PHYSIOLOGY-GUIDED WORLD MODELS FOR INTERVENTION-AWARE CARDIAC SIMULATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Electrocardiogram (ECG)-based models have achieved strong diagnostic performance, yet remain limited in modeling how cardiac dynamics evolve under external interventions. Existing approaches primarily focus on static prediction, with limited capacity to capture ECG changes under different pharmacological conditions. In this work, we propose an ECG World Model for action-conditioned predictive simulation of cardiac electrophysiology. Our framework integrates physiological ordinary differential equation (ODE) priors into latent diffusion dynamics through energy regularization. This structural guidance encourages the generation of physiologically plausible post-intervention ECG trajectories while reducing generative inconsistencies. Building on these simulations, we introduce an uncertainty-aware evaluation strategy that leverages stochastic diffusion sampling to estimate both the expected risk score and its variability, supporting comparative assessment of candidate interventions. We evaluate our method across diverse settings, including controlled drug-response datasets and real-world clinical records. Beyond improvements in waveform prediction, experimental results demonstrate better preservation of clinically relevant ECG intervals and more accurate prediction of the magnitude and direction of drug-induced risk changes. These results highlight the potential of our approach for intervention-aware cardiac simulation and simulation-based assessment of candidate treatments.

1 INTRODUCTION

Recent advances in large-scale foundation models have demonstrated remarkable emergent capabilities across a wide range of domains (Wei et al., 2022; Li et al., 2023a; Gu & Dao, 2024; Shao et al., 2024). However, extending these capabilities to physiologically grounded simulation in high-stakes, dynamic medical domains remains challenging.

> ***The Clinical Imagination Gap:*** *While current AI excels at parsing the clinical past, reliable simulation of intervention-conditioned futures remains challenging. Viewing clinical treatment as a Partially Observable Markov Decision Process (POMDP) motivates extending medical AI from static pattern recognition to action-conditioned physiological simulation.*

This limitation is especially pronounced in high-stakes medical domains, where effective clinical analysis requires not merely pattern recognition, but a mechanistic understanding of how physiological processes respond to interventions. Electrocardiography (ECG), a fundamental tool in cardiology, provides a rich representation of cardiac electrical activity and is routinely used to support diagnosis and monitor disease progression. Despite recent progress in ECG foundation models (McKeen et al., 2025; Xu et al., 2026; Knight et al., 2026), which achieve strong performance on tasks such as arrhythmia classification and disease prediction, these approaches remain largely focused on discriminative or predictive objectives. They often provide limited support for characterizing how physiological states evolve under clinical interventions, such as medications or therapies. This limitation becomes particularly critical in interventional settings. In practice, clinicians must consider multiple candidate medications, requiring prospective *"what-if" reasoning* to anticipate how a patient's cardiac state might evolve under each option. Since real-world trial-and-error carries clinical risks and responses to alternative treatments cannot all be observed for the same patient under identical conditions, prospective simulation offers a route to more individualized analysis. Current ECG-based

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.

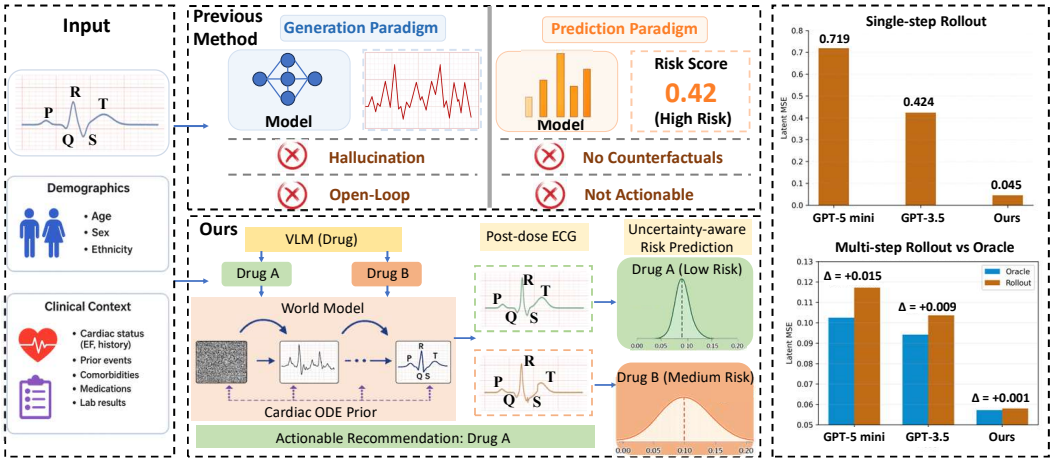

Figure 1: Overview and evaluation paradigm. Left: Existing ECG generation and prediction methods primarily operate in open-loop or static settings, whereas our framework models intervention-conditioned post-dose trajectories and evaluates simulated outcomes. Right: Single- and multi-step evaluations quantify predictive fidelity and rollout–oracle terminal-error differences under intervention-conditioned prediction.

AI systems provide limited support for such prospective reasoning due to their fragmented design. As illustrated in Figure 1 (left), generative models focus primarily on open-loop signal reconstruction, while predictive models estimate risk from static observations without explicitly modeling treatment actions. A more comprehensive framework would integrate (i) candidate pharmacological actions, (ii) simulation of action-conditioned cardiac dynamics, and (iii) evaluation of downstream ECG-derived risk signals. Without explicitly integrating these components, current AI systems remain limited in supporting simulation-driven analysis in cardiology.

To address the aforementioned challenges, we introduce an ECG World Model for action-conditioned predictive simulation of cardiac electrophysiology. As illustrated in Figure 1 (left), rather than simply pipelining disjoint networks, our framework integrates non-linear physiological ordinary differential equation (ODE) priors into latent diffusion models through EPK conditioning and physiology-informed energy regularization. Instead of directly predicting clinical outcomes from static signals, our approach explicitly models the action-conditioned evolution of cardiac states. Given an input ECG representing the current cardiac state, the system first proposes candidate pharmacological actions. Guided by these ODE priors, our diffusion-based dynamics model is trained to encourage physiologically plausible post-intervention ECG states. Subsequently, these simulated trajectories are evaluated using a frozen ECG foundation model to obtain diagnostic scores that serve as risk proxies. Crucially, to bridge the gap between generative simulation and downstream analysis, we introduce an uncertainty-aware evaluation strategy. Instead of relying on a single deterministic estimate, our framework leverages the stochasticity of the diffusion process to sample multiple potential future trajectories. This enables the characterization of both the mean diagnostic risk score and its variability across simulated trajectories for each candidate intervention. As a result, our approach extends ECG analysis from a purely diagnostic setting to a simulation-driven framework for comparative evaluation. Furthermore, the rollout–oracle comparison in Figure 1 (right) shows small additional average terminal errors over the evaluated short chains. Complementary experiments demonstrate improved waveform and interval fidelity and stronger agreement with observed post-dose diagnostic changes, supporting the potential of this framework for intervention-aware cardiac analysis. The main contributions of this paper are summarized as follows:

- **Physiology-Informed World Model** We propose an ECG World Model that integrates non-linear cardiac ODE priors into a latent diffusion process via energy regularization, encouraging physiological consistency in generated trajectories. Our model combines structured physiological priors with learned multi-lead waveform representations.
- **Intervention-Conditioned Simulation** We unify candidate treatment generation, post-intervention ECG simulation, and downstream risk-proxy evaluation into an integrated framework for comparative intervention analysis. This framework enables intervention-aware analysis by modeling the potential physiological responses to different medications.

- **Uncertainty-Aware Risk Evaluation** By leveraging the stochasticity of the diffusion process, our framework characterizes the mean and variability of diagnostic risk proxies for candidate interventions. These quantities support simulation-based candidate comparison, while diagnostic-change fidelity is assessed against observed post-dose ECGs.
- **Comprehensive Evaluation** Our method improves waveform prediction, ECG interval preservation, and diagnostic-change agreement over the evaluated baselines across datasets covering normal and abnormal cardiac states. Additional analyses assess short-chain recursive prediction and reconstruction under missing-lead conditions. The project page is available at `https://anon-researcher-202605.github.io/ECG_World_Model/`.

2 Related Work

World Models for Decision-Making In recent years, World Models (Ha & Schmidhuber, 2018; Matsuo et al., 2022; Zhou et al., 2025b; Chen & Zhu, 2026) have attracted increasing attention for their ability to simulate the dynamic evolution of environments. Recent advances incorporate explicit causal structures, physiological constraints, and high-dimensional control. For example, SPARTAN (Lei et al., 2025) enforces causal stability via time-dependent causal graphs, while PIN-WM (Li et al., 2025) embeds differentiable rigid-body dynamics to ensure physical consistency. In terms of controllability, Ctrl-world (Guo et al., 2025) improves action-conditioned generation, and 3D-VLA (Zhen et al., 2024) integrates 3D perception with embodied actions. Beyond open-loop simulation, the Dreamer family (Wang et al., 2025) enables closed-loop learning of latent dynamics and policy optimization through imagination. Such capabilities are particularly relevant in healthcare settings, where trial-and-error is costly and often irreversible, making the modeling of intervention effects an "ultimate epistemic stress test" for world models (Chen & Zhu, 2026). Early explorations in this domain have recently emerged: VCWorld (Wei et al., 2025) constructs a white-box biological simulator for gene-level interventions, while MeWM (Yang et al., 2025) applies world models to clinical imaging by generating treatment hypotheses, simulating tumor progression, and evaluating outcomes via survival analysis. However, despite these advances in imaging and genomics, the extension of world models to continuous physiological signals (e.g., ECG) for intervention-aware simulation and analysis remains largely unexplored. To address this gap, we propose an ECG World Model for modeling action-conditioned cardiac dynamics.

Static and Conditional ECG Generation Early ECG research focused mainly on signal synthesis and risk prediction (Chen et al., 2024; Lai et al., 2025; Adib et al., 2025), such as diffusion-based data augmentation (Adib et al., 2023), multi-lead reconstruction (Liu et al., 2024), and cardiovascular risk prediction (Zhou et al., 2025a; Prifti et al., 2021). More recent work introduces conditional and multimodal generation: BeatDiff (Bedin et al., 2024) formulates ECG reconstruction as a Bayesian inverse problem, and DiffuSETS (Lai et al., 2025) incorporates clinical text and metadata for personalized ECG synthesis. To model interventions, Cardiac Digital Twins (Camps et al., 2025) and DADM (Shao et al., 2025) leverage mechanistic and ODE-based priors. However, these remain passive, open-loop generators that cannot evaluate clinical outcomes or compare treatments. We bridge this gap with a closed-loop ECG World Model. Given a baseline state, our framework actively proposes candidate interventions, simulates their physiological trajectories, and evaluates their simulated outcomes to compare diagnostic-proxy scores under sampling variability.

3 Preliminaries

3.1 ECG-based Treatment Decision Problem

Electrocardiography (ECG) signals provide a non-invasive measurement of cardiac electrophysiological activity and play an important role in analyzing responses to pharmacological interventions. Given an ECG observation, the goal is to characterize how candidate medications may influence patient-specific cardiac dynamics under different intervention scenarios. Formally, an ECG recording is represented as $x \in \mathbb{R}^{C \times T_{\text{sig}}}$, where C denotes the number of channels and T_{sig} denotes the temporal length of the signal. An encoder $f_{\text{enc}}(\cdot)$ maps the ECG signal to a representation in the latent space: $z = f_{\text{enc}}(x)$, where $z \in \mathbb{R}^d$ captures the underlying physiological state of the patient. Let $\mathcal{A}$ denote the set of candidate pharmacological interventions, where each action $a \in \mathcal{A}$ corresponds to a single medication or a multi-step therapy. Deterministic approaches typically evaluate interventions using a static scoring function $R(z, a)$, where $R(\cdot)$ denotes a clinical risk estimate associated with a simulated outcome. Rather than directly optimizing for a single intervention, our focus is on

modeling and comparing the potential outcomes induced by different candidate interventions. As we demonstrate, deterministic point-estimates of interventions are fragile in safety-critical clinical settings, motivating an uncertainty-aware formulation that captures both the expected risk and the variance across simulated trajectories.

3.2 Denoising Diffusion Probabilistic Models and Energy Regularization

Denoising Diffusion Probabilistic Models (DDPMs) (Ho et al., 2020) offer a powerful paradigm for generative modeling by learning to invert a predefined Markovian noise-addition process. Let z_0 be the initial clean latent representation. The forward process gradually adds Gaussian noise over N steps, producing a sequence of increasingly noisy states $z_1, \ldots, z_N$. The generative reverse process then learns the transition probability $p_\theta(z_{\tau-1}|z_\tau, a)$ to iteratively denoise a Gaussian prior $z_N \sim \mathcal{N}(0, I)$ back to the target distribution, conditioned on an external variable such as an action a, where $\tau \in \{1, \ldots, N\}$ denotes the internal diffusion timestep.

To incorporate structural constraints or external physical knowledge into DDPMs without retraining the core noise-prediction network, *energy regularization* is a mathematically principled technique. Given a condition (e.g., an action a) and an energy function $\mathcal{E}(z_\tau, a)$ that measures the deviation of the intermediate state from the desired constraint, the generative reverse process can be steered using the gradient of this energy. The modified score function is formulated as:

$$\nabla_{z_\tau} \log \tilde{p}(z_\tau|a) = \nabla_{z_\tau} \log p_\theta(z_\tau|a) - \gamma \nabla_{z_\tau} \mathcal{E}(z_\tau, a), \tag{1}$$

where $\gamma > 0$ is a temperature coefficient, which controls the guidance scale. This formulation effectively tilts the original data distribution toward regions that minimize the energy function, allowing for flexible constraint-satisfaction during the step-by-step stochastic generation. Energy-based score modification provides a useful probabilistic interpretation of how physiological constraints can bias diffusion trajectories. In our implementation, however, we do not apply this gradient correction explicitly during inference. Instead, the physiological energy is incorporated as a training-time regularizer on the predicted clean latent state. Consequently, the deployed model uses standard DDPM sampling, with physiological information amortized into the learned denoising dynamics.

3.3 Mechanistic Modeling of Cardiac Electrophysiology

Although deep learning captures statistical patterns, classical electrophysiology provides mechanistic priors. A foundational continuous-time model for generating synthetic ECGs is the McSharry dynamical system (McSharry et al., 2003). It characterizes the intrinsic dynamical properties of the ECG signals using an Ordinary Differential Equation (ODE). The continuous-time evolution of the theoretical ECG signal amplitude $y(\theta)$ along the cardiac phase θ is formulated as:

$$\frac{dy(\theta)}{d\theta} = -\sum_{i \in \{P,Q,R,S,T\}} \alpha_i \Delta\theta_i \exp\left(-\frac{\Delta\theta_i^2}{2b_i^2}\right) - (y(\theta) - y_0), \tag{2}$$

where $\Delta\theta_i = (\theta - \theta_i) \bmod 2\pi$ defines the phase shift corresponding to specific anatomical landmarks (P, Q, R, S, T waves), α_i and b_i control the amplitude and width of these waves, and y_0 represents the baseline wander. Let $\theta \in [0, 2\pi)$ denote the continuous cardiac phase. This phase-space ODE provides a robust structural template for both healthy rhythms and pathological variations. It is important to note that $y(\theta)$ denotes a 1-dimensional theoretical scalar amplitude (such as Lead II). We explicitly use y to distinguish this abstract, single-channel physical template from the raw multi-lead observation matrix $x \in \mathbb{R}^{C \times T_{\text{sig}}}$. Rather than explicitly formulating the complex 3D spatial projections of all 12 leads within the ODE, we leverage $y(\theta)$ purely as a *global temporal rhythm template*. This 1D template captures fundamental mechanistic dynamics, which are subsequently mapped into the latent space by our networks to guide the high-dimensional multi-lead generation.

3.4 Risk-Sensitive Evaluation under Uncertainty

In standard predictive tasks, actions are often evaluated based on a single point-estimate of the expected outcome. However, in safety-critical domains, such as medical interventions, the inherent stochasticity of patient responses requires an uncertainty-aware evaluation. Following standard risk-sensitive decision theory, let $R(z, a)$ be a risk evaluation function. For a stochastic predictive

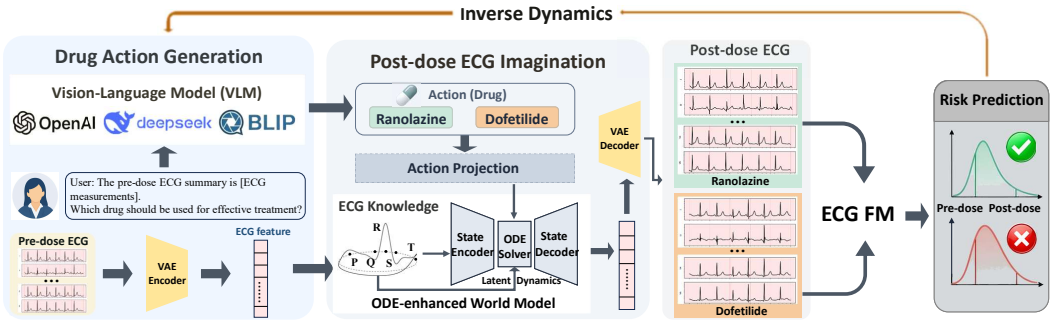

Figure 2: Overview of our intervention evaluation framework. Candidate drugs, optionally proposed by a VLM, are evaluated through ODE-enhanced post-dose ECG simulation and downstream diagnostic-proxy scoring. The reported experiments use fixed candidate sets.

model that produces a distribution of potential future trajectories, K independent samples can be drawn to estimate the empirical mean $\mu_R(a)$ and the standard deviation $\sigma_R(a)$ (derived from variance $\sigma_R^2(a)$) of the clinical risk. To balance the expected risk against its variability, the evaluation is typically formulated using a mean-variance objective:

$$S(a) = \mu_R(a) + \lambda \cdot \sigma_R(a), \tag{3}$$

where $\lambda > 0$ is a risk-aversion hyperparameter. This objective penalizes interventions that have highly variable or unpredictable outcomes, prioritizing those with low expected risk.

