epsilon > 0$ prevents division by zero and clipping truncates extremes. The resulting $\delta(a)$ locates each measurement relative to its age-matched pediatric reference band using only pre label information. Reference construction and lookup rules are given in Appendix A.3; Appendix A.1 summarizes input provenance.

As illustrated in figure 1, PedECG then recontextualizes the fixed adult evidence (h_A, z_A) using developmental context to construct pediatric candidate scores, and subsequently evaluates those candidates with record-specific evidence before final selection.

3.2 TARGET-SIDE REINTERPRETATION UNDER DEVELOPMENTAL CONTEXT

Target-side reinterpretation provides the bridge between reusable adult knowledge and pediatric decision semantics. The same adult feature can carry different relevance at different developmental stages, so the target domain needs a controlled mechanism for changing which source evidence is emphasized, how it is mapped into the pediatric ontology, and where the resulting decision boundary is placed without relearning the source encoder. We therefore introduce age-conditioned gates over morphology and source-domain logits,

$$\widetilde{h}_A = g_h(a) \odot h_A, \qquad \widetilde{z}_A = g_z\big(a, \delta(a)\big) \odot z_A, \tag{2}$$

where $g_h(a) \in (0,1)^{D_h}$ and $g_z(a, \delta(a)) \in (0,1)^{|\mathcal{Y}_A|}$ are learnable sigmoid gates matched in dimension to h_A and z_A, and $\odot$ denotes elementwise multiplication. We then map the gated source evidence into the pediatric label space,

$$r_P(a) = W_h \widetilde{h}_A + W_f \delta(a) + b + \sum_{k=1}^{K_\pi} \pi_k(a) P_k \widetilde{z}_A, \tag{3}$$

where $r_P(a) \in \mathbb{R}^{|\mathcal{Y}_P|}$ is a pediatric evidence score vector rather than a generic hidden representation, W_h, W_f, and b are learnable projection parameters, $\{P_k\}_{k=1}^{K_\pi}$ are low-rank source-to-target bases, and $\pi(a)$ is an age-conditioned softmax mixture over the K_π bases whose weights are nonnegative and sum to one. Age therefore affects which adult dimensions are retained, how source-domain evidence is projected, and where the record lies relative to pediatric physiological references.

Development-conditioned candidate readout. An age-indexed offset $\tau(a) \in \mathbb{R}^{|\mathcal{Y}_P|}$ converts $r_P(a)$ into candidate probabilities $p(a) = \sigma\big(r_P(a) - \tau(a)\big)$, where σ is the elementwise sigmoid and $p_y(a)$ is the raw candidate probability for label $y \in \mathcal{Y}_P$. For fixed adult evidence, the resulting pediatric decision surface is therefore conditional on developmental context. Age and reference-relative measurements can alter both the effective direction through which source evidence is read and the offset at which a pediatric label is activated. This provides the geometric interpretation of target-side reinterpretation analyzed in Sec. 4.4; see Appendix B.2. The candidate module is trained with multilabel BCE and a pairwise ranking objective; full optimization details are given in Appendix A.4. After scoring, we retain the top-K candidates by $p_y(a)$, where the candidate budget K is separate from the mixture count K_π and is fixed on validation. Candidate scores are then validation-remapped per label to calibrated probabilities $\bar{p}_y$ before verification; all calibration and selection parameters are fixed before test evaluation. These remaps support F1 selection rather than clinically calibrated probabilities.

Sec. 4.3 compares this structured recontextualization with a simple concatenation control under the same Top-K, verifier, and fusion protocol. This comparison evaluates two alternative strategies for organizing the same frozen adult evidence and underlying quantitative measurements under a shared Top-K selection, verifier, and fusion protocol: concatenation receives an explicit feature vector that includes age, PEDNORM coordinates, and normalized raw measurements, whereas the structured readout consumes measurements primarily through the reference-relative coordinates $\delta(a)$ and age-conditioned gates/transport. Implementation details of the matched readout variants appear in Appendix B.1.

3.3 Record-Specific Evidence Verification and Fusion

High-recall transfer is useful only if competing candidates can be resolved against evidence from the current record. The structured readout therefore produces a candidate prior that is subsequently refined by record-specific evidence verification. For each retained candidate y, the verifier receives the ECG image x_i, FeatureDB measurements f, demographic context d, the candidate identity, and the calibrated probability $\bar{p}_y$. It estimates a conditional support score $q_y \in (0, 1)$ through a binary verbalizer, implemented as normalized next token probabilities for `Yes` and `No`.

The final diagnostic score combines the calibrated prior and verifier support,

$$s_y = \sigma\big(\beta_0 + \beta_1 \operatorname{logit}(\bar{p}_y) + \beta_2 \operatorname{logit}(q_y) + \beta_3^\top \rho_y\big)\,, \tag{4}$$

where $\operatorname{logit}(u) = \log \frac{u}{1-u}$; $\beta_0, \beta_1, \beta_2$ are scalar fusion coefficients; β_3 is a coefficient vector; and ρ_y collects rank and label covariates detailed in Appendix A.5. A label enters the prediction set $\widehat{Y}$ when s_y exceeds a validation selected threshold. Fusion parameters and the threshold are held fixed for test evaluation. This construction separates candidate retrieval from diagnostic selection, so high recall proposal coverage and within pool disambiguation can be quantified independently; see Sec. 4.2.

PedECG additionally produces structured ECG reports from the same record evidence, organizing rhythm, intervals, axis and voltage, conduction, repolarization, and supporting findings into an interpretable diagnostic summary. Illustrative examples are shown in Appendix E; reporting is kept separate from the scored diagnostic endpoint.

3.4 Training Procedure

The training schedule mirrors the conceptual decomposition of PedECG. Adult waveform encoders and ECG-adapted language backbones are learned first, after which the adult waveform interface is frozen and pediatric adaptation focuses on target-side reinterpretation and record-specific verification. Pediatric adaptation then optimizes the candidate readout and verifier. All model selection, candidate budgets, probability remapping, fusion parameters, and thresholds are fixed on validation before test evaluation. Full optimization and initialization details are provided in Appendix A.4.

4 Experiments

4.1 Experimental Setup

Datasets. MIMIC-IV-ECG-derived adult data provide source-domain waveform supervision, while rendered adult ECGs paired with report text derived from MIMIC-IV-ECG provide domain-specific vision-language supervision [8]. Pediatric adaptation and primary evaluation use the patient-disjoint ZZU-pECG partition [21]: 9,943 training ECGs from 8,150 patients, 1,428 validation ECGs from 1,164 patients, and 2,819 test ECGs from 2,329 patients. All recordings from the same child remain in a single split.

Input protocol and supervision. The diagnostic input is restricted to information available before the pediatric reference label is known: ECG waveform, rendered image, age, sex, waveform derived FeatureDB measurements, deterministic age-indexed PedNorm ranges, and outputs from the frozen adult waveform model. Pediatric gold labels, source report text, and synthetic teacher reports are target-side supervision only and are never serialized into diagnostic prompts. Appendix A.1 lists the provenance of every input field.

Evaluation label space and metrics. The primary endpoint is micro-F1 against ECG-observable reference labels. Our endpoint evaluates agreement with the dataset provided ECG-observable reference labels after the predefined ontology filtering described in Appendix A.5; we do not interpret these labels as independently adjudicated clinical ground truth. All methods use the same alias normalization and observable gold partition; canonical predictions outside that reference partition are retained and count as false positives. We report micro precision and recall alongside F1, and exact set match and at least one recovery in Table 9. Confidence intervals for PedECG use patient-cluster bootstrap resampling with 2,000 draws over 2,329 patients; bootstrap and multi seed conventions are

Table 1: Pediatric diagnostic performance on the patient-disjoint test set.

Method	Recall	Precision	Micro-F1
ECG specialized external baseline			
ECG-R1, native	0.158	0.350	0.218
ECG-R1, shared-candidate matched	0.303	0.380	0.337
General purpose models			
DeepSeek-v4-flash, shared-candidate	0.243	0.242	0.243
DeepSeek-v4-pro, shared-candidate	0.176	0.205	0.189
Pediatric supervised baseline			
Pediatric ResNet1D trained from scratch	0.477	0.460	0.467
Qwen3-VL-4B based systems			
Zero-shot inference	0.287	0.285	0.286
PedECG	**0.527**	0.484	**0.504**

Note. n=2819 ECG records. Gold labels are restricted to the ECG-observable reference partition; predictions are not pruned to this partition and unmatched predictions count as false positives. Shared-candidate matched inputs reuse PedECG Top-15 proposals. PedECG F1 patient-bootstrap 95% CI $[0.491, 0.518]$. Scratch is a 3-seed mean. Matched transfer and readout controls appear in table 2.

detailed in Appendix A.5. Where relevant we also follow reporting recommendations for clinical AI evaluation [29, 30].