4 Methodology

4.1 The Overall Framework

To enable action-aware cardiac analysis, we propose the **ECG World Model**, an end-to-end framework designed to simulate the evolution of cardiac electrophysiology under pharmacological interventions. As illustrated in Figure 2, our framework extends ECG analysis from purely diagnostic settings to a predictive simulation-based paradigm. Specifically, given a pre-dose ECG signal x_k at clinical step k, patient metadata d, and a candidate pharmacological intervention a, dedicated encoders extract the baseline cardiac state s_k, the clinical condition embedding c_d, and the intervention embedding c_a. Building upon these representations, the core of our world model leverages a hybrid dynamics architecture to simulate the *action-conditioned post-dose ECG* x_{k+1}. To mitigate physiological inconsistencies, an *External Physiological Knowledge (EPK)* module first translates the intervention a into a physiology-guided prior e_{epk} by solving the underlying cardiac ODE. This prior is injected into a Denoising Diffusion Probabilistic Model (DDPM), which iteratively transforms a Gaussian noise distribution into a constraint-aware post-dose ECG sample. Once simulated, the post-intervention physiological trajectory x_{k+1} is evaluated by a downstream *ECG Foundation Model*. By performing multiple stochastic simulations across candidate interventions, the framework enables an uncertainty-aware characterization of both expected risk and its variability, supporting comparative evaluation of different intervention scenarios.

4.2 Physiology-Informed ECG World Model

Directly modeling high-dimensional ECG signals is challenging due to complex temporal dependencies. We encode raw ECG signals into a low-dimensional latent space using a pre-trained variational autoencoder (VAE). Given an ECG signal x, the encoder maps it to a latent variable $z = f_{\text{enc}}(x) \in \mathbb{R}^d$. The decoder reconstructs the signal as $\hat{x} = f_{\text{dec}}(z)$. The VAE is trained by minimizing the standard reconstruction loss and the KL divergence. This compact state space enables efficient modeling of cardiac dynamics under pharmacological interventions. To incorporate physiological constraints into latent cardiac dynamics and mitigate generative inconsistencies, we propose an energy-regularized diffusion world model. Let k denote the clinical intervention step. An idealized energy-tilted reference distribution is

$$p^\star(z_{k+1} \mid z_k, a) \propto p_{\text{diff}}(z_{k+1} \mid z_k, a) \exp\big(-\gamma E_\phi(z_{k+1})\big), \tag{4}$$

where p_{diff} is the reference diffusion distribution, $\gamma > 0$ controls the theoretical tilt, and E_ϕ penalizes deviations from a physiological anchor. This formulation motivates regularized training rather than exactly characterizing the deployed sampler (Appendix D).

Latent EPK Energy Given the 1D continuous physiological trajectory derived from the ODE-based cardiac model, we discretize it into a sequence e_{EPK}. To align this temporal rhythm with the latent space, we apply a projection module $\Pi_\phi(\cdot)$, implemented via temporal resampling and a learnable 1×1 convolution, to obtain $z_{\text{EPK}} = \Pi_\phi(e_{\text{EPK}})$. For a fixed clinical transition, let z_0 denote the clean target latent and $\tau \in \{1, \dots, N\}$ the internal diffusion timestep. We define the energy on the predicted clean latent $\hat{z}_0 = \frac{z_\tau - \sqrt{1-\bar{\alpha}_\tau}\epsilon_\theta}{\sqrt{\bar{\alpha}_\tau}}$: $E_\phi(\hat{z}_0) = \|\hat{z}_0 - \Pi_\phi(e_{\text{EPK}})\|_2^2$. This formulation regularizes the denoised latent rather than intermediate noisy states, incorporating physiological structure without waveform decoding to compute the energy term.

Energy-Regularized Training Objective Instead of optimizing a fully normalized energy-based model, we regularize the standard diffusion objective with our latent EPK energy: $\mathcal{L} = \mathbb{E}_{z_0,\epsilon,\tau}\left[\|\epsilon_\theta(z_\tau, \tau) - \epsilon\|_2^2 + \lambda_{\text{EPK}} E_\phi(\hat{z}_0)\right]$. Conditioning arguments are omitted for brevity. Here, λ_{EPK} is a fixed coefficient, with no additional timestep-dependent weighting. The main pharmacological-mapping experiments use $\lambda_{\text{EPK}} = 0.1$. Default inference retains EPK conditioning and uses standard DDPM sampling without explicit energy-gradient correction.

4.3 Uncertainty-Aware Risk Evaluation and Treatment Assessment

Given the current state z_k, we evaluate a finite set of candidate interventions $\mathcal{A} = \{a_1, \dots, a_M\}$, where each action represents a discrete drug or regimen. The framework supports VLM-based candidate proposal from a specified intervention vocabulary. In the reported experiments, we directly use fixed candidate sets: eight predefined regimens for the interval and EPK experiments, and the five most frequent medications in the MIMIC-IV test split for the risk experiments. The candidate sets and selection procedure are detailed in Appendix E.7.

To characterize variability across simulated outcomes, we introduce an *Uncertainty-Aware Evaluation Module* that uses the stochastic generative nature of our model. For each candidate intervention $a \in \mathcal{A}$, conditioned on the current state z_k, we draw S independent, fully denoised next-state samples:

$$\hat{z}_{k+1}^{(s)} \sim p_{\text{WM}}(\cdot \mid z_k, a), \qquad s = 1, \dots, S, \tag{5}$$

where p_{WM} denotes the trained conditional world model sampled using standard DDPM with EPK conditioning. A frozen ECG-FM diagnostic model f_{risk} evaluates the decoded waveforms. Let $R(x)$ denote the scalar diagnostic risk proxy obtained by aggregating its output probabilities. Each generated sample is scored as $R^{(s)}(a) = R\Big(f_{\text{dec}}(\hat{z}_{k+1}^{(s)})\Big)$. The diagnostic labels and aggregation procedure are detailed in Appendix B.10. We estimate the intervention-specific mean score and variance across generated samples as

$$\mu_R(a) = \frac{1}{S}\sum_{s=1}^{S} R^{(s)}(a), \qquad \sigma_R^2(a) = \frac{1}{S-1}\sum_{s=1}^{S}\Big(R^{(s)}(a) - \mu_R(a)\Big)^2. \tag{6}$$

These quantities summarize the expected diagnostic-proxy score under the model and its variability across stochastic simulations. We apply the risk-aware scoring function $S(a) = \mu_R(a) + \lambda\sigma_R(a)$ introduced in Equation 3. The coefficient λ affects candidate ranking but does not enter the diagnostic-change calculation for a fixed recorded medication. The resulting scores support comparative assessment within the specified candidate set, rather than calibrated estimates of clinical risk or treatment efficacy.

4.4 Theoretical Foundations and Mechanistic Interpretation

We interpret physiological regularization through an idealized energy-tilted distribution balancing a learned reference distribution with an ODE-derived anchor. This perspective motivates the practical training objective while distinguishing theoretical properties from empirical behavior. Detailed derivations are provided in Appendix D.

Principled Optimality For a fixed reference density $p_\theta(z \mid a)$ and energy $E_\phi(z)$, the distribution $\tilde{p}(z \mid a) \propto p_\theta(z \mid a) \exp(-\gamma E_\phi(z))$ minimizes

$$\tilde{p}(\cdot \mid a) = \arg\min_q \{\mathrm{KL}(q \,\|\, p_\theta(\cdot \mid a)) + \gamma \mathbb{E}_q[E_\phi(z)]\}, \tag{7}$$

assuming $\gamma > 0$ and a finite, nonzero normalizing constant. The KL term favors agreement with the reference distribution, while the energy term penalizes deviations from the physiological anchor. This establishes a distribution-level trade-off between data fidelity and physiological preference, providing a rationale for regularization rather than an exact equivalence to the implemented training loss.

Score-Based Anchoring and Amortized Guidance At the clean-latent level, the idealized score admits an exact decomposition:

$$\nabla_z \log \tilde{p}(z \mid a) = \underbrace{\nabla_z \log p_\theta(z \mid a)}_{\text{Data-Driven Pull}} - \gamma \underbrace{\nabla_z E_\phi(z)}_{\text{Prior-Driven Pull}}. \tag{8}$$

This identity explains how physiological energy modifies the reference score, but does not directly specify intermediate noisy-state scores. Our implementation incorporates physiological information through EPK conditioning and the training penalty $\lambda_{\text{EPK}} E_\phi(\hat{z}_0)$, which encourages predicted clean latents to remain near the projected ODE anchor. The energy term is evaluated directly in latent space without waveform decoding. Default inference uses standard DDPM sampling without explicit energy-gradient correction, with physiological guidance supplied through conditioning and the regularized model parameters.

Under suitable regularity and boundary conditions, the Fokker–Planck analysis establishes $\tilde{p}$ as stationary for the specified continuous-time Langevin dynamics. This result does not establish exact sampling by the deployed DDPM or automatic fallback under prior mismatch. Robustness is assessed empirically through generated-sample dispersion under prior perturbations and short-chain terminal errors (Appendices B.9 and E.5), rather than inferred from a theoretical safety guarantee.

5 Experiments

5.1 Experiment Setup

Datasets We evaluate our framework using a combination of real-world clinical electronic health records (EHRs) and controlled ECG datasets. To strictly prevent data leakage, each dataset is partitioned at the subject level based on patient IDs (about 8:2 train-test split) (more details in Appendix A): (1) **MIMIC-IV-ECG & Clinical** – By integrating MIMIC-IV-ECG (Gow et al., 2023) with MIMIC-IV-Clinical (Johnson et al., 2023), we align 12-lead ECG recordings with longitudinal patient EHRs. This enables multi-modal conditioning on demographics, lab results, and existing comorbidities. The dataset comprises 17,493 unique patients and 108,980 patient–drug episodes, with 3,090 distinct medications; (2) **Drug-Response ECG Datasets** – To evaluate drug-induced cardiac morphological effects, we utilize the ECGRDVQ (Johannesen et al., 2014) and ECGDMMLD (Johannesen et al., 2016) databases. These provide 7,763 valid high-fidelity ECGs from 43 subjects across eight distinct drug regimens; and (3) **In-house OOD Dataset** – To assess cross-dataset generalization, we test our model, trained primarily on the controlled drug-response cohorts, against an in-house dataset of MIMIC-IV ICU patients administered Verapamil.

Training Details Our world model is built upon a Denoising Diffusion Probabilistic Model (DDPM) backbone (Ho et al., 2020), augmented with an action-conditioned module to model intervention effects. During training, we incorporate an ECG-specific ODE prior as a physiological constraint to guide the diffusion process and capture the underlying cardiac dynamics. We train the model using AdamW with a learning rate of 5×10^{-4} and a batch size of 64 for 200 epochs. A cosine annealing scheduler is applied to adjust the learning rate during training. The main pharmacological-mapping experiments use an energy coefficient $\lambda_{\text{EPK}} = 0.1$. The energy-weight sweep in Appendix B.3 retains its separately reported coefficients, including 0.25. Diagnostic-score evaluation uses $S = 5$ stochastic samples per candidate intervention. Default inference uses standard DDPM sampling, denoted by $K_{\text{res}} = 1$. The $K_{\text{res}} = 8$ configuration is evaluated in the separate 48-pair pilot in Appendix E.6. The action-selection coefficient λ is reported for each selection evaluation.

Architecture Implementation We adopt a modular design to implement the proposed framework, where each component is instantiated using either pre-trained foundation models or task-specific adaptations: (1) Treatment Action Interface – The framework supports VLM-based candidate proposal

(GPT-4o (Hurst et al., 2024), DeepSeek-VL (Lu et al., 2024), BLIP-FlanT5 (Li et al., 2023b)), while the reported experiments directly evaluate the fixed intervention sets specified in Appendix E.7; (2) Physiology-Informed World Model – We model cardiac dynamics using our proposed latent diffusion world model. To ensure physiological consistency and mitigate hallucinations, External Physiological Knowledge (EPK) from cardiac ODEs is injected as both conditioning signals and an energy-based regularization term during the denoising process; (3) Variational Autoencoder (VAE) – To obtain compact representations of the 12-lead ECGs, we use a VAE pre-trained (Lai et al., 2025) on MIMIC-IV-ECG. We keep this module completely frozen to maintain a consistent latent space across all experiments; and (4) Downstream Risk Prediction Model – For risk evaluation, we use a state-of-the-art ECG foundation model (McKeen et al., 2025), and this model is kept strictly frozen to assess simulated post-dose ECGs in a purely zero-shot manner.

Baselines To rigorously evaluate our framework, we conduct comprehensive comparisons across two major dimensions: physiological simulation fidelity and downstream clinical evaluation.

(1) ECG Generative and Dynamics Models To evaluate *waveform quality and physiological consistency* during simulation, we include a wide range of open-loop generative models for ECG signals, covering both GAN-based and diffusion-based approaches. Specifically, we consider classical GAN variants (*WGAN* (Arjovsky et al., 2017), *StyleGAN* (Karras et al., 2019), *ECG ODE-GAN* (Golany et al., 2021), *TTS-CGAN* (Li et al., 2022), and *CECG-GAN* (Yang et al., 2024)), as well as recent diffusion-based models (*Diffusets* (Lai et al., 2025), *SE-Diff* (Wang et al., 2026) and *DADM* (Shao et al., 2025)). We utilize these approaches as purely data-driven baselines to highlight the effectiveness of our proposed method in mitigating generative inconsistencies.

(2) Intervention-Aware Simulation and Comparative Evaluationiv Moving beyond unconditional generation, we evaluate action-conditioned dynamics across *single-step prediction*, *multi-step rollouts*, and *out-of-distribution (OOD) scenarios*. For the downstream *Intervention Ranking Consistency* task, we compare our simulation-driven approach against representative large-scale clinical foundation models and multi-modal LLMs, including *Qwen2.5-VL-7B* (Bai et al., 2025), *GPT-series* (Achiam et al., 2023), *GLM-series* (Glm et al., 2024), and *MedGemma* (Sellergren et al., 2025). These models serve as strong zero-shot baselines that attempt to infer intervention preferences directly from clinical context, rather than through explicit physiological simulation.

5.2 Experimental Results

Table 1: Comparison of methods on latent and signal Post-dose ECG prediction.

Methods	Latent ECG		Signal ECG	
	MSE ↓	MAE ↓	MSE ↓	MAE ↓
Qwen2.5-VL-7B	2.148	0.368	0.195	0.201
GPT-5 mini	0.719	0.166	0.066	0.128
GPT-4o	0.550	0.414	0.578	0.386
GPT-3.5	0.424	0.285	0.265	0.232
GLM-4.5	0.095	0.164	0.062	0.125
MedGemma	5.655	0.899	0.561	0.349
Ours	**0.045**	**0.159**	**0.053**	**0.110**

Evaluation of Emergent Capabilities on Clinical Dynamics We investigate whether the emergent capabilities of general-purpose foundation models extend to continuous cardiac dynamics. Since this cohort primarily consists of abnormal cases, the ECG signals often lack the regular patterns observed in normal populations. Under this setting, we evaluate model performance using mean squared error (MSE) and mean absolute error (MAE). As shown in Table 1, despite their strong reasoning performance in language and vision tasks, these zero-shot models exhibit notable limitations in modeling fine-grained ECG dynamics. This results in substantially higher prediction errors in both the latent physiological space and the decoded signal reconstruction on the MIMIC-IV-ECG benchmark.

Table 2: Performance comparison of different methods on three ECG intervals.