Compared models. ECG-R1 serves as an ECG specialized external baseline; we report both its native protocol and a *shared-candidate* matched input protocol that supplies FeatureDB, PEDNORM, and the same PedECG Top-15 candidates rather than each model's own proposals. DeepSeek-v4-flash and DeepSeek-v4-pro are general purpose multimodal models under that same protocol. None of these external systems receives pediatric fine tuning. External systems therefore provide application-level reference points under native or shared-candidate protocols, while the matched internal controls in Sec. 4.3 are used to study transfer mechanisms. Qwen3-VL-4B-based comparisons use the same VLM backbone for zero-shot inference and for PedECG's verifier/report adapters. Matched transfer and readout controls include a pediatric-only scratch ResNet, a frozen adult linear probe, a randomly initialized frozen probe, and a simple concat readout; see table 2. Alternative VLM adaptation pathways are reported in Appendix B.4.

4.2 MAIN RESULTS

Overall pediatric ECG label prediction. PedECG achieves the strongest overall performance among the evaluated systems in Table 1 under their reported protocols, reaching 0.504 micro-F1 on the locked patient-disjoint test set. The supervised pediatric ResNet reaches 0.467 micro-F1, so the improvement is not restricted to zero-shot or externally trained references. More importantly for the transfer story, the 0.913 candidate recall shows that the adult-derived pathway already retrieves most observable pediatric labels within a modest candidate budget. The dominant challenge is therefore no longer candidate discovery alone, but selecting which transferred hypotheses remain valid under pediatric physiology and the evidence of the current record. This observation motivates the mechanism analyses in Sec. 4.3. Performance by label prevalence and age is reported in Appendices D.3 and C.2; PedECG recovers at least one reference label in 0.700 of records with nonempty observable reference sets.

From candidate coverage to diagnostic commitment. The Top-15 stage deliberately preserves multiple plausible diagnoses rather than collapsing early to a single inherited decision rule. PedECG resolves this ambiguity through record-specific evidence verification and calibrated fusion, turning broad transferred coverage into target-domain diagnostic commitment. The incremental value of this stage is isolated under matched controls in table 3.

4.3 UNDERSTANDING THE TRANSFER MECHANISM

The central experimental question is not only whether PedECG improves pediatric prediction, but *what transfers and what must change*. We organize the analysis around three hypotheses that standard

Table 2: Pediatric from scratch and adult transfer controls for representation reuse and target readout design. Block A evaluates encoder representations; Block B compares readouts under the same downstream pipeline. Results are not directly comparable across blocks.

Method	Recall	Precision	Micro-F1
A. Representation reuse			
Pediatric ResNet1D trained from scratch	0.477	0.460	0.467
Random frozen ResNet1D + linear probe	0.399	0.176	0.245
Adult pretrained frozen ResNet1D + linear probe	0.487	0.446	0.464
B. Readout design			
Naive concatenation	0.526	0.459	0.490
PedECG structured readout	**0.527**	0.484	**0.504**

Note. n=2819. Block A: Scratch and adult probe are 3-seed means, reported as F1 $\pm$0.007 and $\pm$0.003, respectively; random probe is a single seed control. Block B: Both use the same frozen adult evidence, Top-15 candidate budget, verifier architecture, and validation-fitted fusion protocol; measurement context is encoded differently as described in the text.

end-to-end evaluation leaves entangled. Each intervention targets one stage of the pipeline while holding the remaining evaluation protocol fixed whenever possible.

H1: Does pediatric transfer require relearning the waveform representation? A linear probe on frozen adult features reaches 0.464$\pm$0.003 micro-F1 on the patient-disjoint test set, closely matching a ResNet1D trained end-to-end from pediatric data alone at 0.467$\pm$0.007, while an architecture-matched randomly initialized frozen encoder falls to 0.245; see table 2. The result is clear: adult pretraining already provides a representation whose pediatric diagnostic utility is largely preserved without any pediatric encoder updates. Representation reuse is therefore a viable starting point for the target task rather than merely an initialization convenience.

H2: Does preserved adult evidence benefit from a new target-side readout? The concatenation control receives the same frozen adult evidence and underlying FeatureDB measurements as the structured readout, together with an explicit age/PEDNORM feature vector, under a shared Top-15, verifier, and validation-fitted fusion protocol. Structured reinterpretation reaches 0.504 versus 0.490 for naive concatenation, with a paired patient-bootstrap difference of Δ+0.014 and 95% CI $[0.006, 0.023]$; see table 2. This supports the value of organizing transferred evidence through an explicit pediatric readout rather than relying on direct feature concatenation. Additional structured variants in Appendix B.1 remain within a narrow performance range, indicating that the advantage of structured target-side readout is not dependent on a single parameterization. Because aggregate endpoint performance cannot reveal how developmental context changes model behavior, Sec. 4.4 directly probes the learned interpretation under controlled context substitution.

H3: Is candidate transfer sufficient without record-specific validation? Holding the scored Top-15 pool fixed, adding verifier support raises test micro-F1 from 0.445 to 0.504 (Δ+0.059, 95% CI $[0.049, 0.068]$); see table 3, Block A. Replacing the correct ECG image with a blank or mismatched image reduces performance to 0.444 and 0.419, respectively, showing that the verifier is not simply reusing the candidate prior: it depends on evidence from the current ECG (Block B). Together, these controls show that high-recall candidate transfer is not the end of the problem; record-specific validation makes a distinct contribution to final diagnostic selection. Alternative VLM adaptation pathways are reported in Appendix B.4.

4.4 DEVELOPMENTAL CONTEXT CHANGES MODEL INTERPRETATION

We next test the behavioral consequence of target-side developmental context directly: does the pediatric readout change when the physiological reference frame changes while the underlying ECG evidence is held fixed?

Developmental context substitution probe. We hold the waveform x_s, raw measurements f, and frozen adult evidence (h_A, z_A) fixed while substituting age together with the corresponding PEDNORM reference context. On an exploratory 20-record held-out probe described in Appendix C.1, this substitution changes the thresholded candidate set in 19/20 cases and the top ranked label in

Table 3: Evidence verification controls. Block A isolates the incremental value of verifier support under matched covariates; Block B changes only the verifier image under the locked pipeline.

Setting	Test F1	Δ
A. Incremental verifier value under matched covariates		
Matched fusion without q	0.445	—
Matched fusion with q	**0.504**	+0.059
B. Image intervention under locked fusion		
Correct record image	**0.504**	—
Blank image	0.444	−0.061
Mismatched other-record image	0.419	−0.086

Note. n=2819. Blocks A–B share the Top-15 pool. Block A: VAL fit once, TEST apply; ΔF1 95% CI $[0.049, 0.068]$. Block B: locked fusion, no refit; only the verifier image changes.

12/20. The candidate set statistic uses a probe threshold of 0.5 on raw candidate probabilities rather than the locked Top-15/fusion decision.

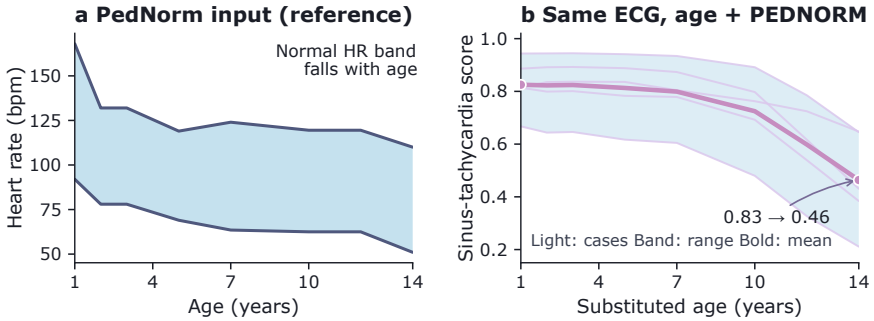


Figure 2: Reference context and learned age response. **(a)** Age-matched PEDNORM heart rate band supplied as model input, taken from a literature table rather than learned. **(b)** Age/PEDNORM substitution with ECG and measurements held fixed. Light curves: five records with top ranked sinus tachycardia at age 1; bold curve: mean; shading shows the across-record range. Scores are pre fusion.

Figure 2 illustrates this substitution for sinus tachycardia: as the age-indexed reference range changes while the ECG and raw measurements remain fixed, the candidate score changes substantially. Developmental context is therefore functionally active in the learned readout rather than serving as passive metadata: the same adult evidence is interpreted differently when the pediatric physiological reference frame changes. The substitution probe is designed to establish this functional responsiveness; it does not by itself quantify the independent endpoint gain or physiological correctness of a substituted trajectory.

Scope and future directions. This study focuses on identifying the transfer mechanism within a single pediatric cohort and ontology. Multi-institution evaluation is the natural next step for testing whether the same preserve–reinterpret–validate decomposition generalizes across sites and acquisition settings. The present substitution analysis establishes functional developmental responsiveness; future matched-architecture studies can isolate the independent contribution of developmental conditioning more directly. Rare-label estimation remains data limited, motivating larger pediatric cohorts and targeted evaluation of low-prevalence diagnoses. Prospective clinical validation and broader demographic fairness analyses will be required before deployment.

5 CONCLUSION

Cross-population medical AI is often evaluated as a single adaptation problem, even when the source representation and the target decision context change for different reasons. In adult-to-pediatric ECG transfer, PedECG makes this distinction explicit. The framework preserves adult-learned cardiac representations, reconstructs their diagnostic use under pediatric physiological context, and validates candidate diagnoses against record-specific evidence before final selection.