Methods	QT_c Interval		PR Interval		$T_{\text{peak}} - T_{\text{end}}$ Interval	
	Acc (%) ↑	Rec (%) ↑	Acc (%) ↑	Rec (%) ↑	Acc (%) ↑	Rec (%) ↑
WGAN	$75.23_{\pm 0.69}$	$77.23_{\pm 0.71}$	$57.67_{\pm 0.64}$	$58.04_{\pm 0.54}$	$85.49_{\pm 0.12}$	$87.35_{\pm 0.16}$
StyleGAN	$72.71_{\pm 1.06}$	$74.29_{\pm 1.49}$	$52.15_{\pm 3.54}$	$55.35_{\pm 1.20}$	$88.71_{\pm 1.96}$	$90.35_{\pm 1.87}$
ECG ODE-GAN	$72.81_{\pm 3.24}$	$73.95_{\pm 2.25}$	$54.63_{\pm 1.49}$	$56.16_{\pm 1.39}$	$87.03_{\pm 0.52}$	$88.19_{\pm 0.79}$
TTS-CGAN	$86.43_{\pm 1.65}$	$87.21_{\pm 0.64}$	$70.09_{\pm 2.08}$	$72.82_{\pm 2.34}$	$83.43_{\pm 1.85}$	$85.48_{\pm 1.57}$
CECG-GAN	$79.92_{\pm 1.27}$	$82.75_{\pm 1.31}$	$61.08_{\pm 1.43}$	$60.62_{\pm 0.71}$	$89.29_{\pm 0.13}$	$91.26_{\pm 0.05}$
Diffusets	$86.81_{\pm 0.98}$	$90.07_{\pm 1.24}$	$72.62_{\pm 1.34}$	$74.20_{\pm 1.54}$	$89.84_{\pm 0.98}$	$92.10_{\pm 0.96}$
SE-Diff	$89.19_{\pm 0.58}$	$89.19_{\pm 0.58}$	$70.83_{\pm 0.42}$	$73.58_{\pm 0.66}$	$92.35_{\pm 0.78}$	$94.22_{\pm 0.81}$
DADM	$89.46_{\pm 1.36}$	$93.69_{\pm 1.49}$	$68.02_{\pm 1.14}$	$72.26_{\pm 1.72}$	$91.63_{\pm 0.76}$	$93.04_{\pm 0.58}$
Ours	$\mathbf{90.55}_{\pm 0.76}$	$\mathbf{96.26}_{\pm 0.82}$	$\mathbf{76.04}_{\pm 0.97}$	$\mathbf{80.84}_{\pm 1.24}$	$\mathbf{94.70}_{\pm 1.13}$	$\mathbf{95.06}_{\pm 0.49}$

Physiological Simulation Fidelity For normal cohorts, we adopt three evaluation metrics to provide a more comprehensive assessment. As shown in Table 2, conventional metrics such as MSE are insufficient to capture clinically relevant characteristics of physiological signals. Instead, we evaluate models based on their ability to preserve key cardiac biomarkers, including QT_c, PR, and $T_{\text{peak}} - T_{\text{end}}$ intervals. A prediction is considered correct only if all indicator labels match the ground truth. Following (Shao et al., 2025), QT_c is computed using the Bazett's formula $QT_c = QT/\sqrt{RR}$, with normal ranges defined as ≤ 450 ms for men and ≤ 470 ms for women. The PR interval is considered normal within 120–200 ms, while $T_{\text{peak}} - T_{\text{end}}$ is normal within 80–113 ms. Under this clinically motivated criterion, our method more accurately preserves physiologically consistent signal characteristics compared to all baselines, particularly under drug-induced variations.

Table 3: Out-of-distribution (OOD) comparison across different methods.

Models	Latent	Signal
Diffusets	0.096	0.150
DADM	0.080	0.119
Ours	**0.068**	**0.084**

OOD Generalization on Unseen Diseases As shown in Table 3, we evaluate out-of-distribution (OOD) generalization against strong diffusion-based baselines (DiffuSETS and DADM). To rigorously test this, all models are trained exclusively on the controlled *Drug-Response ECG Datasets* and evaluated on the *In-house OOD Dataset*, which contains patients with unseen comorbidities and complex underlying pathologies. This setting reflects real-world clinical deployment, where generative models must safely generalize to previously unseen pathological patterns without hallucinating physically impossible cardiac dynamics.

Table 4: Comparison of Core Evaluation Metrics between Models

Core Evaluation Metric	Ours	DADM
ΔRisk Pearson Correlation	**0.620**	0.266
ΔRisk Spearman Correlation	**0.598**	0.255
ΔRisk Sign Agreement	**76%**	58%
MAE	**0.0973**	0.1291
RMSE	**0.1253**	0.1625

Drug-Induced Risk Modeling and Directional Consistency We evaluate diagnostic-change fidelity on 200 subject-disjoint MIMIC-IV test cases covering Propofol, Regular Insulin, Prophylactic Heparin Sodium, Furosemide, and Norepinephrine. Diagnostic scores average 17 label probabilities from a frozen ECG-FM. For each recorded medication, predicted post-dose scores are averaged over $S = 5$ standard-DDPM samples; predicted and observed changes use the same pre-dose reference. Ours improves agreement and reduces errors over DADM (Table 4), achieving Pearson $r = 0.620$, Spearman $\rho = 0.598$, and 76% sign agreement (rounded) in the reported run. Score definitions and all six sampling runs are provided in Appendix B.10 and Table 12, respectively.

Additional Analyses in Appendix We provide additional analyses of (i) EPK-module ablation and short-chain rollout–oracle comparisons (Figures 3 and 4, and Table 5); (ii) diagnostic-change agreement across repeated stochastic sampling runs (Figure 5) and the number of stochastic samples K (Appendix B.4); (iii) the energy coefficient λ_{EPK} with EPK conditioning retained (Table 6); (iv) reconstruction under missing-lead settings in normal (Table 8 and Figure 7) and abnormal conditions (Figure 8); and (v) generated-sample variability and latent dispersion under prior perturbations (Figure 9), where EPK reduces sample dispersion across the evaluated perturbation regimes.

6 Conclusion

In this work, we propose an ECG World Model for intervention-aware cardiac analysis, enabling simulation-driven modeling of drug-induced dynamics. By integrating intervention generation, physiology-informed latent dynamics, and risk evaluation, our framework extends ECG analysis from static prediction toward intervention-conditioned simulation. Experiments demonstrate strong performance in modeling drug responses and generalization across multiple clinical datasets. Several limitations remain. First, although the framework supports continuous dosage representations, current public datasets lack sufficiently dense pharmacokinetic measurements, so we use discrete intervention tokens. Second, the physiology-guided energy regularization relies on classical ODE priors, which may be less suitable for highly irregular rhythms such as severe ventricular fibrillation. Third, the constrained action space provides only an approximation of complex ICU pharmacology and does not fully capture polypharmacy interactions. Finally, the current results are proof-of-concept and require substantially broader clinical validation before real-world use.

REFERENCES

Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023.

Edmond Adib, Fatemeh Afghah, and John J Prevost. Synthetic ecg signal generation using generative neural networks. *PloS one*, 20(3):e0271270, 2025.

Edmonmd Adib, Amanda S Fernandez, Fatemeh Afghah, and John J Prevost. Synthetic ecg signal generation using probabilistic diffusion models. *IEEE Access*, 11:75818–75828, 2023.

Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein generative adversarial networks. In *International conference on machine learning*, pp. 214–223. Pmlr, 2017.

Shuai Bai, Keqin Chen, Xuejing Liu, Jialin Wang, Wenbin Ge, Sibo Song, Kai Dang, Peng Wang, Shijie Wang, Jun Tang, et al. Qwen2. 5-vl technical report. *arXiv e-prints*, pp. arXiv–2502, 2025.

Lisa Bedin, Gabriel Cardoso, Josselin Duchateau, Remi Dubois, and Eric Moulines. Leveraging an ecg beat diffusion model for morphological reconstruction from indirect signals. *Advances in Neural Information Processing Systems*, 37:84409–84446, 2024.

Julia Camps, Zhinuo Jenny Wang, Ruben Doste, Lucas Arantes Berg, Maxx Holmes, Brodie Lawson, Jakub Tomek, Kevin Burrage, Alfonso Bueno-Orovio, and Blanca Rodriguez. Harnessing 12-lead ecg and mri data to personalise repolarisation profiles in cardiac digital twin models for enhanced virtual drug testing. *Medical image analysis*, 100:103361, 2025.

Zhikang Chen and Tingting Zhu. From generative engines to actionable simulators: The imperative of physical grounding in world models. *arXiv preprint arXiv:2601.15533*, 2026.

Zhikang Chen, Danni Yang, Tianrui Cui, Ding Li, Houfang Liu, Yi Yang, Sheng Zhang, Sifan Yang, and Tian-Ling Ren. A novel imbalanced dataset mitigation method and ecg classification model based on combined 1d_cbam-autoencoder and lightweight cnn model. *Biomedical Signal Processing and Control*, 87:105437, 2024.

Team Glm, Aohan Zeng, Bin Xu, Bowen Wang, Chenhui Zhang, Da Yin, Dan Zhang, Diego Rojas, Guanyu Feng, Hanlin Zhao, et al. Chatglm: A family of large language models from glm-130b to glm-4 all tools. *arXiv preprint arXiv:2406.12793*, 2024.

Tomer Golany, Daniel Freedman, and Kira Radinsky. Ecg ode-gan: Learning ordinary differential equations of ecg dynamics via generative adversarial learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 134–141, 2021.

Brian Gow, Tom Pollard, Larry A Nathanson, Alistair Johnson, Benjamin Moody, Chrystinne Fernandes, Nathaniel Greenbaum, Jonathan W Waks, Parastou Eslami, Tanner Carbonati, Ashish Chaudhari, Elizabeth Herbst, Dana Moukheiber, Seth Berkowitz, Roger Mark, and Steven Horng. MIMIC-IV-ECG: Diagnostic Electrocardiogram Matched Subset. *PhysioNet*, September 2023. doi: 10.13026/4nqg-sb35. URL `https://doi.org/10.13026/4nqg-sb35`. Version 1.0.

Albert Gu and Tri Dao. Mamba: Linear-time sequence modeling with selective state spaces. In *First conference on language modeling*, 2024.

Yanjiang Guo, Lucy Xiaoyang Shi, Jianyu Chen, and Chelsea Finn. Ctrl-world: A controllable generative world model for robot manipulation. *arXiv preprint arXiv:2510.10125*, 2025.

David Ha and Jürgen Schmidhuber. World models. *arXiv preprint arXiv:1803.10122*, 2(3), 2018.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

Aaron Hurst, Adam Lerer, Adam P Goucher, Adam Perelman, Aditya Ramesh, Aidan Clark, AJ Ostrow, Akila Welihinda, Alan Hayes, Alec Radford, et al. Gpt-4o system card. *arXiv preprint arXiv:2410.21276*, 2024.

Lars Johannesen, Jose Vicente, Jay W Mason, Carlos Sanabria, Kristin Waite-Labott, Mira Hong, Ping Guo, John Lin, Jens Stampe Sørensen, Loriano Galeotti, et al. Differentiating drug-induced multichannel block on the electrocardiogram: randomized study of dofetilide, quinidine, ranolazine, and verapamil. *Clinical Pharmacology & Therapeutics*, 96(5):549–558, 2014.

Lars Johannesen, Jose Vicente, Jay W Mason, Cassandra Erato, Carlos Sanabria, Kristin Waite-Labott, Mira Hong, John Lin, Ping Guo, Abdul Mutlib, et al. Late sodium current block for drug-induced long qt syndrome: results from a prospective clinical trial. *Clinical Pharmacology & Therapeutics*, 99(2):214–223, 2016.

Alistair EW Johnson, Lucas Bulgarelli, Lu Shen, Alvin Gayles, Ayad Shammout, Steven Horng, Tom J Pollard, Sicheng Hao, Benjamin Moody, Brian Gow, et al. Mimic-iv, a freely accessible electronic health record dataset. *Scientific data*, 10(1):1, 2023.

Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 4401–4410, 2019.

Elizabeth Knight, Evangelos K Oikonomou, Arya Aminorroaya, Aline F Pedroso, and Rohan Khera. Wearable echo fm: An ecg echo foundation model for 1 lead electrocardiography. *European Heart Journal-Digital Health*, pp. ztag049, 2026.

Yongfan Lai, Jiabo Chen, Qinghao Zhao, 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. *Patterns*, 6(10), 2025.

Anson Lei, Bernhard Schölkopf, and Ingmar Posner. Spartan: A sparse transformer world model attending to what matters. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

Chunyuan Li, Cliff Wong, Sheng Zhang, Naoto Usuyama, Haotian Liu, Jianwei Yang, Tristan Naumann, Hoifung Poon, and Jianfeng Gao. Llava-med: Training a large language-and-vision assistant for biomedicine in one day. *Advances in Neural Information Processing Systems*, 36: 28541–28564, 2023a.

Junnan Li, Dongxu Li, Silvio Savarese, and Steven Hoi. Blip-2: Bootstrapping language-image pre-training with frozen image encoders and large language models. In *International conference on machine learning*, pp. 19730–19742. PMLR, 2023b.

Wenxuan Li, Hang Zhao, Zhiyuan Yu, Yu Du, Qin Zou, Ruizhen Hu, and Kai Xu. Pin-wm: Learning physics-informed world models for non-prehensile manipulation. *arXiv preprint arXiv:2504.16693*, 2025.

Xiaomin Li, Anne Hee Hiong Ngu, and Vangelis Metsis. Tts-cgan: A transformer time-series conditional gan for biosignal data augmentation. *arXiv preprint arXiv:2206.13676*, 2022.

Jingwei Liu, Hongyan Li, and Shenda Hong. Synthesis of standard 12-lead ecg from single-lead ecg using shifted diffusion models. In *Joint European conference on machine learning and knowledge discovery in databases*, pp. 271–286. Springer, 2024.

Haoyu Lu, Wen Liu, Bo Zhang, Bingxuan Wang, Kai Dong, Bo Liu, Jingxiang Sun, Tongzheng Ren, Zhuoshu Li, Hao Yang, Yaofeng Sun, Chengqi Deng, Hanwei Xu, Zhenda Xie, and Chong Ruan. Deepseek-vl: Towards real-world vision-language understanding, 2024.

Yutaka Matsuo, Yann LeCun, Maneesh Sahani, Doina Precup, David Silver, Masashi Sugiyama, Eiji Uchibe, and Jun Morimoto. Deep learning, reinforcement learning, and world models. *Neural Networks*, 152:267–275, 2022.

Kaden McKeen, Sameer Masood, Augustin Toma, Barry Rubin, and Bo Wang. Ecg-fm: An open electrocardiogram foundation model. *Jamia Open*, 8(5):ooaf122, 2025.

Patrick E McSharry, Gari D Clifford, Lionel Tarassenko, and Leonard A Smith. A dynamical model for generating synthetic electrocardiogram signals. *IEEE transactions on biomedical engineering*, 50(3):289–294, 2003.

Edi Prifti, Ahmad Fall, Giovanni Davogustto, Alfredo Pulini, Isabelle Denjoy, Christian Funck-Brentano, Yasmin Khan, Alexandre Durand-Salmon, Fabio Badilini, Quinn S Wells, et al. Deep learning analysis of electrocardiogram for risk prediction of drug-induced arrhythmias and diagnosis of long qt syndrome. *European Heart Journal*, 42(38):3948–3961, 2021.

Andrew Sellergren, Sahar Kazemzadeh, Tiam Jaroensri, Atilla Kiraly, Madeleine Traverse, Timo Kohlberger, Shawn Xu, Fayaz Jamil, Cían Hughes, Charles Lau, et al. Medgemma technical report. *arXiv preprint arXiv:2507.05201*, 2025.

Qian Shao, Bang Du, Zepeng Li, Qiyuan Chen, Hongxia Xu, Jimeng Sun, Jian Wu, and Jintai Chen. Generation of drug-induced cardiac reactions towards virtual clinical trials. *arXiv preprint arXiv:2502.07297*, 2025.

Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang, Mingchuan Zhang, YK Li, Y Wu, and Daya Guo. DeepSeekMath: Pushing the limits of mathematical reasoning in open language models. *arXiv preprint arXiv:2402.03300*, 2024.

Boyuan Wang, Xinpan Meng, Xiaofeng Wang, Zheng Zhu, Angen Ye, Yang Wang, Zhiqin Yang, Chaojun Ni, Guan Huang, and Xingang Wang. Embodiedreamer: Advancing real2sim2real transfer for policy training via embodied world modeling. *arXiv preprint arXiv:2507.05198*, 2025.

Xiaoda Wang, Kaiqiao Han, Yuhao Xu, Xiao Luo, Yizhou Sun, Wei Wang, and Carl Yang. Se-diff: Simulator and experience enhanced diffusion model for comprehensive ecg generation. In *The Fourteenth International Conference on Learning Representations*, 2026.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, et al. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35:24824–24837, 2022.

Zhijian Wei, Runze Ma, Zichen Wang, Zhongmin Li, Shuotong Song, and Shuangjia Zheng. Vcworld: A biological world model for virtual cell simulation. *arXiv preprint arXiv:2512.00306*, 2025.

Yuhao Xu, Xiaoda Wang, Yi Wu, Wei Jin, Xiao Hu, and Carl Yang. Ecg-moe: Mixture-of-expert electrocardiogram foundation model. *arXiv preprint arXiv:2603.04589*, 2026.

Yang Yang, Tianyu Lan, Yang Wang, Fengtian Li, Liyan Liu, Xupeng Huang, Fei Gao, Shuhua Jiang, Zhijun Zhang, and Xing Chen. Data imbalance in cardiac health diagnostics using cecg-gan. *Scientific Reports*, 14(1):14767, 2024.

Yijun Yang, Zhao-Yang Wang, Qiuping Liu, Shuwen Sun, Kang Wang, Rama Chellappa, Zongwei Zhou, Alan Yuille, Lei Zhu, Yu-Dong Zhang, et al. Medical world model: Generative simulation of tumor evolution for treatment planning. *arXiv preprint arXiv:2506.02327*, 2025.