Our analyses show that adult waveform representations remain highly reusable after the population shift, while target-side reinterpretation and evidence validation provide explicit mechanisms for

adapting how preserved source knowledge is used in pediatric prediction. This stage-resolved view changes what transfer needs to optimize: a population shift need not always require relearning the underlying representation, but successful reuse depends on identifying which parts of the evidence-to-decision pathway must be reconstructed for the target population.

More broadly, PedECG provides a practical template for extending large-scale medical knowledge beyond the cohorts on which it was learned. By separating reusable physiological representations from population-specific interpretation and record-level validation, it points toward medical AI systems that can leverage data-rich source populations while adapting responsibly to clinically distinct and data-limited groups.

Reproducibility and Ethics

Anonymous code for the developmental candidate module, evaluation helpers, and fusion protocol implementation is available at `https://anonymous.4open.science/r/PedECG-B7E1`. The release includes patient level split identifiers, input provenance audits, label mappings, training configurations, candidate files, and validation locked decision rules. All headline results use the patient-disjoint protocol in section 4.1. Dataset access and ethics follow the corresponding source documentation.

Use of Large Language Models

Large language models were used to assist with manuscript drafting and editing, light code review, and generation of synthetic report targets used only for the structured reporting branch. They were not used to construct diagnostic test labels or diagnostic inputs. All scientific claims, experimental results, references, and final wording were verified by the authors.

References

[1] PR Rijnbeek, Maarten Witsenburg, E Schrama, Jakob Hess, and JA Kors. New normal limits for the paediatric electrocardiogram. *European heart journal*, 22(8):702–711, 2001.

[2] Hung-Chi Lue, Mei-Hwan Wu, Jou-Kou Wang, Ming-Tai Lin, Chun-Wei Lu, Shuenn-Nan Chiu, Chun-An Chen, En-Ting Wu, Ching-Chia Wang, Chun-Min Fu, et al. Normal ecg standards and percentile charts for infants, children and adolescents. *Pediatrics & Neonatology*, 64(3): 256–273, 2023.

[3] András Bratincsák, Chieko Kimata, Blair N Limm-Chan, Kevin P Vincent, Matthew R Williams, and James C Perry. Electrocardiogram standards for children and young adults using z-scores. *Circulation: Arrhythmia and Electrophysiology*, 13(8):e008253, 2020.

[4] Elizabeth V Saarel, Suzanne Granger, Jonathan R Kaltman, L LuAnn Minich, Martin Tristani-Firouzi, Jeffrey J Kim, Kathleen Ash, Sabrina S Tsao, Charles I Berul, Elizabeth A Stephenson, et al. Electrocardiograms in healthy north american children in the digital age. *Circulation: Arrhythmia and Electrophysiology*, 11(7):e005808, 2018.

[5] Awni Y Hannun, Pranav Rajpurkar, Masoumeh Haghpanahi, Geoffrey H Tison, Codie Bourn, Mintu P Turakhia, and Andrew Y Ng. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature medicine*, 25(1):65–69, 2019.

[6] Antônio H Ribeiro, Manoel Horta Ribeiro, Gabriela MM Paixão, Derick M Oliveira, Paulo R Gomes, Jéssica A Canazart, Milton PS Ferreira, Carl R Andersson, Peter W Macfarlane, Wagner Meira Jr, et al. Automatic diagnosis of the 12-lead ecg using a deep neural network. *Nature communications*, 11(1):1760, 2020.

[7] Patrick Wagner, Nils Strodthoff, Ralf-Dieter Bousseljot, Dieter Kreiseler, Fatima I Lunze, Wojciech Samek, and Tobias Schaeffter. Ptb-xl, a large publicly available electrocardiography dataset. *Scientific data*, 7(1):154, 2020.

[8] Brian Gow, Tom Pollard, Larry A Nathanson, Alistair Johnson, Benjamin Moody, Chrystinne Fernandes, Nathaniel Greenbaum, Jonathan W Waks, Parastou Eslami, Tanner Carbonati, et al. Mimic-iv-ecg-diagnostic electrocardiogram matched subset. *Type: dataset*, 2023.

[9] Nils Strodthoff, Patrick Wagner, Tobias Schaeffter, and Wojciech Samek. Deep learning for ecg analysis: Benchmarks and insights from ptb-xl. *IEEE journal of biomedical and health informatics*, 25(5):1519–1528, 2020.

[10] Zhongwei Wan, Che Liu, Xin Wang, Chaofan Tao, Hui Shen, Jing Xiong, Rossella Arcucci, Huaxiu Yao, and Mi Zhang. Meit: Multimodal electrocardiogram instruction tuning on large language models for report generation. In *Findings of the association for computational linguistics: ACL 2025*, pages 14510–14527, 2025.

[11] Jiarui Jin, Haoyu Wang, Xingliang Wu, Xiaocheng Fang, Xiang Lan, Zihan Wang, Deyun Zhang, Bo Liu, Yingying Zhang, Xian Wu, et al. Ecg-r1: protocol-guided and modality-agnostic mllm for reliable ecg interpretation. *arXiv preprint arXiv:2602.04279*, 2026.

[12] Yubao Zhao, Jiaju Kang, Tian Zhang, Puyu Han, and Tong Chen. Ecg-chat: A large ecg-language model for cardiac disease diagnosis. In *2025 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6. IEEE, 2025.

[13] Konstantinos C Siontis, Peter A Noseworthy, Zachi I Attia, and Paul A Friedman. Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. *Nature Reviews Cardiology*, 18(7):465–478, 2021.

[14] Sulaiman Somani, Adam J Russak, Felix Richter, Shan Zhao, Akhil Vaid, Fayzan Chaudhry, Jessica K De Freitas, Nidhi Naik, Riccardo Miotto, Girish N Nadkarni, et al. Deep learning and the electrocardiogram: review of the current state-of-the-art. *Ep Europace*, 23(8):1179–1191, 2021.

[15] 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. *Advances in Neural Information Processing Systems*, 38:94421–94455, 2026.

[16] Jintai Chen, Shuai Huang, Ying Zhang, Qing Chang, Yixiao Zhang, Dantong Li, Jia Qiu, Lianting Hu, Xiaoting Peng, Yunmei Du, et al. Congenital heart disease detection by pediatric electrocardiogram based deep learning integrated with human concepts. *Nature Communications*, 15(1):976, 2024.

[17] Joshua Mayourian, William G La Cava, Akhil Vaid, Girish N Nadkarni, Sunil J Ghelani, Rebekah Mannix, Tal Geva, Audrey Dionne, Mark E Alexander, Son Q Duong, et al. Pediatric ecg-based deep learning to predict left ventricular dysfunction and remodeling. *Circulation*, 149(12):917–931, 2024.

[18] Hiroki Mori, Kei Inai, Hisashi Sugiyama, and Yoshihiro Muragaki. Diagnosing atrial septal defect from electrocardiogram with deep learning. *Pediatric cardiology*, 42(6):1379–1387, 2021.

[19] Shuhei Toba, Yoshihide Mitani, Yusuke Sugitani, Hiroyuki Ohashi, Hirofumi Sawada, Mami Takeoka, Naoki Tsuboya, Kazunobu Ohya, Noriko Yodoya, Takato Yamasaki, et al. Deep learning-based analysis of 12-lead electrocardiograms in school-age children: a proof of concept study. *Frontiers in Cardiovascular Medicine*, 12:1471989, 2025.

[20] David M Leone, Donnchadh O'Sullivan, and Katia Bravo-Jaimes. Artificial intelligence in pediatric electrocardiography: a comprehensive review. *Children*, 12(1):25, 2024.

[21] Jian Tan, Haoyi Fan, Jiawei Luo, Yanjie Zhou, Ning Wang, Xizheng Wang, Guizhi Liu, Chengyu Liu, and Zongmin Wang. A pediatric ecg database with disease diagnosis covering 11643 children. *Scientific Data*, 12(1):867, 2025.

[22] Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10):1345–1359, 2010.

[23] Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario March, and Victor Lempitsky. Domain-adversarial training of neural networks. *Journal of machine learning research*, 17(59):1–35, 2016.

[24] Mingsheng Long, Yue Cao, Jianmin Wang, and Michael Jordan. Learning transferable features with deep adaptation networks. In *International conference on machine learning*, pages 97–105, 2015.

[25] Jason Yosinski, Jeff Clune, Yoshua Bengio, and Hod Lipson. How transferable are features in deep neural networks? In *Advances in neural information processing systems*, volume 27, 2014.

[26] Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, and Trevor Darrell. DeCAF: A deep convolutional activation feature for generic visual recognition. In *International conference on machine learning*, pages 647–655, 2014.

[27] Simon Kornblith, Jonathon Shlens, and Quoc V Le. Do better ImageNet models transfer better? In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2661–2671, 2019.

[28] Shenda Hong, Meng Wu, Yuxi Zhou, Qingyun Wang, Junyuan Shang, Hongyan Li, and Junqing Xie. Encase: An ensemble classifier for ecg classification using expert features and deep neural networks. In *2017 Computing in cardiology (cinc)*, pages 1–4. IEEE, 2017.