Haoyu Zhen, Xiaowen Qiu, Peihao Chen, Jincheng Yang, Xin Yan, Yilun Du, Yining Hong, and Chuang Gan. 3d-vla: A 3d vision-language-action generative world model. *arXiv preprint arXiv:2403.09631*, 2024.

Pan Zhou, Zhao Yang, Yiming Hao, Fangfang Fan, Wenlang Zhao, Ziyu Wang, Qiuju Deng, Yongchen Hao, Na Yang, Lizhen Han, et al. A hybrid algorithm-based ecg risk prediction model for cardiovascular disease. *European Heart Journal-Digital Health*, 6(3):466–475, 2025a.

Xin Zhou, Dingkang Liang, Sifan Tu, Xiwu Chen, Yikang Ding, Dingyuan Zhang, Feiyang Tan, Hengshuang Zhao, and Xiang Bai. Hermes: A unified self-driving world model for simultaneous 3d scene understanding and generation. *arXiv preprint arXiv:2501.14729*, 2025b.

A DATASETS

We utilize a combination of real-world clinical electronic health records and controlled ECG datasets to comprehensively evaluate our framework.

MIMIC-IV-ECG & Clinical Cohort Our primary dataset is constructed by integrating the MIMIC-IV-ECG (Gow et al., 2023) and MIMIC-IV-Clinical (Johnson et al., 2023) databases. MIMIC-IV-ECG contains a large collection of 12-lead electrocardiogram (ECG) waveform recordings, while MIMIC-IV-Clinical provides comprehensive electronic health record (EHR) data, including patient demographics, diagnoses, laboratory measurements, procedures, and clinical outcomes. We link the two datasets through shared patient identifiers, enabling the alignment of ECG signals with corresponding clinical context. In particular, medication records are aligned with each patient and their corresponding pre-dose ECG observations, forming patient–drug pairs that reflect real-world treatment trajectories. The ECG data capture fine-grained cardiac dynamics, while explicit exposure to drugs allows the world model to learn patient-specific state transitions. This multimodal dataset serves as our primary testbed for evaluating *treatment consistency*, allowing us to compare model-inferred intervention preferences with clinician-prescribed treatments. To prevent data leakage, we perform dataset splitting at the patient level based on unique subject identifiers, ensuring that no patient appears in both training and test sets. The cohort includes 17,493 unique patients, of which 13,994 are assigned to the training set and 3,499 to the test set. At the episode level, the dataset comprises 108,980 samples in total, with 87,652 training episodes and 21,328 test episodes, consistent with the total number of processed records. In addition, the dataset includes 3,090 distinct medications, providing a diverse set of intervention signals for modeling drug-induced physiological responses.

Drug-Response ECG Datasets. To evaluate drug-induced ECG changes, we use the ECGRDVQ database and the ECGDMMLD database. The experimental cohort used in this study comprises 7,763 12-lead ECG recordings from 43 subjects across eight drug regimens: Dofetilide, Lidocaine+Dofetilide, Mexiletine+Dofetilide, Moxifloxacin+Diltiazem, Placebo, Quinidine, Ranolazine, and Verapamil. We partition this cohort at the subject level into 6,143 recordings from 34 subjects for training and 1,620 recordings from 9 subjects for testing. No subject appears in both splits. These datasets are used to evaluate the fidelity of action-conditioned ECG generation.

In-house OOD Evaluation Dataset To further evaluate the effectiveness of our approach in realistic clinical settings, we adopt a cross-dataset generalization strategy. Specifically, our model is trained on Drug-response ECG datasets, which consist of ECG recordings from healthy individuals following drug administration. To assess the model's ability to generalize to more clinically relevant and challenging out-of-distribution (OOD) scenarios, we construct an in-house evaluation dataset based on MIMIC-IV-ECG. This dataset contains ECG recordings from ICU patients after the administration of Verapamil, reflecting real-world clinical conditions with underlying diseases and complex physiological variability. Therefore, this dataset provides a rigorous benchmark for evaluating the robustness of our physiology-guided energy regularization against severe distribution shifts.

B EXPERIMENTS

B.1 ABLATION STUDIES: THE ROLE OF ENERGY REGULARIZATION

To rigorously isolate the contribution of our proposed physiological constraints, we conduct an ablation study comparing the full ECG World Model against an unconstrained diffusion variant (i.e., removing the EPK-informed Energy Regularization).

Mitigating Error Accumulation in Multi-Step Rollouts As illustrated in Figure 3, our model consistently outperforms the unconstrained variant across all rollout horizons. In the one-step setting, energy regularization reduces the latent MSE from 0.061 to 0.045. As the rollout horizon increases, the unconstrained model exhibits significant error accumulation and deviates from physiologically plausible cardiac trajectories, while our method maintains stable predictions. This effect is further confirmed in Table 5, where we compare autoregressive rollout performance with oracle predictions. The gap (Δ) between rollout and oracle errors remains minimal for our model (+0.0008), whereas

general-purpose models exhibit substantially larger deviations. These results suggest that our approach effectively mitigates temporal error accumulation and improves long-horizon consistency.

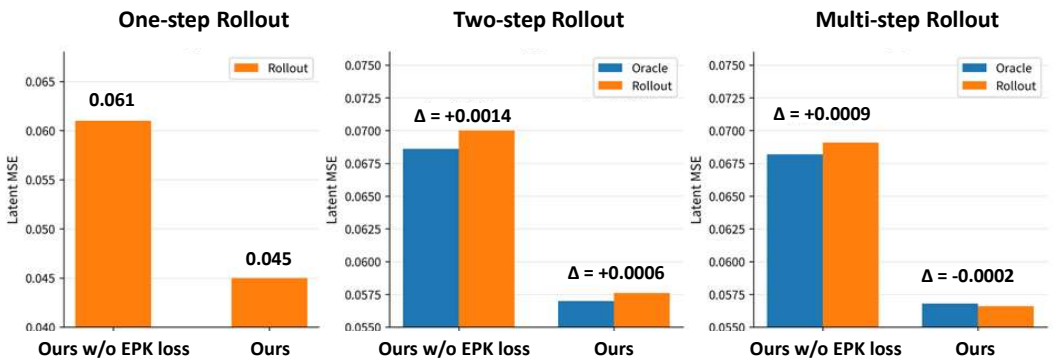

Figure 3: The overview of rollout settings used in this study. One-step rollout evaluates single-step prediction on the Drug-response ECG dataset, while two-step and multi-step rollouts assess short- and long-horizon predictions on the MIMIC-IV-ECG dataset.

Table 5: Comparison of autoregressive multi-step rollout and oracle predictions across models using MIMIC-IV-ECG dataset. Δ denotes the difference between rollout and oracle latent signal mse.

Model	Oracle ↓	Rollout ↓	Δ
GPT-4o	0.1124	0.1245	+0.0121
GPT-5-mini	0.1025	0.1172	+0.0147
GPT-3.5	0.0942	0.1036	+0.0094
Ours	**0.0572**	**0.0580**	**+0.0008**

Fidelity Across Latent and Signal Spaces Figure 4 further shows that removing the energy regularization mechanism degrades predictive fidelity in both the latent space and the high-dimensional signal space. Our fully regularized model achieves consistently lower MSE and MAE across both domains. These ablation results support our core premise that purely data-driven diffusion models are prone to temporal inconsistencies during multi-step intervention simulation. Incorporating physiological priors via energy regularization plays a key role in improving temporal consistency and predictive stability.

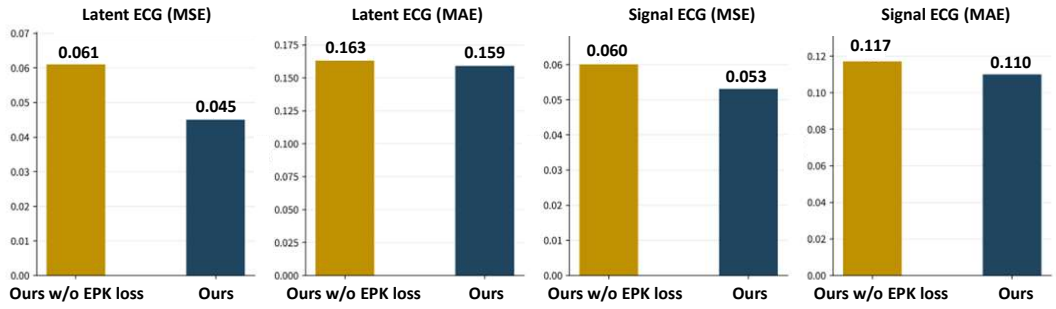

Figure 4: Ablation study on the Drug-response ECG dataset. We compare our full model with a variant without EPK loss across both latent and signal spaces using MSE and MAE metrics, demonstrating the effectiveness of the EPK loss in improving prediction quality.

B.2 Repeated-Sampling Evaluation of Diagnostic-Change Fidelity

We evaluate diagnostic-change fidelity across six stochastic sampling runs on the same 200 test cases. For each case, the predicted post-dose diagnostic score for the recorded medication is averaged over $S = 5$ standard-DDPM samples. Predicted and observed diagnostic changes are computed relative to the same pre-dose reference using the frozen diagnostic backend.

The risk-aversion coefficient λ is used only for candidate selection through the score $\mu_R(a) + \lambda\sigma_R(a)$. It does not enter the diagnostic-change calculation for the recorded medication. Although the six

Table 6: Architecture-preserving ablation of the physiological energy coefficient λ_{EPK}. EPK conditioning remains enabled in every row; only the energy-loss coefficient is varied. The coefficients and numerical results are retained from this weight-sweep protocol, including the $\lambda_{\text{EPK}} = 0.25$ setting.

λ_{EPK}	Latent Space MSE ↓	Latent Space MAE ↓	Signal Space MSE ↓	Signal Space MAE ↓
0	0.049	0.167	**0.052**	0.117
0.25	**0.045**	**0.159**	0.053	**0.110**
0.5	0.049	0.165	0.058	0.116
0.75	0.047	0.163	0.053	0.113
1.0	0.049	0.166	0.058	0.118

evaluation files were associated with different selection coefficients, each run regenerated its DDPM samples. Differences in the reported correlations and sign agreement therefore reflect stochastic sampling variation rather than an isolated effect of λ.

As shown in Figure 5, positive diagnostic-change agreement is observed across all six runs. Pearson correlation is 0.592 ± 0.021, Spearman correlation is 0.577 ± 0.022, and sign agreement is $(71.3 \pm 2.2)\%$. These values report the mean and sample standard deviation across the six runs, quantifying evaluation-sampling variability rather than training-seed variability. The results support diagnostic-change fidelity across repeated stochastic sampling under the frozen evaluation backend. Complete run-level results and the evaluation definitions are provided in Appendix E.4.

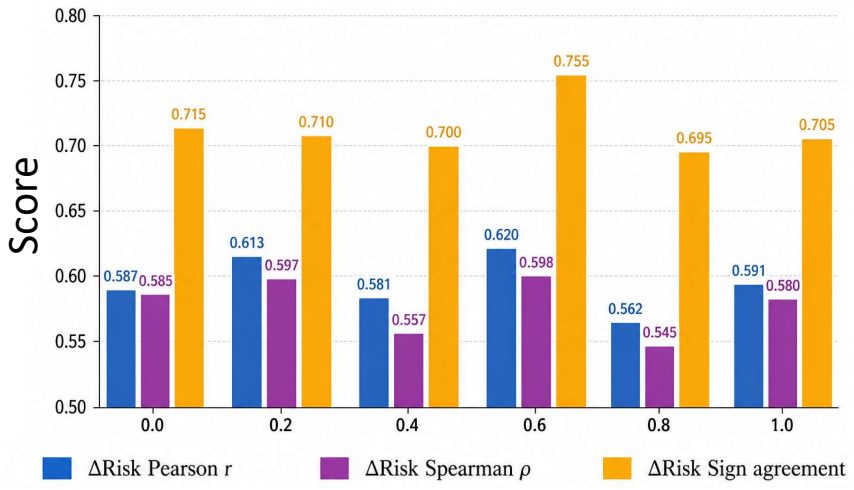

Figure 5: Diagnostic-change agreement across six stochastic sampling runs on the same 200 test cases. Each run regenerates standard-DDPM samples for the recorded medication. Bars report Pearson r, Spearman ρ, and sign agreement between predicted and observed diagnostic-score changes. The horizontal axis indexes sampling runs, not a risk-aversion ablation: λ is used only for candidate selection and does not enter these diagnostic-change metrics.

B.3 Architecture-Preserving Energy-Weight Ablation

We vary the physiological energy coefficient λ_{EPK} while retaining the EPK architecture and conditioning pathway. In particular, $\lambda_{\text{EPK}} = 0$ removes the energy penalty but does not disable EPK conditioning. This setting is therefore distinct from the EPK-module removal in Figure 4.

Table 6 shows a non-monotonic dependence on the energy coefficient. Within this sweep, $\lambda_{\text{EPK}} = 0.25$ gives the lowest latent MSE, latent MAE, and signal MAE. The zero-weight setting achieves slightly lower signal MSE, 0.052 compared with 0.053. Larger coefficients do not consistently improve the reported metrics.

Table 7: Effect of the number of stochastic samples K on reconstruction errors in latent and signal spaces.

K	Latent Space		Signal Space	
	MSE ↓	MAE ↓	MSE ↓	MAE ↓
1	0.046	0.162	0.056	0.121
3	0.045	0.159	0.053	0.110
5	0.045	0.157	0.052	0.106

These results characterize the effect of energy-loss weighting with EPK conditioning retained. The zero-weight latent MSE of 0.049 and the module-removal result of 0.061 correspond to different configurations and should not be interpreted as duplicate measurements of the same ablation. The main pharmacological-mapping experiments use $\lambda_{\text{EPK}} = 0.1$ under their separately reported protocol; this does not relabel the 0.25 setting in the present sweep.

B.4 Effect of the Number of Stochastic Samples

We examine how the number of stochastic samples K affects reconstruction performance and sampling cost. Table 7 reports latent-space and waveform-space errors for $K \in \{1, 3, 5\}$.

Increasing K from 1 to 3 reduces all four reported errors. In particular, signal MSE decreases from 0.056 to 0.053, while signal MAE decreases from 0.121 to 0.110. Increasing K to 5 provides further modest improvements, with latent MSE unchanged at the reported precision. These results indicate diminishing reconstruction gains as the sampling budget increases.

We use $K = 3$ for this reconstruction evaluation setting to balance prediction accuracy and computational cost. The diagnostic-score evaluation in Section 4.3 separately uses $S = 5$ samples per intervention to compute the mean and variability of diagnostic-proxy scores. The reconstruction metrics reported here do not directly assess uncertainty calibration.

B.5 Robustness under Missing-Lead Settings

Table 8 shows that our full model consistently outperforms the variant without EPK loss across all missing-lead settings. Improvements are observed in both latent and signal spaces, with lower MSE and MAE under every masking pattern. In particular, the performance gap becomes more pronounced under more challenging missing-lead conditions, indicating that EPK-based regularization enhances the model's robustness to incomplete observations. By incorporating physiological priors, our method is better able to recover coherent cardiac dynamics even when input signals are partially missing.

Table 8: Reconstruction performance under missing-lead settings. Comparison between Ours and Ours w/o EPK loss test sets using 50 samples (lower is better).

Pattern	Ours				Ours w/o EPK loss			
	Latent		Signal		Latent		Signal	
	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓
miss-I	0.0451	0.1650	0.0413	0.1013	0.0485	0.1698	0.0480	0.1095
miss-II	0.0454	0.1650	0.0420	0.1017	0.0484	0.1694	0.0484	0.1107
miss-III	0.0449	0.1644	0.0405	0.1010	0.0476	0.1683	0.0467	0.1089
miss-aVR	0.0453	0.1641	0.0399	0.0999	0.0488	0.1704	0.0487	0.1113
miss-aVL	0.0426	0.1602	0.0357	0.0946	0.0489	0.1707	0.0483	0.1109
miss-aVF	0.0443	0.1632	0.0381	0.0984	0.0502	0.1724	0.0496	0.1136
miss-V1	0.0442	0.1625	0.0405	0.0999	0.0494	0.1712	0.0491	0.1112
miss-V2	0.0458	0.1653	0.0412	0.1016	0.0487	0.1701	0.0490	0.1107
miss-V3	0.0436	0.1611	0.0404	0.0991	0.0482	0.1695	0.0473	0.1097
miss-V4	0.0433	0.1607	0.0395	0.0972	0.0495	0.1702	0.0496	0.1119
miss-V5	0.0442	0.1622	0.0421	0.0995	0.0495	0.1720	0.0490	0.1127
miss-V6	0.0445	0.1629	0.0401	0.0979	0.0510	0.1743	0.0523	0.1183

B.6 Counterfactual Drug-Response Simulation

Figure 6 demonstrates the model's ability to perform counterfactual simulation of drug responses from a single pre-dose ECG. Different candidate drugs lead to distinct post-dose ECG patterns and clinical indicators, reflecting their heterogeneous physiological effects. This enables a direct comparison of candidate interventions at the individual level, allowing the model to analyze how different treatments may alter cardiac dynamics before actual administration. This capability is valuable for personalized and risk-aware clinical analysis.

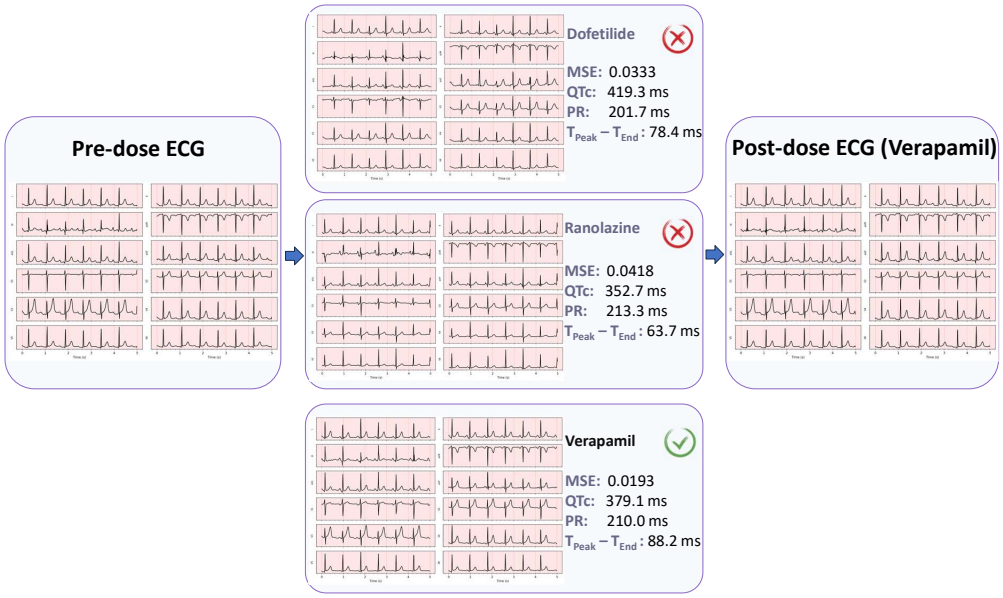

Figure 6: Counterfactual drug-response simulation with the ECG World Model. Given a normal pre-dose ECG, the model simulates post-dose outcomes under different candidate drugs, enabling direct comparison of drug-specific cardiac responses via generated ECG signals and clinical biomarkers.

B.7 Reconstruction under Single and Multiple Missing Leads

To evaluate the robustness of our approach against data corruption—a prevalent issue in continuous clinical monitoring caused by electrode detachment or sensor malfunction—we investigated the model's performance under simulated single and multiple missing lead conditions. As illustrated in Figure 7, the model demonstrates exceptional fault tolerance. Even when predicting the post-dose state from pre-dose ECGs with one or two missing leads, the framework successfully reconstructs the absent signals simultaneously. Crucially, the imputed waveforms exhibit high morphological fidelity compared to the ground-truth 12-lead ECGs, preserving vital structural components (e.g., QRS complexes and T-waves) and overall rhythm dynamics without introducing perceptible artifacts or signal distortion.

This robust performance under extreme data degradation scenarios serves as compelling evidence for the efficacy of our proposed world model paradigm. Conventional direct mapping networks typically suffer severe performance degradation or catastrophic failure when confronted with incomplete input modalities, as they heavily rely on complete point-to-point correspondences. In contrast, our world model does not merely perform surface-level signal interpolation. Instead, it has internalized the intrinsic spatio-temporal dynamics and structural dependencies of the cardiovascular system's electrical activity.

By establishing a robust latent representation of this physiological "world," the model treats the missing leads as partial observations. It effectively leverages the spatial correlations inherent in the remaining functional leads to accurately infer the complete, unobserved physiological state. The ability to synthesize coherent, high-fidelity 12-lead ECGs from severely corrupted inputs—whether a single lead or multiple leads are compromised—underscores that the model has learned representations

that support coherent reconstruction under the evaluated perturbations. This intrinsic understanding provides empirical evidence of robustness within the evaluated settings.

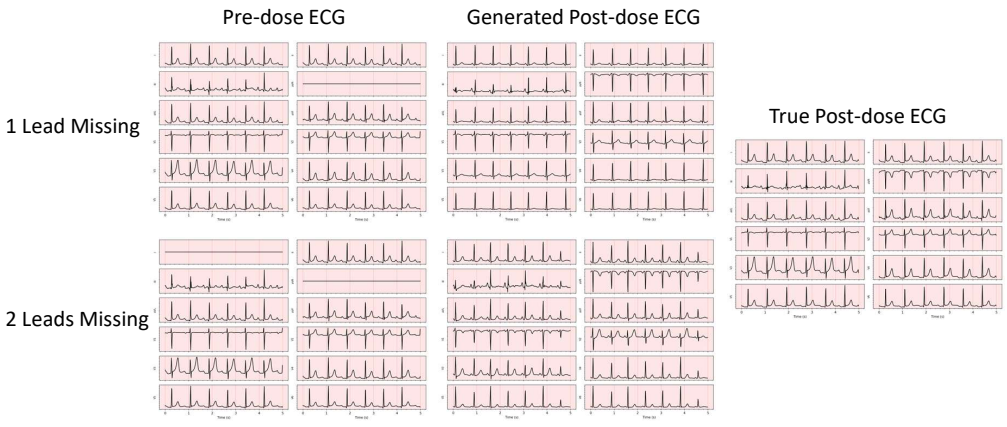

Figure 7: Robustness of the predictive world model under varying degrees of data corruption. The figure demonstrates the model's ability to simultaneously forecast post-dose states and impute missing sensor data. **Left Panel (Pre-dose ECG)**: Serves as the conditioned input, with simulated sensor failures visible as flatlines. The top row illustrates a 1 Lead Missing scenario, while the bottom row demonstrates a more severe 2 Leads Missing scenario. **Middle Panel (Generated Post-dose ECG)**: Shows the model's output. Despite the corrupted pre-dose inputs, the model successfully reconstructs the missing leads while accurately predicting the morphological changes induced by the pharmacological intervention. **Right Panel (True Post-dose ECG)**: Provides the ground-truth reference. A comparison reveals that the generated waveforms—including the previously missing leads—exhibit high structural fidelity and correctly preserve inter-lead spatial correlations, indicating that the learned representation preserves useful cross-lead dependencies under the evaluated missing-lead conditions.

B.8 Reconstruction under Single-Lead Missing in Abnormal Patients

Under the single-lead missing setting, the reconstruction task becomes particularly challenging for abnormal patients, where critical diagnostic information may be distributed across leads. As shown in Figure 8, our method is able to accurately infer the missing information and recover post-dose ECG waveforms that remain highly consistent with the ground truth. Notably, key pathological patterns are well preserved, indicating that the model effectively captures cross-lead dependencies and underlying physiological dynamics.

For instance, in a representative case with complex cardiovascular conditions—subendocardial infarction (initial episode of care), acute on chronic systolic heart failure, and coronary artery dissection—the patient had a pre-dose potassium level of 3.9 mmol/L and was administered Furosemide (Lasix) combined with Heparin Sodium. Despite the presence of multi-factorial cardiac abnormalities and pharmacological intervention, the reconstructed ECG successfully preserves clinically relevant waveform characteristics, further demonstrating the robustness of our approach under challenging pathological scenarios.

In contrast, DADM struggles in this setting, producing oversmoothed waveforms with diminished abnormal features. This suggests that it fails to leverage inter-lead relationships when partial observations are provided, resulting in degraded reconstruction quality. Overall, these results highlight the advantage of our approach in handling incomplete inputs and maintaining clinically meaningful signal structures under challenging abnormal conditions.

B.9 Stability and Graceful Degradation

As illustrated in Figure 9, our empirical evaluation reveals a consistent ordering across all three perturbation regimes (Healthy, Heavy Jitter, and Latent Gaussian): EPK on $<$ EPK$\times 0.25 <$ EPK off. This trend is strictly preserved across all evaluated metrics, Mean Pairwise ℓ_2 Distance, Mean

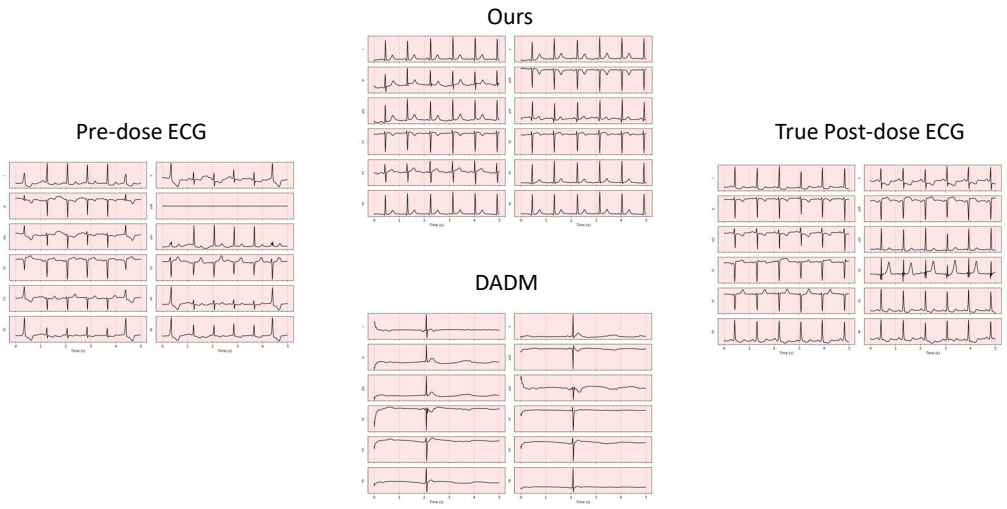

Figure 8: Reconstruction under single-lead missing in an abnormal exemplar patient. Given pre-dose ECGs with one lead removed, our method reconstructs post-dose waveforms that closely match the ground truth, effectively preserving pathological characteristics. In contrast, DADM produces oversmoothed signals and fails to recover clinically relevant patterns, leading to noticeable deviations from the true post-dose ECG.

Distance to Centroid, and global Latent Standard Deviation. The simultaneous reduction in inter-sample variability and latent dispersion confirms that the EPK mechanism effectively enhances overall generation stability.Crucially, as the system is subjected to increasing prior mismatch, moving from the well-aligned Healthy regime to unstructured Latent Gaussian noise, the degradation in stability is smooth and bounded. While all configurations experience moderate performance drops due to the growing divergence between the ODE prior and the true data dynamics, the EPK-enabled models do not exhibit abrupt spikes or catastrophic collapse. Furthermore, the intermediate performance of the weak prior (EPK $\times$ 0.25) indicates that the influence of the prior is continuous and tunable via scaling, rather than acting as a binary constraint.Together, these observations provide strong evidence that our model exhibits graceful degradation. Rather than failing catastrophically when the prior becomes unreliable, the system smoothly transitions toward data-driven behavior. It maintains a strict stability advantage over the unguided baseline (EPK off), demonstrating robust performance even under severe prior misalignment.

B.10 Clinical Risk Formulation and Label Aggregation

To quantitatively assess diagnostic changes in simulated pharmacological responses, we map generated ECG waveforms to a scalar diagnostic risk proxy using a frozen multi-label ECG-FM model, denoted as f_{risk}.

Diagnostic Model and 17-Dimensional Target Labels The diagnostic model is formulated around the MIMIC-IV-ECG machine reads task as a multi-label classification problem with 17 binary targets. Each target indicates the presence of a particular interpretation in the machine-generated diagnostic report. The categories are: Poor data quality, Sinus rhythm, PVC, Tachycardia, Ventricular tachycardia, SVT with aberrancy, Atrial fibrillation, Atrial flutter, Bradycardia, Accessory pathway conduction, AV block, 1st degree AV block, Bifascicular block, RBBB, LBBB, Myocardial infarction, and Electronic pacemaker.

Aggregation into a Scalar Risk Proxy For an ECG waveform x, the model f_{risk} outputs 17 logits. Let $\ell_j(x)$ denote the logit for label j. Applying a sigmoid to each logit gives the corresponding diagnostic probability:

$$p_j(x) = \text{sigmoid}(\ell_j(x)), \qquad j = 1, \ldots, 17.$$

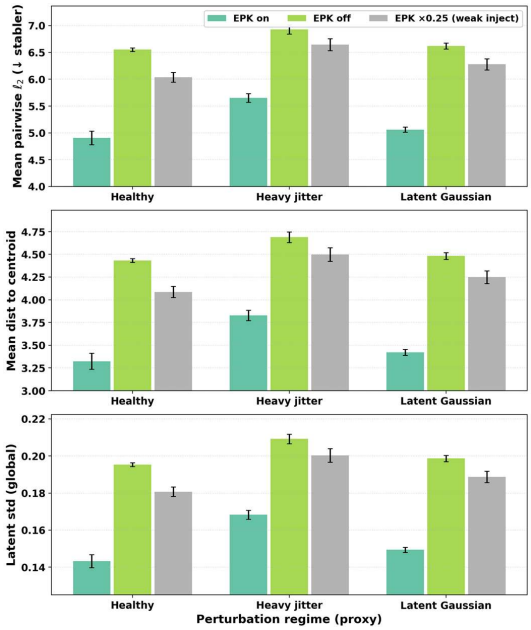

Figure 9: Empirical stability of the EPK mechanism under increasing prior mismatch. We compare full prior injection (EPK on), weak prior (EPK $\times$ 0.25), and no prior (EPK off) across three perturbation regimes: Healthy, Heavy Jitter, and Latent Gaussian. Stability is quantified via three complementary metrics: Mean Pairwise ℓ_2 Distance, Mean Distance to Centroid, and global Latent Standard Deviation (lower is better for all). Across all regimes, enabling EPK significantly reduces generation variability and latent dispersion. Notably, under severe prior mismatch, EPK-enabled models exhibit graceful degradation—stability metrics degrade smoothly and remain bounded without catastrophic failure, consistently outperforming the baseline without prior injection.

The implemented aggregation is the arithmetic mean of these probabilities:

$$R(x) = \frac{1}{17} \sum_{j=1}^{17} p_j(x). \tag{9}$$

The same frozen model and aggregation rule are applied to generated, observed post-dose, and pre-dose ECGs.

For each candidate intervention a, we generate $S = 5$ next-state samples using standard DDPM with EPK conditioning:

$$\hat{z}_{k+1}^{(s)} \sim p_{\mathrm{WM}}(\cdot \mid z_k, a), \qquad s = 1, \ldots, S.$$

Scoring each decoded waveform yields

$$R^{(s)}(a) = R\Big(f_{\mathrm{dec}}(\hat{z}_{k+1}^{(s)})\Big). \tag{10}$$

We summarize the scores using their empirical mean and variance:

$$\mu_R(a) = \frac{1}{S} \sum_{s=1}^{S} R^{(s)}(a), \qquad \sigma_R^2(a) = \frac{1}{S-1} \sum_{s=1}^{S} \Big(R^{(s)}(a) - \mu_R(a)\Big)^2. \tag{11}$$

These quantities characterize the mean diagnostic score under the world model and its variability across stochastic simulations.

For the recorded medication a_{true}, predicted and observed diagnostic-score changes use the same pre-dose reference:

$$\Delta_{\mathrm{wm}} = \mu_R(a_{\mathrm{true}}) - R(x_{\mathrm{pre}}), \tag{12}$$

$$\Delta_{\mathrm{real}} = R(x_{\mathrm{post}}) - R(x_{\mathrm{pre}}). \tag{13}$$

The risk-aversion coefficient λ affects candidate selection through $\mu_R(a) + \lambda\sigma_R(a)$, but does not enter these diagnostic-change calculations. Comparing Δ_{wm} with Δ_{real} assesses diagnostic-change fidelity under a shared backend. Because the label set includes heterogeneous findings, the aggregate is treated as an evaluation proxy rather than a calibrated measure of overall clinical severity.

C Derivation of the ECG World Model

This appendix provides a detailed derivation of the proposed ECG world model. We first introduce the notation and assumptions, then derive the diffusion process and its reverse parameterization, and finally present the physiological prior, training objective, and inference procedure.

C.1 Notations and Basic Assumptions

C.1.1 Time-Scale Separation

We distinguish between two time scales:

- **MDP timestep**: indexed by $k = 0, 1, \ldots, L$. At step k, the true latent cardiac state is denoted by $z_k \in \mathbb{R}^d$, the model prediction is $\hat{z}_k$, and the applied medical action is $a_k \in \mathcal{A}$. The environment evolves according to
$$z_{k+1} = \mathcal{T}(z_k, a_k) + \eta_k, \qquad \eta_k \sim \mathcal{N}(0, \sigma_\eta^2 I_d), \tag{14}$$
where I_d is the $d \times d$ identity matrix and η_k is independent of the past conditioned on (z_k, a_k).
- **Diffusion timestep**: indexed by $\tau = 0, 1, \ldots, N$, where $\tau = 0$ denotes the clean latent sample and $\tau = N$ denotes pure noise. For a fixed MDP step k, the diffusion model generates z_{k+1} by denoising from Gaussian noise, conditioned on the current estimate $\hat{z}_k$ and the action a_k.