[29] Gary S Collins, Karel GM Moons, Paula Dhiman, Richard D Riley, Andrew L Beam, Ben Van Calster, Marzyeh Ghassemi, Xiaoxuan Liu, Johannes B Reitsma, Maarten Van Smeden, et al. Tripod+ ai statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *bmj*, 385, 2024.

[30] Baptiste Vasey, Myura Nagendran, Bruce Campbell, David A Clifton, Gary S Collins, Spiros Denaxas, Alastair K Denniston, Livia Faes, Bart Geerts, Mudathir Ibrahim, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: Decide-ai. *Nature medicine*, 28(5):924–933, 2022.

A Reproducibility and Evaluation Details

A.1 Input Provenance and Leakage Audit

Table 4: Provenance and inference availability of diagnostic inputs.

Input	Origin	Gold label dependent?	Available at inference?
ECG waveform	raw record	no	yes
Rendered ECG image	raw waveform rendering	no	yes
Age / sex	record metadata	no	yes
FeatureDB measurements	engineered waveform features [28]	no	yes
PEDNORM ranges	age-indexed literature table	no	yes
Adult evidence h_A, z_A	frozen adult only model	no	yes
Synthetic pediatric report	training-side remapped/LLM-assisted targets	yes	no; supervision only
Pediatric gold labels	dataset reference	yes	no; supervision only

A.2 Adult Source Pretraining

The frozen encoder E_A and source head C_A are trained once on MIMIC-IV-ECG-derived adult waveforms [8] before any pediatric adaptation, then held fixed. Training uses multilabel binary cross entropy on a 32-label adult diagnostic ontology extracted from the adult export manifest, a ResNet1D encoder with channel widths $[128, 192, 256, 320]$, depths $[2, 2, 3, 3]$, and embedding dimension 512, AdamW with learning rate 3×10^{-4}, weight decay 10^{-4}, batch size 256, gradient clipping 1.0, and up to 20 epochs with early stopping on validation at patience 3. The training configuration caps the adult train split at $100{,}000$ records and the validation split at $5{,}000$ records for throughput; per lead z-score normalization and an on disk waveform cache are used. No pediatric ECGs enter this stage. The resulting (h_A, z_A) interface is the only adult waveform pathway consumed by the pediatric candidate module.

A.3 PEDNORM Construction

PEDNORM is a deterministic, age-indexed reference table used only as pre label context. Numeric bands are taken from a curated pediatric reference JSON derived primarily from Lue et al. interval/percentile tables [2], with related pediatric reference literature motivating the developmental framing [1, 3, 4]. Chronological age in days is mapped to one of twelve inclusive age bins by table lookup; ages beyond the last bin are clamped to that bin. For each covered measurement j, the band endpoints $[l_j(a), u_j(a)]$ are the tabulated lower/upper percentiles; the implementation uses $p02$/$p98$ when present. Under Eq. 1, δ_j is scaled by the full band width $u_j(a) - l_j(a)$, so the endpoints map to $\pm\frac{1}{2}$ rather than ±1; values are clipped to $[-5, 5]$. FeatureDB keys with a matching band use Eq. 1; keys without a band, notably frontal axis, use a fixed scale of axis$/90$ with clipping; missing measurements or empty bands contribute zeros. PEDNORM never depends on pediatric gold labels and is recomputed whenever a substituted age is supplied.

A.4 Locked Main Run Configuration

Table 5 summarizes the fixed configuration that produces the headline test result of 0.504 micro-F1 on n=2819 records. Checkpoint names and repository paths are omitted here; the anonymous code release pins the corresponding artifacts.

Missing measurements are zero imputed in FeatureDB/PEDNORM blocks. Records without usable age use a neutral age embedding in age aware analyses only; the main fusion rule does not take age as a covariate. Alias normalized strings outside the ECG-observable gold partition remain eligible as false positives. Empty ECG-observable gold sets contribute zero true positives and do not enter the at least one hit denominator (n=2708).

Table 5 specifies the primary diagnostic run, not the external baselines or alternative adaptation paths. Table 3 and Table 7 hold scored candidates fixed while changing the decision rule or verifier image; nested refits are separate from the locked main rule. Table 8 (Appendix B.4) changes the adaptation

Table 5: Locked configuration for the primary PedECG diagnostic result.

Component	Setting
Data split	Patient-disjoint train/val/test on ZZU-pECG
Adult encoder E_A+C_A	Frozen ResNet1D waveform encoder
Candidate prior	Age-conditioned dynamic readout; vocabulary of 104 ECG-observable labels
Capacity	K_π=4; hidden 128; age embedding 32
Age features	$[\log(1+d),\ y,\ \sin\phi,\ \cos\phi]$ with d in days, y=$d/365.25$, ϕ=$2\pi\ \min(y/18, 1)$
Training	BCE + λ_{rank}=0.5 pairwise ranking with $\leq$ 32 negatives; λ_{gate}=λ_τ=0.01; 20 epochs
Selection	Top-15 with high confidence k=5 and escape ranks 6–15; validation probability remapping
Vision language model	Qwen3-VL-4B-Instruct with LoRA at r=16, α=32, dropout 0.05
Pediatric adapters	Diagnostic LoRA, then a one epoch Yes/No verifier initialized from it
Verification	Support score q_y from Yes/No next token probabilities
Fusion	L2 logistic on $\text{logit}(q)$, $\text{logit}(\bar{p})$, escape flag for ranks 6–15, four rank buckets, and 32 label biases
Lock	Class balanced fit on validation; decision threshold 0.70; applied once on test with no refit
Endpoint	Gold restricted to ECG-observable labels; predictions not pruned to that partition

pathway, whereas Appendix B.1 varies the candidate readout parameterization with verifier weights held fixed. Each changed candidate pool fits its own validation fusion rule before test application.

Training and data flow overview. Table 6 summarizes the staged adaptation path used by the locked diagnostic run. Supervision and trainable parameters are confined to the listed pediatric stages; the adult waveform encoder remains frozen during pediatric training, and reporting never feeds generated text back into diagnosis.

Table 6: Staged training and data flow for the locked PedECG diagnostic pathway.

Stage	Split	Input	Supervision	Trainable / output
Adult waveform pretraining	adult train	waveform x_s	adult labels	$E_A, C_A \rightarrow (h_A, z_A)$
Pediatric candidate readout	ped. train	(h_A, z_A), age, $\delta(a)$ from f	ECG obs. labels	readout $\rightarrow p(a)$
Pediatric diagnostic LoRA	ped. train	image + structured context	diagnosis targets	LoRA init
Verifier	ped. train	image, f, candidate, $\bar{p}_y$	Yes/No support	verifier LoRA $\rightarrow q_y$
Calibration / fusion	validation	candidate/verifier scores	val. labels	remap + logistic lock
Test	test	pre-label evidence only	none	frozen $\rightarrow \widehat{Y}$

Candidate module architecture details. Age features pass through a two layer GELU MLP ($4 \rightarrow 32 \rightarrow 32$). Gates g_h and g_z are likewise two layer GELU MLPs ending in a sigmoid; g_h sees only the age embedding, while g_z concatenates age embedding with PEDNORM coordinates when enabled. Source to target mixture bases use low-rank factors with rank $\max(4, \min(32, \lfloor d_z/2 \rfloor, \lfloor |\mathcal{Y}_P|/4 \rfloor))$. The candidate module is optimized with a candidate task loss

$$\mathcal{L}_{\text{cand}} = \mathcal{L}_{\text{BCE}} + \lambda_{\text{rank}}\mathcal{L}_{\text{rank}} + \lambda_{\text{gate}}\mathcal{L}_{\text{gate}} + \lambda_\tau\mathcal{L}_\tau, \tag{5}$$

where $\mathcal{L}_{\text{BCE}}$ is multilabel binary cross entropy and $\lambda_{\text{rank}} > 0$ balances a pairwise ranking term. The ranking term is a softplus margin between each positive logit and up to 32 randomly sampled negatives within the same record, with λ_{rank}=0.5 in the locked run. Gate regularization $\mathcal{L}_{\text{gate}}$ penalizes mean gate values away from 0.5; offset smoothness $\mathcal{L}_\tau$ compares τ, gates, and mixture weights at neighboring ages during training, with λ_{gate}=λ_τ=0.01. Candidate level calibration fits a per label affine map $\bar{p}_y = \text{Cal}_y(p_y)$ on validation logits when both positive and negative counts meet the minimum of 10. Escape candidates are ranks 6–15 from the age recalibrated prior; the verifier

scores every retained Top-15 label independently with a Yes/No verbalizer and does not generate free form diagnoses.

In the locked fusion rule of Eq. 4, the covariate vector ρ_y contains an escape indicator for candidates ranked 6–15 when the high confidence prior retains the Top-5, four rank buckets, and indicators for 32 fixed validation selected labels. Adult waveform pretraining and ECG adapted language backbone initialization precede pediatric adaptation; the main verifier is initialized from a pediatric diagnosis adapted checkpoint, with direct adult ECG adaptation evaluated separately in Appendix B.4. Reference context ablations remove the relevant features consistently from training and inference within the ablated component.