C.1.2 VAE and Latent-Space Assumptions

We assume an encoder $f_{\mathrm{enc}} : \mathbb{R}^{C\times T_{\mathrm{sig}}} \to \mathbb{R}^d$ and a decoder $f_{\mathrm{dec}} : \mathbb{R}^d \to \mathbb{R}^{C\times T_{\mathrm{sig}}}$ such that $f_{\mathrm{dec}}(f_{\mathrm{enc}}(x)) \approx x$. For analysis, we assume that the decoder is locally bi-Lipschitz on a compact subset of the physiological manifold $\mathcal{M} \subset \mathbb{R}^d$:

$$\exists\, 0 < C_1 \le C_2, \quad \forall z, z' \in \mathcal{M}, \qquad C_1\|z - z'\|_2 \le \|f_{\mathrm{dec}}(z) - f_{\mathrm{dec}}(z')\|_2 \le C_2\|z - z'\|_2. \tag{15}$$

This assumption explicitly connects the bounded multi-step latent-space errors to waveform-space structural divergence in our analysis.

C.2 Diffusion Model Basics

For simplicity, we omit the explicit condition $(\hat{z}_k, a_k)$ in this subsection, although all distributions are conditioned on them.

C.2.1 Forward Diffusion Process

Let $\beta_1, \beta_2, \ldots, \beta_N \in (0, 1)$ be a variance schedule, and define

$$\alpha_\tau = 1 - \beta_\tau, \qquad \bar{\alpha}_\tau = \prod_{i=1}^{\tau} \alpha_i.$$

The forward diffusion process is the Markov chain

$$q(z_{k+1}^{(\tau)} \mid z_{k+1}^{(\tau-1)}) = \mathcal{N}\Big(z_{k+1}^{(\tau)};\ \sqrt{\alpha_\tau}\, z_{k+1}^{(\tau-1)}, \beta_\tau I_d\Big), \qquad \tau = 1, \ldots, N, \tag{16}$$

with $z_{k+1}^{(0)} = z_{k+1}$. By iterating the recursion, we obtain the standard closed form

$$q(z_{k+1}^{(\tau)} \mid z_{k+1}^{(0)}) = \mathcal{N}\Big(z_{k+1}^{(\tau)};\ \sqrt{\bar{\alpha}_\tau}\, z_{k+1}^{(0)}, (1 - \bar{\alpha}_\tau) I_d\Big). \tag{17}$$

Derivation The result follows by induction. For $\tau = 1$, the claim is immediate from the definition. Assume it holds for $\tau - 1$. Then

$$\begin{aligned} z_{k+1}^{(\tau)} &= \sqrt{\alpha_\tau}\, z_{k+1}^{(\tau-1)} + \sqrt{\beta_\tau}\, \epsilon_\tau \\ &= \sqrt{\alpha_\tau}\left(\sqrt{\bar{\alpha}_{\tau-1}}\, z_{k+1}^{(0)} + \sqrt{1-\bar{\alpha}_{\tau-1}}\, \epsilon'\right) + \sqrt{\beta_\tau}\, \epsilon_\tau \\ &= \sqrt{\bar{\alpha}_\tau}\, z_{k+1}^{(0)} + \sqrt{\alpha_\tau(1-\bar{\alpha}_{\tau-1})}\, \epsilon' + \sqrt{\beta_\tau}\, \epsilon_\tau, \end{aligned}$$

where $\epsilon', \epsilon_\tau \sim \mathcal{N}(0, I_d)$ are independent. Since the last two terms are independent Gaussians, their sum is Gaussian with covariance

$$\alpha_\tau(1-\bar{\alpha}_{\tau-1}) + \beta_\tau = 1 - \bar{\alpha}_\tau,$$

which proves Equation 17.

C.2.2 Reverse Process and Parameterization

The exact reverse posterior has the form

$$q\left(z_{k+1}^{(\tau-1)} \mid z_{k+1}^{(\tau)}, z_{k+1}^{(0)}\right) = \mathcal{N}\left(z_{k+1}^{(\tau-1)};\, \tilde{\mu}_\tau(z_{k+1}^{(\tau)}, z_{k+1}^{(0)}), \tilde{\beta}_\tau I_d\right), \tag{18}$$

where

$$\tilde{\beta}_\tau = \frac{1-\bar{\alpha}_{\tau-1}}{1-\bar{\alpha}_\tau}\, \beta_\tau, \tag{19}$$

and

$$\tilde{\mu}_\tau(z_{k+1}^{(\tau)}, z_{k+1}^{(0)}) = \frac{\sqrt{\bar{\alpha}_{\tau-1}}\beta_\tau}{1-\bar{\alpha}_\tau}\, z_{k+1}^{(0)} + \frac{\sqrt{\alpha_\tau}(1-\bar{\alpha}_{\tau-1})}{1-\bar{\alpha}_\tau}\, z_{k+1}^{(\tau)}. \tag{20}$$

We parameterize the learned reverse process by predicting the clean latent sample. Let

$$\hat{z}_\theta = \hat{z}_\theta\left(z_{k+1}^{(\tau)}, \tau, \hat{z}_k, a_k\right)$$

denote the network prediction of $z_{k+1}^{(0)}$. Then the learned reverse transition is

$$p_\theta\left(z_{k+1}^{(\tau-1)} \mid z_{k+1}^{(\tau)}, \hat{z}_k, a_k\right) = \mathcal{N}\left(z_{k+1}^{(\tau-1)};\, \mu_\theta\left(z_{k+1}^{(\tau)}, \tau, \hat{z}_k, a_k\right), \tilde{\beta}_\tau I_d\right), \tag{21}$$

with

$$\mu_\theta\left(z_{k+1}^{(\tau)}, \tau, \hat{z}_k, a_k\right) = \tilde{\mu}_\tau\left(z_{k+1}^{(\tau)}, \hat{z}_\theta\left(z_{k+1}^{(\tau)}, \tau, \hat{z}_k, a_k\right)\right). \tag{22}$$

Equivalently, if the network predicts noise ϵ_θ, then the clean estimate is

$$\hat{z}_\theta = \frac{1}{\sqrt{\bar{\alpha}_\tau}}\left(z_{k+1}^{(\tau)} - \sqrt{1-\bar{\alpha}_\tau}\, \epsilon_\theta(z_{k+1}^{(\tau)}, \tau, \hat{z}_k, a_k)\right), \tag{23}$$

and substituting this expression into Equation 22 yields the standard DDPM reverse parameterization.

C.3 Construction of the External Physiological Prior (EPK)

ECG ODE The ECG waveform $y(\theta)$, viewed as a function of the cardiac phase angle θ, is governed by

$$\frac{dy}{d\theta} = -\sum_{i\in\{P,Q,R,S,T\}} \alpha_i \Delta\theta_i \exp\left(-\frac{\Delta\theta_i^2}{2b_i^2}\right) - (y - y_0), \tag{24}$$

where $\Delta\theta_i = (\theta - \theta_i) \bmod 2\pi$, and $\Theta = \{\alpha_i, b_i, \theta_i, y_0\}_{i\in\{P,Q,R,S,T\}}$ denotes the physiological parameters controlling the amplitude, width, center position, and baseline of the wave components.

Parameter Modulation under Interventions Given a pharmacological action $a \in \mathcal{A}$, we define a modulation map

$$\Theta_a = \mathcal{M}(a, \Theta), \tag{25}$$

where $\mathcal{M}$ adjusts the ODE parameters according to the expected physiological effect of the intervention. In our implementation, $\mathcal{M}$ is a rule-based heuristic informed by pharmacological literature. It scales selected baseline parameters, such as the wave widths $\{b_i\}$, to construct regimen-dependent physiological templates.

Estimating the Initial Phase from the Current State Given the estimated latent state $\hat{z}_k$ at MDP step k, we decode it into an ECG waveform:

$$\hat{x}_k = f_{\mathrm{dec}}(\hat{z}_k). \tag{26}$$

A deterministic detector h extracts an initial phase anchor from the waveform:

$$\theta_{\mathrm{start}} = h(\hat{x}_k) \in [0, 2\pi). \tag{27}$$

In practice, h uses Pan–Tompkins R-peak localization to establish the phase anchor.

ODE Integration and Projection Given a prediction horizon $\Delta\theta_{\mathrm{pred}}$, let $y^{(a)}(\theta)$ denote the physiological waveform obtained by integrating Equation 24 under the modulated parameters Θ_a. Here, y denotes the waveform-level amplitude, distinct from the latent state z used by the VAE and diffusion model. Starting from θ_{start}, the trajectory evolves along the phase coordinate as

$$\begin{aligned} y^{(a)}(\theta) &= y(\theta_{\mathrm{start}}) + \int_{\theta_{\mathrm{start}}}^{\theta} \mathcal{F}\Big(y^{(a)}(u), u; \Theta_a\Big)\, du, \\ \theta &\in [\theta_{\mathrm{start}}, \theta_{\mathrm{start}} + \Delta\theta_{\mathrm{pred}}], \end{aligned} \tag{28}$$

where $\mathcal{F}$ denotes the right-hand side of Equation 24.

We numerically integrate the ODE and sample the resulting trajectory at L uniformly spaced phase points:

$$e_{\mathrm{EPK}} = \left\{ y^{(a)}\left(\theta_{\mathrm{start}} + \frac{l}{L}\Delta\theta_{\mathrm{pred}}\right) \right\}_{l=1}^{L} \in \mathbb{R}^L. \tag{29}$$

The physiological sequence is mapped into the latent space through the projection module Π_ϕ:

$$z_{\mathrm{EPK}}(\hat{z}_k, a_k) = \Pi_\phi\big(e_{\mathrm{EPK}}(\hat{z}_k, a_k)\big) \in \mathbb{R}^d. \tag{30}$$

The projection consists of temporal resampling to the latent sequence length followed by a learnable 1×1 convolution. Its parameters ϕ are jointly optimized with the diffusion model parameters θ, while the pre-trained VAE remains frozen. The projected anchor supplies the target for the latent physiological energy term; EPK conditioning is provided through a separate embedding pathway.

Remark For fixed model parameters, z_{EPK} depends on the current estimated state $\hat{z}_k$ and action a_k, but not on the diffusion timestep τ. It therefore provides a common physiological anchor for the next-step clean-latent predictions across diffusion timesteps. The decoding used for phase initialization is distinct from evaluating the energy penalty, which operates directly in latent space.

C.4 Derivation of the Training Loss

C.4.1 Standard Diffusion Loss

Let $z_{k+1}^{(0)}$ denote the clean next-step latent sample. We sample a diffusion timestep $\tau \sim \mathrm{Uniform}\{1, \ldots, N\}$ and construct

$$z_{k+1}^{(\tau)} = \sqrt{\bar{\alpha}_\tau}\, z_{k+1}^{(0)} + \sqrt{1 - \bar{\alpha}_\tau}\, \epsilon, \qquad \epsilon \sim \mathcal{N}(0, I_d).$$

The denoiser predicts the injected noise:

$$\hat{\epsilon} = \epsilon_\theta\left(z_{k+1}^{(\tau)}, \tau, \hat{z}_k, a_k\right),$$

where additional conditioning inputs, including the EPK embedding, are omitted for brevity. The implemented diffusion loss is the unweighted noise-prediction objective:

$$\mathcal{L}_{\mathrm{data}}(\theta) = \mathbb{E}\left[\|\hat{\epsilon} - \epsilon\|_2^2\right]. \tag{31}$$

The expectation is taken over training transitions, diffusion timesteps, and Gaussian noise.

Interpretation. The corresponding clean-latent estimate is derived from the predicted noise:

$$\hat{z}_0 = \hat{z}_\theta\Big(z_{k+1}^{(\tau)}, \tau, \hat{z}_k, a_k\Big) = \frac{z_{k+1}^{(\tau)} - \sqrt{1-\bar{\alpha}_\tau}\,\hat{\epsilon}}{\sqrt{\bar{\alpha}_\tau}}.$$

Here, $\hat{z}_\theta$ denotes the derived clean estimate, not a separate prediction network. This estimate is used for physiological regularization, while the data loss in Equation 31 is evaluated in noise space.

C.4.2 Physiological Regularization Term

We regularize the predicted clean sample rather than the noisy intermediate state. The forward process introduces Gaussian noise for diffusion training; applying the physiological penalty to $\hat{z}_0$ instead of $z_{k+1}^{(\tau)}$ targets the denoised representation.

We define the physiological residual as

$$\mathcal{R}_{\text{phys}}(\theta, \phi) = \|\hat{z}_0 - z_{\text{EPK}}(\hat{z}_k, a_k)\|_2^2 . \tag{32}$$

The anchor is obtained through the EPK projection:

$$z_{\text{EPK}}(\hat{z}_k, a_k) = \Pi_\phi(e_{\text{EPK}}(\hat{z}_k, a_k)) .$$

Connection to the EPK Energy. This squared residual is the latent energy introduced in the main text: $E_\phi(\hat{z}_0) = \mathcal{R}_{\text{phys}}(\theta, \phi)$. It measures deviation from the intervention-conditioned physiological anchor directly in latent space, without waveform decoding to evaluate the penalty.

The physiological loss uses a fixed coefficient:

$$\mathcal{L}_{\text{phys}}(\theta, \phi) = \lambda_{\text{EPK}}\, \mathbb{E}\left[\mathcal{R}_{\text{phys}}(\theta, \phi)\right]. \tag{33}$$

There is no additional timestep-dependent weighting. The main pharmacological-mapping experiments use $\lambda_{\text{EPK}} = 0.1$, while the energy-weight ablation uses its separately reported coefficients.

Training convention During training, we use teacher forcing by setting $\hat{z}_k = z_k$. The physiological anchor is constructed from this current state and the recorded action a_k. For a fixed transition and fixed projection parameters, the anchor does not depend on the diffusion timestep τ, although the predicted clean latent does.

C.4.3 Joint Objective

The final training objective combines the noise-prediction loss and physiological regularization:

$$\begin{aligned} \mathcal{L}_{\text{ECGWM}}(\theta, \phi) &= \mathcal{L}_{\text{data}}(\theta) + \mathcal{L}_{\text{phys}}(\theta, \phi) \\ &= \mathbb{E}\left[\|\hat{\epsilon} - \epsilon\|_2^2 + \lambda_{\text{EPK}}\, \|\hat{z}_0 - z_{\text{EPK}}(\hat{z}_k, a_k)\|_2^2\right]. \end{aligned} \tag{34}$$

The diffusion parameters θ and projection parameters ϕ are jointly optimized, while the pre-trained VAE remains frozen.

Summary Equation 34 combines conditional noise prediction with physiological anchoring of the derived clean latent. The first term learns denoising from data, while the second encourages agreement with the projected intervention-conditioned prior.

C.5 Inference-Time Behavior under Physiological Regularization

Although physiological regularization embeds domain knowledge during training, the diffusion model remains stochastic at inference time.

In this work, we do not perform explicit energy-based resampling during inference. Instead, we follow the standard diffusion sampling procedure, where the next-step latent state is generated by iteratively denoising from Gaussian noise using the learned reverse process:

$$\hat{z}_{k+1} \sim p_\theta(\cdot \mid \hat{z}_k, a_k). \tag{35}$$

Implicit knowledge injection The physiological constraint introduced by the EPK module is incorporated implicitly through the energy-regularized training objective. Specifically, the learned model p_θ is biased toward generating samples that are close to the physiological anchor $z_{\mathrm{EPK}}(\hat{z}_k, a_k)$ in latent space.

Effective generative distribution As a result, the effective generative behavior can be interpreted as approximately sampling from a distribution that favors physiologically consistent states:

$$p_{\theta,\phi}(z_{k+1} \mid \hat{z}_k, a_k) \approx p_\theta(z_{k+1} \mid \hat{z}_k, a_k), \tag{36}$$

where the influence of the physiological prior is already absorbed into the learned parameters θ during training.

Closed-loop generative operator We denote the resulting generative process by

$$\hat{z}_{k+1} \sim \mathcal{G}_\theta(\cdot \mid \hat{z}_k, a_k), \tag{37}$$

where $\mathcal{G}_\theta$ represents the diffusion model trained with physiological regularization.

Remark Although explicit energy-based reweighting or resampling could further enforce physiological constraints at inference time, we find that incorporating the EPK prior during training is sufficient to produce stable and physiologically plausible trajectories in practice.

C.6 Multi-Step Prediction Error Analysis

C.6.1 Error Definition and Recurrence

Define the prediction error at step k as $\delta_k = \hat{z}_k - z_k$. Using Equation 14 and the learned generative operator, we obtain

$$\delta_{k+1} = \mathcal{G}_\theta(\hat{z}_k, a_k) - \mathcal{T}(z_k, a_k) - \eta_k, \tag{38}$$

Taking conditional expectation and using $\mathbb{E}[\eta_k] = 0$ and $\mathbb{E}[\|\eta_k\|_2^2] = \sigma_\eta^2 d$, we get

$$\mathbb{E}\big[\|\delta_{k+1}\|_2^2\big] = \mathbb{E}\Big[\big\|\mathcal{G}_\theta(\hat{z}_k, a_k) - \mathcal{T}(z_k, a_k)\big\|_2^2\Big] + \sigma_\eta^2 d. \tag{39}$$

C.6.2 Contraction Assumption

Assumption 1 (Contractive closed-loop dynamics)**.** *There exist $\kappa \in (0,1)$ and $\Delta_{\mathrm{sys}} \geq 0$ such that for any $z, \hat{z} \in \mathbb{R}^d$ and any action $a \in \mathcal{A}$,*

$$\mathbb{E}\Big[\big\|\mathcal{G}_\theta(\hat{z}, a) - \mathcal{T}(z, a)\big\|_2^2\Big] \leq \kappa \|\hat{z} - z\|_2^2 + \Delta_{\mathrm{sys}}. \tag{40}$$

Scope of the contraction assumption. The condition $\kappa < 1$ is an assumption of the error analysis, rather than a property established by the training objective. Our short-chain rollout experiments show limited additional terminal error under the evaluated conditions, but do not verify global contraction or guarantee long-horizon stability.