A.5 Label Space, Aliases, and Metrics

Primary metrics retain only ECG-observable gold labels: structural and etiologic findings such as septal defects, cardiomyopathy, myocarditis, and Kawasaki disease, electrolyte/drug effects, and catch all "Otherwise" labels are dropped from gold by a fixed exclusion list of exact matches and substring rules. The locked candidate vocabulary contains 104 pediatric diagnosis strings already restricted to the observable space; dataset-provided reference sets may still include clinical associated labels, which are filtered at scoring time. Predictions are alias normalized with a fixed synonym map but are *not* pruned to the observable partition, so out of partition predictions count as false positives. Empty observable gold contributes no true positives and is excluded from the at least one hit denominator of n=2708. The 32 fusion label bias identities are the most frequent labels among validation scored candidates and are fixed in the locked rule before any test application; the nested fusion audit freezes that same list.

For record i, let Y_i be the reference label set and $\widehat{Y}_i$ the prediction. Define

$$T_i = |Y_i \cap \widehat{Y}_i|, \qquad F_i = |\widehat{Y}_i \setminus Y_i|, \qquad M_i = |Y_i \setminus \widehat{Y}_i|. \tag{6}$$

Micro precision, micro recall, and micro-F1 aggregate these counts over records after alias normalization. Reference labels are filtered to the ECG-observable partition; structural and etiologic labels such as septal defects, myocarditis, and electrolyte syndromes are excluded from gold. Predictions are not retrospectively pruned to remove false positives outside it. Exact set match requires $F_i = M_i = 0$. At least one hit is $\mathbf{1}[T_i \geq 1]$, reported on nonempty gold records. Support stratified scores in table 10 restrict counts to the labels in each support bin; age bin scores use record ages and the same locked predictions. Reported bootstrap intervals resample patients for PedECG, within strata for the age plot. Where reported, paired patient bootstrap intervals quantify differences between fixed trained predictors. We do not perform separate hypothesis tests, and these intervals do not capture retraining variability. Where three-seed results are reported, metrics are arithmetic means over three independently trained runs, and $\pm$ denotes the standard deviation across seeds; the same averaging convention is used for supplementary record-level metrics when multi-seed runs are available. Top-15 candidate recall is the fraction of ECG-observable gold-label occurrences contained in the candidate pool before verification and fusion. As used in this paper, "calibration" denotes validation fitting of candidate probability remaps and/or fusion coefficients and thresholds for F1 selection, not a claim of well calibrated clinical probabilities.

A.6 External Baseline Protocols

ECG-R1-8B-RL is evaluated under its native input specification and a shared-candidate matched input configuration that supplies FeatureDB, PEDNORM, and the same PedECG Top-15 candidates used by the locked diagnostic run, rather than candidates generated by ECG-R1 itself. DeepSeek matched input runs use the same shared-candidate file. Neither external system receives pediatric fine tuning. The paired protocols distinguish differences in evidence availability from target domain parameter adaptation.

B Transfer Mechanism Analyses

B.1 Structured Readout Sensitivity

Across five structured target-side readout variants (static transport/offset, conditioned transport only, conditioned offset only, without candidate-side PedNorm, and the locked full readout), test micro-F1 ranges from 0.501 to 0.506 under a shared Top-K/verifier/fusion protocol, compared with 0.490 for simple concatenation and 0.499 for an age-free concatenation baseline. Specifically, static, transport-only, offset-only, no-reference-context, and full readouts achieve 0.505, 0.503, 0.506, 0.501, and 0.504 micro-F1, respectively. The locked PedECG configuration reaches 0.504. These results indicate robustness to the particular structured readout parameterization rather than a uniquely optimal transport or offset design. We therefore do not treat test-set differences among structured variants as a basis for post-hoc model selection.

B.2 Context Conditional Decision Surface

Let $\phi_A = [h_A; z_A]$. For a fixed measurement vector f, write $\delta_f(a)$ for the reference-relative coordinates. Equation 3 becomes $r_P(a) = A(a, \delta_f(a))\phi_A + W_f\delta_f(a) + b \in \mathbb{R}^{|\mathcal{Y}_P|}$. With age-indexed offset $\tau(a)$, the candidate margin is $m(a) = r_P(a) - \tau(a)$ and $p(a) = \sigma(m(a))$. The map depends on both age and measurement context because the source logit gate uses $\delta_f(a)$. Conditional on this context, the zero margin surface for label y is

$$B_y(a; f): \quad \underbrace{\big(A(a, \delta_f(a))^\top e_y\big)}_{w_y^{\mathrm{eff}}(a;f)}{}^\top \phi_A + \underbrace{e_y^\top W_f \delta_f(a) + e_y^\top b - \tau_y(a)}_{b_y^{\mathrm{eff}}(a;f)} = 0, \tag{7}$$

where e_y selects label y. Age and reference-relative measurements can change both the effective orientation and offset, even when adult evidence is fixed. This is a context conditional surface in the candidate module, not a single age only hyperplane or the final fused diagnostic boundary.

B.3 Verifier Fusion Audit

The exploratory with/without-q refits freeze the locked bias label set and feature order, fit L2 logistic regression (C=1, balanced class weights) on validation scores, select a decision threshold on validation micro-F1 over $\{0.05, \ldots, 0.95\}$, and apply once on test. They do not replace the locked main fusion rule. Patient cluster bootstrap (2,000 draws over 2,329 patients) on the test F1 difference yields 95% CI $[0.049, 0.068]$ around Δ+0.059; this interval is not interchangeable with the locked model F1 CI. Table 3 reports the matched with/without-q contrast and the image intervention; Table 7 expands the diagnostic prior/verifier breakdown.

Table 7: Full prior/verifier diagnostic breakdown under a fixed Top-15 candidate pool.

Evidence configuration	Val F1	Test Rec.	Test F1
Developmental prior (thr. on p)	0.471	0.507	0.456
Evidence verification (thr. on q)	0.381	0.359	0.373
Calibrated prior w/o q (nested)	0.461	0.458	0.445
Matched fusion w/ q (refit)	0.502	0.526	0.504
PedECG fusion (locked)	**0.503**	**0.527**	**0.504**

Note. Shared candidate pool; nested rows are VAL refit rules. Locked row is the main fusion rule.

The image ablation in table 3, Block B, reuses the locked Top-15 scored pool and FeatureDB prompts, substitutes only the verifier image, and applies the locked fusion rule once per condition without refitting. The blank condition feeds a constant white PNG. The mismatched condition uses a fixed circular shift over the sorted TEST record id list: record i receives the ECG image of record i+1, wrapping at the end, so the image is never the record's own image. This is a single deterministic permutation rather than an average over random shuffles.

B.4 Adult VLM Adaptation Path Ablations

Table 8: Transfer path ablations on the patient-disjoint test set.

Setting	Recall	Precision	Micro-F1
No adult ECG VLM	0.485	0.493	0.489
Adult ECG VLM + verifier	0.488	0.437	0.461
Adult ECG VLM free gen	0.209	0.238	0.223
PedECG	0.527	0.484	0.504

Note. Distinct adaptation paths on the same endpoint (n=2819), not a controlled data size ablation.

Adult ECG VLM adaptation alone does not improve the pediatric endpoint and can degrade performance under direct transfer (0.461 vs 0.489). The final gain of PedECG should therefore not be attributed to adult VLM initialization in isolation, but to the complete target-side adaptation and evidence integration procedure.

C Developmental Sensitivity Analyses

C.1 Developmental Context Substitution Probe Protocol

The main text substitution probe substitutes chronological age and the corresponding PedNorm reference context while holding the waveform x_s, raw FeatureDB measurements f, and frozen adult evidence (h_A, z_A) fixed. We evaluate n=20 TEST records for which precomputed age sweep JSON artifacts were available from a prior mechanism audit. The probe uses eight substituted ages $\{1, 2, 3, 5, 7, 10, 12, 14\}$ years and a raw candidate threshold of 0.5 on sigmoid probabilities rather than the locked Top-15/fusion rule. Across ages, the thresholded candidate set changes in $19/20$ records and the top ranked label changes in $12/20$. This collection is an exploratory held out probe rather than a random sample of the full test set; inclusion follows availability of precomputed sweep files and is not conditioned on the locked fusion endpoint. Aggregate flip rates therefore demonstrate model responsiveness under fixed ECG evidence, not a causal estimate of developmental benefit or a population estimate of endpoint gain. The sinus tachycardia age response curves in figure 2 further restrict visualization to the five records whose raw Top-1 label at the first sweep age is sinus tachycardia, an output conditioned selection for illustration.

C.2 Age Stratified Locked Performance

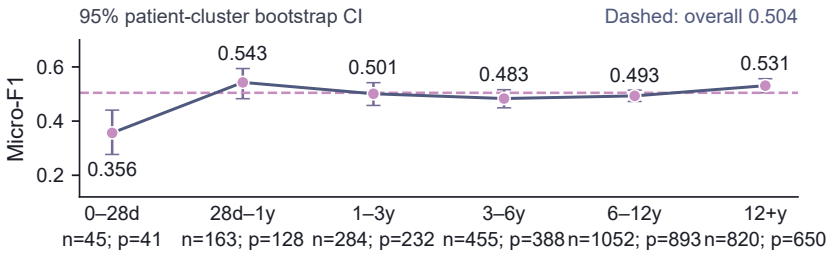


Figure 3: Diagnostic performance across developmental stages under one locked decision rule. Error bars indicate patient-cluster bootstrap 95% confidence intervals. n: records; p: patients. Age bins use lower inclusive, upper exclusive intervals; patients may contribute records to more than one bin. These are descriptive strata with different case and label mixtures, not a test that conditioning improves performance at a given age.

Figure 3 stratifies all 2819 test records. Non neonatal strata have micro-F1 of approximately 0.48–0.54, whereas the neonatal estimate is lower at 0.356 with n=45. Errors are not confined to one childhood interval, and neonatal transfer remains a particularly uncertain setting.

C.3 Age Response Sweep

Figure 2 reports the main text age response visualization. Panel (a) shows the age-matched PedNorm heart rate band supplied as model input, taken from a literature table rather than learned. Panel (b)

substitutes age and the corresponding PEDNORM context while holding the ECG and FeatureDB measurements fixed; light curves are the five records with top ranked sinus tachycardia at age 1, the bold curve is the mean, and shading shows the across-record range rather than a confidence interval. Scores are pre fusion candidate probabilities and do not establish physiological correctness.

D ADDITIONAL DIAGNOSTIC RESULTS

D.1 EXACT SET AND PARTIAL RECOVERY METRICS

Table 9 reports complementary record-level recovery metrics for the primary comparison systems. Exact-set match is deliberately stringent: a record is counted as correct only when the entire predicted label set exactly matches the ECG-observable reference set. PedECG reaches 0.177 exact-set match while achieving the highest at-least-one-hit rate of 0.700. Together with the primary micro-F1 results in Table 1, these metrics indicate that PedECG's main advantage lies in recovering and selecting relevant pediatric diagnoses rather than maximizing exact agreement of the complete multilabel set.

Table 9: Complementary record-level recovery metrics for the primary comparison systems on the patient-disjoint test set.

Method	Exact set	At least one hit
ECG-R1 (native)	0.057	–
ECG-R1, matched inputs	0.085	–
DeepSeek-v4-flash, matched inputs	0.051	–
DeepSeek-v4-pro, matched inputs	0.037	–
Zero-shot Qwen	0.057	–
Pediatric only ResNet1D (scratch)	0.175	0.634
Frozen adult h_A linear probe	0.144	0.639
Random init frozen ResNet1D probe	0.000	0.650
PedECG	0.177	**0.700**

Note. Adaptation-path ablations are reported separately in Appendix B.4 and are not repeated here. Exact-set match is a secondary, stringent record-level metric; micro-F1 remains the prespecified primary endpoint. At least one hit uses nonempty gold records (n=2708). Scratch/adult probe: 3-seed means. Random probe: seed 0. Dashes: unavailable.

D.2 COMMON LABEL PRECISION AND RECALL

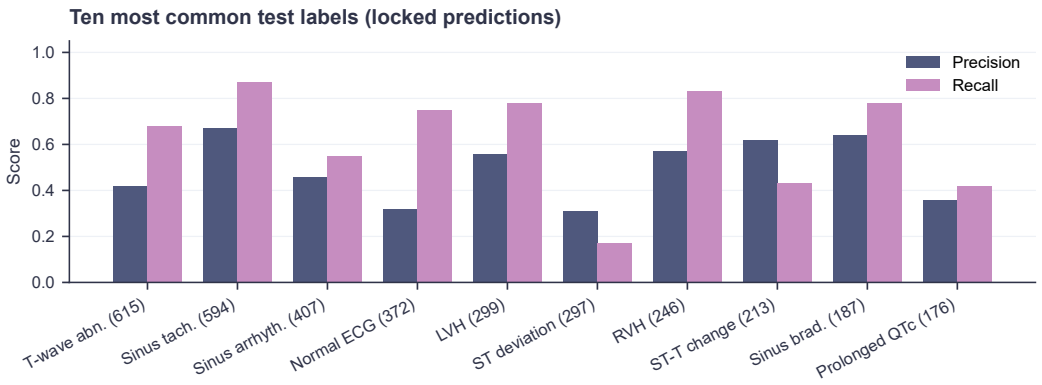


Figure 4: Precision and recall for the ten ECG-observable labels with the highest test support, under the locked predictions. Numbers in parentheses are positive counts.

D.3 Performance by Label Support

Table 10: Diagnostic performance stratified by label prevalence.

Test support	Labels	Positives	Precision	Recall	Micro-F1
At least 50	22	4489	0.488	0.568	0.525
20–49	6	208	0.495	0.226	0.310
10–19	11	132	0.320	0.121	0.176
5–9	14	98	0.214	0.061	0.095
1–4	25	51	0.125	0.039	0.060
Five most common	5	2287	0.475	0.728	0.575
All labels	78	4978	0.484	0.527	0.504

Note. Micro averaging restricted to each support bin using locked predictions from table 1.

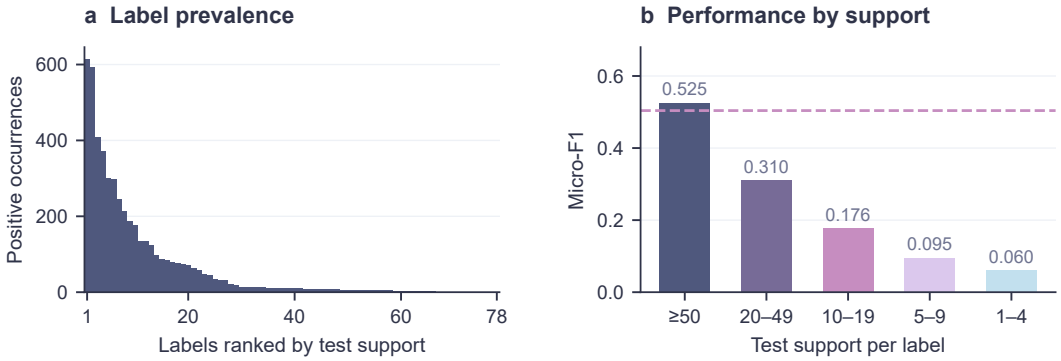


Figure 5: Label prevalence and diagnostic performance. **(a)** Test set support for the 78 ECG-observable labels, ranked from most to least frequent. **(b)** Micro-F1 of the locked predictions restricted to each support bin. The dashed line is overall micro-F1 (0.504).

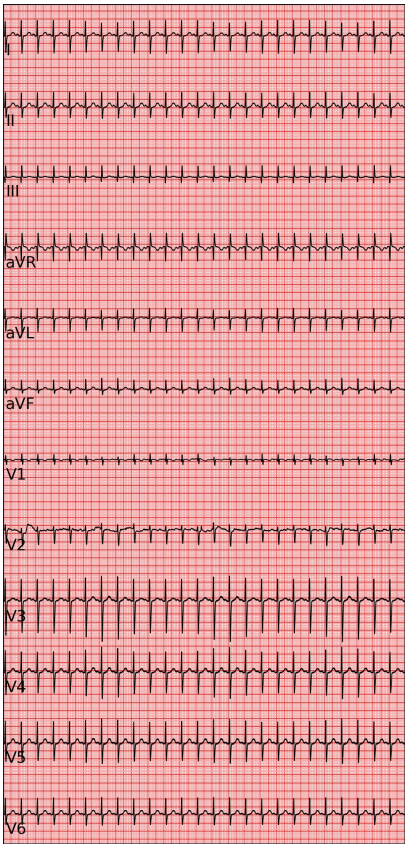
E Structured ECG Reporting Examples

PedECG additionally generates structured ECG reports from the same record-level evidence used by the diagnostic pathway. Training targets for the reporting adapter are constructed only on pediatric training records by remapping available source-report text into a fixed eight-section schema (rhythm/rate, intervals, axis/voltage, conduction/arrhythmia, ST–T, key evidence, image quality, conclusion); language models may assist drafting, consistent with the LLM-use statement. These targets supervise only the reporting adapter and never enter the candidate readout, verifier, fusion stage, validation selection for the diagnostic endpoint, or test labels.

We present three illustrative test examples below to show the organization and content of these reports. For each record we show the ECG and the structured outputs of PedECG, ECG-R1 (matched inputs), DeepSeek-v4-flash, and DeepSeek-v4-pro under the same shared-candidate matched inputs used in the main experiments. The examples demonstrate reporting capability and format rather than constitute a quantitative evaluation of report quality; we make no physician-verified report superiority claim. Display text may include light language cleanup for reference-range wording without adding clinical claims or changing the retained answer labels.

Cases A–C below show the ECG, then PedECG, ECG-R1 with matched inputs, DeepSeek-v4-flash, and DeepSeek-v4-pro under the same shared-candidate matched inputs used in the main experiments.

EXAMPLE A: P00244_E01, SINUS TACHYCARDIA



PedECG

[Rhythm and Rate]

Sinus rhythm, heart rate approximately 150 bpm.

[Intervals]

PR~50 ms, QRS~50–60 ms, QT~280–300 ms, QTc~380–400 ms (all visually estimated from lead II and V5). For a 3-year-old, the PR interval is below the age-specific reference range (107–151 ms), and the QRS duration is below the V5 reference range (64–108 ms), while the uncorrected QT is within the 3–6-year reference range.