Extension to Waveform Space Crucially, utilizing the bi-Lipschitz assumption established in Equation 15, the expected error in the physical continuous waveform space is strictly bounded by the latent space error:

$$\mathbb{E}\Big[\big\|\hat{x}_{k+1} - x_{k+1}\big\|_2^2\Big] \leq C_2^2\, \mathbb{E}\big[\|\delta_{k+1}\|_2^2\big]\,. \tag{41}$$

This formalizes that an EPK-induced contractive latent dynamic ($\kappa < 1$) strictly prevents compounding divergence in the macroscopic physiological signal space.

C.7 Algorithm Pseudocodes

Both algorithms retain EPK conditioning. We include e_{EPK} as a conditioning argument, with its embedding transformation and other clinical conditioning inputs omitted for brevity.

Algorithm 1 Training Algorithm for ECG World Model

Require: Dataset $\mathcal{D}$, frozen VAE $(f_{\mathrm{enc}}, f_{\mathrm{dec}})$, EPK module, coefficient λ_{EPK}, diffusion schedule $\{\beta_\tau\}_{\tau=1}^{N}$

Initialize diffusion parameters θ and EPK projection parameters ϕ
repeat
  Sample an MDP transition $(z_k, z_{k+1}^{(0)})$ and action a_k from $\mathcal{D}$
  Set $\hat{z}_k \leftarrow z_k$ {Teacher forcing}
  Construct the physiological sequence: $e_{\mathrm{EPK}} \leftarrow e_{\mathrm{EPK}}(\hat{z}_k, a_k)$
  Sample $\tau \sim \mathrm{Uniform}\{1, \ldots, N\}$ and $\epsilon \sim \mathcal{N}(0, I_d)$
  Construct the noisy latent:

$$z_{k+1}^{(\tau)} = \sqrt{\bar{\alpha}_\tau}\, z_{k+1}^{(0)} + \sqrt{1-\bar{\alpha}_\tau}\, \epsilon$$

  Predict the injected noise:

$$\hat{\epsilon} = \epsilon_\theta\Big(z_{k+1}^{(\tau)}, \tau, \hat{z}_k, a_k, e_{\mathrm{EPK}}\Big)$$

  Compute the predicted clean latent:

$$\hat{z}_0 = \frac{z_{k+1}^{(\tau)} - \sqrt{1-\bar{\alpha}_\tau}\, \hat{\epsilon}}{\sqrt{\bar{\alpha}_\tau}}$$

  Compute the physiological anchor: $z_{\mathrm{EPK}} \leftarrow \Pi_\phi(e_{\mathrm{EPK}})$
  Compute the joint loss:

$$\mathcal{L} = \|\hat{\epsilon} - \epsilon\|_2^2 + \lambda_{\mathrm{EPK}} \|\hat{z}_0 - z_{\mathrm{EPK}}\|_2^2$$

  Update θ and ϕ to minimize $\mathcal{L}$, keeping the VAE frozen
until the training budget is reached

Algorithm 2 Inference Algorithm for ECG World Model

Require: Initial state z_0, action sequence $a_0, \ldots, a_{L-1}$, trained model parameters θ, frozen VAE $(f_{\mathrm{enc}}, f_{\mathrm{dec}})$, EPK module, diffusion schedule $\{\beta_\tau\}_{\tau=1}^{N}$

$\hat{z}_0 \leftarrow z_0$
for $k = 0$ to $L-1$ **do**
  Construct the physiological sequence: $e_{\mathrm{EPK}} \leftarrow e_{\mathrm{EPK}}(\hat{z}_k, a_k)$
  Sample $z_{k+1}^{(N)} \sim \mathcal{N}(0, I_d)$
  for $\tau = N$ down to 1 **do**
    Sample using the standard DDPM reverse transition:

$$z_{k+1}^{(\tau-1)} \sim p_\theta\Big(\cdot \mid z_{k+1}^{(\tau)}, \tau, \hat{z}_k, a_k, e_{\mathrm{EPK}}\Big)$$

  end for
  Set $\hat{z}_{k+1} \leftarrow z_{k+1}^{(0)}$
end for
Output: $\hat{z}_1, \ldots, \hat{z}_L$

The training loss uses a fixed physiological coefficient without additional timestep-dependent weighting. The projection Π_ϕ consists of temporal resampling and a learnable 1×1 convolution. Default inference uses standard DDPM sampling ($K_{\mathrm{res}} = 1$), without explicit energy-gradient correction or energy-based resampling. Physiological information is supplied through EPK conditioning and the parameters learned under the regularized training objective.

D Detailed Proofs for Energy-Regularized Diffusion

In this section, we provide the complete mathematical proofs for the theoretical claims.

Let $p_\theta(z)$ denote the base empirical data distribution learned by the diffusion model, and $E(z)$ denote the parameterized energy function derived from the physiological ODE. Our target energy-regularized distribution is defined as:

$$p^*(z) = \frac{1}{Z} p_\theta(z) \exp(-\gamma E(z)), \tag{42}$$

where $Z = \int p_\theta(z) \exp(-\gamma E(z)) dz$ is the intractable partition function, and $\gamma > 0$ is the temperature hyperparameter.

D.1 Proof of the Variational Perspective (Gibbs Posterior)

Proposition 1. The target distribution $p^*(z)$ is the unique minimizer of the following variational free-energy objective:

$$\mathcal{J}(q) = \mathrm{KL}(q \| p_\theta) + \gamma \mathbb{E}_q[E(z)]. \tag{43}$$

Proof. We can expand the objective functional $\mathcal{J}(q)$ using the definitions of Kullback-Leibler (KL) divergence and expectation:

$$\begin{aligned} \mathcal{J}(q) &= \int q(z) \log \frac{q(z)}{p_\theta(z)} dz + \int q(z) \big(\gamma E(z)\big) dz \\ &= \int q(z) \left(\log q(z) - \log p_\theta(z) + \gamma E(z) \right) dz. \end{aligned} \tag{44}$$

By adding and subtracting the log partition function $\log Z$, we can rewrite the terms inside the integral:

$$\mathcal{J}(q) = \int q(z) \left(\log q(z) - (\log p_\theta(z) - \gamma E(z) - \log Z) - \log Z \right) dz. \tag{45}$$

Notice that from Equation 42, we have $\log p^*(z) = \log p_\theta(z) - \gamma E(z) - \log Z$. Substituting this back into the equation yields:

$$\begin{aligned} \mathcal{J}(q) &= \int q(z) \left(\log q(z) - \log p^*(z) \right) dz - \log Z \int q(z) dz \\ &= \mathrm{KL}(q \| p^*) - \log Z. \end{aligned} \tag{46}$$

Since $\log Z$ is a constant independent of the distribution q, minimizing $\mathcal{J}(q)$ is strictly equivalent to minimizing $\mathrm{KL}(q \| p^*)$. According to Gibbs' inequality, the KL divergence is non-negative ($\mathrm{KL} \geq 0$) and achieves its unique global minimum of zero if and only if $q(z) = p^*(z)$ almost everywhere. Therefore, $p^* = \arg\min_q \mathcal{J}(q)$. ■

D.2 Proof of Score Function Equivalence

Proposition 2. The score function of the energy-regularized distribution $p^*(z)$ is given by:

$$\nabla_z \log p^*(z) = \nabla_z \log p_\theta(z) - \gamma \nabla_z E(z). \tag{47}$$

Proof. Starting from the definition of the target distribution in Equation 42, we take the natural logarithm of both sides:

$$\log p^*(z) = \log p_\theta(z) - \gamma E(z) - \log Z. \tag{48}$$

Next, we apply the gradient operator ∇_z with respect to the latent variable z. Since the partition function Z is a constant scalar with respect to z, its gradient is zero ($\nabla_z \log Z = 0$). Thus, we immediately obtain:

$$\nabla_z \log p^*(z) = \nabla_z \log p_\theta(z) - \gamma \nabla_z E(z). \tag{49}$$

This confirms that incorporating the ODE energy constraint during generation requires only a straightforward additive modification to the score field learned by the base diffusion model. ■

D.3 Proof of Stationary Distribution via the Fokker-Planck Equation

Proposition 3. Consider the modified Langevin sampling dynamics governed by the Stochastic Differential Equation (SDE):

$$dz_t = \underbrace{\left(\nabla_z \log p_\theta(z_t) - \gamma \nabla_z E(z_t)\right)}_{\text{drift term } f(z_t)} dt + \sqrt{2}\, dW_t, \tag{50}$$

where W_t is a standard Wiener process. *(Note: Here, t denotes the continuous integration time of the Langevin process, which is distinct from the MDP sequence step k and the discrete diffusion step τ.)* The target distribution $p^*(z)$ is a valid stationary distribution of this process.

Proof. Let $p(z,t)$ denote the probability density of the state z at time t. The time evolution of $p(z,t)$ under the SDE in Equation 50 is described by the Fokker-Planck equation (Kolmogorov forward equation):

$$\frac{\partial p(z,t)}{\partial t} = -\nabla_z \cdot \big(f(z)p(z,t)\big) + \nabla_z \cdot \big(\nabla_z p(z,t)\big). \tag{51}$$

We can express the right-hand side in terms of the probability flux $J(z,t)$:

$$\frac{\partial p(z,t)}{\partial t} = -\nabla_z \cdot J(z,t), \quad \text{where} \quad J(z,t) = f(z)p(z,t) - \nabla_z p(z,t). \tag{52}$$

A distribution $p(z)$ is a stationary distribution if the probability density does not change over time, i.e., $\frac{\partial p(z)}{\partial t} = 0$. A sufficient condition for this is that the probability flux vanishes everywhere: $J(z) = 0$.

We now evaluate the probability flux at our target distribution $p^*(z)$. Using the result from Proposition 2, we can rewrite the drift term $f(z)$ as:

$$f(z) = \nabla_z \log p_\theta(z) - \gamma \nabla_z E(z) = \nabla_z \log p^*(z). \tag{53}$$

Substituting $p^*(z)$ and $f(z) = \nabla_z \log p^*(z)$ into the expression for the probability flux:

$$\begin{aligned} J^*(z) &= f(z)p^*(z) - \nabla_z p^*(z) \\ &= \big(\nabla_z \log p^*(z)\big)p^*(z) - \nabla_z p^*(z). \end{aligned} \tag{54}$$

By applying the standard calculus identity for the derivative of a logarithm, $\nabla_z \log p^*(z) = \frac{\nabla_z p^*(z)}{p^*(z)}$, we have $\big(\nabla_z \log p^*(z)\big)p^*(z) = \nabla_z p^*(z)$. Therefore:

$$J^*(z) = \nabla_z p^*(z) - \nabla_z p^*(z) = 0. \tag{55}$$

Since the probability flux $J^*(z)$ is identically zero for all z, the divergence $\nabla_z \cdot J^*(z)$ is also zero. Consequently, $\frac{\partial p^*}{\partial t} = 0$.

This establishes $p^*(z)$ as a valid stationary target distribution for the modified Langevin dynamics. Note that rigorously proving unique asymptotic global convergence from an arbitrary initialization requires standard additional structural assumptions on the energy landscape (e.g., ergodicity, log-Sobolev inequalities, or strict positivity), which we assume hold practically within the continuous data support. ■

E Additional Evaluation of Physiological Fidelity and Predictive Consistency

We complement the main experiments with analyses of pharmacological mapping, twelve-lead interval fidelity, diagnostic risk changes, autoregressive rollout, and inference-time energy sampling. These evaluations distinguish the effect of the drug-to-ODE mapping from removal of the EPK architecture, assess interval values beyond normality agreement, and examine whether generated trajectories preserve observed diagnostic changes and remain stable during recursive prediction.

E.1 Millisecond-Level Interval Fidelity against Official DiffuSETS

To assess numerical fidelity beyond normality categories, we compare the full EPK world model with the official DiffuSETS `unet_all` checkpoint on the same 1,822 evaluation pairs. For an observed interval u and generated interval $\hat{u}$, value accuracy requires

$$|\hat{u} - u| \leq \max(25\,\text{ms},\, 0.05|u|)\,. \tag{56}$$

The DiffuSETS intervals are re-scored from saved outputs without regenerating the ECGs. The full EPK results aggregate three training seeds: 17, 42, and 2026.

Table 9: Interval-value accuracy (%) and signal MSE on 1,822 evaluation pairs. Value accuracy follows Equation 56 and is distinct from normality accuracy. Full EPK results summarize three training seeds.

Method	QTc ↑	PR ↑	TpTe ↑	Signal MSE ↓
DiffuSETS official	17.7	44.2	66.5	0.087
Ours, full EPK	$\mathbf{55.8 \pm 0.2}$	$\mathbf{78.4 \pm 0.7}$	$\mathbf{93.4 \pm 0.3}$	$\mathbf{0.057 \pm 0.003}$

As shown in Table 9, Ours improves value accuracy over the official DiffuSETS checkpoint by 38.1, 34.2, and 26.9 percentage points for QTc, PR, and TpTe, respectively, while reducing signal MSE from 0.087 to 0.057. These gains demonstrate stronger numerical interval fidelity and waveform reconstruction for the complete EPK world model under this evaluation, beyond agreement on normality categories.

The distinction between the two accuracy definitions is material. On this split, which contains no reference QTc abnormalities, the same saved DiffuSETS outputs yield 88.53% QTc normality accuracy but 17.7% value accuracy. Likewise, recall of normal intervals does not measure abnormality sensitivity; PR balanced accuracy in the evaluated EPK runs is approximately 50%. We therefore report normality agreement and interval-value accuracy as different endpoints, without interpreting near-perfect normality recall as evidence of reliable abnormality detection.

E.2 EPK Architecture, Template Variation, and Cross-Lead Structure

Architecture-preserving controls. The EPK module uses 128-dimensional embeddings for cross-attention conditioning and a learned energy projection implemented as a 1×1 convolution from one input channel to four output channels, with weight shape $[4, 1, 1]$. Writing the pre-projection ODE sequence as e_{EPK}, the projected anchor is $z_{\text{EPK}} = \Pi_\phi(e_{\text{EPK}})$. Our architecture-preserving controls retain both EPK pathways and vary the drug-to-ODE mapping or the energy regularization weight λ_{EPK}. Disabling `use_epk` removes the projection and is therefore not equivalent to changing the mapping or setting only the energy weight to zero.

Regimen-dependent template variation. Across eight drug regimens, the pharmacological mapping produces nonzero variation in the pre-projection ODE templates, whereas the identity mapping produces the same template for every regimen (Table 10). The permutation control preserves the template set while changing its assignment to regimens, and therefore retains the same aggregate variance and pairwise distance as the pharmacological mapping. These diagnostics establish that the pharmacological map supplies regimen-dependent templates rather than a repeated identity template. We evaluate the pharmacological–identity difference separately through generated intervals; the permutation statistics here serve only as a template-variation control.

Projected anchors and multi-lead generation. The projected anchor has a mean cross-drug variance of 1.66×10^{-8}, indicating low cross-drug variability on the recorded scale; the separate EPK conditioning stream remains active. We therefore assess the practical benefit at the generated ECG level rather than inferring it from template diversity alone.

The one-dimensional McSharry prior supplies temporal structure, while the world-model UNet and VAE decoder produce lead-specific morphology. The complete world model achieves a cross-lead correlation-matrix Frobenius error of 2.69 relative to the observed ECG, compared with 6.39 when decoding the projected ODE template alone. This corresponds to a 57.9% lower error and shows that

Table 10: Pre-projection EPK template variation across eight drug regimens. Permutation preserves aggregate template variation by construction; these statistics characterize template diversity, not the correctness of drug–template assignments.

Mapping	Mean variance	Mean pairwise ℓ_2 distance
Pharmacological	0.004979	1.091
Identity	0	0
Permuted	0.004979	1.091

the learned generator captures cross-lead relationships more accurately than template-only decoding. This supports the role of the learned generator in complementing the temporal prior with multi-lead structure.

E.3 Twelve-Lead Interval Effects of Pharmacological Mapping

We compare the pharmacological and identity mappings under the `dadm_table1` evaluation protocol, using the respective `unet_best.pth` checkpoints from the `map_pharm_1` and `map_identity_1` runs. Both variants retain the EPK conditioning stream and energy projection. Accuracy in this diagnostic measures agreement between the normal/abnormal categories of generated and observed intervals, rather than millisecond matching. Table 11 reports all 12 leads and three interval metrics, with $\Delta = \mathrm{Acc}_{\mathrm{pharm}} - \mathrm{Acc}_{\mathrm{id}}$ expressed in percentage points.