[Axis and Voltage]

QRS axis approximately +60 to +80 degrees (normal for age). Limb and precordial voltages are within expected pediatric limits without clear evidence of chamber hypertrophy.

[Conduction and Arrhythmia]

Atrioventricular and intraventricular conduction appear intact, with no evidence of AV block, bundle branch block, or pre-excitation.

[ST-T and Ischemia/Infarction]

ST segments are isoelectric without significant elevation or depression in limb or precordial leads. T waves are upright in I, II, and V3–V6 and show normal age-appropriate morphology without ischemic patterns.

[Key Evidence]

Regular P waves preceding each narrow QRS complex in lead II with a rate above the age-specific reference support sinus tachycardia. Short PR timing and narrow QRS morphology are observed in lead II as accompanying interval findings; the retained diagnostic conclusion is sinus tachycardia. Normal R-wave progression and T-wave morphology in V3–V6 argue against ventricular hypertrophy.

Answer: Sinus tachycardia

ECG-R1 (matched inputs)

[Rhythm and Rate]

The ECG demonstrates a regular rhythm with a heart rate of 150 bpm, consistent with sinus tachycardia. The P waves are present and upright in leads II, III, and aVF, indicating a sinus origin. The PR interval is within normal limits, and the QRS complexes are narrow, ruling out ventricular tachycardia.

[Intervals]

The PR interval is 50 ms, which is within the normal range for this age group. The QRS duration is 54 ms, also within normal limits. The QT interval is 330 ms, which is unremarkable. The QTc interval is 521 ms, which is prolonged.

[Axis and Voltage]

The frontal plane axis is normal, with positive QRS complexes in leads I and II. The QRS voltage is within normal limits, with no evidence of left or right ventricular hypertrophy.

[Conduction and Arrhythmia]

There is no evidence of atrioventricular block, as the PR interval is consistent and within normal limits. The QRS complexes are narrow, ruling out bundle branch blocks. The rhythm is regular, with no premature beats or pauses.

[ST-T and Ischemia/Infarction]

There is no ST-segment elevation or depression, and the T waves are normal in morphology. There are no pathologic Q waves or other signs of ischemia or infarction.

[Key Evidence]

The ECG is consistent with sinus tachycardia, with no evidence of conduction abnormalities, chamber enlargement, or acute ischemic changes.

[Conclusion]

Answer: Sinus tachycardia

DeepSeek-v4-flash

[Rhythm and Rate]

The rate is 150 bpm, above the age reference, and the rhythm is consistent with sinus tachycardia. No premature beats or pauses are confirmed.

[Intervals]

PR interval is 50 ms, below the reference range, consistent with a short PR interval. QRS duration is 54 ms, below reference, and uncorrected QT is 330 ms, within reference. QTc by Bazett is 521 ms, but no pediatric QTc reference band is available, so this is not classified as prolonged.

[Axis and Voltage]

Axis and voltage data are not available in the numeric measurement block, and the tracing cannot be visually confirmed. Therefore, no definitive axis deviation or ventricular hypertrophy is diagnosed.

[Conduction and Arrhythmia]

No conduction block or arrhythmic pattern is confirmed. The short PR interval may reflect enhanced atrioventricular conduction, but preexcitation cannot be verified.

[ST-T and Ischemia/Infarction]

No definitive ST-segment or T-wave abnormality is established from the available data.

[Key Evidence]

Heart rate 150 bpm above reference; PR interval 50 ms below reference; QRS duration 54 ms below reference; uncorrected QT 330 ms within reference; QTc 521 ms without available reference band.

[Conclusion]

Sinus tachycardia and a short PR interval.

Answer: Sinus tachycardia; Short PR interval

DeepSeek-v4-pro

[Rhythm and Rate]

Sinus tachycardia at a heart rate of 150 bpm, above the 69–119 bpm reference for age 3–6 years.

[Intervals]

PR interval is 50 ms, below the age-specific reference of 107–151 ms; QRS duration is 54 ms, below reference of 64–108 ms; uncorrected QT is 330 ms, within reference of 288–374 ms. QTc is 521 ms by Bazett, but no age-specific PedNorm QTc reference band is available for classification.

[Axis and Voltage]

Axis and voltage criteria are not provided in the measurement block; no hypertrophy or axis deviation label can be reliably assigned from the available data.

[Conduction and Arrhythmia]

There is a short PR interval without evidence of high-grade AV block, preexcitation, or ectopy in the displayed rhythm. No conduction block or arrhythmia other than sinus tachycardia is identified.

[ST-T and Ischemia/Infarction]

No definitive ST-segment deviation, T-wave abnormality, or ischemia can be confirmed from the available tracing; T-wave polarity is not reliably assessable from the provided machine data.

[Key Evidence]

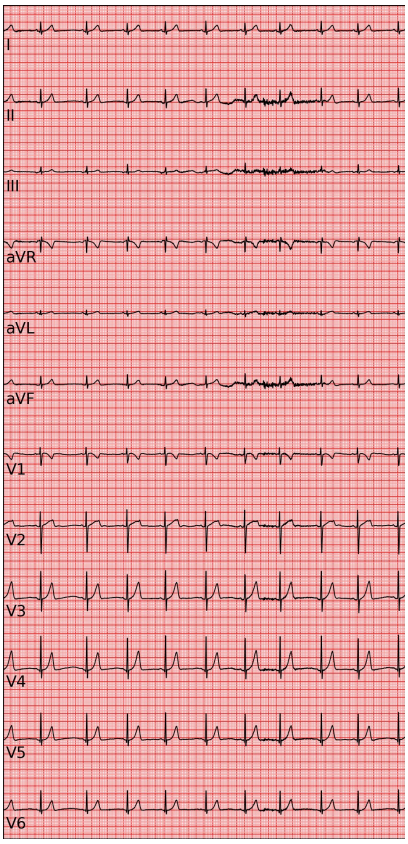
Heart rate is 150 bpm, above the 3–6 years reference; PR is 50 ms, below its reference; QRS is 54 ms, below its reference; uncorrected QT is 330 ms, within reference. QTc is 521 ms but has unavailable reference status and is not interpreted against an uncorrected QT limit.

[Conclusion]

Sinus tachycardia with a short PR interval; no other diagnostic label from the allowed list is supported by the available evidence.

Answer: Sinus tachycardia; Short PR interval

EXAMPLE B: P02886_E01, SINUS BRADYCARDIA / ARRHYTHMIA



PedECG

[Rhythm and Rate]

Sinus rhythm with sinus bradycardia and sinus arrhythmia, heart rate approximately 80 bpm.

[Intervals]

PR~100 ms, QRS~50–60 ms, QT~320–340 ms, QTc~340–360 ms.

[Axis and Voltage]

QRS axis approximately +60 to +70 degrees (normal for age). Limb and precordial lead voltages are within expected pediatric limits, with no clear pattern of ventricular hypertrophy.

[Conduction and Arrhythmia]

Atrioventricular and intraventricular conduction appear intact, with no evidence of block, pre-excitation, or bundle branch block.

[ST-T and Ischemia/Infarction]

ST segments are isoelectric without significant elevation or depression in limb or precordial leads. T waves are upright in I, II, and V3–V6 and show normal age-appropriate morphology, with no ischemic or infarction pattern.

[Key Evidence]

Regular P waves preceding each QRS in lead II with variable RR intervals support sinus bradycardia and sinus arrhythmia. Short PR intervals in some beats (e.g., ~100 ms) are consistent with sinus bradycardia rather than AV block. Normal QRS duration and QT/QTc in leads II and V5 argue against conduction delay or ventricular hypertrophy. Normal R-wave progression from V1 to V6 and absence of ST elevation/inversion argue against ventricular hypertrophy or ischemia.

Answer: Sinus arrhythmia; Sinus bradycardia

ECG-R1 (matched inputs)

[Rhythm and Rate]

The ECG demonstrates a sinus arrhythmia, characterized by a variable heart rate with a normal P-wave morphology and a normal PR interval. The heart rate is within the normal range, and there is no evidence of atrial fibrillation or flutter.

[Intervals]

The PR interval is within the normal range, and the QRS duration is normal. The QT interval is uncorrected, but the QTc interval is within the normal range.

[Axis and Voltage]

The ECG shows a normal cardiac axis, and the QRS voltage is within the normal range. There is no evidence of left or right ventricular hypertrophy.

[Conduction and Arrhythmia]

The ECG demonstrates a normal conduction pattern, with no evidence of bundle branch blocks or fascicular blocks. The rhythm is sinus arrhythmia, and there are no premature complexes or other arrhythmias present.

[ST-T and Ischemia/Infarction]

The ECG shows no ST-segment elevation or depression, and there are no T-wave inversions or abnormalities. The ECG is normal, with no evidence of ischemia or infarction.

[Key Evidence]

The ECG is normal, with no evidence of cardiac disease or abnormalities. The sinus arrhythmia is a normal variant, and the other findings are within the normal range.

[Conclusion]

Answer: Sinus arrhythmia

DeepSeek-v4-flash

[Rhythm and Rate]

Normal sinus rhythm at 83 bpm, within the age-based reference range. No significant arrhythmia is present.

[Intervals]

PR 114 ms and uncorrected QT 327 ms are within reference. QRS is narrow at 56 ms, below the V5 reference lower bound, without a corresponding conduction abnormality. QTc is 343 ms and is not compared against the uncorrected QT reference.

[Axis and Voltage]

Frontal axis and precordial voltages are within normal limits for age. No ventricular hypertrophy or axis deviation pattern is evident.

[Conduction and Arrhythmia]

No conduction block, preexcitation, or clinically significant ectopy is present. The narrow QRS argues against bundle-branch block or intraventricular conduction delay.

[ST-T and Ischemia/Infarction]

No significant ST-segment deviation, T-wave abnormality, or ischemic pattern is apparent.

[Key Evidence]

Heart rate, PR interval, and uncorrected QT interval are within age-adjusted reference ranges. QRS duration is narrow and below the reference lower bound but is not associated with any pathological conduction finding.

[Conclusion]

No clinically significant ECG abnormality is identified; the ECG is normal for age.

Answer: Normal ECG

DeepSeek-v4-pro

[Rhythm and Rate]

Heart rate is 83 bpm, within the 6–9 year reference range. The rhythm is sinus with phasic variation, consistent with sinus arrhythmia.

[Intervals]

DeepSeek-v4-pro (cont.)

PR interval is 114 ms, within reference. QRS duration is 56 ms in V5, below the PedNorm reference range but narrow. Uncorrected QT is 327 ms, within reference; QTc by Bazett is 343 ms, with no pediatric QTc reference band available for comparison.

[Axis and Voltage]

No definitive abnormal axis or ventricular hypertrophy/high-voltage criteria are identified from the available tracing.

[Conduction and Arrhythmia]

No AV block, bundle-branch block, ventricular preexcitation, ectopy, or tachyarrhythmia is seen.

[ST-T and Ischemia/Infarction]

No convincing ST-segment deviation, T-wave abnormality, or ischemia/infarction pattern is identified.

[Key Evidence]

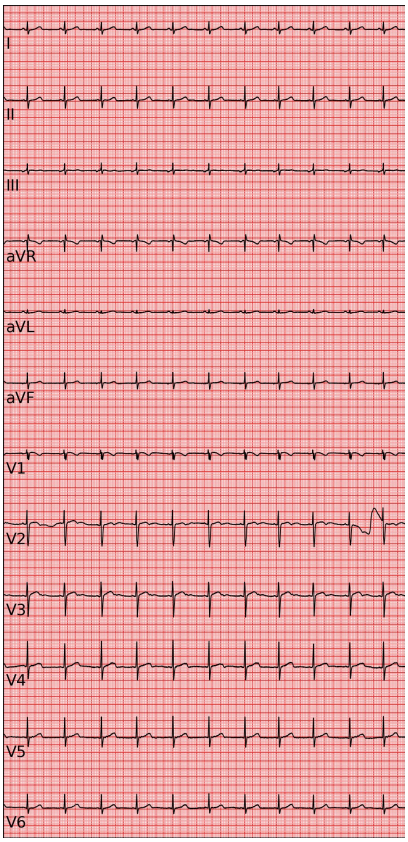
HR 83 bpm and PR 114 ms are within reference; QRS 56 ms is below the V5 PedNorm reference; uncorrected QT 327 ms is within reference. QTc 343 ms is not judged against an unavailable pediatric reference band. The highest-ranked rhythm candidate is sinus arrhythmia.

[Conclusion]

Sinus arrhythmia with otherwise normal findings for age.

Answer: Sinus arrhythmia

EXAMPLE C: P06798_E01, PROLONGED QTC INTERVAL



PedECG

[Rhythm and Rate]

Heart rate is approximately 67 bpm, within the normal range for a 9-year-old.

PedECG (cont.)

[Intervals]

PR is approximately 80 ms, shorter than the expected pediatric reference. QRS duration is approximately 70 ms, within normal limits. QT interval is approximately 460 ms, above the age-specific reference range; QTc is approximately 500 ms.

[Axis and Voltage]

QRS axis is approximately +60 degrees, normal for age. Limb and precordial voltages are within expected pediatric limits, with no clear evidence of chamber hypertrophy.

[Conduction and Arrhythmia]

Atrioventricular and intraventricular conduction appear intact, with no evidence of bundle branch block or pre-excitation.

[ST-T and Ischemia/Infarction]

ST segments are isoelectric without significant elevation or depression in limb or precordial leads. T waves are upright in inferior and lateral leads and show normal age-appropriate morphology, with no ischemic pattern.

[Key Evidence]

Regular sinus P waves precede each QRS with variable RR intervals in lead II, showing RR variability compatible with sinus arrhythmia. A short PR interval is observed in lead II with normal QRS duration, without evidence of AV block. The prolonged QT interval together with the elevated QTc supports the retained diagnostic conclusion of prolonged QTc. Absence of ST elevation or depression in leads II, aVF, and V4–V6 argues against acute ischemia. Lead II is the primary tracing for rate and interval assessment, with V5 providing additional confirmation.

Answer: Prolonged QTc interval

ECG-R1 (matched inputs)

[Rhythm and Rate]

The ECG demonstrates a regular rhythm with a heart rate of 67 bpm, which is within the normal range for a pediatric patient. The P waves are present and uniform, indicating a sinus origin. The rhythm is regular, ruling out sinus arrhythmia.

[Intervals]

The PR interval is 84 ms, which is within the normal range for a pediatric patient. The QRS duration is 76 ms, also within the normal range. The QT interval is 460 ms, which is borderline prolonged. The QTc interval is 498 ms, which is significantly prolonged.

[Axis and Voltage]

The cardiac axis is normal, as the QRS complexes are predominantly positive in lead I and negative in lead aVR. The QRS voltage is normal, with no evidence of left or right ventricular hypertrophy.

[Conduction and Arrhythmia]

There is no evidence of atrioventricular block, as the P waves are consistently followed by QRS complexes. The QRS complexes are narrow, ruling out ventricular tachycardia. The rhythm is regular, with no premature beats or pauses.

[ST-T and Ischemia/Infarction]

There is no ST segment elevation or depression, and the T waves are normal in morphology. There are no signs of ischemia or infarction.

[Key Evidence]

The key evidence is the prolonged QTc interval, which is a significant finding in a pediatric patient. The normal cardiac axis, QRS voltage, and ST-T segments rule out other significant abnormalities.

[Conclusion]

The ECG is consistent with a sinus rhythm in a pediatric patient with a prolonged QTc interval. This is a significant finding that requires further evaluation and monitoring.

Answer: Sinus rhythm; Prolonged QTc interval

DeepSeek-v4-flash

[Rhythm and Rate]

The median ventricular rate is 67 bpm, within the 6-9-year reference range. The underlying rhythm appears to be sinus with likely sinus arrhythmia; no atrial or ventricular ectopy is confirmed.

DeepSeek-v4-flash (cont.)

[Intervals]

The PR interval is 84 ms in lead II, below the reference range (100.0-161.5 ms), indicating short PR. The QRS duration is 76 ms, within normal limits. The uncorrected QT interval is 460 ms, above the age-specific reference (287.0-385.0 ms); QTc is 498 ms, but no valid pediatric QTc reference is available, so QTc prolongation is not separately labeled.

[Axis and Voltage]

No axis deviation or voltage-based chamber enlargement is established from the available measurement data.

[Conduction and Arrhythmia]

No bundle-branch block, AV block, or pre-excitation pattern can be confirmed. The shortened PR interval is the main conduction finding.

[ST-T and Ischemia/Infarction]

T-wave abnormality is suspected from the candidate evidence, but no definite ST-segment deviation or infarction pattern is diagnosed.

[Key Evidence]

HR 67 bpm: within reference. PR 84 ms: below reference. QRS 76 ms: within reference. Uncorrected QT 460 ms: above reference. QTc 498 ms: reference unavailable.

[Conclusion]

The ECG shows sinus arrhythmia, a short PR interval, and QT prolongation; a T-wave abnormality is possible.

Answer: Sinus arrhythmia; Short PR interval; Prolonged QT interval; T-wave abnormality

DeepSeek-v4-pro

[Rhythm and Rate]

The ventricular rate is 67 bpm, within the age-specific reference range for a 6–9-year-old male.

[Intervals]

PR interval is 84 ms in lead II, below the pediatric reference range. QRS duration is 76 ms in V5, within the reference range. Uncorrected QT interval is 460 ms, above the pediatric reference range; QTc is 498 ms but is not judged against a pediatric reference band here.

[Axis and Voltage]

Axis and voltage measurements are not provided in the feature block, so no definitive hypertrophy or axis abnormality can be established from the available numbers.

[Conduction and Arrhythmia]

Short PR interval is present; QRS duration is normal and no atrioventricular block or preexcitation pattern is established from the available data.

[ST-T and Ischemia/Infarction]

ST-T and ischemia assessment is limited by image quality; no definite ST-segment or T-wave abnormality is established from the available data.

[Key Evidence]

PR interval is below reference and uncorrected QT interval is above reference; heart rate and QRS duration remain within reference limits.

[Conclusion]

Short PR interval and prolonged uncorrected QT interval are supported by the provided pediatric reference comparisons; no other allowed labels are established.

Answer: Short PR interval; Prolonged QT interval