Table 11: Twelve-lead interval normality accuracy (%) for identity (Id.) and pharmacological (Pharm.) EPK mappings. Positive Δ values favor pharmacological mapping and are highlighted in bold. Differences are taken from the original evaluation output rather than recomputed from the displayed rounded accuracies.

	QTc			PR			TpTe		
Lead	Id.	Pharm.	Δ	Id.	Pharm.	Δ	Id.	Pharm.	Δ
I	99.69	99.63	-0.06	97.08	96.02	-1.06	98.50	98.19	-0.31
II	99.69	99.63	-0.06	95.59	92.73	-2.86	98.94	98.07	-0.87
III	98.70	99.19	$\mathbf{+0.50}$	76.78	77.75	$\mathbf{+0.97}$	99.42	99.30	-0.13
aVR	98.27	98.03	-0.24	52.05	52.27	$\mathbf{+0.22}$	81.56	82.27	$\mathbf{+0.71}$
aVL	99.19	99.13	-0.06	94.79	93.48	-1.30	97.08	96.83	-0.25
aVF	99.44	99.25	-0.19	75.79	83.36	$\mathbf{+7.57}$	98.94	98.62	-0.31
V1	98.20	98.20	0.00	72.86	69.42	-3.44	55.29	56.45	$\mathbf{+1.16}$
V2	97.95	97.95	0.00	70.02	70.40	$\mathbf{+0.38}$	79.59	82.17	$\mathbf{+2.58}$
V3	98.13	98.07	-0.07	98.20	98.57	$\mathbf{+0.37}$	82.02	85.21	$\mathbf{+3.19}$
V4	99.50	99.31	-0.19	95.65	94.64	-1.01	92.48	88.65	-3.83
V5	99.32	99.13	-0.19	95.03	93.85	-1.18	97.33	95.21	-2.11
V6	99.69	99.50	-0.19	93.54	91.80	-1.74	99.07	98.32	-0.75

The largest improvement occurs for PR in lead aVF, where accuracy increases from 75.79% to 83.36% (+7.57 percentage points). Additional gains appear for QTc and PR in lead III, TpTe in aVR and V1, and PR and TpTe in V2/V3. In particular, TpTe accuracy improves by 2.58 and 3.19 percentage points in V2 and V3, respectively. The mapping effects are heterogeneous: TpTe accuracy decreases by 3.83 and 2.11 percentage points in V4 and V5, and several other endpoints favor identity. Thus, this comparison identifies specific lead–interval benefits from pharmacological modulation rather than uniform twelve-lead superiority. QTc accuracy near 99% on most leads reflects normality agreement under a low abnormal rate; numerical fidelity is assessed separately using Equation 56.

E.4 Agreement with Observed Post-Dose Diagnostic Changes

This evaluation examines whether generated ECGs preserve the magnitude and direction of diagnostic-score changes observed after medication administration, complementing the waveform and interval evaluations.

Table 12: Agreement between simulated and observed diagnostic-score changes on the same 200 test cases across six stochastic sampling runs. Each run regenerates five DDPM samples per recorded medication. Recorded λ values apply only to action selection and do not enter these metrics. The final row reports the mean and sample standard deviation across runs.

Run	Recorded λ	Pearson r	Spearman ρ	Sign agreement (%)
1	0.0	0.587	0.585	71.5
2	0.2	0.613	0.597	71.0
3	0.4	0.581	0.557	70.0
4	0.6	0.620	0.598	75.5
5	0.8	0.562	0.545	69.5
6	1.0	0.591	0.580	70.5
Mean $\pm$ SD	–	0.592 ± 0.021	0.577 ± 0.022	71.3 ± 2.2

Evaluation protocol. We use the frozen `ecg_fm_diagnostic` backend described in Appendix B.10 to score both generated and observed ECGs. Given the predicted probabilities of 17 diagnostic labels, we define

$$R(x) = \frac{1}{17} \sum_{j=1}^{17} p_j(x). \tag{57}$$

For each case i and its recorded medication a_i^{true}, we generate $S = 5$ standard-DDPM samples with EPK conditioning and compute

$$\mu_{R,i}\big(a_i^{\text{true}}\big) = \frac{1}{S} \sum_{s=1}^{S} R\Big(f_{\text{dec}}\Big(\hat{z}_i^{(s)}(a_i^{\text{true}})\Big)\Big). \tag{58}$$

Predicted and observed diagnostic changes are measured relative to the same pre-dose ECG:

$$\Delta_{\text{wm},i} = \mu_{R,i}\big(a_i^{\text{true}}\big) - R(x_{\text{pre},i}), \tag{59}$$

$$\Delta_{\text{real},i} = R(x_{\text{post},i}) - R(x_{\text{pre},i}). \tag{60}$$

We assess their agreement using Pearson correlation, Spearman correlation, and sign agreement. The observed change is computed from the recorded post-dose ECG, providing a reference independent of the generated trajectories.

Consistency across repeated stochastic sampling. We evaluate the same `unet_best.pth` checkpoint from `ewm_world_8` on the same 200 test cases across six sampling runs. Each run independently regenerates its DDPM samples. Although the evaluation files record different action-selection coefficients λ, this coefficient does not enter the diagnostic-change calculation for a fixed recorded medication. The results therefore characterize sensitivity to stochastic sampling rather than the effect of λ.

As shown in Table 12, all six runs exhibit positive agreement with observed diagnostic changes: Pearson correlation ranges from 0.562 to 0.620, Spearman correlation from 0.545 to 0.598, and sign agreement from 69.5% to 75.5%. Across runs, the corresponding means are 0.592 ± 0.021, 0.577 ± 0.022, and $(71.3 \pm 2.2)\%$, respectively. The relatively small between-run variation indicates that the observed diagnostic-change agreement is maintained across repeated stochastic generation, rather than being confined to a single sampling run.

Relationship to the main comparison. Table 4 reports the correlation and sign agreement from Run 4, in which our model outperforms DADM on both correlation metrics and directional agreement. The six-run results provide additional evidence that our model consistently preserves diagnostic-change information under repeated sampling. These standard deviations quantify sampling-related variation for a fixed checkpoint and test cohort; they do not estimate training-seed variability or population-level confidence intervals.

Interpretation. Together with the waveform and interval results, these findings support the ability of the ECG World Model to preserve diagnostically relevant changes under recorded medication

conditions. Because predicted and observed changes share the pre-dose reference and use the same frozen diagnostic backend, their correlations should be interpreted as diagnostic-change fidelity within this evaluation protocol. They do not establish calibrated clinical risk, causal treatment effects, or the benefit of selecting an alternative medication.

E.5 Short-Chain Autoregressive Rollout Stability

For each subject–admission pair, identified by `subject_id` and `hadm_id`, we retain the longest recorded chain. Autoregressive execution recursively generates the preceding states. The oracle instead computes the final prediction using the true preceding latent. Let $\overline{E}_{\mathrm{roll}}$ and $\overline{E}_{\mathrm{oracle}}$ be the corresponding mean final-step latent MSEs. We report

$$\begin{aligned} \Delta_{\mathrm{roll}} &= \overline{E}_{\mathrm{roll}} - \overline{E}_{\mathrm{oracle}}, \\ \mathrm{Retention} &= 100\,\frac{\overline{E}_{\mathrm{oracle}}}{\overline{E}_{\mathrm{roll}}}. \end{aligned} \tag{61}$$

Retention is a ratio of mean errors, not a classification accuracy or a mean of per-chain error ratios.

Table 13: Final-step autoregressive rollout relative to oracle prediction. The 1,273-chain evaluations use `unet_best`; the separate 100-chain Ours evaluation uses `unet_500`. Qwen2.5-VL is an auxiliary untrained latent-prediction control, with parsing limitations discussed in the text.

Model	Checkpoint	Chains	Δ_{roll}	Retention (%)
Ours (`ewm8`)	`unet_best`	1,273	$+0.000882$	98.7
Ours (`ewm4`)	`unet_best`	1,273	-0.000171	100.3
Ours (`ewm8`)	`unet_500`	100	-0.000030	100.0
Qwen2.5-VL-7B	—	100	$+0.002576$	93.8

On 1,273 hospitalization chains, `ewm8` achieves 98.7% retention with an additional mean latent MSE of only 0.000882. The `ewm4` checkpoint has a slightly negative error difference of -0.000171 and 100.3% retention. Both results indicate a small average terminal-error difference between recursive and oracle execution in this evaluation. A negative difference is possible because the oracle supplies the true preceding input to the same learned predictor, rather than defining a mathematical lower bound on prediction error.

The evaluated horizons are short: for the 1,273-chain `ewm8` run, chain length in rows has median 2, mean 2.16, and maximum 6. Approximately 53% of chains have positive error differences, while the mean remains below 0.001 latent MSE. Accordingly, these results demonstrate limited additional terminal error over the observed hospitalization chains, not an absence of absolute prediction error or a guarantee for substantially longer horizons.

Auxiliary language-model control. Under the separate 100-chain protocol, `ewm8` retains 100.0% compared with 93.8% for Qwen2.5-VL-7B. However, Qwen is not trained on these ECG latents and incurs 296 parse-failure steps. Its 1,273-chain run has 3,768 parse failures, 99.1% retention, and a median error difference of zero. These failures limit interpretation of the language-model comparison; the primary stability evidence is the Ours rollout–oracle comparison, not superiority over a trained ECG-specific dynamics baseline.

E.6 Inference-Time Energy Sampling

Our default inference procedure uses standard DDPM sampling with EPK conditioning from the model trained with physiological energy regularization. To examine an additional inference-time use of this prior, we evaluate EPK energy resampling on 48 evaluation pairs, using the same pharmacological checkpoint and 200 denoising steps. The resampling parameter K_{res} is distinct from the sample counts used for reconstruction evaluation and diagnostic-score estimation.

As shown in Table 14, EPK energy resampling reduces latent MSE by approximately 9.4% and signal MSE by 8.3% relative to standard DDPM. These results demonstrate improved reconstruction under the evaluated resampling configuration. Because the comparison changes both the sampling

Table 14: Inference-time energy resampling on 48 evaluation pairs using the same pharmacological checkpoint and 200 denoising steps. Energy resampling is an optional inference procedure with additional sampling cost.

Sampler	Latent MSE ↓	Signal MSE ↓
Standard DDPM ($K_{\text{res}} = 1$)	0.0543	0.0676
EPK energy resampling ($K_{\text{res}} = 8$)	**0.0492**	**0.0620**

Table 15: Fixed candidate medications for the MIMIC-IV risk experiments. Counts denote occurrences in the full test split; IDs are internal medication identifiers in the processed data.

Medication	Test count	Medication ID
Propofol	1,045	7
Insulin - Regular	936	25
Heparin Sodium (Prophylaxis)	848	22
Furosemide (Lasix)	736	14
Norepinephrine	726	19

procedure and its computational budget, it does not isolate the benefit of energy-based selection at matched cost.

A separate direct energy-gradient sweep on the same 48 pairs yields only a small change in latent MSE, from 0.05023 to 0.05022 as the coefficient increases from 0 to 0.1. Thus, resampling provides the clearer empirical benefit among the inference-time configurations tested here. The energy-tilted distribution serves as a theoretical reference; neither the training regularizer nor this resampling pilot establishes exact sampling from that distribution. Standard DDPM remains the default inference procedure for the reported main experiments.

Summary. The additional evaluations support complementary aspects of physiology-informed ECG simulation: improved interval-value accuracy relative to the official DiffuSETS checkpoint, lead-specific benefits from pharmacological mapping, improved cross-lead structure relative to template-only decoding, agreement with observed diagnostic-score changes, and limited additional terminal error over the evaluated short hospitalization chains. The resampling pilot further suggests that additional inference-time computation can improve reconstruction. Together, these findings support the model's predictive fidelity across waveform, interval, diagnostic-change, and short-chain prediction evaluations.

E.7 Intervention Sets and Candidate Selection

Eight-regimen set for interval and EPK experiments. The controlled drug-response experiments use eight predefined regimen labels implemented in `dataset/ecg_medication_8class.py`: Ranolazine, Dofetilide, Verapamil, Quinidine, Placebo, Moxifloxacin+Diltiazem, Mexiletine+Dofetilide, and Lidocaine+Dofetilide. Labels 0–4 are drawn from `ecg_data`, and labels 5–7 from `ecg_data1`, corresponding to the controlled drug-response cohorts described in Appendix A. Combination regimens are encoded as distinct categorical actions. This intervention vocabulary is predefined and is not derived from MIMIC-IV medication frequencies.

Five-medication set for risk experiments. The risk experiments use a separate candidate set constructed from `ewm_world_vae_mortality.pt`. The function `compute_topk_test_only_drug_texts` counts medication names only among records with `split=test`. For each record, the name is taken from the label's `text` field, falling back to `drug_name` when `text` is absent. The procedure scans 21,328 test records containing 1,263 distinct medication-name strings and selects the five most frequent names using `most_common(5)`. Training records do not contribute to this ranking. Table 15 reports the resulting candidate set. Each name is mapped to the first matching `medication_id` found in the same data file.

Evaluation scope. Candidate ranking is restricted to this fixed five-medication set. Top-1 concordance measures whether the selected candidate matches the recorded medication, with a uniform

random-choice baseline of $1/5 = 0.20$ when the recorded medication belongs to the candidate set. The candidate vocabulary is determined from test-split medication frequencies, rather than specified independently of the test set. No diagnostic outcomes are used for candidate selection. Accordingly, this evaluation measures restricted-set medication concordance and does not establish treatment optimality or clinical safety.

F Impact Statements

This work proposes an ECG World Model for simulation-based analysis of cardiac interventions, enabling AI systems to explore potential physiological responses under different treatment scenarios. By combining a physiological prior with data-driven generative modeling, the framework supports hypothesis generation and in-silico evaluation of treatment effects.

Several limitations should be noted. First, the model is trained on observational and simulated data, which may not fully capture real-world complexity, variability, or rare edge cases, potentially leading to deviations from true clinical outcomes under distributional shifts. Second, the physiological prior is simplified and may be misspecified, limiting its ability to represent full cardiac dynamics, especially in pathological regimes. Third, evaluation focuses on aggregate stability metrics rather than clinical endpoints, and does not fully explain the interaction between prior-driven and data-driven components. Finally, the model lacks calibrated uncertainty estimates and formal safety guarantees, constraining its use in high-stakes settings.

Accordingly, the system should not be used as a standalone clinical tool, but rather as a decision-support framework requiring expert validation. Generated scenarios should be treated as exploratory hypotheses rather than definitive predictions.

Despite these limitations, the model exhibits stable and bounded behavior across varying levels of prior mismatch, with variability increasing smoothly rather than collapsing. The influence of the prior is also continuous and controllable, as partial prior injection yields intermediate performance. These results suggest a degree of robustness under imperfect assumptions.

Overall, this work takes a step toward more robust and interpretable generative models for physiological simulation, while underscoring the need for further validation, mechanistic understanding, and safety-aware design prior to real-world deployment.

G Devices

In the experiments, we conducted all methods on a local Linux server equipped with an AMD EPYC 7742 64-Core Processor (128 logical threads). All methods are implemented using the PyTorch framework, and all models are trained on NVIDIA A800-SXM4-80G GPUs (80GB HBM2e memory).

H AI Use Statement

In preparing this paper, we used large language models (LLMs) solely as an assistive tool for language polishing and writing improvements. The models were not involved in research ideation, experimental design, data analysis, or drawing scientific conclusions. The authors take full responsibility for the contents of this paper. For experimental evaluation, we additionally conducted testing using large language models such as GPT.

I Reproducibility Statement

We document the model formulation, physiological prior construction, training objective, and inference procedure, together with dataset sources, cohort construction, and subject-disjoint training and test splits. The reported experimental settings include the main optimization hyperparameters, physiological regularization settings, stochastic sampling configurations, frozen pretrained components, and computing environment. We also provide training and inference pseudocode, mathematical derivations with their underlying assumptions, and additional ablation and robustness evaluations.

Reproducing the experiments requires access to the corresponding ECG and clinical datasets, pre-trained components, and model services used in the comparisons. Evaluations should preserve the reported cohort definitions, subject-level split protocol, model configurations, and sampling settings. Source data should be obtained and used in accordance with the applicable access and usage conditions.

J ETHICS STATEMENT

This work investigates intervention-aware cardiac simulation using existing ECG and clinical datasets, including MIMIC-IV-ECG, MIMIC-IV-Clinical, ECGRDVQ, and ECGDMMLD. The out-of-distribution evaluation cohort is constructed from MIMIC-IV-ECG records. The proposed framework is a proof-of-concept for hypothesis generation and simulation-based analysis, rather than a standalone clinical tool or a substitute for professional clinical judgment. Access to and reuse of the source data should respect the corresponding data-use conditions and patient-confidentiality requirements.

Generated ECG trajectories and intervention comparisons should be interpreted as exploratory model predictions, not definitive evidence of treatment efficacy or safety. The downstream risk scores are proxies derived from ECG diagnostic predictions rather than directly observed clinical outcomes. Similarly, variability across stochastic simulations does not constitute a formal guarantee of clinical safety. Simplified physiological priors, discrete intervention representations, incomplete modeling of drug interactions, and distribution shifts may limit the reliability of simulated responses. Broader clinical validation and expert assessment are therefore required before any real-world application, and model outputs should not be used independently to prescribe medications or alter patient care.